ADVANCING HUMAN SOFT TISSUE PATHOLOGY ASSESSMENT USING ARTIFICIAL INTELLIGENCE IN MEDICAL IMAGING

Zakia Khatun
December 2024
Department of Engineering, School of Technology
Reykjavík University

Ph.D. Dissertation

Advancing Human Soft Tissue Pathology Assessment Using Artificial Intelligence in Medical Imaging

by

Zakia Khatun

Dissertation submitted to the Department of Engineering, School of Technology
at Reykjavík University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

December 2024

Thesis Committee:

Paolo Gargiulo, Supervisor
Professor, Reykjavík University, Iceland

Francesco Tortorella, Supervisor
Professor, Università degli Studi di Salerno, Italy

Halldór Jónsson Jr., Co-advisor
Professor, Chief Surgeon & Chairman, Landspitali University Hospital, Iceland

Cristiana Corsi, Examiner
Associate Professor, Università di Bologna, Italy

ISBN 978-9935-539-47-2 Electronic version
ISBN 978-9935-539-46-5 Print version

ORCID Zakia Khatun 0009-0007-7303-338X
https://orcid.org/0009-0007-7303-338X

The undersigned hereby certify that they recommend to the Department of Engineering, School of Technology, Reykjavík University, that this dissertation entitled **Advancing Human Soft Tissue Pathology Assessment Using Artificial Intelligence in Medical Imaging**, submitted by **Zakia Khatun**, be accepted as partial fulfilment of the requirements for the degree of **Doctor of Philosophy (Ph.D.) in Engineering**

08.12.2024

date

Paolo Gargiulo, Supervisor
Professor, Reykjavík University, Iceland

Francesco Tortorella, Supervisor
Professor, Università degli Studi di Salerno, Italy

Halldór Jónsson Jr., Co-advisor
Professor, Chief Surgeon & Chairman, Landspitali University Hospital, Iceland

Cristiana Corsi, Examiner
Associate Professor, Università di Bologna, Italy

08.12.2024

.....

date

.....

Zakia Khatun
Doctor of Philosophy

Advancing Human Soft Tissue Pathology Assessment Using Artificial Intelligence in Medical Imaging

Zakia Khatun

December 2024

Abstract

Artificial Intelligence (AI) has revolutionized various fields by automating complex tasks, uncovering patterns in large datasets, and making accurate predictions. Machine learning, particularly deep learning, plays a pivotal role in enabling AI to emulate human intelligence for tasks such as image segmentation, classification, and recognition. In the realm of medical imaging, modalities like Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Ultrasound are indispensable tools for diagnosing pathologies, detecting abnormalities, guiding treatment plans, and monitoring disease progression. The integration of AI with medical imaging offers the potential to enhance the accuracy and efficiency of these processes.

In particular, AI's application in the assessment of tendon-related conditions, such as tendinopathy, presents a promising avenue for improving patient care. Tendinopathy can significantly impact a patient's quality of life, and early detection is crucial to optimize treatment outcomes. This thesis focuses on the development of advanced AI-driven methods for analyzing human soft tissue pathologies, with a primary emphasis on tendon segmentation, pathology detection (classification), and tendon reflex response assessment. By automating the analysis of tendons and other human soft tissues, these methods aim to reduce human error and variability, thereby enabling more consistent and reliable clinical decisions. Ultimately, the goal is to support earlier, more accurate diagnoses and interventions, leading to better patient outcomes and more personalized treatment strategies.

This thesis begins with a study that analyzes MRI and CT scans from 47 participants to investigate the relationships between the tendons, cartilage, and muscles in the knee. This study has two primary objectives: first, to predict knee cartilage degeneration, and second, to predict patellar tendinopathy. For both objectives, predictions are made using features extracted solely from the patellar tendon and quadriceps, rather than directly from the cartilage itself. This approach explores the potential of using features from surrounding tissues as indirect predictors of knee-related pathologies. This study demonstrates that both knee cartilage degeneration and patellar tendinopathy can be predicted using these features from adjacent structures, highlighting the importance of surrounding tissues as potential indicators of pathology. Traditional machine learning models are employed to identify the most relevant features for each prediction task, highlighting their importance in the diagnosis of these conditions. This foundational research deepens our understanding of the interrelationships between knee soft tissues, contributing to more accurate diagnostic approaches in musculoskeletal health and enhancing clinical decision-making and treatment strategies.

A central focus of this thesis is the development of an end-to-end tendon segmentation module. This system integrates a superpixel-based coarse segmentation step that serves as a foundation for the final, more precise segmentation. In this approach, the segmentation task is framed as a superpixel classification problem. To achieve this, two distinct approaches are developed: (1) Random Forest (RF) and Support Vector Machine (SVM) classifiers for superpixel categorization, and (2) a Graph Convolutional Network (GCN) for transforming superpixels into graph structures for node classification. The RF and SVM classifiers demonstrate exceptional performance, achieving Area Under the Curve (AUC) scores of 0.992 and 0.987, respectively, with high sensitivity, indicating their effectiveness in accurately classifying superpixels. Although the GCN approach yields slightly lower performance, it showcases the potential of deep learning methods for improving segmentation by leveraging the structural relationships between superpixels. The findings suggest that both traditional machine learning and deep learning techniques offer promising avenues for advancing tendon segmentation, with superpixel-based methods offering a pathway to more reliable and automated segmentation in medical imaging.

Another key component of this thesis is the development of an end-to-end tendon pathology detection module, utilizing the same MRI dataset. This module adopts a graph-based approach, where superpixels are treated as nodes and connected by edge relationships. Each MRI scan is transformed into a graph, with the task framed as a graph classification problem to determine the presence or absence of pathology. To achieve this, a Graph Echo State Network (GESN) is employed. Known for its ability to efficiently represent data without the need for iterative backpropagation, the GESN leverages both temporal and structural dependencies in the data, enhancing classification performance. In this study, the GESN outperforms traditional machine learning models, achieving a mean accuracy of 0.953 and sensitivity of 0.943. These results underscore the potential of the GESN to significantly enhance diagnostic accuracy, offering a powerful tool for early detection and clinical decision-making in tendon pathology assessment. Moreover, the GESN's ability to handle complex, high-dimensional data suggests its broad applicability to other medical imaging tasks, further expanding its potential clinical utility.

The final study of this thesis explores the impact of demographic factors, including age, height, weight, and gender, on reflex response times in healthy individuals. This analysis is based on electromyography (EMG) recordings from 40 participants. The results reveal that elderly individuals, particularly those who are taller, heavier, and male, exhibit delayed reflex onsets. Even after normalizing for height, older participants still demonstrate slower reflex responses. These findings highlight the role of demographic factors in neuromuscular reflexes, aiding in the diagnosis and early detection of related disorders.

In conclusion, this research demonstrates the potential of AI, particularly superpixel-based and graph-based models, to advance tendon pathology assessment and exploratory tendon reflex studies, leading to better patient outcomes and musculoskeletal health management.

Keywords: Medical Imaging, Soft Tissue Diagnosis, Segmentation, Classification, Superpixel, Graph Neural Network (GNN)

Framfarir í mati á mjúkum vefjaskemmdum hjá mönnum með aðstoð gervigreindar í læknisfræðilegri myndgreiningu

Zakia Khatun

desember 2024

Útdráttur

Gervigreind (AI) hefur umbreytt mörgu með því að sjálfvirknivæða flókin verkefni, bera kennsl á sérkennandi mynstur í stórum gagnasöfnum og veita þannig nákvæmari spár. Vélanám, sérstaklega djúpnám, gerir AI kleift að líkja eftir mannlegri greind í verkefnum eins og myndasneiðingu, flokkun og kennslum stórra gagnasafna. Í læknisfræðilegum myndrannsóknum eru aðferðir eins og segulómun (MRI), sneiðmynd (CT) og ómun (US) lykilatriði við greiningu á vandamálum í stoðkerfinu, uppgötvun frávika, leiðbeiningum um meðferð og eftirlit með framvindu sjúkdóma og heilsu almennt. Samþætting AI við læknisfræðilegar myndrannsóknir getur bætt mat og meðferð á sinatengdum vandamálum eins og margvíslegum sinasjúkdómum, sem geta haft veruleg áhrif á lífsgæði. Snemmbær greining er því mikilvæg til að hámarka árangur meðferðar. Þess vegna er ritgerðarmarkmið mitt að þróa hátækni aðferðir til að greina mjúkvefjasjúkdóma, sérstaklega skiptingu sina, flokkun vandamála og meta mat á sinaviðbrögðum. Með því að sjálfvirknivæða greiningu á mjúkvefjavandamálum miða þessar aðferðir að því að draga úr mannlegum mistökum og breytileika, sem styður við áreiðanlegri klínískar ákvarðanir og snemmbærari inngripum.

Fyrsta rannsóknin greinir MRI- og CT- myndir 47 einstaklinga til að kanna tengsl milli sina, brjósks og vöðva. Rannsóknin sýnir að hægt er að spá fyrir um hrörnun í liðbrjóski í hnjám með því að nota eiginleika einungis úr Fjórhöfða læris og hnéskeljarsin. Hefðbundin vélanámslíkön eru notuð til að bera kennsl á lykileiginleika sem tengjast bæði spám um sinasjúkdóma og brjóskhrörnun. Þessi grunnrannsókn eykur skilning á samböndum mjúkvefja, sem stuðlar að betri greiningaraðferðum í stoðkerfisheilsu.

Annað lykilatriði þessarar ritgerðar er þróun kerfi fyrir sneiðingu sina. Kerfið inniheldur grófa sneiðingu byggða á ofurdílum til að auðvelda loka sneiðinguna. Tvær aðferðir eru bornar saman: (1) slembiskógur (RF) og stuðningsvigurvél (SVM) fyrir flokkun ofurdíla og (2) vafið tauganet (GCN) til að breyta ofurdílum í samsafn línurita. RF og SVM- aðferðirnar ná svæðum undir línuferli (AUC) með gildum upp á 0,992 og 0,987, í sömu röð og með mikilli næmni. GCN- aðferðin, þótt hún sé aðeins minna áhrifarík, sýnir möguleika ofurdíla til að bæta sneiðinguna.

Þriðja rannsóknin einbeitir sér að þróun á kerfi fyrir flokkun sinasjúkdóma með notkun sama MRI- gagnasafns. Línuritalíkan, þar sem ofurdílar virka sem hnútar, er útfært með endurómunartauganeti (GESN) fyrir flokkun. GESN, sem lýsir gögnum á skilvirkan hátt án endurtekinna bakútreikninga, skarar fram úr hefðbundnum vélanámslíkönum og nær meðalnákvæmni upp á 0,953 og næmni upp á 0,943, sem undirstrikar möguleika þess til að bæta greiningarnákvæmnina.

Síðasta rannsóknin skoðar hvernig aldur, hæð, þyngd og kyn hafa áhrif á viðbragðstíma hjá heilbrigðum einstaklingum með notkun á rafvöðvariti (EMG) frá 40 einstaklingum. Niðurstöður sýna að eldri einstaklingar, sérstaklega þeir sem eru hærri, þyngri og karlkyns, hafa seinkun á viðbragðstíma. Jafnvel með hæðarjöfnun sýna eldri þátttakendur hægari viðbrögð. Yngri þátttakendur hafa lengri heildarviðbragðstíma, líklega vegna hæðar, mynstur sem er samræmt á milli kynja. Þessar niðurstöður leggja áherslu á áhrif lýðfræðilegra þátta á taugavöðvaviðbrögðum og geta hjálpað við greiningu á tauga-vöðvasjúkdómum.

Að lokum sýnir þessi ritgerð möguleika AI, sérstaklega ofurdíla og línuritsbyggðra líkana, til að auka greiningu á sinasjúkdómum og kanna sinaviðbrögð, sem leiðir til betri útkomu fyrir sjúklinga og betri stjórnun á stoðkerfisheilsu.

Efnisorð: Læknisfræðileg myndgreining, Greining mjúkvefja, Sneiðing, Flokkun, Ofurdílar, Graf tauganet (GNN)

"To my perseverance, guided by God, supported by my family, uplifted by my guardian angel Kespu, and to Kristin and Paola for embracing me as one of their own."

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Paolo Gargiulo and Prof. Francesco Tortorella, for their support throughout my PhD journey. Prof. Paolo's guidance has been pivotal, offering not only academic supervision but also invaluable insights and encouragement that have significantly influenced both my professional and personal development. I am deeply grateful for his mentorship and the trust he has placed in me. Similarly, the profound wisdom of Prof. Tortorella and his expert guidance have been crucial in shaping my research, leaving a lasting impact on my work. I am deeply grateful to Dr. Halldór Jónsson Jr., a clinician of remarkable humility and integrity, whose support has been indispensable throughout my journey. His expertise and guidance have been instrumental in shaping the direction of my thesis, and without him, my work would have taken a different course. I would also like to extend my sincere thanks to Ms. Maria Tsirilaki, whose clinical expertise and support as a radiologist played a crucial role in advancing my research. I am also grateful to Prof. Þórður Helgason, Sara, and Jón for their invaluable support during my company secondment period. I am deeply appreciative of Prof. Hélder Almeida Santos and Ms. Pauliina, who served as exceptional coordinators. Their support, kindness, care, and unwavering dedication went far beyond their roles, ensuring I was always informed of my rights and advocating for me whenever necessary. Additionally, I am grateful to Prof. Arnau Oliver, Prof. Saniat Rahman Zishan, and Prof. Chowdhury Akram Hossain for their insightful contributions, which have significantly influenced both my academic journey and professional development.

On a personal note, I am deeply indebted to my family for their unwavering support, which has been instrumental in helping me reach this milestone. My beloved mother, Selina Parveen, and my brilliant father, A B M Saifulla, have been my pillars of strength, offering me boundless love and encouragement, and making countless sacrifices along the way. I am also incredibly grateful to my sisters, Jannat and Jolly, who have always cheered me on and stood by me, and to my baby brother, Ahiyan, for his constant love. Without their love and support, I couldn't have made it this far. And of course, my dear husband, Tushar, whose patience, support, and sacrifices have been beyond words. His love, and belief in me have been a source of strength throughout this journey. Thanks to my in-laws as well. Finally, I want to acknowledge Kespu, whose presence in my life has brought immense joy and happiness. Truly an angel, Kespu has been a source of peace, making some of my most cherished moments possible.

I am deeply grateful to my dear friend Kristin Eyjolfsdottir for always being there for me, offering her love and companionship at every turn. Embracing me as family, Kristin has made my journey more enjoyable. From our first mountain hike to witnessing the Northern Lights, Iceland has truly felt like home. I am also very grateful to Magnús Bergmann Magnússon, for his care and generosity, which have greatly en-

riched my experience. A special thanks to dear Paola Cirillo for her warmth and love, which have had a profound impact on my life. The welcoming atmosphere and the many shared moments, from wonderful meals to thoughtful gestures, will always be cherished. The laughter, celebrations, and goodbyes will remain unforgettable. My gratitude also extends to Nonna, Zia Mariella, Lele, Monia, Pietro, and Tip for their kindness and for making me feel like part of the family. These cherished memories will always be carried with me.

I would also like to thank my friends and colleagues for the countless enjoyable moments, conversations, and support throughout this journey. Special thanks to Alif, Shimul, Sohan, Sergio, Jhon, Esteban, Nadia, Maëlys, Lara, Simona, Vittorio, Yuri, Daniela, Seyma, Viktoría, Romain, Marco, Umberto, Stefano, Halldór, Gianmarco, Riccardo, Damiano, Carmine, Gabri, Arnar, Eda, Lorena, Federica, Mattia, and everyone else I've had the pleasure of sharing this experience with. A special mention goes to the administrative staff, particularly Ms. Verity, Ms. Alisha, and Ms. Sigrún, for their exceptional support.

Finally, I would like to express my sincere gratitude to the funder of this project. This work was funded by EU H2020-MSCA-ITN-EJD-2020, Grant Agreement ID: 955685, Project Name: Perspectives For Future Innovation in Tendon Repair (P4-FIT).

Preface

This thesis presents my original research on the assessment of human soft tissue pathologies through the application of artificial intelligence (AI) to medical imaging. The significance of this work lies in the development of automated systems that enhance the understanding of soft tissue-related pathologies. By minimizing human error and variability, these systems aim to support more reliable clinical decision-making and enable earlier interventions, ultimately improving patient outcomes.

My motivation for pursuing this research topic stems from both my academic expertise and the urgent societal and economic healthcare demands associated with the increasing incidence of tendinopathy. This research is funded by the EU H2020-MSCA-ITN-EJD-2020, Grant Agreement ID: 955685, with the project titled *Perspectives for Future Innovation in Tendon Repair* (P4-FIT). As one of the 15 Early Stage Researchers involved in this project, I have been dedicated to advancing technologies for the study of human soft tissue pathologies through medical imaging and AI. Within the larger framework of the COMPUT P4 FIT work package of this project, I contributed to the vision of creating innovative computational AI approaches for medicine.

This thesis is structured into the following chapters:

1. **Introduction**: An overview of the research background, the significance of the field, key challenges, study objectives, literature review, critiques of existing works, and the motivational theories that have guided the development of the systems.

2. **Tendon, Cartilage, and Muscle Interactions**: Exploring the mechanisms of tendinopathy and cartilage degeneration to better understand the interactions between tendon, cartilage, and muscle.

3. **Medical Image Segmentation**: Developing fully automated AI-driven methods for medical image segmentation, specifically designed to accurately delineate the Achilles tendon, thereby improving the analysis and diagnosis of tendon-related pathologies.

4. **Medical Image Classification**: Developing a fully automated AI-based system for medical image classification, focusing on the detection of tendon-related pathologies to improve diagnostic accuracy and support clinical decision-making.

5. **Tendon Reflex Responses**: Analyzing patellar tendon reflex responses with EMG to evaluate neuromuscular function. The study explores the influence of factors such as age, height, weight, and gender on reflex response times in healthy individuals, providing insights for future research on neuromuscular conditions.

6. **Conclusion**: This chapter provides an overall conclusion to the entire thesis.

This thesis aims to enhance our understanding of human soft tissue pathologies, facilitating the development of automated methods for their identification and assessment. By advancing diagnostic capabilities, it seeks to inspire further research and innovation in the field, ultimately improving clinical practices and patient outcomes.

Publications

The following two articles have already been published: the first in the "IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering" and the second one in a Q1 ranked journal named "Computer Methods and Programs in Biomedicine" (Impact Factor: 4.9, CiteScore: 12.3).

[1] Z. Khatun, M. Tsirilaki, A. Lindemann, et al., "The role of muscle and tendon in predicting cartilage degeneration and tendinopathy", in 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 2022, pp. 289–294. doi: 10.1109/MetroXRAINE54828.2022.9967653.

[2] Z. Khatun, H. Jónsson jr., M. Tsirilaki, et al., "Beyond pixel: Superpixel-based mri segmentation through traditional machine learning and graph convolutional network", Computer Methods and Programs in Biomedicine, vol. 256, p. 108 398, 2024, issn: 0169-2607. doi: https://doi.org/10.1016/j.cmpb.2024.108398. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260724003912.

The following article is being prepared for submission to the Q1 ranked journal "Medical Image Analysis" (Impact Factor: 10.7, CiteScore: 22.1).

[3] Z. Khatun, H. Jónsson jr., F. Tortorella, et al., "Graph Echo State Networks for MRI-based Tendon Pathology Classification", Medical Image Analysis.

The following article has been accepted for publication in the "Frontiers in Neurology" journal (Impact Factor: 2.7, CiteScore: 4.9), under the "Neuromuscular Disorders and Peripheral Neuropathies" section.

[4] Z. Khatun, S. Kristinsdóttir, A. Thórisdóttir, et. al., "Assessing Neuromuscular System via Patellar Tendon Reflex Analysis Using EMG in Healthy Individuals", Frontiers in Neurology.

Contents

List of Abbreviations

CT	Computed Tomography
MRI	Magnetic Resonance Imaging
US	Ultrasound
ROI	Region of Interest
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
RF	Random Forest
SVM	Support Vector Machines
LDA	Linear Discriminant Analysis
XGBoost	eXtreme Gradient Boosting
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
GNN	Graph Neural Network
GCN	Graph Convolutional Network
RC	Reservoir Computing
GESN	Graph Echo State Network
EMG	Electromyography
sEMG	Surface Electromyography
PTR	Patellar Tendon Reflex

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

1.1.1 Basic Facts

Clinical Insights

Tendons are tough, inelastic bands of connective tissues composed of a dense, highly organized matrix of parallel collagen fibers maintained by resident tenocytes. They connect skeletal muscles to bones and other structures, enabling the transmission and absorption of force. While all tendons share these fundamental characteristics, they vary significantly in morphological, molecular, and mechanical properties depending on their specific functional roles [3].

Anatomy, Structure, and Function of Achilles Tendon

- Anatomy

The Achilles tendon, named after the Greek mythological character Achilles, is also referred to as the calcaneal tendon due to its bony attachment to the calcaneus, or heel bone. As the largest and strongest tendon in the human body, it links the posterior three-headed calf muscle, the triceps surae (the lateral and medial gastrocnemii, and the soleus muscles) to the heel bone [4].

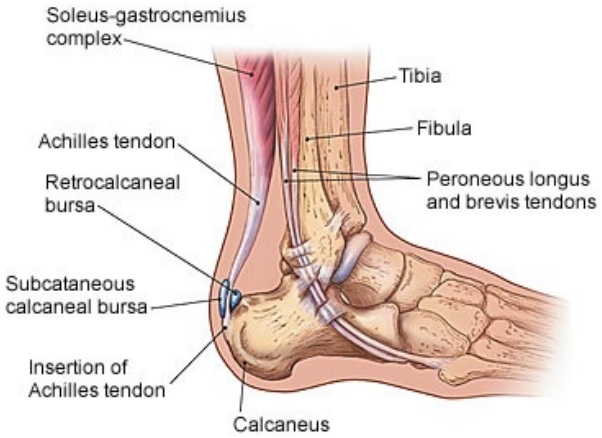


Figure 1.1: Basic Anatomy of Achilles tendon [5]

On its path towards the heel, the Achilles tendon twists medially by about 90 degrees, inserting into the mid-portion of the posterior surface of the calcaneus. The

superficial gastrocnemius muscle originates from the posterior aspect of the medial and lateral condyles of the distal femur. Beneath it lies the deeper soleus muscle, which arises from the posterior surface of the fibula and the medial border of the tibia. Together, these muscles form the superficial posterior compartment of the leg.

- Structure

The Achilles tendon is predominantly composed of type II fast-twitch fibers, providing the elasticity required for rapid propulsion. Structurally, it consists primarily of type I collagen, which imparts strength, and elastin, contributing to its flexibility. The tendon fibers are arranged in a spiral pattern, allowing for concentrated stress distribution and potentially enhancing mechanical efficiency. Unlike tendons with true synovial sheaths, the Achilles tendon is surrounded by a paratenon—a loose connective tissue sheath that encases the tendon. The paratenon can stretch up to 2–3 cm during movement, facilitating optimal gliding action [6].

- Function

The contraction of the gastrocnemius and soleus muscles generates force that is transmitted through the Achilles tendon, resulting in plantar flexion of the foot. This movement facilitates lower limb locomotion and propulsion in activities such as walking, running, and jumping. The Achilles tendon endures some of the highest loads in the human body, with tensile forces reaching up to ten times the body's weight during these actions. Its unique characteristics enable these functions by efficiently transmitting muscle-generated force to the heel, allowing for plantar flexion (pointing the foot downward) and playing a crucial role in ambulation and mobility [7].

Pathologies

Tendons can undergo degenerative and traumatic processes. It is possible that there may be several pathways that can result in tendon degeneration. The following mechanisms have all been proposed as initiators of tendinopathy: hypoxia, ischemic damage, oxidative stress, hyperthermia, impaired apoptosis, inflammatory mediators, fluoroquinolones, and matrix metalloproteinase imbalance [8] [9] [10]. Tendinopathy can be viewed as a failure of the cell matrix to adapt to a variety of stresses as a result of an imbalance between matrix degeneration and synthesis [11]. The vulnerable tendons are those of the rotator cuff, the long head of the biceps brachii, the wrist extensors and flexors, the adductors, the posterior tibial tendon, the patellar tendon, and the Achilles tendon. At least 50% of tendon problems are secondary to overload. In particular, Achilles tendon injuries are common in football, tennis, badminton, and jumping and have a prevalence in running athletes of 11%, but 1/3 of patients with this pathology do not practice intense physical activity. Pathology of the patellar tendon is common in jumping sports such as basketball, tennis, football, hockey, and volleyball [12]. Since the Achilles tendon is a subject that is both comparatively more at risk to be affected and also a tendon that is of interest to us, from now on we will mostly focus on Achilles tendon-related pathologies. Below is a brief description of some of Achilles tendon pathologies.

- Tendinopathy

Tendinopathy can be described by where it localizes, either midportion or at its distal attachment. Swelling and ecchymosis are not the expected presentation. Man-

agement of Achilles tendinopathy varies by the clinician; this includes but is not limited to orthotics, physical therapy, topical steroids and iontophoresis, soft tissue treatments, extracorporeal shock wave therapy, anti-inflammatory medications, peritendinous corticosteroid injections, platelet-rich plasma injections, and surgery. Many authors describe tendinopathy as a problematic, frequent pathology that can be seen in up to 30% of medical care consultations involving musculoskeletal disorders [13]. In the case of mid-portion tendinopathy, a common musculoskeletal condition that causes discomfort and disability, the Achilles tendon degenerates in the mid-portion. It is a widespread problem that affects up to 50% of athletes and 30% of the general population, according to [14]. Previous studies have shown that Achilles tendinopathy is associated with age, gender, body mass index, exercise (over-training or lack of training), medications, systemic diseases, and anatomical abnormalities of the feet and ankles [15]. Tendinopathy in the upper extremities (shoulder, elbow, hand) is mostly related to repetitive motion exercises during work, throwing sports excluded. Meanwhile, tendons most vulnerable to repetitive motion exercises in the lower extremities (foot, knee, hip) are injured mainly associated with sports and training. Interestingly, however, all tendons of the human body may undergo tendinopathic changes under certain circumstances [16]. Furthermore, tendinopathy may be classified on the basis of the specific structures involved such as tendinopathy of tendon body, tendinopathy of adjacent tissues, pantendinopathy, insertional tendinopathy, and tendons rupture [17].

- Tendinosis

Tendinosis is the chronic, diffuse thickening of the tendon and histological evidence of degeneration in the absence of acute injury or clinical signs of inflammation. The pathophysiology of Achilles tendon tendinosis appears to be repetitive microtrauma with a relative lack of healing due to the limited vasculature. The Achilles tendon thickens in response to this activity, which is visible on imaging such as MRI and ultrasound and may result in a palpable nodule at the Achilles tendon. The common sign of tendinosis is the thickening of the midportion of the Achilles tendon with an increased signal [18]. Thickening of the whole Achilles tendon rarely occurs. Whereas, Insertional Tendinopathy is a common cause of heel pain and operative treatment is more often indicated than midportion tendinopathy. If the bony prominence on the posterosuperior part of the calcaneus causes impingement of the Achilles tendon, the condition is called Haglund's syndrome [19].

Injuries

Ruptures represent a spectrum of partial-thickness to complete-thickness tears. Tendon ruptures occur when there is a sudden force placed on the Achilles tendon, typically an eccentric load or explosive plyometric contraction. Other movements, such as pivoting, may cause a rupture. The majority of ruptures occur during recreational sports and athletic events; however, ruptures can occur in non-athletes and sedentary individuals. Due to the strength and elasticity of the tendon, patients often complain of a sudden pop during rupture. Commonly, patients feel as if they were struck or pushed at the level of the ankle. The provider may appreciate swelling, bruising, and/or a visible defect in the tendon when compared to the unaffected leg. In partial tears, they may have pain with intact plantar flexion. In the Thompson test, squeezing the calf should normally elicit plantar flexion, but this motion may become diminished

or absent in tendon rupture. Conversely, the presence of a plantar-flexion mechanism does not definitively exclude a partial tear or rupture. In full rupture, tendon rupture can be either acute, in the case of an individual experiencing a rupture from a single high-load impact (for example, an Achilles tendon rupture associated with sudden or violent dorsiflexion of ankle or lunge), or chronic, when the tendon is weakened due to tendinopathy or aging, and the tendon ruptures at lower loads [20]. It is a common sports injury [21]. Achilles tendon rupture may occur in the proximal part, in the midportion, or in the distal part of the tendon. The most common site of a rupture is at the mid-portion of the tendon at 5–6 cm from the insertion, although ruptures may occur at the musculotendinous junction or close to the insertion on the calcaneus [22]. However, the Achilles tendon commonly tears at the mid-substance [23]. Partial ruptures are rather uncommon and are often misinterpreted as aggravated Achilles tendinopathy, and are not always considered a differential diagnosis. Most partial tears are caused by an overload of the tendon tissue and occur in areas afflicted with tendinopathic tissue changes. In histopathological analysis, degenerative changes, such as collagen disorientation and fiber separation, with increased mucoid ground substance, fibrin deposits, hypercellularity, and neovascularization are present in partial ruptures [24]. Since these are also current findings in tendinosis, it has been advocated that partial ruptures are rather microlesions representing an advanced stage of tendinosis. However, granulation tissue, fibroblastic and myofibroblastic proliferation, and hemorrhage are more frequently associated with partial ruptures than with tendinopathy [24]. Some patients have an altered signal within the midsubstance of the tendon but without the thickening associated with classical tendinopathy. The retrocalcaneal bursa may also have a higher signal, although there may be a paucity of clinical features [25]. Conversely, some patients have tendon thickening without a typical focal intratendinous signal.

Clinical Symptoms, and Sign

While Achilles tendon is the strongest in the body, it is prone to injury and is the most commonly ruptured tendon. Achilles tendinopathy is a common presentation among running and jumping athletes. Two-thirds of Achilles tendon injuries in competitive athletes are paratenonitis and one-fifth are insertional complaints (bursitis and insertion tendinosis). The remaining afflictions consist of pain syndromes of the myotendineal junction and tendinopathies. The majority of Achilles tendon injuries from the sport occur in males, mainly because of their higher rates of participation in sport. About 75% of total and the majority of partial tendon ruptures are related to sports activities usually involving abrupt repetitive jumping and sprinting movements. Mechanical factors and a sedentary lifestyle play a role in the pathology of these injuries. Achilles tendon overuse injuries occur at a higher rate in older athletes than most other typical overuse injuries [26]. The Achilles tendon is vulnerable to injuries due to areas of limited blood supply and its subjection to strong forces. The tendon becomes increasingly stiff as individuals age. Intrinsic risk factors for rupture include subtalar hyperpronation, excessive forefoot or hindfoot varus or valgus, the inflexibility of the triceps surface at the musculotendinous junction, leg length discrepancy, muscle imbalance, muscle fatigue, excess weight or obesity, and age. Extrinsic risk factors include primarily incorrect running and jumping techniques, improper footwear, and improper training. Certain medications have been implicated, such as steroids and fluoroquinolones. Overall, the incidence of Achilles tendon rupture is 12 per 100,000

individuals. The typical age is between 40 and 50 years old. The male-to-female ratio ranges from 2 to 1 to 12 to 1. Management of Achilles tendon rupture varies. For active individuals or professional athletes, surgical intervention is often a therapeutic option. In individuals with partial tears, older individuals, or people who are poor surgical candidates, conservative therapy with immobilization, rest, and progressive physical therapy may be a consideration [6]. Typical histological features of chronically inflamed paratendineal tissue of the Achilles tendon are profound proliferation of loose, immature connective tissue and marked obliterative and degenerative alterations in the blood vessels. These changes cause continuing leakage of plasma proteins, which may have an important role in the pathophysiology of these injuries. The chronically inflamed paratendineal tissues of the Achilles tendon do not seem to have enough capacity to form a mature connective tissue [26].

Achilles tendinopathy is usually diagnosed by medical history and physical examination. Patients are often present with pain, swelling, and limited function in and around the Achilles tendon. Pain is associated with inflammation, but in tendinopathy, there is little patient evidence of inflammation. Pain probably originated from a combination of mechanical and biochemical factors. Tendon degeneration associated with a mechanical breakdown of collagen could theoretically explain the pain, but clinical and surgical observations have challenged this point of view [27].

The Role of Medical Imaging

Tendon-related pathologies are typically diagnosed from a physical exam and medical history. On the other hand, medical imaging is typically used to improve surgical treatment planning, and in some cases, to determine the degree of tendon injury. These images have so far rarely been used for diagnostic studies. Very few studies that have tried to determine whether the tendon was healthy or not have gone largely unnoticed. In particular, we would like to investigate the feasibility of using MRI for diagnosis, which will be covered in more detail in our research proposal section.

Ultrasonography is the most commonly used imaging modality for the assessment of an injured or painful tendon. Ultrasound (US) is a cheap and easily accessible modality; however, it is operator-dependent, whereas Magnetic resonance imaging (MRI) depends on an optimal protocol. The degree of bursitis and the presence of a tendon lesion within the insertion may also be determined effectively on MRI. Postoperative complications, infection, and abscess formation, and tendon or graft rupture, can be assessed on MRI with contrast. Ultrasound is also as good as MRI in the detection of tendinopathy and full-thickness tears [28]. However, clinical findings in tendinopathy correlate more with MRI than with ultrasound[29]. MRI is superior to ultrasound in the detection of partial tears [30]. However, one imaging modality (MRI or US) can be preferred over another depending on the condition of Achilles tendon disease.

From a variety of plane angles, sagittal and axial sections are the most helpful in assessing the Achilles tendon. In terms of MRI sequences, the following distinct sequences are helpful for highlighting various entities:

- T1 weighted image
  Presents lower signal for more water content (as in edema, tumor, infarction, inflammation, infection, hyperacute or chronic hemorrhage), for fat and paramagnetic substances show higher signal, and Achilles tendon appears dark.

- T2 weighted images
  Good for showing fluid accompanying tendon pathologies (shows higher signal), lower signal for tumor, lower signal for fat (note that this only applies to standard Spin Echo (SE) sequences and not the more modern Fast Spin Echo (FSE) sequence (also referred to as Turbo Spin Echo, TSE), which is the most commonly used technique today), lower signal for paramagnetic substances, bones and Achilles tendon appears to be dark.

- Fat-suppressed T1 or T2 images
  Due to short relaxation times, fat has a higher signal. This high signal, easily recognized on MRI, may be useful to characterize a lesion.

- Proton Density Weighted images
  Good for joint disease and injury identification. Shows high signal from meniscus tears, tendon appears dark and bone appears gray.

Modification of classical diffusion tensor imaging (DTI) MRI may reveal tendinopathy-induced microstructural disorganization in the very early stages. The process of regeneration, healing, remodeling, and progress or treatment response are possible applications of DTI [31]. Direct imaging of water within the tendon may be a favorable way of diagnosing tendon disorders. T2-mapping is a method that shows promise for monitoring tendon healing. The T2 value of a normal healthy tendon is lower than that of a damaged tendon. The concentration of water is higher after trauma and then decreases with time, thus the healing process may be quantified using T2 mapping [32]. The use of fat suppression sequences allows the ability to detect focal lesions. Sequences with contrast are indicated in postoperative investigations and suspicion of infection, arthritis, or tumor. MRI may reveal inflammatory changes with minor symptoms long before the clinical manifestations of seronegative spondyloarthropathy. The most common non-traumatic focal lesion of Achilles tendon is Achilles tendon xanthoma, which is manifested by an intermediate or slightly higher signal on T1- and T2-weighted images compared to that in the normal Achilles tendon. Other tumors of Achilles tendon are very rare, whereas the involvement of the tendon from tumors in adjacent structures is more frequent [33].

The conditions under which MRI is preferable are - Paratendonitis (the US can be under diagnostic), partial rupture (MRI is superior in the diagnosis of partial thickness tear [30]), midportion tendinopathy (may be overestimated in ultrasound [34]) with a partial tear, insertional tendinopathy (both Ultrasound and MRI play complementary roles), monitoring effects following different treatment, infection (MRI shows soft tissue and bone involvement [35]), tumors and tumor-like lesions (better differentiation of malignant and benign changes), etc. The conditions where the US is preferred over MRI are - tendon elongation (easier simultaneous comparison with the healthy side) and insertional tendinopathy (both Ultrasound and MRI play complementary roles).

1.1.2 Medical Image Features

Feature extraction is a vital step in the field of data analytics and machine learning, especially in the fields of pattern recognition, data mining, and image analysis. It is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. A characteristic of these large data sets is the

large number of variables that require a lot of computing resources to process. Feature extraction is the name for methods that select and/or combine variables into features, effectively reducing the amount of data that must be processed, while still accurately and completely describing the original data set. The process of feature extraction is useful when it is needed to reduce the number of resources needed for processing without losing important or relevant information. Feature extraction can also reduce the amount of redundant data for a given analysis. Also, the reduction of the data and the machine's efforts in building variable combinations (features) facilitate the speed of learning and generalization steps in the machine learning process [36]. These extracted features function as inputs to various algorithms, models, or analytical techniques for tasks such as classification, clustering, and more. Different machine learning algorithms may require data in specific formats or representations. Feature extraction can help prepare data in a way that is compatible with your chosen algorithm, allowing for seamless integration into the modeling process.

Radiomics is a rapidly evolving field of research concerned with the extraction of quantitative metrics - the so-called radiomic features - within medical images. Through mathematical extraction of the spatial distribution of signal intensities and pixel interrelationships, radiomics quantifies textural information by using different analysis methods [37]. Radiomics is a heavily discussed topic in nuclear medicine and in medical imaging in general. Although the term is not strictly defined, radiomics generally aims to extract quantitative, and ideally reproducible, information from diagnostic images, including complex patterns that are difficult to recognize or quantify by the human eye [38]. This approach is broadly employed in the field of oncology, where it allows the development of a decision support system, capable of overcoming the performance of the existing tools, by exploiting both medical imaging data and other patient characteristics when they are available. The core assumption behind radiomics is that biomedical images store disease-specific information that is imperceptible by the human eye, so not accessible through a visual inspection of such images [37].

First, radiomics may be used to capture tissue and lesion properties such as shape and heterogeneity and, on serial imaging, their changes over time, such as during treatment or surveillance. Second, radiomic data are mineable, meaning that in sufficiently large datasets, they may be used to discover previously unknown markers and patterns of disease evolution, progression, and treatment response. This so-called population-imaging approach [39] either may use unstructured data from different modalities (e.g., PET, CT, and MRI) acquired for a specific but possibly unrelated diagnostic purpose in broadly defined groups or may use - as in the German National MRI Cohort Study, a single imaging test in a large cohort for a multicentric longitudinal observational study [40]. Such radiomic data can be combined with clinical, laboratory, histologic, genomic, or other data, using unsupervised machine learning.

Radiomic features can be roughly subdivided into statistical, including histogram-based and texture-based; model-based; transform-based; and shape-based [41].

- Histogram Features
  The simplest statistical descriptors are based on the global gray-level histogram and include gray-level mean, maximum, minimum, variance, and percentiles [41]. Because these features are based on single-pixel or single-voxel analyses, they are called first-order features. More sophisticated features include skewness and kurtosis, which describe the shape of the intensity distribution of data: skewness

reflects the asymmetry of the data distribution curve to the left (negative skew, below the mean) or right (positive skew, above the mean), whereas kurtosis reflects the tailedness of a data distribution relative to a gaussian distribution due to outliers. Other features include histogram entropy and uniformity. Notably, these differ from their cooccurrence matrix counterparts of the same name.

- Shape-Based Features
  Shape-based features describe geometric properties of ROIs. Many shape-based features are conceptually much simpler than other radiomic features, such as 2D and 3D diameters, axes, and their ratios. Surface- and volume-based approaches found on the use of meshes (i.e., small polygons such as triangles and tetrahedrons) are more complex. Features include compactness and sphericity, which describe how the shape of an ROI differs from that of a circle (for 2D analyses) or a sphere (for 3D analyses), and density, which relies on the construction of a minimum oriented bounding box enclosing the ROI [42].

- Texture Features

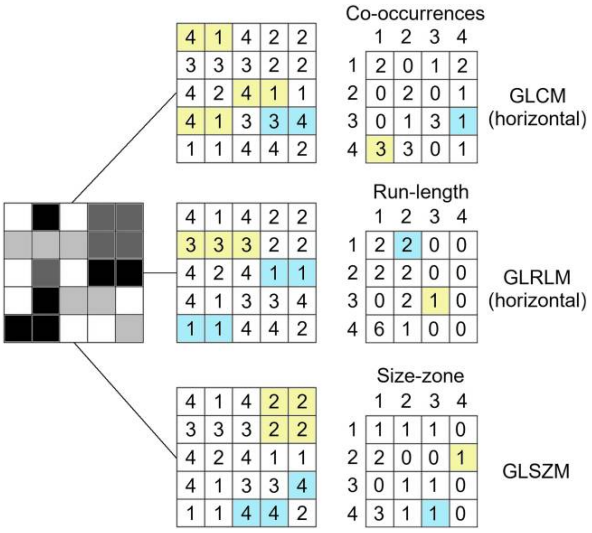


Figure 1.2: Texture Features

1. Gray-Level Cooccurrence Matrix (GLCM)
   First described by Haralick et al. [43], the GLCM is a second-order gray-level histogram. GLCM captures spatial relationships of pairs of pixels or voxels with predefined gray-level intensities, in different directions (horizontal, vertical, or diagonal for a 2D analysis or 13 directions for a 3D analysis), and with a predefined distance between the pixels or voxels. GLCM features include entropy, a measure of gray-level inhomogeneity or randomness; angular second moment (also called uniformity or energy), which reflects gray-level homogeneity or order; and contrast, which emphasizes gray-level differences between pixels or voxels belonging to a pixel or voxel pair [44].
2. Gray-Level Run-length Matrix (GLRLM)
   The GLRLM, described by [45], provides information about the spatial distribution of runs of consecutive pixels with the same gray level, in one or more directions, in 2 or 3 dimensions. GLRLM features include fraction, which assesses the percentage of pixels or voxels within the ROI that are part

of the runs and therefore reflects graininess; long- and short-run emphasis (inverse) moments, which are weighted toward the presence of numbers of long and short runs, respectively; and gray-level and run-length nonuniformity, which assesses the distribution of runs over different gray levels and run lengths, respectively [45].

3. Gray-Level Size Zone Matrix (GLSZM) and Gray-Level Distance Zone Matrix (GLDZM)
   Described by [46], the GLSZM is based on a similar principle to the GLRLM, but here, counts of the number of groups (so-called zones) of interconnected neighboring pixels or voxels with the same gray level form the basis for the matrix. A more homogeneous texture will result in a wider and flatter matrix. GLSZM is not computed for different directions but may be computed for different pixel or voxel distances that define the neighborhood. GLSZM features may be calculated in 2 dimensions (8 neighboring pixels) or 3 dimensions (26 neighboring voxels) and, following GLRLM definitions, include fraction (percentage of pixels or voxels that are part of the zones), large- and small-zone emphasis, and others.
4. Neighborhood Gray-Tone Difference Matrix (NGTDM)
   Proposed by [47], the NGTDM quantifies the sum of differences between the gray level of a pixel or voxel and the mean gray level of its neighboring pixels or voxels within a predefined distance. Key features include NGTDM coarseness, busyness, and complexity. Coarseness reflects the gray-level difference between the central pixel or voxel and its neighborhood and thus captures the spatial rate of changes in gray-level intensities; that is, an ROI consisting of larger areas with relatively uniform gray levels (i.e., a lower rate of spatial intensity changes) will have a high coarseness value. Busyness, on the other hand, reflects rapid gray-level changes between the central pixel and voxel and its neighbors (i.e., a high spatial frequency of intensity changes), so that an ROI comprising many small areas with markedly different gray levels will have greater busyness.

1.1.3 Computer Vision

Computer vision has been expanded into the vast area of field ranging from recording raw data to the extraction of image patterns and information interpretation. It has a combination of concepts, techniques, and ideas from digital image processing, pattern recognition, artificial intelligence, and computer graphics. Most of the tasks in computer vision are related to the process of obtaining information on events or descriptions, from input scenes (digital images) and feature extraction [48].

- Image Segmentation
  Image segmentation can be defined as a specific image processing technique that is used to divide an image into two or more meaningful regions. Image segmentation can also be seen as a process of defining boundaries between separate semantic entities in an image. From a more technical perspective, image segmentation is a process of assigning a label to each pixel in the image such that pixels with the same label are connected with respect to some visual or semantic property. Image segmentation subsumes a large class of finely related problems in computer vision [49].

- Image Classification
  Image classification is the process of categorizing and labeling groups of pixels or vectors within an image based on specific rules. The categorization law can be devised using one or more spectral or textural characteristics [50].

1.1.4 Problem Definition and Rationale

This research initiative stems from a collaborative effort with two highly qualified orthopedic specialists and a comprehensive review of existing literature. Our focus centers on leveraging medical imaging data, particularly magnetic resonance imaging (MRI), to address the diagnostic challenges associated with Achilles tendinopathy and tendon rupture. This area of study aligns with the overarching goal of this EU-funded collaborative PhD program: to explore human soft tissues and investigate the interplay between machine learning and medical imaging in advancing their analysis.

While several studies have focused on distinguishing between healthy and pathological tendons, our research seeks to go beyond this basic goal. We are dedicated to supporting the orthopedic community by developing a comprehensive pipeline that facilitates not only the segmentation of the region of interest, specifically the Achilles tendon but also the evaluation of tendon conditions in individuals through classification as either healthy or pathological. Moreover, we aim to investigate the intrinsic features of pathological Achilles tendons to identify the underlying factors contributing to these conditions, which remain unclear to date. Our research is distinctive in that it does not operate in isolation; we are committed to collaborating closely with medical specialists to bridge the gap between our findings and the clinical expertise of orthopedic professionals. In summary, our goal is to develop a comprehensive pipeline for tendon segmentation and pathology identification/classification. We also aim to provide insights into the relationship between our findings and specialists' understanding of responsible features that can assist in treatment planning. Additionally, we will explore the relationship between tendon-related pathology and cartilage degeneration. We will also investigate various demographic factors to assess neuromuscular function through the patellar tendon reflex.

1.1.5 Importance of the Field

Tendinopathy, a widespread and challenging musculoskeletal disorder, affects a considerable amount of our population. It affects up to 30% of medical consultations related to musculoskeletal problems [13]. Regardless of the severity of tendinopathy, its symptoms can profoundly impact the patient's perceived quality of life. Medical imaging techniques such as ultrasound (US), magnetic resonance imaging (MRI), and X-rays play a crucial role in the diagnosis and evaluation of tendinopathy. These imaging modalities provide valuable information about structural and pathological changes within tendons. The ability to accurately identify and study tendon injuries allows healthcare professionals to make informed clinical decisions. Machine learning, on the other hand, is a subset of artificial intelligence that offers the potential to transform tendinopathy research and diagnosis. It can process large amounts of image data and extract complex features that might not be very clear to the naked human eye. By training machine learning models, it can be possible to develop algorithms that can detect, classify, and predict tendinopathy. The combination of medical images and

machine learning has the potential to revolutionize tendinopathy-related research and clinical practice. Also, from the point of view of accuracy and early diagnosis can lead to more effective interventions and personalized treatment plans. Furthermore, understanding and monitoring pathology progression and evaluation of treatment outcomes can be significant, ultimately improving patient care and reducing healthcare costs. Overall, the synergy between medical imaging and machine learning offers potential in tendinopathy tendinopathy-related research field. By utilizing these technologies, we can better address the clinical and social challenges posed by tendinopathy, ultimately improving the quality of life of those affected.

1.1.6 Challenges of the Field

Research related to tendinopathy presents several challenges due to the complexity of the condition and the lack of knowledge of the exact reasons underlying the pathologies. Some of the key challenges in tendinopathy research include:

- Tendon structure is typically disorganized in tendinopathic conditions, however, several authors have shown a discrepancy between medical imaging results and clinical symptoms [51]. Frequently, asymptomatic subjects present tendon disorganization signs when explored through MRI or US, whereas subjects with tendon pain present an image of apparently healthy tendons [52]. It is possible to notice changes in a patient's clinical symptoms (such as an improvement or worsening of their feeling of pain) that do not correspond to changes in the medical image structure [53].

- The normal Achilles tendon is composed of three twisted sub tendons; two from the medial and lateral heads of the gastrocnemius muscle and one from the soleus muscle [54]. These twisted sub-tendons are separated by thin high signal septae, which is a potential pitfall in MRI because they mimic a tendon tear.

- Pain is associated with inflammation, but in tendinopathy, there is little patient evidence of inflammation. Pain probably originated from a combination of mechanical and biochemical factors. Tendon degeneration associated with a mechanical breakdown of collagen could theoretically explain the pain, but clinical and surgical observations have challenged this point of view of individuals to increased risk of rupture [27].

- In samples from 397 ruptured Achilles tendons, Kannus et al. found no evidence of inflammation under light and electron microscopy [55]. Arner et al. also did not find neutrophilic infiltration in Achilles tendons on the first day after rupture, and they concluded that any inflammation seen at a later stage occurred subsequent to the rupture. Tendinopathy shows features of disordered healing, and inflammation is not typically seen. The role played by inflammation in tendon rupture is less clear [12]. The underlying pathophysiological process that results in tendinopathy also remains undetermined. Tendinopathic changes are often clinically silent, and the only manifestation may be a rupture; however, they may also coexist with symptomatic paratendinopathy [56]. Therefore, tendinopathy is currently diagnosed as a clinical hypothesis based on the patient's symptoms and physical context. There is no defined model of the patient regarding the level of activity, tendon loading, pain symptoms, or functional capability [57].

- Tendinopathy can affect different tendons in our body. And each of them could have unique characteristics and underlying causes. This heterogeneity can make it difficult to find a generalized solution/finding.
- It is challenging to distinguish postoperative changes from infections based only on MRI (especially in early stages).
- Achilles tendon in rheumatoid arthritis: Some patients may have an altered signal in the mid-substance of the tendon but without any tendon thickness. Conversely, some patients have the opposite case [25].
- Tendinopathy can mimic the symptoms of other musculoskeletal conditions, making it critical to distinguish it from conditions such as muscle strains or bursitis.
- Slice thickness together with the position of the tendon in the scanner contribute to the challenge of determining the actual distance that the tendon ends are separated [19].

1.2 Architectures

1.2.0.1 Classical Machine Learning Models

Classical machine learning is a branch of machine learning that encompasses traditional algorithms that were developed prior to the advent of deep learning. These algorithms are based on mathematical and statistical models and are widely used for various tasks, including regression and classification [58]. Below are some of the models in classical machine learning:

- K-nearest Neighbor
  K-NN is a simple and widely used algorithm for classification and regression tasks [59]. The algorithm operates on the principle of proximity, where it determines the class or value of a query point by analyzing the k nearest data points in the feature space. These nearest neighbors are identified based on a chosen distance metric, with Euclidean distance being one of the most commonly used metrics due to its intuitive representation of geometric proximity. For classification, the algorithm assigns the query point to the most frequent class among its k nearest neighbors, leveraging majority voting.
- Support Vector Machine
  The SVMs are supervised machine learning techniques that make a non-probabilistic binary classifier by assigning new examples to one class or the other. More specifically, the kernel support vector machines (SVM) is a nonlinear classifier where the representations are built from pre-specified filters. This is in contrast to the deep learning paradigm in which good representations are learned from data. Consequently, the kernel SVM are sample efficient learning methods that are more adequate for medical imaging applications with a small training sample size. In addition, the training phase of the kernel SVM involves tuning the hyperparameters of the SVM classifier only, which can be carried out quickly and efficiently. Contrary to deep learning models, the kernel SVM is a transparent learning model whose theoretical foundations are grounded in the extensive statistical machine learning literature [60].

- Random Forest
  Random forests or random decision forests are an ensemble learning method that are used to build predictive models by combining decisions from a sequence of base models. Ensemble methods use multiple learning models to gain better predictive results. In the case of a random forest, the model creates an entire forest of random uncorrelated decision trees to arrive at the best possible answer. Such methods are often called Bootstrap Aggregation or bagging and are used to overcome a bias-variance trade-off problem. In general, learning error can be explained in terms of bias and variance. The bias in learning error is reduced by averaging results from respective trees, and while the predictions of a single tree are highly sensitive to its training set, the mean of individual trees is not sensitive, as long as the trees are not correlated. If trees are independent from each other, then the central limit theorem would ensure variance reduction. Random forest uses an algorithm which selects a random subset of the features at the process of splitting each candidate to reduce the correlation of the trees in a bagging sample [61].

- Kernel-based Regression
  Kernel-based regression uses a kernel function to map the input data into a high-dimensional feature space where linear regression can be applied, allowing for the modeling of complex non-linear relationships between the features and the target variable.

- Linear Regression
  Linear regression is perhaps one of the most well-known methods in statistics and machine learning, whose theoretical performance is studied extensively. Despite its simple framework, its concept is still a basis for other advanced techniques. In linear regression, the model is determined by linear functions whose unknown parameters are estimated from data. Simply put, linear regression is related to finding a linear equation that represents the model well. Linear regression models are often fitted using minimization of the l-norm (ex., 2-norm minimization is the least square approach) [62].

- Graph Search algorithms and Graph Cuts
  Graph search algorithms and graph cuts are widely used in computer vision and image segmentation. In graph-based methods, the data is represented as a graph, where nodes correspond to data points, and edges define relationships (e.g., similarity). Graph search aims to find optimal paths or regions in the graph, while graph cut methods partition the graph by cutting edges, minimizing a cost function such as energy. The cut divides the nodes into distinct groups, useful for segmenting images into meaningful regions by minimizing within-group similarity and maximizing between-group dissimilarity [63].

- K-Means Clustering
  Clustering is a method to divide a set of data into a specific number of groups. It's one of the popular method is k-means clustering. In k-means clustering, it partitions a collection of data into a k-number group of data. It classifies a given set of data into k number of disjoint clusters. K-means algorithm consists of two separate phases. In the first phase, it calculates the k centroid and in the second phase, it takes each point to the cluster which has the nearest centroid from the

respective data point. There are different methods to define the distance of the nearest centroid and one of the most used methods is Euclidean distance. Once the grouping is done it recalculate the new centroid of each cluster and based on that centroid, a new Euclidean distance is calculated between each center and each data point and assigns the points in the cluster which have minimum Euclidean distance. Each cluster in the partition is defined by its member objects and by its centroid. The centroid for each cluster is the point to which the sum of distances from all the objects in that cluster is minimized. So K-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid, over all clusters [64].

- Markov Random Field
  Another segmentation method using the classical machine learning concept is the Markov random field (MRF) segmentation. MRF is itself a conditional probability model, where the probability of a pixel is affected by its neighboring pixels. MRF is a stochastic process that uses the local features of the image [65] [66]. It is a powerful method to connect spatial continuity due to prior contextual information. This provides useful information for segmentation.

1.2.0.2 Deep Learning Models

This section provides an overview of prominent deep neural network architectures used by the computer vision community, including convolutional neural networks, recurrent neural networks and long short-term memory, encoder-decoder and autoencoder models, and generative adversarial networks.

- Convolutional Neural Networks (CNNs)
  CNNs are among the most successful and widely used architectures in the deep learning community, especially for computer vision tasks. CNNs were initially proposed by Fukushima [67] in his seminal paper on the "Neocognitron", which was based on Hubel and Wiesel's hierarchical receptive field model of the visual cortex. Subsequently, Waibel et al. [68] introduced CNNs with weights shared among temporal receptive fields and backpropagation training for phoneme recognition, and LeCun et al. [69] developed a practical CNN architecture for document recognition (fig. 1.3).

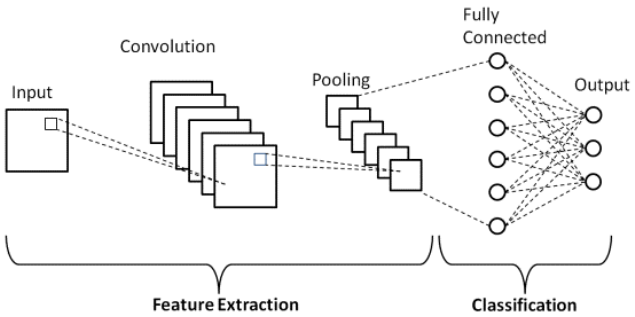


Figure 1.3: Basic CNN Architecture [70]

CNNs usually include three types of layers: i) convolutional layers, where a kernel (or filter) of weights is convolved to extract features; ii) nonlinear layers, which apply (usually element-wise) an activation function to feature maps, thus

enabling the network to model nonlinear functions; and iii) pooling layers, which reduce spatial resolution by replacing small neighborhoods in a feature map with some statistical information about those neighborhoods (mean, max, etc.). The neuronal units in layers are locally connected; that is, each unit receives weighted inputs from a small neighborhood, known as the receptive field, of units in the previous layer. By stacking layers to form multiresolution pyramids, the higher-level layers learning features from increasingly wider receptive fields. The main computational advantage of CNNs is that all the receptive fields in a layer share weights, resulting in a significantly smaller number of parameters than fully-connected neural networks. Some of the most well known CNN architectures include AlexNet [71], VGGNet [72], and ResNet [73].

- Recurrent Neural Networks (RNNs) and LSTM
  RNNs are commonly used to process sequential data, such as speech, text, videos, and time-series [74]. Referring to fig. 1.4, each time step of the model collects the input x_t and the hidden state h_{t-1} from the previous step, and outputs a target value o_t and the next hidden state h_{t+1}. RNNs are typically problematic for long sequences as they cannot capture long-term dependencies in many real-world applications and often suffer from gradient vanishing or exploding problems. However, a type of RNN known as the Long Short-Term Memory (LSTM) is designed to avoid these issues [75]. The LSTM architecture (fig. 1.5) includes three gates (input gate, output gate, and forget gate) that regulate the flow of information into and out of a memory cell that stores values over arbitrary time intervals.

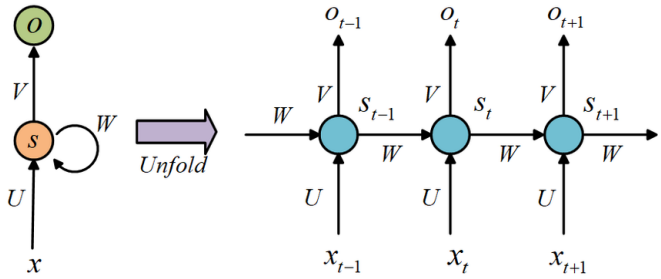


Figure 1.4: Basic RNN Architechture [76]

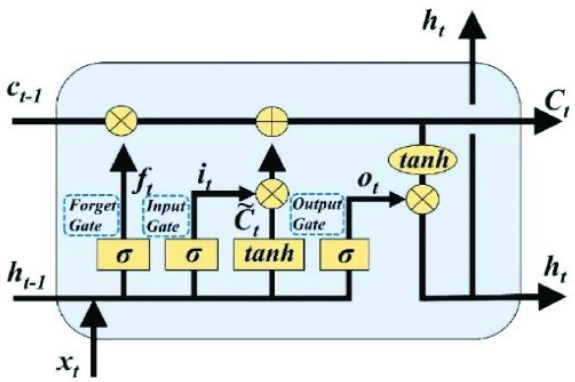


Figure 1.5: Basic LSTM Module [77]

- Encoder-Decoder and Auto-Encoder Models
  Encoder-decoders are a family of models that learn to map data points from an input domain to an output domain via a two-stage network (fig. 1.6): The encoder, performing an encoding function z = g(X), compresses the input x into a latent-space representation z, while the decoder y = f(z) predicts the output y from z [78]. The latent, or feature (vector), representation captures the semantic information of the input useful in predicting the output. Such models are popular for sequence-to-sequence modeling in Natural Language Processing (NLP) applications as well as in image-to-image translation, where the output could be an enhanced version of the image (such as in image deblurring, or super-resolution) or a segmentation map. Autoencoders are a special case of encoder-decoder models in which the input and output are the same.

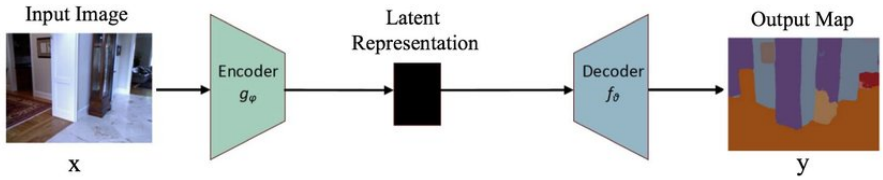


Figure 1.6: Simple Encoder-Decoder Architecture [79]

- Generative Adversarial Networks (GANs)

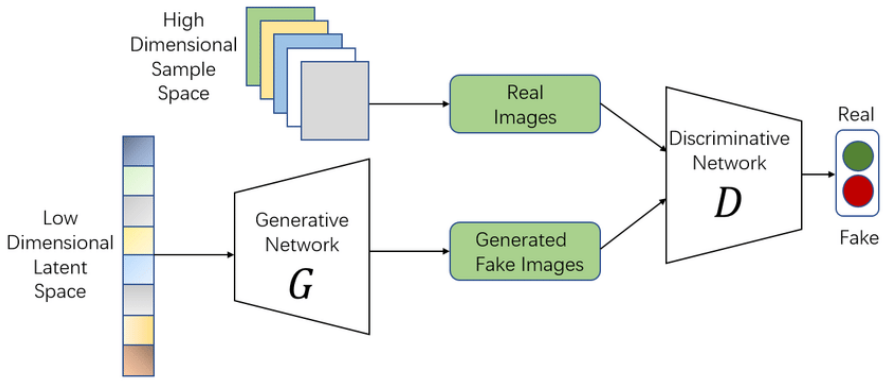


Figure 1.7: Basic GAN Architecture [80]

GANs are a newer family of deep learning models [81]. They consist of two networks, a generator and a discriminator (fig. 1.7). In the conventional GAN, the generator network G learns a mapping from noise z (with a prior distribution) to a target distribution y, which is similar to the "real" samples. The discriminator network D attempts to distinguish the generated "fake" samples from the real ones. The GAN may be characterized as a minimax game between G and D, where D tries to minimize its classification error in distinguishing fake samples from real ones, hence maximizing a loss function, and G tries to maximize the discriminator network's error, hence minimizing the loss function.

1.3 Image Segmentation

Image segmentation is the process of clustering an image into several coherent sub-regions according to the extracted features, e.g., color, or texture attributes, and classifying each sub-region into one of the pre-determined classes. Segmentation can also be viewed as a form of image compression, which is a crucial step in inferring knowledge from imagery.

1.3.1 Applications of Image Segmentation

The following are some notable applications of image segmentation:

- Medical Imaging
  In medical imaging, segmentation techniques can help early detection and better treatment planning. Radiologists can use image segmentation to perform various tasks, such as tumor detection, disease diagnosis etc. They can analyze and diagnose conditions more accurately by segmenting organs or abnormalities from surrounding tissues in MRI, CT scans, or other imaging modalities.

- Autonomous Driving
  Image segmentation is essential for autonomous driving systems to perceive and understand the surrounding environment. This allows self-driving cars to determine roads, people, other vehicles, and obstacles. This allows them to determine where to drive and also to avoid accidents.

- Object Recognition
  Segmentation improves object recognition by isolating objects from the background, making it easier to identify and classify individual items in complex scenes.

- Remote Sensing
  Image segmentation is also used in satellite image analysis. This can be applied to land cover classification, environmental monitoring, and urban planning. Researchers use image segmentation to analyze changes over time, monitor natural resources and plan sustainable development. Because of the large-scale imagery, identifying those categories becomes automatically computable and analyzed through image segmentation.

- Retail and E-commerce
  With image segmentation, retailers can identify items, categorize them, and extract relevant attributes for cataloging and recommendation systems. This improves data quality by reducing ambiguity and increasing the speed of product identification and inventory tracking, enhancing the overall efficiency of e-commerce operations [82].

1.3.2 Machine Learning Approaches

Broadly, segmentation techniques are divided into two categories (i.e., supervised and unsupervised). In the unsupervised segmentation paradigm, only the structure of the image is leveraged. In particular, unsupervised segmentation techniques rely on

the intensity or gradient analysis of the image via various strategies, such as thresholding, graph cut, edge detection, and deformation, to delineate the boundaries of the target object in the image. Such approaches perform well when the boundaries are well-defined. Nevertheless, gradient-based segmentation techniques are prone to image noise and artifacts that result in missing or diffuse organ/tissue boundaries. Graph-based models such as Markov random fields are another class of unsupervised segmentation techniques that are robust to noise and somewhat alleviate those issues but often come with a high computational cost due to employing an iterative scheme to enhance the segmentation results in multiple steps [83].

In contrast, supervised segmentation methods incorporate prior knowledge about the image-processing task through training samples [84]. Atlas-based segmentation methods are an example of supervised models that attracted much attention in the 1990s [85][86]. These methods, such as probabilistic atlases and statistical shape models, can capture the organs' shape well and generate more accurate results than unsupervised models. Support vector machine (SVM), random forest (RF), and k-nearest neighbor clustering are also supervised segmentation techniques studied rigorously in the past decade. However, the success of such methods in delineating fuzzy boundaries of organs in radiological images is limited.

1.3.2.1 Image Segmentation using Classical Machine Learning

Nock et al. explored a statistical basis for a process often described in computer vision: image segmentation by region merging following a particular order in the choice of regions [87]. The authors exhibited a particular blend of algorithmics and statistics whose segmentation errors are shown as limited from both the qualitative and quantitative standpoint. In the study of Dhanachandra et. al, image segmentation is performed in terms of the K -means clustering algorithm and subtractive clustering method [64]. The proposed algorithm consists of partial contrast stretching, subtractive clustering, k-means clustering and median filter. Mostly the medical images that are used for segmentation have low contrast. So contrast stretching is used to improve the quality of the image. After improving the quality of image, subtractive clustering algorithm is used to generate the centers, based on the potential value of the image. Number of centre is generated based on number of cluster k. This centre is used as initial centre in k-means algorithm. Using the k-means algorithm, the image is segmented into k numberof cluster. After the segmentation of image, the image can still contain some unwanted region or noise. These noises are removed by using the median filter. Otsu presented a nonparametric and unsupervised method of automatic threshold selection for image segmentation [88]. An optimal threshold is selected by the discriminant criterion, namely, to maximize the separability of the resultant classes in gray levels. The procedure is very simple, utilizing only the zeroth- and the first-order cumulative moments of the gray-level histogram. Kass et al. utilized the active contour-based model (snakes) which is an energy-minimizing spline guided by external constraint forces and influenced by image forces that pull it toward features such as lines and edges [89]. Snakes are active contour models: they lock onto nearby edges, localizing them accurately. The scale-space continuation can be used to enlarge the capture region surrounding a feature. Snakes provide a unified account of several visual problems, including detection of edges, lines, and subjective contours; motion tracking; and stereo matching. The authors used snakes for interactive interpretation, in which user-imposed constraint forces guide the snake near features of interest. Boykov et

al. addressed the problem of minimizing a large class of energy functions that occur in early vision. The major restriction is that the energy function's smoothness term must only involve pairs of pixels. The authors proposed two algorithms that use graph cuts to compute a local minimum even when very large moves are allowed [90]. The first move we consider is an α - β swap: for a pair of labels α, β this move exchanges the labels between an arbitrary set of pixels labeled and another arbitrary set labeled β. The first algorithm generates a labeling such that there is no swap move that decreases the energy. The second move it is considered as an α-expansion: for a label α, this move assigns an arbitrary set of pixels the label α. The second algorithm, which requires the smoothness term to be a metric, generates a labeling such that there is no expansion move that decreases the energy. Moreover, this solution is within a known factor of the global minimum. The authors experimentally demonstrated the effectiveness of our approach on image restoration, stereo and motion. Plath et al. proposed a method for multi-class image segmentation which comprises two aspects for coupling local and global evidence. First, the authors formulated a Conditional Random Field that coupled local segmentation labels in a scale hierarchy. This idea followed largely previous work which was extended to the multi-class case and the use of SVMs to provide the local patch evidence. Second, the authors used global image classification information to decide prior to the segmentation process which segment labels were considered possible in a given new image. Experiments showed that without this use of global classification, the segmentation performed poorly. However, with a hypothetical optimal classification, the segmentation outperformed the best state-of-the-art segmentation algorithms [91]. Held et al. described a fully-automatic three-dimensional (3-D)-segmentation technique for brain magnetic resonance (MR) images by means of Markov random fields (MRF's) [92] to address the following three practical issues on MR images simultaneously. Their segmentation algorithm captures three key features that are practical obstructers to MR image segmentation (i.e., nonparametric distributions of tissue intensities, neighborhood correlations, and signal inhomogeneities): a) Nonparametric distribution of tissue intensities are modeled by Parzen-window statistics, b) Neighbor tissue correlations are dealt with MRF to manage the noisy MR data, c) Signal inhomogeneities are also described by a priori MRF. Then, the statistical model is optimized by simulated annealing or iterated conditional modes. They offered the segmentation of simulated MR images with respect to noise, inhomogeneity, smoothing, and optimization method. The accuracy was measured by error rate and the error rates in most cases were less than 10%.

1.3.3 Deep Learning Approaches

Depending on whether labels of the training dataset are present, deep learning can be roughly divided into supervised, unsupervised, and semi-supervised learning. In supervised learning, all training images are labeled, and the model is optimized using the image-label pairs. For each testing image, the optimized model will generate a likelihood score to predict its class label [69]. For unsupervised learning, the model will analyze and learn the underlying patterns or hidden data structures without labels. If only a small portion of training data is labeled, the model learns input-output relationship from the labeled data, and the model will be strengthened by learning semantic and fine-grained features from the unlabeled data. This type of learning approach is defined as semi-supervised learning [93].

1.3.3.1 Image Segmentation using Deep Learning

In recent years, deep learning (DL) models have yielded a new generation of image segmentation models with remarkable performance improvements, often achieving the highest accuracy rates on popular benchmarks [94]. Long et al. [95] proposed Fully Convolutional Networks (FCNs), a milestone in DL-based semantic image segmentation models. An FCN includes only convolutional layers, which enables it to output a segmentation map whose size is the same as that of the input image. To handle arbitrarily-sized images, the authors modified existing CNN architecture, such as VGG16 and GoogLeNet, by removing all fully connected layers such that the model outputs a spatial segmentation map instead of classification scores. FCNs have been applied to a variety of segmentation problems, such as brain tumor segmentation, instance-aware semantic segmentation, skin lesion segmentation, and iris segmentation. While demonstrating that DNNs can be trained to perform semantic segmentation in an end-to-end manner on variable-sized images, the conventional FCN model has some limitations - it is too computationally expensive for real-time inference, it does does not account for global context information in an efficient manner, and it is not easily generalizable to 3D images. Since FCN ignores potentially useful scene-level semantic context. To exploit more context, several approaches incorporated into DL architecture probabilistic graphical models, such as Conditional Random Fields (CRFs) and Markov Random Fields (MRFs). Schwing and Urtasun [96] proposed a fully-connected deep structured network for image segmentation. They jointly trained CNNs and fully-connected CRFs for semantic image segmentation, and achieved encouraging results on the challenging PASCAL VOC 2012 dataset. Zheng et al. [97] proposed a similar semantic segmentation approach. In related work, Lin et al. [98] proposed an efficient semantic segmentation model based on contextual deep CRFs. Liu et al. proposed a semantic segmentation algorithm that incorporates rich information into MRFs, including high-order relations and a mixture of label contexts. Other kinds of DL-based segmentation models use some kind of encoder-decoder architecture. Noh et al. [99] introduced semantic segmentation based on deconvolution (a.k.a. transposed convolution). Their model, DeConvNet, consists of two parts, an encoder using convolutional layers adopted from the VGG 16-layer network and a multilayer deconvolutional network that inputs the feature vector and generates a map of pixel-accurate class probabilities. The latter comprises deconvolution and unpooling layers, which identify pixel-wise class labels and predict segmentation masks. Badrinarayanan et al. [100] proposed SegNet, a fully convolutional encoder-decoder architecture for image segmentation. Similar to the deconvolution network, the core trainable segmentation engine of SegNet consists of an encoder network, which is topologically identical to the 13 convolutional layers of the VGG16 network, and a corresponding decoder network followed by a pixel-wise classification layer. The main novelty of SegNet is in the way the decoder upsamples its lower-resolution input feature map(s); specifically, using pooling indices computed in the maxpooling step of the corresponding encoder to perform nonlinear up-sampling. A limitation of encoder-decoder based models is the loss of fine-grained image information, due to the loss of resolution through the encoding process. Other than recovering high-resolution representations as is done in DeConvNet, SegNet, and other models, HRNet maintains high-resolution representations through the encoding process by connecting the high-to-low resolution convolution streams in parallel and repeatedly exchanging the information across resolutions. There are four stages: the 1st stage consists of high-resolution convolutions, while the 2nd/3rd/4th stage repeats

2-resolution/ 3-resolution/4-resolution blocks. Several recent semantic segmentation models use HRNet as a backbone. Several other works adopt transposed convolutions, or encoder-decoders for image segmentation, such as Stacked Deconvolutional Network (SDN) [101], Linknet [102], WNet [103], and locality-sensitive deconvolution networks for RGB-D segmentation [104]. Several models inspired by FCNs and encoder-decoder networks were initially developed for medical/biomedical image segmentation, but are now also being used outside the medical domain. Ronneberger et al. [105] proposed the U-Net for efficiently segmenting biological microscopy images. The U-Net architecture comprises two parts, a contracting path to capture context, and a symmetric expanding path that enables precise localization. The U-Net training strategy relies on the use of data augmentation to learn effectively from very few annotated images. V-Net, proposed by Milletari et al. [106] for 3D medical image segmentation, is another well known FCNbased model. The authors introduced a new loss function based on the Dice coefficient, enabling the model to deal with situations in which there is a strong imbalance between the number of voxels in the foreground and background. Multiscale analysis, a well established idea in image processing, has been deployed in various neural network architectures. One of the most prominent models of this sort is the Feature Pyramid Network (FPN) proposed by Lin et al. [107], which was developed for object detection but was also applied to segmentation. The inherent multiscale, pyramidal hierarchy of deep CNNs was used to construct feature pyramids with marginal extra cost. To merge low and high resolution features, the FPN is composed of a bottom-up pathway, a top-down pathway and lateral connections. The concatenated feature maps are then processed by a 3*3 convolution to produce the output of each stage. Finally, each stage of the top-down pathway generates a prediction to detect an object. For image segmentation, the authors use two multilayer perceptrons (MLPs) to generate the masks. Zhao et al. [108] developed the Pyramid Scene Parsing Network (PSPN), a multiscale network to better learn the global context representation of a scene (Fig. 16). Multiple patterns are extracted from the input image using a residual network (ResNet) as a feature extractor, with a dilated network. These feature maps are then fed into a pyramid pooling module to distinguish patterns of different scales. They are pooled at four different scales, each one corresponding to a pyramid level, and processed by a 1*1 convolutional layer to reduce their dimensions. The outputs of the pyramid levels are up-sampled and concatenated with the initial feature maps to capture both local and global context information. Finally, a convolutional layer is used to generate the pixel-wise predictions. The Regional CNN (R-CNN) and its extensions have proven successful in object detection applications. In particular, the Faster R-CNN [109] architecture uses a region proposal network (RPN) that proposes bounding box candidates. The RPN extracts a Region of Interest (RoI), and an RoI Pool layer computes features from these proposals to infer the bounding box coordinates and class of the object. Some extensions of R-CNN have been used to address the instance segmentation problem; i.e., the task of simultaneously performing object detection and semantic segmentation. Dilated convolutions have been popular in the field of real-time segmentation, and many recent publications report the use of this technique. Some of the most important include the DeepLab family [110], multiscale context aggregation [111], Dense Upsampling Convolution and Hybrid Dilated Convolution (DUC-HDC) [112], densely connected Atrous Spatial Pyramid Pooling (DenseASPP) [113], and the Efficient Network (ENet) [114]. While CNNs are a natural fit for computer vision problems, they are not the only possibility. RNNs are useful in modeling the short/long term dependencies among pixels to (potentially) improve

the estimation of the segmentation map. Using RNNs, pixels may be linked together and processed sequentially to model global contexts and improve semantic segmentation. However the natural 2D structure of images poses a challenge. Visin et al. [115] proposed an RNN-based model for semantic segmentation called ReSeg. Attention mechanisms have been persistently explored in computer vision over the years, and it is not surprising to find publications that apply them to semantic segmentation. Chen et al. [116] proposed an attention mechanism that learns to softly weight multiscale features at each pixel location. They adapt a powerful semantic segmentation model and jointly train it with multiscale images and the attention model. GANs have been applied to a wide range of tasks in computer vision, not excluding image segmentation. Luc et al. [117] proposed an adversarial training approach for semantic segmentation in which they trained a convolutional semantic segmentation network (Fig. 32), along with an adversarial network that discriminates between ground-truth segmentation maps and those generated by the segmentation network. They showed that the adversarial training approach yields improved accuracy on the Stanford Background and PASCAL VOC 2012 datasets. Souly et al. [118] proposed semi-weakly supervised semantic segmentation using GANs.

1.3.4 Our Problem Statement in Segmentation

Tendon pathologies, such as tendinopathy, tendinosis, and tears, significantly impact musculoskeletal function and quality of life. Accurate diagnosis and treatment planning rely heavily on precise imaging and analysis of tendon structures. However, traditional imaging techniques often pose challenges in delineating tendon boundaries due to the complex and heterogeneous nature of the surrounding tissues. As a result, automated and robust segmentation of tendons from medical images has become a critical need in medical research and clinical practice.

Building on the need for accurate tendon analysis, existing methods of tendon segmentation face limitations such as reliance on manual delineation, variability in interpretation between clinicians, and inconsistencies in segmenting tendons across different imaging modalities. These challenges hinder the efficiency and reproducibility of tendon pathology studies. Therefore, this thesis focuses on developing an automated segmentation pipeline, to provide robust and consistent tendon segmentation. The ultimate goal is to integrate these segmentation techniques into clinical workflows, reducing diagnostic errors and improving treatment outcomes for tendon-related conditions. Additionally, the automated methods developed in this research could offer new insights into tendon structure and pathology, paving the way for further advancements in other studies related to other musculoskeletal disorders. This solution aims to bridge the gap between advanced computational techniques and clinical applicability, ultimately contributing to improved patient care.

1.3.4.1 Problem Specific Studies

This section summarizes the findings from our systematic literature review in the MEDLINE database using the search keywords: **tendon AND segmentation AND machine learning**. The last time this search was done was in July 2022. 3 of the results from this search, out of a total of 13, were disqualified from the abstract reading phase since those 3 did not contain tendon segmentation.

The automated segmentation of the supraspinatus (SSP) tendon using ultrasound images was proposed by Gupta et al. [119] in 2014. The automated segmentation of the SSP tendon was accomplished using image processing techniques. The image processing methods integrated curvelet transform and mathematical concepts of logical and morphological operators along with area filtering. The study used a database of 116 ultrasound (US) images of the supraspinatus tendon, both with and without pathological conditions. The suggested algorithm's procedures for automatically segmenting the SSP tendon are as follows:

- Step 1: Selecting an ultrasound image of the SSP tendon from the database
- Step 2: Applying image enhancement using Rayleigh adaptive contrast enhancement followed by despeckling using anisotropic diffusion method
- Step 3: Decomposition of the enhanced image by real-valued curvelet coefficients using the wrapping function
- Step 4: Morphological, Logical, and Area filtering operation to remove the regions that do not belong to the SSP tendon
- Step 5: Polynomial curve fitting to smooth tendon and recover lost boundary points and construct the mask as per radiologist requirement

The quantitative analysis for the assessment of the proposed methodology: segmentation was assessed using three metrics: 1) false positive rate (FPR), 2) true positive rate (TPR) and 3) Accuracy (Jaccard index) whose values are 0.913, 0.086, and 0.956 respectively. It should be noted that these results were not compared to any other segmentation method. An adaptive texture-based Active Shape Model (ASM) was suggested by Chuang et al. [120] to segment tendon and synovium sheath. In order to alter the contribution of energy terms based on image attributes at various points, modified weights were applied throughout the segmentation phase. The segmented tendon area's wavelet and co-occurrence texture features were then used to determine the pathology. Chuang's dataset was used in this study which consisted of ultrasound images, of which 74 are finger tendon images and 57 are synovial sheath images. Due to the characteristics of the bottom side of the tendon boundary, all images were split into clear and fuzzy boundary groups. The traditional active shape model (ASM) contains information on tendon shape from training images and obtains a more applicable tendon boundary than the active contour model (ACM). However, the energy term which only considers the gradient information is easily affected by the speckle noise. By adding the texture information in energy terms as in the texture-based ASM, the segmentation results are improved and less affected by the noise. Since the top boundary of the tendon is surrounded by a narrow band of synovial fluid, and the boundaries on both sides of the tendon are usually blurry, the texture-based ASM using equivalent weights of energy terms is less effective in segmentation. By adopting different weights of energy terms at different positions, the proposed adaptive texture-based active shape model (ATASM) was able to segment the upper and lower boundaries of the tendon more precisely. After adopting the adaptive weighting mechanism, a genetic algorithm was then applied to optimize the final shape parameters by maximizing the objective function, which included texture and gradient energies. Chuang et al. [120] compared the segmentation accuracy of tendons with clear and fuzzy boundary

groups with those of the other three segmentation methods (active contour model, active shape model, and texture-based active shape model). The average mean absolute distance (MAD) values for tendon segmentation are 3.14 and 3.34 pixels, and the average dice similarity coefficient (DSC) values are 0.91 and 0.90 for the two groups (clear and fuzzy). Both the average MAD and DSC values of the proposed method outperformed the performance of the other three segmentation methods for both groups. By merging a semantic segmentation CNN, 3D fully connected Conditional Random Field and 3D simplex deformable modeling, Zhou et al. [121] developed a segmentation pipeline. The core of the segmentation method, which performs high-resolution pixel-wise multi-class tissue classification for 12 different joint types, is a convolutional encoder-decoder network. In order to regularize contextual relationships between voxels within the same tissue class and between distinct classes, the 3D fully connected 3D conditional random field (CRF) was used. The output from the 3D CRF was adjusted in 3D simplex deformable modeling to maintain the overall shape and a suitable smooth surface for joint structures. On 3D fast spin-echo (3D-FSE) MR image data sets, the approach was assessed. As a dataset, a sagittal frequency selective fat-suppressed 3D fast spin-echo (3D-FSE) sequence was performed on the knee joint of 20 subjects with knee osteoarthritis (12 males and 8 females with an average age of 58 years) using a 3T scanner. 13 classes was created for each 3D-FSE image slice of each subject which are background, femur, femoral cartilage, tibia, tibial cartilage, patella, patellar cartilage, meniscus, quadriceps, and patellar tendons, muscle, synovial fluid-filled joint effusion and Baker's cyst, infrapatellar fat pad, and other non-specified tissues. The 3D image volume of the knee joint was dissembled into a stack of 2D image slices. In the training phase, the 2D image slices were used as the input of the convolutional encoder-decoder architecture (CED network) and were compared to the corresponding pixel-wise class labels in the training data. The training loss of the CED network was determined by multi-class cross-entropy loss weighted by the inverse of class occurring frequency. The weights of the CED network kept getting updated and recorded until the number of iterations reaches a predefined maximum step. Thereafter, the weights of the network used during the testing phase were selected from the iteration wherein the training loss was the lowest among all. In the testing phase, the well-trained CED network was used as a front-end segmentation classifier to segment the testing 2D images and to generate tissue class probabilities for each pixel. To effectively assign the labels to voxels with similar image intensity values and to take into account the 3D contextual relationships among voxels, a fully connected 3D conditional random field (CRF) was applied to fine-tune the segmentation results from the CED networks. In the fully connected 3D CRF process, a maximum aposteriori (MAP) inference was defined over the 3D volume of the whole knee joint. In the model, the probability results for each label from the CED network were used to generate the unary potential on each voxel, and the original 3D knee image volume was used to calculate the pairwise potentials on all pairs of voxels. The iterative CRF optimization is carried out by minimizing the Gibbs energy. In particular, the processed labels representing cartilage and bone are discretized into 3D simplex meshes by using the marching cube algorithm and sent to the 3D simplex deformable process. In the deformable process, each individual segmentation object was refined to preserve smooth tissue boundaries and maintain overall anatomical geometry. The final 3D segmentation was obtained by merging all the segmentation objects. It should be noted here that the later model (3D deformable model) did not participate in tendon segmentation. For tendon segmentation, it was clearly demonstrated in the results that the performance of the combined model

of convolutional encoder-decoder architecture with a fully connected 3D conditional random field performed better than using convolutional encoder-decoder architecture alone and also outperformed the two state-of-the-art 3D CNNs, deepMedic [122] and V-Net [123]. Martins et al. [125] proposed a segmentation approach of the extensor tendon in ultrasound images of the second metacarpophalangeal joint (MCPJ). Their segmentation approach was based on an active contours framework, preceded by a phase symmetry pre-processing and with prior knowledge energies to automatically identify the extensor tendon. The dataset consisted of images of the dorsal view of the second metacarpophalangeal joint extracted from 83 patients resulting in a total of 175 images. The set of images was unbalanced, with more cases without pathology (137) than with pathology (38). Martins et al. looked for the number of points, N, knowing that high values would allow more complex shapes but might also promote local minimums and increase the computational time. On the other hand, low values would make the line less deformable and it might not be able to fit all needed shapes. Given this trade-off, the value of N was supposed to be the lowest possible without compromising the results. It should be noted that local minimums could be reduced using additional constraints, such as the E_{slope} and $E_{\text{concavity}}$ (internal energy). To study N, the authors extracted points equally spaced from the manual segmentation. Those points were then used to create a new approximated tendon using spline interpolation. The error was calculated using the MHD between the interpolated and reference tendons. Regarding external energy, pre-processing techniques were used to reduce the noise and enhance the objects of interest. The authors proposed the use of phase symmetry to enhance the extensor tendon. As previously shown, the phase symmetry had several parameters, being the most important the angle span and the bandwidth. In addition to internal and external energies, area constraints were also taken into account. The position of the tendon was expected to be somewhere between the skin line and the metacarpus and phalanx bones. In order to incorporate this knowledge in the model, authors proposed to use the area between the tendon and the skin line, A_{Up}, and the area between the bones and the tendon, A_{Down}, as references. It was expected that

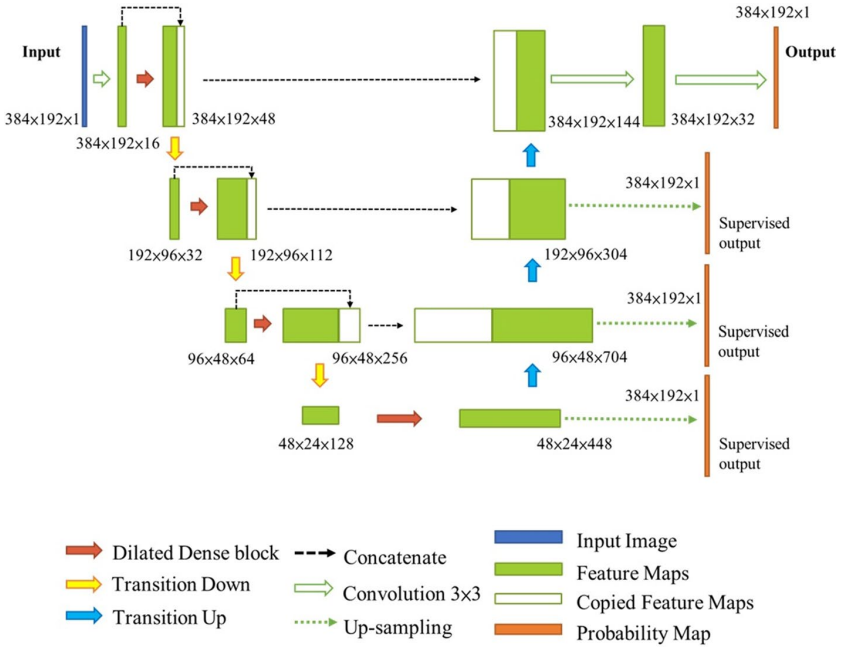


Figure 1.8: Segmentation Network proposed by [124]

those areas would have similar values, with some variations due to natural anatomical differences. One way to quantify that was to calculate the percentage of the area above the tendon, referred to as A_{Ratio}.

$$A_{Ratio} = A_{Up}/(A_{Up} + A_{Down})$$

In order to promote better results and faster convergence, the authors decided to impose some additional constraints. The lower and upper bound limits were used to reduce the search space of the points N. Several energy metric configurations were compared using the Modified Hausdorff Distance and results showed that this segmentation was not only possible but exhibited errors smaller than 0.5mm with a confidence of 95% with the phase symmetry pre-processing and energies based on the line neighborhood, area ratio, slope, and concavity measurements. It should be mentioned here that the authors did not compare their segmentation performance with any other method. Kuok et al.[124] proposed a unique finger tendon segmentation technique, which may also be applied to synovial sheath segmentation to produce a more thorough description for analysis. In this study, a hybrid of effective convolutional neural network techniques was applied, resulting in a deeply supervised dilated fully convolutional DenseNet (D2FC-DN), which demonstrated excellent segmentation performance on the tendon and synovial sheath. Two datasets of ultrasound images of fingers at the A1 pulley, captured in the transverse view, were used in this study. The first set was Chuang's dataset which consists of ultrasound images, of which 74 are finger tendon images and 57 are synovial sheath images. The second set consisted of eight groups of images captured for building a model of finger tissue for a surgical training system, termed the modeling-building (MB) dataset. A total of 1035 images were acquired, with approximately 90–200 images in each group. *Dilated convolution and Deeply supervised nets*: The dilated convolution [126] applied multi-scale information in segmentation tasks without losing resolution. After applying this technique, the receptive field of the convolution can be exponentially expanded. Deeply supervised nets [127] minimize classification error, while the learning process of hidden layers is direct and transparent. Deeply supervised nets introduce a common objective to the individual hidden layers instead of the overall loss function and then use the layer-wise training strategy to enhance the classification capability of the feature maps. The dilated convolution and deep supervision are CNN techniques that can further improve the details of segmentation. However, a CNN structure that combines both of these is rarely seen. Authors attempted to integrate the concepts of the deeply supervised net and dilated convolution into the FC-DenseNet, resulting in a new CNN termed "Deeply supervised dilated FC-DenseNet (D2FC-DN)". Their work proposed a hybrid CNN technique method for finger tissue segmentation, in which the FC-DenseNet was the basic structure, and dilated convolution was combined to allow multi-scale feature map information without losing resolution. Then, deeply supervised learning was applied to increase the transparency of the hidden layers to ensure that each hidden layer output was trained under supervision. In the experiments conducted in their work, the new CNN was used to segment the finger tendon and synovial sheath from transverse ultrasound images. Fig. 1.8 represents the segmentation architecture proposed by the authors. For the evaluation of tendon and synovial sheath segmentation, the proposed D2FCDN got the best performance when compared to the other CNN methods. D2FC-DN outperformed the other deep models because it combined several effective CNN techniques to achieve their advantages and improve the segmentation results. A region-adaptive network (RAN) was proposed by Zhou et al. [128] to local-

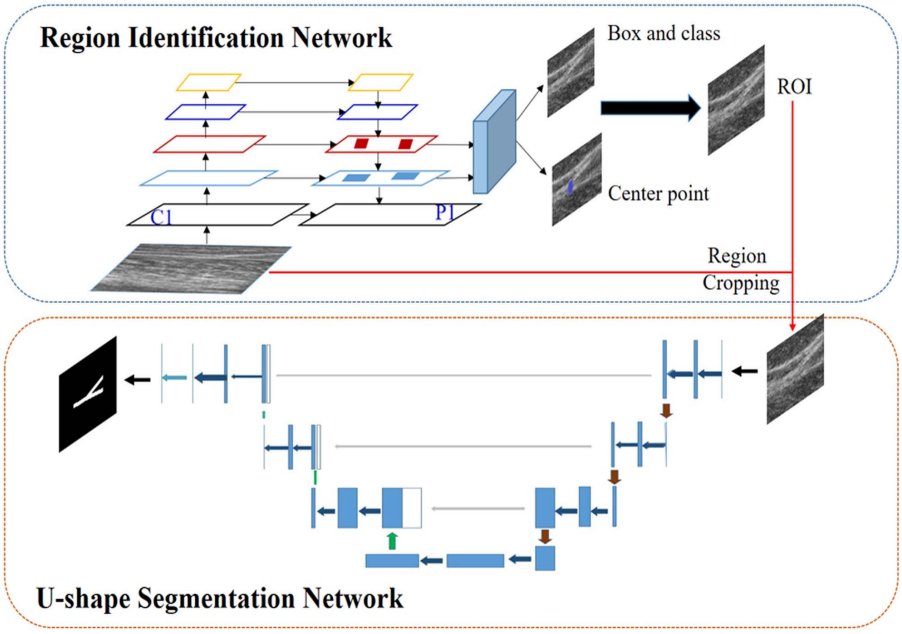


Figure 1.9: Segmentation Network proposed by [128]

ize the myotendinous junction (MTJ) region and to segment it in a single shot. This myotendinous junction consists of muscles and tendons play a crucial role in force generation and kinetic energy mechanism during human movement [129]. Here the model learned about the salient information of MTJ with the help of a composite architecture. Herein, a region-based multitask learning network explored the region containing MTJ, while a parallel end-to-end U-shaped path extracted the MTJ structure from the adaptively selected region for combating data imbalance and boundary ambiguity. A data set consisting of 10 ultrasound sequences (10 healthy adults) was built. Each sequence consisted of 800 to 1200 images with an image resolution of 0.11 mm/pixel. Fig. 1.9 represents the proposed network by Zhou et al. [128], a novel region-adaptive network, called RAN. In this new model, authors advanced the proposal-based Faster RCNN architecture by integrating a landmark regression branch for adaptively selecting the MTJ region more accurately, followed by a U-shaped segmentation network that extracts the multilevel semantic information from the adaptive region. The backbone of the region proposal network (RPN) augmented the ResNet-101 equipped with the feature pyramid network (FPN), which has been widely used in the proposal-based network to generate a sparse set of ROIs efficiently [130]. The authors used transfer learning to fine-tune the ResNet-101 backbone pretrained on a large-scale MS COCO data set, exploring the underlying characteristics more efficiently. In addition, the authors initialized the ResNet-101 backbone C1–C5 with the parameters well-trained on the MS COCO but initialized other layers from uniform distributions. The output of the myotendinous junction RPN built on the ResNet-101 backbone was the bounding box position and center-point landmark to indicate the MTJ anchor with high localization accuracy. Later the adaptive region was cropped from the original ultrasound image for achieving better-balanced data distribution and for obtaining feature maps from different layers, thus enhancing the segmentation of the salient MTJ structure. The authors also applied a hybrid loss of the Dice function and the cross-entropy function to achieve a better balance between the foreground and the background, preserving more segmentation details in the adaptive region [131]. In addition, to better

tackle the problem for successive ultrasound images, authors utilized prior anatomical knowledge of MTJ to encode the mutual dependency in sequential images, overcoming the ambiguity of noisy ultrasound images. *Hybrid loss function proposed by [128]*: In the region identification step, to improve the region identification, a multitasking loss function was defined.

$$L = L_{cls} + \alpha L_{box} + \beta L_{cen}$$

where L_{cls} and L_{box} are the classification loss and bounding-box regression loss, which are the same as those defined in [130]. On the other hand, the authors chose to predict the center using the center offsets regression for each well-aligned anchor box, which was similar to the bounding-box regression. The center-point regression, L_{cen} was then derived for exploring the spatial information of the object in a unified scheme. By contrast, the feature maps obtained by the fully connected layers (FCLs) share the parameters at different spatial locations and may lose part of the position information. Unlike the FCLs, the multitask loss, therefore, became a multi-point regression to encode more position clues for enhancing the identification of the objects. α and β denoted the weights for L_{box} and L_{cen}, respectively. The regression losses, L_{box} and L_{cen}, were only activated for positive anchors and are deactivated otherwise. In the segmentation network, the hybrid loss used is the following:

$$L_{hybrid} = m.L_{DSC} + w.L_{cross}$$

The authors applied a hybrid loss composed of the Dice loss and the cross-entropy function to achieve a better segmentation of MTJ. The Dice loss function handled the data imbalance problem, while the cross-entropy loss forces the model to learn and train better on poorly classified pixels. Where L_{hybrid} denoted the total hybrid loss which is the combination of L_{DSC} and L_{cross}. $m \in [0, 1]$ and $w \in [0, 1]$ represented the weights of the Dice loss and cross-entropy loss respectively. 0.5 was set as the corresponding weights for m and w, respectively, because of the relatively balanced distribution between positive samples and negative samples in the small cropped region. The main objective of this study [132] was to evaluate if deep learning is a feasible approach for the automated detection of supraspinatus tears on MRI. A 3-stage pipeline was developed comprised of a slice selection network based on a pre-trained residual neural network (ResNet); a segmentation network based on an encoder-decoder network (U-Net); and a custom multi-input convolutional neural network (CNN) classifier. A total of 200 shoulder MRI examinations performed between January 2015 and August 2019 were obtained containing a full-thickness tear, partial-thickness tear, or intact supraspinatus tendon. Only T2-weighted fat-saturated sequences were included for use in the model, given their optimal sensitivity for supraspinatus tears. The model was based on a two-dimensional (2D) convolutional neural network framework with a three-network design (1. Slice Selection Network, 2. Segmentation Network, 3. Classification Network). The purpose of the slice selection network was to identify image slices where the supraspinatus was visible to center the model on the region of interest. Binary ground truth labels of "included" and "excluded" were manually created for each slice using the anterior and posterior coronal boundaries of the supraspinatus muscle body as reference. The network design was a transfer-learning approach based on a fine-tuned 50-layer ResNet with pre-trained weights from the ImageNet dataset. The output was the probability for each image slice to contain supraspinatus; these values were aggregated to select the 7 central slices with the highest probability scores. The

selected image slices were used as input to the subsequent segmentation network. The segmentation network generated an image mask for each selected MR slice. The network was based on a randomly initialized U-Net that returned a probability mask for the supraspinatus region. This was automatically post-processed to remove any small non-connected regions and horizontally mirror images to have a uniform orientation. The authors of [132] claim that their study was the first published implementation of deep learning for segmentation of the supraspinatus on coronal oblique MR images. Segmentation of coronal oblique images of the the supraspinatus is more challenging given the narrow structure and anatomical variation between patients, which likely contributed to the relative underperformance of their model compared to the upper accuracy bounds for sagittal segmentation quoted in the literature. Ramakrishnan et al. [133] developed a novel concept for the computerized segmentation of ultrasound images of connective tissue based on movement. To demonstrate this concept, a central cross-section of the patellar tendon was imaged in the axial plane while the movement was generated by manually pulling and pushing the skin close to the imaging area. Maps of the internal movement were created for four representative pairs of consecutive images using normalized cross-correlation. Thresholding followed by a series of morphological operations (k-clustering, blob extraction, curve fitting) enabled the extraction of the superficial most tendon boundary. Tendons and ligaments are both connective tissues characterized by high concentrations of collagen with a structured orientation. Their unique internal structure and composition also lead to a relatively distinctive presentation in ultrasound imaging as well as causing specific imaging artifacts (e.g. dependence of tissue echogenicity on the relative orientation between probe and collagen fibers) [134]. Tendons are connected to muscle and bone and their role is to facilitate movement while ligaments connect bones together to offer support. Even though their function is different their structure and appearance in ultrasound are similar. There is a high acoustic impedance difference between connective tissue and the soft tissue that surrounds it leading to the creation of hyperechoic (bright) regions around tendons/ligaments. However, this bright region could expand within the connective tissue as well as in the surrounding soft tissue making the exact segmentation of the interface between tissues very challenging [133]. Besides, the production of accurate training datasets is made very challenging by the speckle noise, discontinued boundaries, and low contrast of ultrasound images [135]. To address this inherent problem of low image quality in ultrasound, Ramakrishnan et al. [133] proposed to include information on movement in the segmentation for connective tissue. Instead of analyzing a static image, movement-based segmentation analyses a series of images of a moving object to segment based on a movement differential between the targeted object and its background. Even though the analysis of videos can pose its own challenges [136] it also enables using motion as a feature for segmentation when the region of interest is unable to be differentiated from its surroundings based on color, intensity, texture, and gradient features [137]. Therefore, the application of such segmentation approaches relies on accurate methods for motion estimation. At the same time, the application of movement-based segmentation methods in ultrasound would also require reliable and reproducible methods for generating distinctive movement patterns in the targeted tissue [133]. The proposed method involved the manual generation of movement in the tissue and the use of normalized cross-correlation for the mapping of movement between consecutive ultrasound images. The movement was generated by pulling and pushing the skin near the imaged tissues. This movement pattern generated visible differences in displacement close to the skin surface between the patellar

tendon and the subcutaneous tissue along the Y axis (i.e. the direction of the pull/-push movement) enabling the accurate segmentation of the tendon's superficial-most boundary. However, the used external movement pattern did not appear capable of generating substantial displacement deeper in the tissue, which would be a barrier for accurate movement-based segmentation of the tendon's deeper-most boundary. The segmentation of the entire tendon would require movement patterns that affect equally the entire tendon outline (e.g. generated by a contraction of the in-series muscle). The authors came to the conclusion that movement-based segmentation was utilized by itself to show the viability and potential benefit of this strategy [133].

1.3.4.2 Criticism on Existing Studies

While there are studies that have attempted to segment anatomical structures using advanced techniques, such as convolutional-encoder-decoder architectures, the number of studies focusing specifically on tendon segmentation is limited. The existing literature shows that many approaches depend on the expertise of musculoskeletal specialists for accurate grading and diagnosis, which introduces a degree of variability and may lead to suboptimal outcomes. For example, grading systems based on expert opinion require significant time and resources, which can hinder the scalability and reproducibility of findings across different clinical settings. Moreover, the reliance on handcrafted features for segmentation poses challenges. Techniques like active contour and active shape models are often compromised by the presence of speckle noise and imaging artifacts inherent in ultrasound images. These limitations make it difficult to produce accurate training datasets, further complicating the segmentation process. Consequently, there is a pressing need for automated systems that can extract relevant features without human intervention, reducing the potential for inter- and intra-observer variability. The literature also emphasizes the importance of tendon thickness as a crucial feature in assessing tendon-related pathology. However, many studies do not adequately address how to automate thickness measurements, which could enhance diagnostic accuracy. Current clinical management of tendinopathy often relies on a combination of subjective assessments and imaging results, underscoring the necessity for a more integrated approach that combines clinical and imaging data. Additionally, while radiomics offers promising avenues for enhancing diagnostic accuracy by extracting high-dimensional features from medical images, its application in tendon pathology remains underexplored. The focus should be on maximizing the robustness and repeatability of these features, as variations in imaging settings can significantly impact texture values in heterogeneous tissues.

In conclusion, there is a critical need to prioritize automated tendon segmentation and the integration of imaging data with machine learning techniques. A significant consideration in this matter is the availability of data, which poses a challenge for training deep learning models. To address this limitation, it is crucial to design a pipeline that can effectively extract more meaningful features from the available data, especially when working with limited datasets. By optimizing data utilization, this approach will enhance the performance and reliability of the diagnostic models developed for tendon related studies. By incorporating the concept of superpixels, we can enhance the representation of tendon structures beyond traditional pixel-based approaches, allowing for more meaningful segmentations that capture the underlying anatomy more effectively. Additionally, by leveraging graph-based methods instead of solely relying on images, we can model the relationships between anatomical structures

and their spatial contexts more accurately. Overcoming these challenges will lead to the development of more reliable diagnostic models, ultimately enhancing clinical outcomes. More detailed discussions on superpixels and graph-based methodologies will be provided in later sections.

1.3.4.3 Inspiring Studies Shaping Our Proposal

Superpixels

Recently, superpixel segmentation has attracted a lot of interest in computer vision as it provides a convenient way to compute image features and reduces the complexity of subsequent image-processing tasks. Many superpixel segmentation algorithms have been proposed in recent years [138], [139], [140], [141], [142], [143], [144]. The value of superpixel lies in its applications in many fields including object recognition [145], image segmentation [146], object tracking [147], video segmentation [148], classification [149], and reconstruction [150]. Each method has its own advantages and disadvantages.

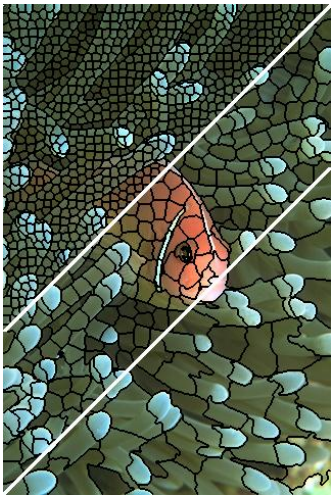


Figure 1.10: Sample Superpixels

- Basics of Superpixels
  Pixels serve as the fundamental building component for most image-processing tasks. However, pixels are a result of the discrete representation of images, not of natural entities. Segmentation is frequently used to separate interesting and unattractive items in any image. As a result, it divides the image into a number of sets composed of homogeneous regions with comparable properties. For many applications in the area of computer vision and image processing, the superpixel generation has attracted a lot of attention in the past years. Ren and Malik ([151]) first proposed the idea of the superpixel as perceptually homogeneous regions. In comparison to pixel representation, superpixel representation decreases the number of image dependencies and offers better support to identify regions depending on image properties [152]. Superpixels are perceptual groupings of pixels or over-segmented segments of any image. This over-segmentation aims to partition the image so that no superpixel is split by an object border, but that objects can be split into several superpixel. In later processing steps, the outlines of the object can be extracted from the superpixel borders in this way [153].

- Categories of Superpixel Generation Algorithms
  Superpixel generation algorithms can be categorized into graph-based methods and gradient-ascent-based methods. In graph-based methods, every pixel is handled as a node in the graph, while the similarity between neighboring pixels is proportional to edge weights between the two nodes. By minimizing the cost function, superpixels are created and defined over the graph. superpixels are generated by minimizing an energy function defined on the graph. The graph-based algorithms include Efficient graph-based image segmentation [154], superpixels and super-voxels in an energy optimization framework [155], Normalized cuts and image segmentation [156], Graph cut textures [157], Entropy Rate superpixels (ERS) [155], Lazy Random Walks superpixel (LRW) [158], and other state-of-the-art methods.

  The gradient ascent algorithm is based on the initial pixel clustering and corrects the clustering outcomes iteratively using the gradient method until the convergence requirements are satisfied, at which point it produces a number of regions of superpixels [159]. Some of the algorithms that utilize the gradient ascent technique include quick-shift and kernel methods for mode seeking [160], turbopixels method which is fast superpixels using geometric flows [161], mean-shift which is a robust approach toward feature space analysis [160], watersheds in digital spaces which is considered as an efficient algorithm based on immersion simulations [162][163][164] and a simple linear iterative clustering (SLIC) and g-SLIC algorithms which address a lot of requirements and outperforming other state-of-the-art algorithms.

  According to the literature (discussed more in detail later), superpixels can be crucial from a number of perspectives in the case of medical image segmentation. It should also be highlighted that Simple Linear Iterative Clustering (SLIC), along with other methods that were heavily influenced by SLIC, has been widely employed in most of the superpixel generation algorithms. The SLIC algorithm generates superpixels from an input image by grouping its pixels based on both their spatial proximity and their intensity similarity. In addition to having various superpixel generating algorithms, the generation of variable-size superpixels, where the variable size depends on predefined criteria based on the pathological context, may be of significance for particular studies. For those superpixels of variable size, various features can be further extracted and applied to classification. Here, it is essential to precisely define the predefined criteria which should be further verified by the specialists. When the objective is that each superpixel should be able to consist of small-sized areas of interest (ROIs), as in the [165] study, the variable superpixel sizes are of greatest interest. This is crucial because the features that will be retrieved from those superpixels will accurately reflect the characteristics of the ROIs. Regarding superpixels merging based on the similarities between the superpixels, there is a good chance that the newborn curves of the superpixels are the contour of the regions of interest. However, there are times when the edge is faint or even disconnected, so in those situations incorporating the information from the contour map can be helpful. superpixels can be used for local boundary optimization, followed by some post-processing techniques [166]. Super pixels have also been used in the literature ([167]) to create explanatory maps for prediction, and because segmentation performance

is so crucial, it is crucial to have robust segmentation. superpixel segmentation can be performed on activation maps, according to existing prior research, but only when the superpixel size is of constant values for the horizontal and vertical pixel window size which constrains the size of each superpixel. Changing these windows can make it easier to collect biomarkers of any size, from tiny nodules to massive infection patterns, but computational costs should also be taken into consideration.

- Superpixel based Studies
  The literature found using the search terms **superpixels AND medical image segmentation** are highlighted in this section as a result of our systematic literature review using the MEDLINE database. The last time this search was performed was in September 2022. There were 17 total results from this search, 5 of which were disqualified from the abstract reading stage. The remaining 12 articles' key points are highlighted below.

  Xu et al. [168] developed a method using a machine learning algorithm based on variable-size super pixel segmentation that efficiently utilized the full dataset to improve the discrimination between early glaucomatous and healthy eyes. This study was conducted with healthy, glaucoma suspects and glaucomatous eyes with a wide range of disease severity selected from the Pittsburgh Imaging Technology Trial (PITT). 192 eyes of 96 subjects (44 healthy, 59 glaucoma suspects,s and 89 glaucomatous eyes) were enrolled to test the glaucoma discrimination performance. As a very first step, the 3D Optical Coherence Tomography (OCT) image data was converted into a 2D feature map by minimizing the blood vessel effect and at the same time generated a retinal nerve fiber layer (RNFL) thickness map. After that, Variable size/ shape superpixels were automatically mapped on the 2D feature map by grouping homogeneous neighboring pixels using a ncut algorithm [169]. 100 super pixels were initially segmented on the feature map. The size of each super pixel was automatically adjusted with the pre-defined criteria based on the pathologic contexts of glaucoma. To be more sensitive to RNFL thinning (glaucomatous damage), smaller super pixels were assigned to thinner RNFL. Each initially segmented super pixel was recursively partitioned into N more super pixels, while N was a function of mean, standard deviation, size, and deviation to the normative database of the given super pixel compared with the global mean and standard deviation of the 2D feature map. The segmented feature map provided a qualitative analysis with more natural representation, in which damaged areas tended to have smaller super pixels, while normal regions had larger super pixels. To summarize the map as quantitative disease indices, a total of 68 super pixel features were extracted and used as the inputs of machine learning classifier analysis. Three thresholds, Tmin, T1, and T2 were set to 50, 133, and 400 pixels respectively based on the experiments to obtain two sub-groups of super pixel with large size and small size. The features were calculated from all super pixels together with two sub-groups. Glaucoma classification was performed by implementing LogitBoost adaptive boosting algorithm, [170] which was designed as a supervised two-class machine classifier. Xu et al. [168] came into the conclusion that the super pixel processing with machine classifier analysis generated from 3D SD-OCT data has the potential to improve the early detection of glaucomatous structural damages. This paper

[171] proposed a novel integrated spatio-colour-texture based graph partitioning method for segmentation of nuclear arrangement in tubules with a lumen or in solid islands without a lumen from digitized Hematoxylin–Eosin stained breast histology images. A new similarity-based super pixel generation method was proposed which was integrated with texton representation to form spatio-colour-texture map of Breast Histology Image. Then a new weighted distance based similarity measure was used for generation of graph and final segmentation using normalized cuts method was performed. The ground truth (GT) image database of about 100 Breast Histology images (60 Normal and Benign, and 40 malignant), created with the help of expert pathologist was used. The very first step of the method was the superpixels generation. The proposed algorithm was an adaption of 'Simple Linear Iterative Method' (SLIC) algorithm which combines pixels based on their location and intensity similarity. This was done in RGB colour space considering intensity of pixel, its location, and label assigned to each initial super pixel in a 6 dimensional feature vector. Belsare et al. [171] showed that their proposed method segmented breast histology images more correctly than state of art methods in computer vision. Zhu et al. [172] proposed a novel lung cancer detection method for CT images based on the superpixels and the level set segmentation methods. The dataset used was a set of lung CT image data provided by Shanghai Chest Hospital, which included 104 groups of the lung CT images of the patients. For lung region segmentation, SLIC (Simple Linear Iterative Clustering) super pixels was utilized to segment lung field considering the CT image-forming process in the proposed method. The similarity relationship of the superpixels was calculated, and the result was formed as a similarity matrix. The matrix was utilized to decide whether or not merging the superpixels into one by thresholding its elements. Since there was the strong possibility that the newborn curves of superpixel would be the contour of the lung parenchyma, but sometimes the edge can be a weak or even disconnection, hence the authors introduced UCM (Ultra-metric Contour Map) to obtain the ground truth edges. An iterative procedure of merging super pixels, to assist that all the super pixels merged its similar neighbor as much as possible. And an energy function was employed to define the strength of each curve. Gaonkar et al. [165] developed a multi-parametric ensemble learning technique to automatically detect and segment lumbar vertebral bodies using MR images of the spine. Spine imaging data were used to illustrate their techniques since low back pain is an extremely common condition and a typical spine clinic evaluates patients that have been referred with a wide range of scanning parameters. The dataset used contained lumbar sagittal cases of 48 sagittal T2-MR scans and 15 T1-MR scans. The input images were first segmented into super pixels following [173]. Later for each superpixel in the imaging data, the ratios were extracted that quantified the shape, the position, and the orientation of the superpixel, which are the following:

- Computing the centroid of the superpixel and dividing them by the corresponding maximal dimension of the image
- Computing orientation of each superpixel and its eccentricity
- Computing the ratio of squared perimeter and area

- Computing the ratio of the area of the superpixel to the area of its convex hull and
- The ratio of the area of the superpixel to the area of the bounding box

And since the superpixel generation stage is very crucial as the ideal superpixel would be in such a size that would contain each vertebra so authors proposed running multiple algorithms or even the same algorithm at several parameter settings which would effectively overcome this issue, thus yielding a multi-parameter ensemble. The extracted features were later used in a random forest classifier to classify whether each superpixel contain the vertebra body or not. To refine the superpixels picked out by the random forest, a limited post-processing protocol was applied that was designed to force the superpixels identified to exactly correspond to grayscale edges in the original intensity image. The authors compared their methodology with three other methods (watershed transform-driven segmentation[174] [175], the related random walker based segmentation [176] and active shape modeling) and their proposed method was found to be superior and robust enough that was able to identify and segment vertebral bodies on CT even after training only on MRI images. A novel fully automatic pipeline was proposed by Yang et al. [177] to achieve accurate and objective segmentation of the heart and of scar tissue within the heart acquired in patients with atrial fibrillation (AF). 37 cases with fair image quality were retrospectively entered into this study including 11 preablation (included 65% of preablation cases) and 26 postablation scans (included 92% of post-ablation cases). For artial scarring segmentation, authors proposed to use a Simple Linear Iterative Clustering (SLIC) based superpixel method, which had been successfully applied to solve various medical image analysis problems. Based on local k-means clustering, the SLIC method iteratively groups pixels into superpixels. The clustering proximity is estimated in both intensity and spatial domains. Superpixel-based Fuzzy c-means clustering (SPOFCM), which took into account the influence of spatially adjacent and similar superpixels, was proposed by Kumar et al. [178]. Additionally, a crow search method was used to maximize the power of influential degrees. The dataset used was the renowned Mammographic Image Analysis Society (MIAS) database which included mammogram images with 1024 $\times$ 1024 pixels, and from this, the actual single-spectrum image data was selected, whereas from the hospitals, the 512 $\times$ 512 pixels comprising breast MRIs were collected, and from this, the actual multi-spectral images was selected. Further, from the OASIS database, 512 $\times$ 512 pixel-sized brain images were collected (http://www.oasis-brains.org/). Firstly wavelet transform was performed on the input image and then Simple Linear Iterative Clustering based segmentation was performed. Next, on each superpixel its neighbouring superpixels are searched. To the exterior of neighbouring super pixels, the number of superpixels was found. The membership for each superpixel and the cluster centres were updated iteratively. The Crow Search Algorithm (CSA) is used for optimizing the influential degree, $\alpha 1$ and $\alpha 2$. By optimizing the influential degree, the segmentation performance eventually improved. This segmentation outcome was compared to renowned segmentation techniques such as including k-means, entropy thresholding (ET), Fuzzy c-means clusterring (FCM), FCM with spatial constraints (FCMS) and Kernel FCM. Results on multi-spectral MRIs and actual single-spectrum mammograms indicated that the proposed algorithm can provide

a better performance for suspicious lesion or organ segmentation in computer-assisted clinical applications. Since the Diffusion tensor imaging (DTI) data is a tensor space, its segmentation is different from ordinary MRI images. Based on the existing deep learning model, an improved image semantic segmentation method based of superpixels and conditional random fields was proposed by Mu et al. [179]. The self-built dataset used consisted of 7000 medical DTI image data from several hospitals and some data have been labeled by imaging specialists. The public data set was from the medical imaging data set DeepLesion released by NIHCC [166], which contained more than 32,000 lesion annotations from more than 10,000 cases. To extract rough features in the images, a fully convolutional network (FCN) model was implemented which performed the classification at the pixel level whose classification label was the label of a single pixel. Next, the Simple Linear Iterative Clustering superpixel segmentation algorithm was used to segment the input image and obtain the superpixel segmentation image. Then combining the two kinds of segmentation images obtained in the first two steps, the proposed optimization algorithm was used to optimize the local edges of the rough features obtained in the first step. After edge optimization, the segmentation results had more accurate edge information than the results of FCN model, and also reduced the diffusion error in the process of up-sampling. Later, a post-processing step was also performed by utilizing a fully connected conditional random field model. The authors concluded that the use of superpixels for local boundary optimization can improve the segmentation accuracy. Signoroni et al [167] utilized superpixel based segmentation for generating explainability maps with an original technique, to visually help the understanding of the network activity on the lung areas. The dataset contained almost 5000 Chest X-Rays (CXRs), which was assumed representative of all possible manifestations and degrees of severity of the COVID-19 pneumonia. To predict the pneumonia severity of a given chest X-Ray (CXR), authors proposed a novel architecture where different blocks cooperate, in an end-to-end scheme, to segment, align, and predict the Brixia score. Later, to evaluate whether the network was predicting the severity score on the basis of correctly identified lung areas, a method was required which was capable of generating explainability maps with a sufficiently high resolution. The creation of the explainability map starts with the input image division into N superpixels. Starting from the input image, N image replicas were created in which a single superpixel i (from 1 to N) was masked to zero. Here, p_0 was the probability map that the model produced starting from the original image. Instead p_i the probability map produced from the i^{th} replica. Then it was accumulated for the differences between all the superpixel masked predictions p_i and the original prediction p_0. In the case of explainability, despite the automated AI-driven scoring is not meant to eliminate radiologists' evaluations, authors have shown that it can be used to aid and streamline the reporting workflow, and improving the timeliness of first evaluation, by proposing a preliminary interpretation of findings. Prakash et al. [180] proposed a deep learning based framework to enhance the diagnostic values of chest X-ray images for improved clinical outcomes. It was realized as a variant of the conventional SqueezeNet classifier with segmentation capabilities, which was trained with deep features extracted from the Chest X-ray images of a standard dataset for binary and multi class classification. Here Gradient based class activations coupled with super pixel segmentation provide best classification and segmenta-

tion abilities of the SqueezeNet based model. The training and testing data were constructed from the Kaggle CXR public database, comprising 219 COVID-19 positive images, 1345 viral pneumonia images and 1341 normal images. First, a deep transfer learning employing the pre-trained SqueezeNet model, built by extending the SqueezeNet classification model with Grad-CAM and super pixel pooling. In this framework, the Grad-CAM was applied on the image feature maps from the final convolutional layer to construct the activation map. This heat map was further given as input to the Super Pixel Pooling Layer (SPP) for segmentation and a normalized class score in the range [0 1] was computed from the active super pixels. For the computation of the class score for the segmented heat map, for each superpixel, an activation score was assigned corresponding to the mean color of the super pixel as 6-Red, 5- Orange, 4-Red-Orange, 3-Maroon, 2-Yellow, 1-Olive, 0- Gray, Blue, Black. The class score was computed as the summation of the activation scores and normalized in the range [0 1] by min–max normalization. The final softmax layer assigns the target class label to the input images based on the class score of each image and the input from the Global Average Pooling layer of the network. The proposed method showed that the SqueezeNet model incorporating a superpixel pooling layer for segmentation of the infected region paves the way for further research on the investigation of the symptoms in the segmented regions and combinations of occurrences in CXR images, for standardizing them as potential biomarkers of COVID-19. The proposed framework [181] for COVID19 detection and segmentation of infections from chest X-ray images is a two-stage cascaded framework with classifier and segmentation sub-network models. The classifier was modeled as a fine-tuned residual SqueezeNet network, and the segmentation network was implemented as a fine-tuned SegNet semantic segmentation network. The segmentation task is enhanced with a bioinspired Gaussian Mixture Model-based super pixel segmentation. Two different datasets were employed where the Kaggle chest X-rays dataset contains 1143 COVID19-positive images, 1345 viral pneumonia images and 1341 normal images and The Cancer Imaging Archive (TCIA) dataset contains 221 chest X-rays and their corresponding infection masks. A pre-trained SqueezeNet was used which strengthened with residual connections for faster learning and fine-tuned with the Kaggle Chest X-ray (CXR) images, preserving the original weights. This approach significantly improved the prospective of the classifier network to learn discerning features from CXR images for improved COVID19 classification. TCIA dataset was employed exclusively for testing the classifier model, and training and testing the segmentation sub-network. Once the classifier was trained, Gaussian mixture model (GMM) based superpixel segmentation was performed on the activation maps obtained from the classifier. Next, those superpixels were classified as background or RoI based on the entropy. The random numbers for Levy fight generation are obtained with the McCulloh's algorithm to select the GMM super pixels in each iteration. The super pixels with high entropy were added to the RoI and the others are discarded. Later, the pre-trained SegNet semantic segmentation network was trained to segment the opacities from the RoIs segmented from the CAMs. The SegNet was trained with these images superimposed with the segmented CAMs and the corresponding segmentation masks as labels. The performance of the segmentation model was enhanced with ROI segmentation of the CAMs employing GMM super pixels. This research paved a new approach, integrating the classification

and segmentation tasks for COVID-19 management. Li et al. [182] proposed a medical image segmentation algorithm based on Simple non-iterative Clustering (SNIC). The experiments were conducted in Berkeley Segmentation dataset and the medical CT image dataset. Firstly, the feature map of the image was obtained by extracting the texture information of it with a feature extraction algorithm. Secondly, the image size was reduced to a quarter of the original image size by downscaling. Then, the SNIC superpixel algorithm with texture information and adaptive parameters was used to segment the downscaling image to obtain the superpixel mark map. Here, Compared with SLIC (Simple Linear Iterative Clustering), which requires multiple iterations to converge superpixel centers, SNIC (simple non-iterative clustering) can update cluster centers online in one iteration. Starting from the initial seed point, the SNIC algorithm uses a priority queue to select the next pixel to add to a superpixel. Finally, the superpixel labeled image was restored to the original size through the idea of the nearest neighbor algorithm. Although the proposed method had no significant increase in the evaluation coefficient of segmentation, it was evident from the qualitative point of view that the proposed method was more suitable for the detailed segmentation of vertebrae, and the regions of the same category were clustered together, which had a good segmentation effect and had a faster image segmentation speed. In [183], Li et al. proposed a novel network enhanced by a self-attention module for intelligent dental plaque segmentation. The algorithm needed to conduct self-attention at the superpixel level and fuse the superpixels' local-to-global features. 2,884 labeled images using 3 different endoscopes was used, and each image's size was 700 $\times$ 700 pixels. The newly designed network architecture afforded the simultaneous fusion of multiple-scale complementary information guided by the powerful deep learning paradigm. The critical fused information included the statistical distribution of the plaque color, the heat kernel signature (HKS) based local-to-global structure relationship, and the circle-LBP-based local texture pattern in the nearby regions centering around the plaque area. Here HKS algorithm was used as a feature filter, of which, we use the RGB, position, and LBP information of the superpixels as the initial heat points. To further refine the fused multiple-scale features, an attention module was devised based on CNN, which could focalize the regions of interest in plaque more easily, especially for many challenging cases. Extensive experiments and comprehensive evaluations confirmed that, for a small-scale training dataset, this method could outperform state-of-the-art methods. Meanwhile, the user studies verified the claim that this method is more accurate than conventional dental practices conducted by experienced dentists.

Graph Convolutional Network

- Basics of Graph
  A graph $G = (V, E)$ is defined by a set of nodes V and a set of edges E between these nodes. It is denoted that an edge going from node $u \in V$ to node $v \in V$ as $(u, v) \in E$. In simple graphs, where there is at most one edge between each pair of nodes, no edges between the node and itself, and where the edges are all undirected, i.e., $(u, v) \in E \iff (v, u) \in E$. A convenient way to represent graphs is through an adjacency matrix $A \in \mathbb{R}^{|V|\times|V|}$. To represent a graph with an adjacency matrix, the nodes in the graph are ordered so that every node

indexes a particular row and column in the adjacency matrix. The presence of edges is represented as entries in this matrix: $A[u, v] = 1$ if $(u, v) \in E$ and $A[u, v] = 0$ otherwise. If the graph contains only undirected edges, then A will be a symmetric matrix, but if the graph is directed (i.e., edge direction matters), then A will not necessarily be symmetric. Graphs can also be classified as weighted and unweighted. In weighted graphs, each edge has a weight or cost associated with it, which represents the strength or distance of the relationship between the nodes whereas unweighted graphs have no associated weights.

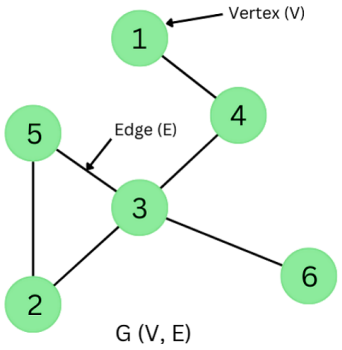


Figure 1.11: Graph

Graphs naturally arise in many real-world applications, including social analysis [184], fraud detection [185] [186], computer vision [187], and many more. By representing the data as graphs, the structural information can be encoded to model the relations among entities and furnish more promising insights underlying the data. However, it is often very challenging to solve the learning problems on graphs, because (1) many types of data are not originally structured as graphs, such as images and text data, and (2) for graph-structured data, the underlying connectivity patterns are often complex and diverse. On the other hand, representational learning has achieved great success in many areas. Thereby, a potential solution is to learn the representation of graphs in a low-dimensional Euclidean space, such that the graph properties can be preserved. Deep learning models on graphs (e.g., graph neural networks) have recently emerged in machine learning and other related areas and demonstrated superior performance in various problems [188].

- Machine Learning on Graphs
  Machine learning on graphs, also known as graph machine learning or graph neural networks (GNN), is a rapidly growing field within the broader domain of machine learning and Artificial Intelligence. It focuses on developing algorithms and models to extract meaningful information from structured data represented as graphs or networks. In graph-based datasets, entities (nodes) are connected by relationships (edges), making them a natural representation for various real-world scenarios. Machine learning on graphs has various applications. It is used in recommendation systems to model user-object interactions, in fraud detection to detect anomalous behavior, in bioinformatics to predict protein-protein interactions, in social network analysis, and in many other fields where data has a relational structure. Overall, it is a dynamic field that addresses complex real-world problems by leveraging the rich structural information present in graph data. It continues to evolve with the development of new models and algorithms

and has the potential to make significant contributions to a wide range of applications.

- Graph Neural Networks (GNN)
  Neural Message Passing
  During each message-passing iteration in a GNN, a hidden embedding $h_u^{(k)}$ corresponding to each node $u \in V$ is updated according to information aggregated from u's graph neighborhood $N(u)$ (Figure 5.1). This message-passing update can be expressed as follows:

$$h_u^{(k+1)} = \text{UPDATE}^{(k)}\left(h_u^{(k)}, \text{AGGREGATE}^{(k)}\left(\{h_v^{(k)}, \forall v \in N(u)\}\right)\right) \tag{5.4}$$

$$= \text{UPDATE}^{(k)}\left(h_u^{(k)}, m_{N(u)}^{(k)}\right) \tag{5.5}$$

  where $\text{UPDATE}^{(k)}$ and $\text{AGGREGATE}^{(k)}$ are arbitrary differentiable functions (i.e., neural networks), and $m_{N(u)}^{(k)}$ is the message that is aggregated from u's graph neighborhood $N(u)$ [189].

  The basic intuition behind the GNN message-passing framework is straightforward: at each iteration, every node aggregates information from its local neighborhood, and as these iterations progress each node embedding contains more and more information from further reaches of the graph. To be precise: after the first iteration (k = 1), every node embedding contains information from its 1-hop neighborhood, i.e., every node embedding contains information about the features of its immediate graph neighbors, which can be reached by a path of length 1 in the graph; after the second iteration (k = 2) every node embedding contains information from its 2-hop neighborhood; and in general, after k iterations, every node embedding contains information about its k-hop neighborhood. Through this information exchange strategy, each individual node within the graph undergoes a process of receiving and aggregating information from its immediate local neighborhood. As iterations of messages passing go on, each node's embedding gradually accumulates a wealth of knowledge derived from even the most distant regions within the graph. Consequently, a graph neural network possesses the ability to construct, for each single node present in the graph, a representation that does not depend exclusively on the intrinsic attributes of that node but is also influenced by the topology of the entire structure of the graph. This information can, therefore, be effectively exploited in various prediction tasks, ranging from node-level predictions to edge-level predictions and even graph-level predictions.

- Graph Convolutional Network - GCN
  One of the most popular baseline graph neural network models - the graph convolutional network (GCN) - employs the symmetric-normalized aggregation as well as the self-loop update approach. This approach was first outlined by [190] and has proved to be one of the most popular and effective baseline GNN architectures.

  Learning on graph-structured data is an efficient variant of convolutional neural networks which operates directly on graphs. The choice of convolutional architecture is motivated via a localized first-order approximation of spectral graph

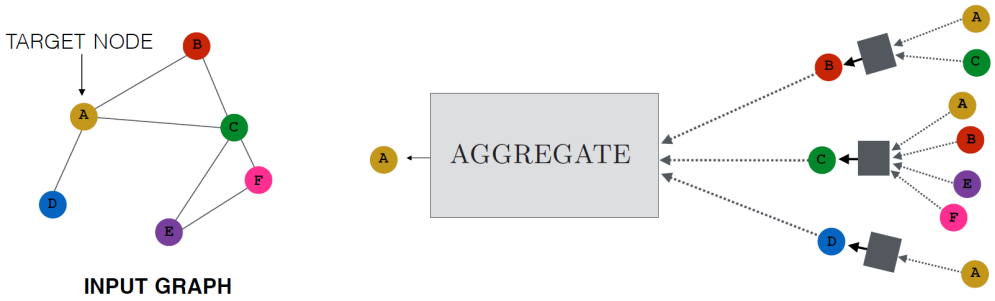


Figure 1.12: Message Passing

convolutions. This type of model linearly scale in the number of graph edges and learns hidden layer representations that encode both local graph structure and features of nodes [191]. Before the arrival of Graph Convolutional Network (GCN), the problem of classifying nodes in a graph was considered where labels are only available for a small subset of nodes. This problem was framed as graph-based semi-supervised learning, where label information is smoothed over the graph via some form of explicit graph-based regularization [192], e.g. by using a graph Laplacian regularization term in the loss function:

$$L = L_0 + \alpha L_{reg}$$

Where,

$$L_{reg} = \sum_{i,j} A_{ij} ||f(X_i) - f(X_j)||^2 \tag{1}$$

Here, A is an adjacency matrix. The formulation of Eq. 1 relies on the assumption that connected nodes in the graph are likely to share the same label. This assumption, however, might restrict modeling capacity, as graph edges need not necessarily encode node similarity, but could contain additional information.

In GCN, a simple and well-behaved layer-wise propagation rule is introduced for neural network models that operate directly on graphs and it can be motivated from a first-order approximation of spectral graph convolutions [193]. The layer-wise propagation rule for a multi-layer Graph Convolutional Network (GCN) is the following:

$$H^{l+1} = \alpha(\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}H^l W^l) \tag{2}$$

Here, $\widetilde{A}$ = A + I is the adjacency matrix of the un-directed graph G with added self-connections. I is the identity matrix, $\widetilde{D}^{-1/2}$ is the degree matrix (normalized), H^l is a stack of feature vectors of every node and W^l is a layer-specific trainable weight matrix, here feature vectors for every node are projected to a lower dimensional space. α denotes a non-linear differentiable activation function.

Eq. 2 is a representation of all the other feature vectors of the connected nodes. Basically, 1 hop neighborhood is considered and the projected features are aggregated and summed up, but before summing the matrix should be normalized

in order to avoid exploding or vanishing gradients. Since GCN is motivated by spectral methods but itself, not a spectral method but rather a spatial method, here it uses a 1 hop neighborhood which reduces computation intensive and approximates the maximum eigenvalue λ_{max} to 2 with an expectation that the parameters of the neural network will adapt to this change in scale during training. And the last approximation is θ instead of using 2 Chebyshev coefficient vectors ($\theta 1$ and $\theta 2$). So, the final equation turn into the following:

$$g_\theta * x = \theta(I + D^{-1/2}AD^{-1/2})x \tag{3}$$

Since I + $D^{-1/2}AD^{-1/2}$ now has eigenvalues in the range [0,2]. Repeated application of this operator can therefore lead to numerical instabilities and exploding/-vanishing gradients when used in a deep neural network model. To alleviate this problem, renormalization trick was introduced where I + $D^{-1/2}AD^{-1/2}$ became $\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}$ [191].

- Applications of GCN in Images
  Image classification is of great importance in many real-world applications. By some carefully hand-crafted graph construction methods (e.g., kNN similarity graphs) or other supervised approaches, the unstructured images can be converted to the structured graph data and thereby are able to be applied to graph convolutional networks (eg. [194] [195] [196]). Another application on images is visual question answering which explores the answers to the questions on images. In addition, as images often contain multiple objects, understanding the relationships (i.e., visual relationships) among the objects helps to characterize the interactions among them, which makes visual reasoning a hot topic in computer vision [188].

- Node classification
  A large social network dataset with millions of users is given. However, it is known that a significant number of these users are actual bots. The identification of these bots could be important for various reasons: advertising to bots may not be desired by a company, or bots may be in violation of the social network's terms of service. Manually examining each user to determine if they are a bot would be prohibitively expensive. Ideally, a model that can classify users as bots (or not) should be developed, given only a small number of manually labeled examples. This is a classic example of node classification, where the goal is to predict the label y_u—which could be a type, category, or attribute—associated with all the nodes $u \in V$, when we are only given the true labels on a training set of nodes $V_{\text{train}} \subset V$. Node classification is perhaps the most popular machine learning task on graph data, especially in recent years. Examples of node classification beyond social networks include classifying the function of proteins in the interactome [197] and classifying the topic of documents based on hyperlinks or citation graphs [190]. Often, we assume that we have label information only for a very small subset of the nodes in a single graph (e.g., classifying bots in a social network from a small set of manually labeled examples). However, there are also instances of node classification that involve many labeled nodes and/or that require generalization across disconnected graphs (e.g., classifying the function of proteins in the interactomes of different species).

At first glance, node classification appears to be a straightforward variation of standard supervised classification, but important differences exist. The most important difference is that the nodes in a graph are not considered independent and identically distributed (i.i.d.). In the construction of supervised machine learning models, the assumption is typically made that each datapoint is statistically independent from all other datapoints. Otherwise, the dependencies between all input points need to be modeled. Additionally, the assumption is made that the datapoints are identically distributed; otherwise, there is no guarantee of model generalization to new datapoints. Node classification completely breaks this i.i.d. assumption. Instead of modeling a set of i.i.d. datapoints, an interconnected set of nodes is modeled. In fact, the key insight behind many of the most successful node classification approaches is the explicit leveraging of connections between nodes. One particularly popular idea is the exploitation of homophily, which is the tendency for nodes to share attributes with their neighbors in the graph [198]. For example, people tend to form friendships with others who share the same interests or demographics. Based on the notion of homophily we can build machine-learning models that try to assign similar labels to neighboring nodes in a graph [199]. Beyond homophily there are also concepts such as structural equivalence, which is the idea that nodes with similar local neighborhood structures will have similar labels, as well as heterophily, which presumes that nodes will be preferentially connected to nodes with different labels. When we build node classification models we want to exploit these concepts and model the relationships between nodes, rather than simply treating nodes as independent datapoints.

- GCN based Studies
  This section contains literature reviews obtained using the keywords - **MRI AND Graph Convolutional Network AND Segmentation**. This search was last performed in October 2022 in the MEDLINE database. In total, 48 publications were found and of these, 7 in particular are relatively more appropriate for our study and are listed below. This particular section of the literature review is more generalized to have a broader perspective on the existing work.

  Aiming for automated Pancreas segmentation, Cai et al. [200] proposed a graph-based decision fusion process combined with deep convolutional neural networks (CNN). Abdominal MRI scans captured from 78 subjects were used in this study. For pancreatic detection and boundary segmentation, two types of CNN models were utilized respectively: (1) the tissue detection step to differentiate pancreas and non-pancreas tissue with spatial intensity context; (2) the boundary detection step to allocate the semantic boundaries of the pancreas. Both detection results of the two networks are fused together as the initialization of a conditional random field (CRF) framework to obtain the final segmentation output. Fully Convolutional Network (FCN) was chosen to localize the pancreas position but since it was not effective on precisely delineating pancreatic boundaries due to its upsampling convolutional operations so Holistically-nested edge detection (HED) was used further. HED [201] improved a regular FCN by adding deep supervision at all convolutional layers against down-sampled maps of the final desirable labeling output. In this way, the training losses were calculated and propagated back per layer. However, HED might fail to capture all weak boundaries so both

FCN and HED outputs were considered as priors and decision fusion was conducted using a principled CRF model. By minimizing the CRF energy function, the decision fusion method was able to remove most false positive segments from the original pancreas segmentation areas. Later, an undirected graphical model with weighted edges was developed for decision fusion. The first group was generated from watershed transformation of the semantic pancreas gradient maps through HED outputs. These regions aligned well to the HED detected pancreas boundaries. The second group was produced by superpixel segmentation on the detected FCN regions. The adjacent nodes were linked with weighted edges, which would reflect the likelihood that the two nodes belonging to the same category (pancreas or non-pancreas). In other words, edges with low similarity would encourage the connected two nodes to take different categories. To conduct CRF learning, a stochastic gradient decent algorithm was applied. The results were not strictly comparable due to the lack of common evaluation dataset but overall performed quiet well. Močnik et al. [202] developed an automatic multi-modal method for parotid glands segmentation from pre-registered CT and MRI images. A collection of planning images from 44 different patients were used, scheduled for Head and Neck radiotherapy. For every patient, a pair of 3D CT and MR images of the Haed and Neck region was available. First, the MRI images were registered to the accompanying CT images using Elastix and ANTs (two different state-of-the-art registration procedures). Then Patches of intensity values from both image modalities, centered around randomly sampled voxels from the reference domain, served as positive or negative samples in the training of the convolutional neural network (CNN) classifier. The trained CNN accepted a previously unseen (registered) image pair and classified its voxels according to the resemblance of its patches to the patches used for training. The final segmentation was refined using a graph-cut algorithm, followed by the dilate-erode operations [202]. The proposed method showed improved results compared to a related CNN-based segmentation method which used CT modality only. Authors demonstrated that CNNs have the ability to successfully incorporate additional image modality to aid the segmentation in cases where image registration between image modalities offers satisfactory results. Tian et al. [203] aimed to develop a robust interactive segmentation method for accurate segmentation of the prostate from MR since this type of segmentation is extensively applied in many clinical applications in prostate cancer diagnosis and treatment. 140 subjects of prostate MRI from three datasets (PROMISE12 (50 subjects), 32 International Symposium on Biomedical Imaging 2013 (ISBI2013) (49 subjects), and in-house (41 subjects) dataset) were used for the training model. Authors proposed an interactive segmentation method based on a graph convolutional network (GCN) to refine the automatically segmented results. An atrous multiscale convolutional neural network (CNN) encoder was proposed to learn representative features to obtain accurate segmentations. Based on the multiscale feature, a GCN block was presented to predict the prostate contour in both automatic and interactive manners. To preserve the prostate boundary details and effectively train the GCN, a contour matching loss was proposed. The GCN took the output feature from the last convolutional layer of the CNN encoder applied on the cropped image as its input. Next to observe the contour of prostate, a contour block was proposed. The output feature of the contour block was concatenated with the output feature of the CNN encoder to produce an en-

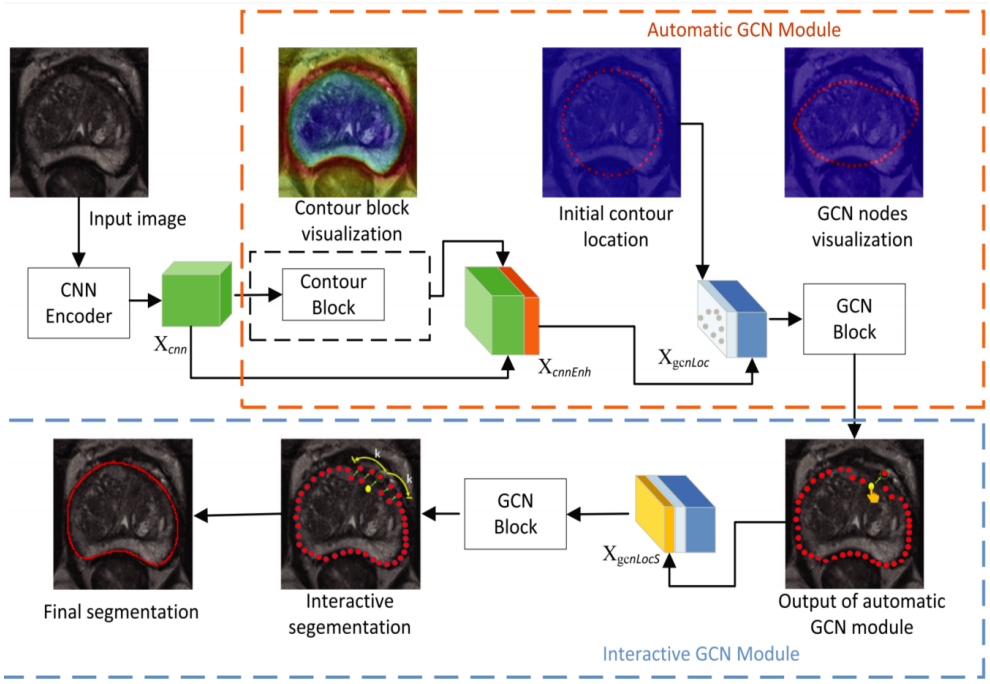


Figure 1.13: Segmentation Network proposed by [203]

hanced feature representation. The enhanced feature was followed by four 3 * 3 convolutional layers that aimed to form a GCN feature. To encode the locations of the interaction vertices, the initial coordinates of the interaction vertices were concatenated with the GCN feature to form a location-aware feature (X_{gcnLoc}). That feature was then fed to a GCN block to predict the locations of the vertices that could obtain the output of the automatic GCN module. The interactive GCN module next followed the automatic GCN module. The interactive GCN aimed to improve the segmentation accuracy by introducing user interactions. The users selected inaccurate interaction points and dragged them to the correct positions. The shifts of these points were then concatenated with feature X_{gcnLoc} to produce a shift-encoded feature which was later fed into the GCN block to obtain the shifts of k neighboring vertices of the current interaction point. To emphasize the superiority of the proposed method, a statistical significance experiment was performed on the whole gland in terms of the DSC (Dice similarity coefficient) and ASD (average symmetric surface distance). Authors also did a t-test for reporting the improvements. The analyses of the tests showed that there was a statistically significant difference on the entire gland ($P < 0.05$) both in terms of DSC and ASD. A unique two-stage system called SpineParseNet was proposed by Pang et al. [204] to perform automatic spine parsing for volumetric MR images. The dataset consisted of T2-weighted volumetric MR images of 215 subjects. The proposed framework contained a 3D GCN for 3D coarse segmentation and 2D ResUNet for 2D segmentation refinement. The framework of the 3D GCSN which included a deep convNet for extracting low-level and high-level image representations, a semantic feature extractor, and a decoder for generating the coarse probability maps. The proposed SpineParseNet outperformed a few other state-of-the-art networks when performance was compared, and it was able to accurately parse spines in volumetric MR images. Eschenburg et al. [205] proposed a general purpose cortical segmentation method that, given

resting-state connectivity features readily computed during conventional MRI pre-processing and a set of corresponding training labels, can generate cortical parcellations for new MRI data. The data used in this study came from the Human Connectome Project (HCP) [206] and from the Midnight Scan Club (MSC) [207]. The proposed methodology was an approach to perform cortical segmentation, a node classification problem using graph neural networks. This study examined the performance of common variants of graph neural networks in a whole-brain cortical classification setting and explored their ability to generalize to new datasets using functional magnetic resonance imaging (fMRI). Graph neural networks are prone to over-fitting of model parameters and over-smoothing of learned embeddings as network depth increases [208]. One approach to alleviate this over-smoothing is to adaptively learn optimized network depths for each node in the graph, a method describe as Jumping-knowledge networks. In the proposed methodology, authors utilized Graph attention network (GAT) network employing the jumping-knowledge mechanism. As an input the network took graph adjacency structure and node wise feature matrix which did output a node-by-label logit matrix. Each of the GATConv block was composed of multiple attention heads. Aggregation function which took as input the embeddings from each GATConv block, learned a convex combination of the layer-wise embeddings. This method was compared with three different graph neural network variants to a baseline fully-connected network. Authors found that, in all cases, graph neural networks consistently and significantly outperformed a baseline neural network that excluded adjacency information. 3D shape reconstruction helps visualize the spatial structure of 3D objects and is relevant to several applications such as, computer-aided diagnosis, surgical planning, image-guided interventions, and computational simulations [209]. Chen et al. [210] proposed a deep learning architecture, coined Mesh Reconstruction Network (MR-Net), which tackles this problem. All experiments conducted to validate MR-Net are performed using 7,870 stacks of 2D contours, available from the manual delineation of short axis view cardiac MR images (at the end of systole and diastole), within the UK Biobank dataset. MR-Net enabled accurate 3D mesh reconstruction in real-time despite missing data and with sparse annotations. Using 3D cardiac shape reconstruction from 2D contours defined on short-axis cardiac magnetic resonance image slices as an exemplar. This work was a hybrid graph convolutional neural network for 3D mesh reconstruction, MR-Net, which approached the problem as a template deformation task conditioned on the sparse point cloud data (stacked 2D contours). Due to the large proportion of missing inter-slice information, 3D shape reconstruction from sparse 2D contours was a challenging task. A template mesh was randomly selected from the training dataset to supply the missing information in the reconstruction process. The input PCs (point clouds) of contours serve as the guidance of template deformation. This method consistently outperformed state-of-the-art techniques for shape reconstruction from unstructured point clouds. Brain tumor segmentation, which is considered one of the most difficult medical image segmentation challenges, is crucial for radiation therapy planning and therapy response monitoring [211]. Ma et al. [212] proposed a novel approach for brain tumor segmentation called the dual graph reasoning unit (DGRUnit) which is a combination of two parallel graph reasoning modules. The BraTS2018 dataset includes a training dataset of 285 glioma patients with four modalities (T1, T1ce, T2 and Flair) and a validation set of

66 patients with hidden ground truth. The proposed methodology consisted of two parallel graph reasoning modules which were a spatial reasoning module and a channel reasoning module. The spatial reasoning module modeled the long-range spatial dependencies between distinct regions in an image using a graph convolutional network (GCN). The channel reasoning module used a graph attention network (GAT) to model the rich contextual inter-dependencies between different channels with similar semantic representations. The spatial reasoning module modeled the spatial relationships between the different regions in the feature map with the following steps.

- To construct the graph-structured data for the GCN from the original grid-like feature map
- Next a self-attention mechanism was used to generate the weights between different nodes as edges
- To feed the data into a graph convolutional layer to model and determine the relationships between the spatial nodes
- The final step was to project the output features from the GCN back to the original space as the output of the spatial reasoning module

Similar to the spatial graph reasoning module, the channel graph reasoning module first projected the feature map to a node feature in an embedded channel node space. Specifically, multiple channels were converted in the feature map to a set of node features. The node features were then fed into the GAT to determine the contextual interdependence. Then, the updated features from the GAT were projected back to the original space as the output of the channel reasoning module. The final output of the DGRUnit was a pointwise summation of the original feature map that included the outputs of the spatial reasoning module and channel reasoning module. The proposed methodology was compared to several state-of-the-art methods, experimental results showed that the proposed approach significantly improved both visual inspection and quantitative metrics for brain tumor segmentation tasks.

The literature listed below is included because it seems relevant to our methodology and is interesting (even if it is outside our search keywords). Wan et al. [213] attempted to create an image classification system for hyperspectral images, which consisted of a high quantity of valuable information and hundreds of contiguous bands. The system was evaluated on three datasets, the Indian Pines (145 $\times$ 145 pixels with a spatial resolution of 20 m $\times$ 20 m and has 220 spectral channels covering the range from 0.4 µm to 2.5 µm), the University of Pavia (610$\times$340 pixels with a spatial resolution of 1.3 m $\times$ 1.3 m and has 103 spectral channels in the wavelength range from 0.43 µm to 0.86 µm), and the Kennedy Space Center (224 bands, 614 $\times$ 512 pixels with a spatial resolution of 18 m with spectral coverage ranging from 0.4 µm to 2.5 µm). In the proposed framework, atfirst the original hyperspectral image was segmented in superpixels by SLIC algorithm [214], where a local region of the hyperspectral image was exhibited which contains eight superpixels. The circles and green lines shown in fig. 1.14 represent the graph nodes and edges, respectively, where different colors of the nodes represented different land-cover types. Different from the commonly

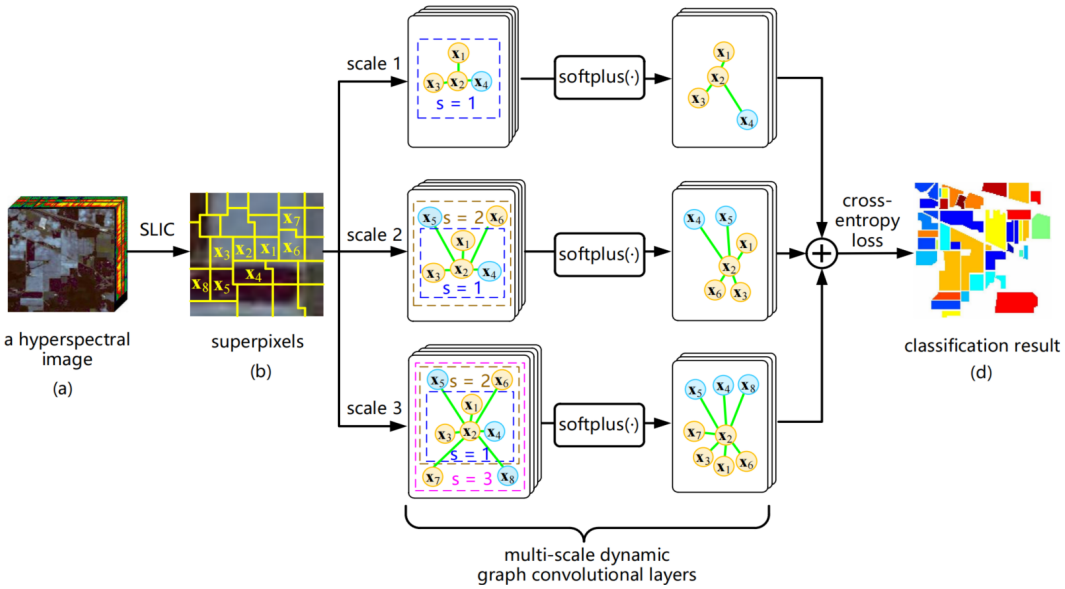


Figure 1.14: Framework proposed by [213]

used GCN models which work on a fixed graph, here the graph was enabled to be dynamically updated along with the graph convolution process so that these two steps can be benefited from each other to gradually produce the discriminative embedded features as well as a refined graph. The main idea was to find an improved graph by fusing the information of current data embeddings and the graph used in the previous layer. Specifically, at each scale, the edge weight was updated gradually along with the convolution on graph nodes, so that the graph could be dynamically refined. Here two dynamic graph convolutional layers were employed for each scale, where soft plus was utilized as the activation function. The classification result was acquired by integrating the multi-scale outputs, and the cross-entropy loss was used to penalize the label difference between the output and the seed superpixels. This framework was tested on three typical benchmark datasets and demonstrated the superiority of the proposed network to other state-of-the-art methods in both qualitative and quantitative aspects.

1.4 Image Classification

Image classification is a computer vision task where the goal is to assign a label or category to an entire image based on its content. This process involves training a model, typically a deep neural network, on a labeled dataset where images are associated with predefined classes. The model learns to extract relevant features and patterns from the images during training, enabling it to predict the class of new, unseen images.

1.4.1 Applications of Image Classification

There are numerous image classification applications across different industries. Here are some of the most prominent ones:

- Healthcare
  Image classification is used in diagnostics to discover conditions from a range of medical images such as X-rays, MRIs, and CT scans. It acts as a tool for doctors to help in the early detection of ailments. Beyond detection, it is also useful for treatment planning and disease monitoring.

- Autonomous Vehicles
  Self-driving cars rely on super-fast camera analysis. They see the world around them by identifying objects like cars, people, signs, and lanes to navigate safely and avoid accidents. Image classification also plays a pivotal role when different lighting conditions make obstacle avoidance a challenge.

- Agricultural Monitoring
  In the agriculture sector, the image classification method has many uses, including crop monitoring, disease detection, yield estimation, and others. Through remote sensing using drone or satellite technology, the image classification apps direct the farmers to recognize plant health issues, optimize irrigation, and increase crop yield.

- Retail
  Retail stores use cameras that can identify products via image classification. This helps them stock shelves, target ads to the right customers, and make shopping easier. Image classification also plays a major role in self-checkout systems. It accurately identifies and categorizes objects scanned by customers to avoid mistakes or theft. Image classification also enables personalized recommendations of products by analyzing customer behaviour and preferences.

- Security and Surveillance
  Image classification plays a key role in making security cameras smarter. They can now recognize faces, objects, and even unusual activity, helping keep people safe in businesses and public areas [215].

1.4.2 Machine Learning Approaches

Machine learning approaches for image classification leverage algorithms and models to categorize images into predefined classes by learning from labeled data.

1.4.2.1 Image Classification using Classical Machine Learning

Su et al. proposed a novel computer-aided method for the generation of a region of interest (ROI), segmentation, and classification of breast tumors without manual intervention [216]. The method detected the ROI using a self-organizing map neural network, incorporating local texture and position features. A modified Normalized Cut approach was then employed to partition the ROI into clusters and establish initial boundaries. A regional-fitting active contour model adjusted these boundaries for final segmentation. The system extracted three texture features and five morphologic features from each breast tumor, using Affinity Propagation clustering for classification into malignant and benign categories without a training process. Validation on 132 cases (67 benign and 65 malignant) demonstrated that the system outperformed traditional methods in detection and classification while exhibiting the lowest computational complexity. Xu et al. demonstrated that the HOG+SVM (Histogram of

Oriented Gradients + Support Vector Machines) algorithm with a Gaussian kernel function achieved the highest classification accuracy of 91.32% for pure cloth classification, outperforming the Small VGG network [217]. The authors concluded that for end-users with standard computing processors, the traditional machine learning algorithm HOG+SVM is recommended for classifying pure clothing images. For the classification of dressed clothing images, they suggested using a more efficient and less computationally intensive model, such as the Small VGG network. Abdulkadir et al. proposed a hybrid method for the classification of white blood cells. Image processing (IP) and machine learning (ML) are used to determine and classify the WBCs in blood smear images [218]. Sajjad et al. proposed a smartphone-based cloud-assisted resource-aware framework for the localization of white blood cells (WBCs) within microscopic blood smear images using a trained multi-class ensemble classification mechanism in the cloud [219]. In their framework, the nucleus was first segmented, followed by the extraction of texture, statistical, and wavelet features. The detected WBCs were then categorized into five classes: basophil, eosinophil, neutrophil, lymphocyte, and monocyte. Experimental results on several benchmark databases validated the effectiveness and efficiency of the proposed system in comparison to other state-of-the-art schemes. Tantikitti et al. utilized blood smear images captured under a digital microscope with 400× magnification specifications, applying various image processing techniques, including color transformation, image segmentation, edge detection, feature extraction, and classification of white blood cells [220]. In their study, they introduced white blood cell count as a novel feature for cell differentiation, which could classify dengue viral infections in patients using decision tree methods. Wan et al. used a dataset comprising 895 breast ultrasound images [221]. They applied traditional machine learning models, including seven commonly used CAD algorithms, along with three content-based radiomic features: Hu Moments, Color Histogram, and Haralick Texture. A convolutional neural network (CNN) model was also utilized, with AutoML Vision trained on the Google Cloud Platform. The study found a significant difference in diagnostic performance among all traditional machine learning classifiers ($P<0.05$). Random Forest achieved the best performance in differentiating between benign and malignant breast lesions, which was statistically comparable to the performance of the CNN and AutoML Vision based on Cochran's Q test ($P>0.05$). Shrivatsava et al. developed an effective and rapid system for automatically recognizing and classifying rice plant diseases. Here, color was a significant feature, and an SVM classifier with a polynomial order-2 kernel function was used for classification [222]. However, this system considers fewer classes. Jadhav et al. suggested a more efficient model for detecting soybean diseases by employing ML algorithms [223]. The k-means algorithm was used to detect soybean disease, and a Multiclass Support Vector Machine (M-SVM) was utilized for classification purposes. In this system, the pre-processing method was used on the region of interest (ROI) of the raw images, and the standard deviation and mean of the RGB channels were determined to extract texture and color features using the gray-level co-occurrence matrix (GLCM). Akmal et al. introduced an automatic plant disease detection model to identify diseases in corn and potato leaves [224]. Three essential features, namely local ternary patterns (LTP), segmented fractal texture analysis (SFTA), and histogram-oriented gradient (HOG) were utilized. To minimize the curse of dimensionality, principal component analysis (PCA) and score values based on entropy skewness were employed. Although this system depends on the chosen features, it reduces the accuracy when there are not enough features. Hossain et al. introduced a color- and texture-based method using a KNN classifier to detect

and classify plant leaf diseases. Texture features were extracted for classification from diseased leaf images. Moreover, the employed KNN Classifier classifies plant diseases, such as canker, leaf spot, alternate, and bacterial blight. As a result, better accuracy was attained, but the size of the dataset was small [225]. Sharif et al. presented a hybrid model for citrus disease detection and classification based on feature selection and weighted segmentation [226]. Initially, an optimized weighted segmentation approach was employed to extract the lesion spots of citrus, and then the most suitable features were chosen using a hybrid feature selection technique. Finally, a multiclass SVM (M-SVM) was used to classify citrus diseases. However, this method is more complex, and there are no proper diseased parts.

1.4.3 Deep Learning Approaches

Deep learning approaches for image classification utilize complex neural network architectures to automatically extract and learn hierarchical features from raw images, enabling the accurate categorization of visual data into various classes.

1.4.3.1 Image Classification using Deep Learning

Sharma et al. conducted image classification using multiple convolutional neural networks (CNNs). They analyzed the prediction accuracy of three different CNNs on popular training and test datasets, specifically CIFAR10 and CIFAR100, focusing on 10 classes from each dataset. The study aimed to assess the accuracy and consistency of predictions across these networks on the same datasets. A detailed comparison of the networks' performance across various object classes was presented [227]. Han et al. proposed a novel two-phase method that combined CNN transfer learning with web data augmentation [228]. They demonstrated that the useful feature representations of pre-trained networks could be efficiently transferred to the target task, and the original dataset could be augmented with valuable Internet images for classification. This approach significantly reduced the need for large training datasets and effectively expanded the training data, contributing to a substantial reduction in overfitting of deep CNNs on small datasets. Additionally, they applied Bayesian optimization to address hyper-parameter tuning during network fine-tuning. Their solution was tested on six public small datasets, and extensive experiments showed that it enabled popular deep CNNs, particularly ResNet, to outperform state-of-the-art models. The results proved that their method is a valuable tool for using deep CNNs on small datasets. In their study, Hussain et al. proposed a CNN architecture model, specifically Inception-v3, to determine its accuracy and efficiency with new image datasets using transfer learning [229]. The retrained model was evaluated, and its results were compared to several state-of-the-art approaches. Maggiori et al. addressed the pixel-wise classification of high-resolution aerial imagery [230]. Although convolutional neural networks (CNNs) are gaining increasing attention in image analysis, adapting them to produce fine-grained classification maps remained challenging due to the well-known trade-off between recognition and localization. While CNNs excelled at recognizing meaningful objects, they often sacrificed spatial precision. To tackle this issue, the authors proposed an architecture that learned features at different levels of detail and a function to combine them. By efficiently and flexibly integrating local and global information, their approach outperformed previous techniques. Lu et al. proposed a customized CNN network for lung image patch classification [231]. The authors designed a fully

automatic neural-based machine learning framework to extract discriminative features from training samples while simultaneously performing classification. Their method was not problem-specific, allowing for easy application in other imaging domains. They incorporated random neural node dropout and utilized a single convolutional layer architecture to reduce the number of parameters in the CNN model, thereby avoiding the overfitting problem. For flexible experimentation, they implemented their own neural network toolkit, which included CNN and RBM, with performance acceleration using Advanced Vector Extensions (AVX). Fu et al. proposed a network called M-Net for glaucoma detection. M-Net mainly consists of a multi-scale input layer, a U-shape convolutional network, a side-output layer, and a multi-label loss function. The multi-scale input layer constructs an image pyramid to achieve multiple-level receptive field sizes. The U-shape convolutional network was employed as the main body network structure to learn the rich hierarchical representation, while the side-output layer acts as an early classifier that produces a companion local prediction map for different scale layers. Finally, a multi-label loss function was proposed to generate the final segmentation map. Here, the glaucoma detection was solved in terms of segmentation [232]. In the field of ultrasound imaging of breast nodules, Chen et al. combined the detection methods of texture features and morphology, using the AlexNet model to find the nodules in the image and predicting the benign and malignant, and the AUC value reached 0.9325 [233]. For the automatic classification of fetal facial ultrasound images, Yu et al. integrated CNN and a random two-coordinate descent optimization algorithm, achieving 96.98% accuracy. At the same time, the authors used the ResNet model in the automatic recognition of melanoma in dermoscopic images based on deep aggregation features. The AUC reached more than 80% [234] [235]. Here, the authors presented a new analysis of Stochastic Dual Coordinate Ascent (SDCA) showing that this class of methods enjoyed strong theoretical guarantees that were comparable to or better than Stochastic Gradient Descent. To study lung cancer pathology, Zhang et al. established the early computer diagnosis system for lung cancer to detect lung cancer pathological sections so that several major types of lung cancer can be detected [236]. Qaiser et al. established a CNN-based intelligent image diagnosis system for breast cancer pathology study with a judgment result of pathological sections (cancer and non-cancer tissue) as high as 0.833, four cancer classifications (normal tissue, benign tumor, original The accuracy of the results of both cancer and invasive cancer turned into as high as 0.778, which reached the level of pathologists [237]. Fu et al. solved the boundary detection problem by proposing vessel segmentation using a fully convolutional neural network (CNN) and conditional random fields. Here, the binary classification result is generated by a fully convolutional network and conditional random field to generate a vessel probability map [238]. An intelligent image diagnosis system constructed by Ehteshami Bejnordi et al. The diagnosis of gastric cancer pathological images can reach 69.9% accuracy compared with pathologists [239]. Here, the authors applied multi-scale superpixel classification to detect epithelial regions in whole-slide images (WSIs). Subsequently, spatial clustering was utilized to delineate regions representing meaningful structures within the tissue such as ducts and lobules. A region-based classifier employing a large set of features including statistical and structural texture features and architectural features was then trained to discriminate between DCIS and benign/normal structures. Yoshida et al. reached 0.556 in three classifications of gastric cancer pathology (normal tissues, adenomas, cancer cells) [240].

1.4.4 Our Problem Statement in Classification

The detection of tendon pathologies, including tendinopathy, tendinosis, and tears, are critical for effective patient management and treatment. Accurate classification is essential for determining the presence or absence of pathology, guiding therapeutic interventions, and predicting patient outcomes. Misdiagnosis or delayed diagnosis can result in inadequate treatment, prolonged pain, and increased healthcare costs, significantly impacting the quality of life for affected individuals. Traditional diagnostic approaches often rely on subjective assessments by healthcare professionals, leading to inconsistencies and variability in diagnoses. On top of that, the complex anatomy and diverse presentation of tendon pathologies complicate the classification process, emphasizing the need for an automated, end-to-end system that can accurately classify whether a subject is pathological or not.

Existing classification methods may struggle with challenges such as imaging noise, variations in tendon structure, and the presence of confounding pathologies. These limitations can compromise the efficiency and reliability of assessments in clinical settings. Therefore, this thesis aims to develop a robust end-to-end classification system that effectively differentiates between pathological and non-pathological conditions using advanced imaging data, such as MRI, combined with artificial intelligence techniques.

The proposed classification framework will enhance diagnostic accuracy and facilitate timely and appropriate treatment decisions by integrating automated detection and classification into clinical workflows. By bridging the gap between advanced computational techniques and practical clinical applications, this research aspires to improve patient care outcomes. Ultimately, the integration of this end-to-end classification system has the potential to standardize diagnoses, reduce errors, and provide valuable insights that can inform future research and innovations in musculoskeletal health.

1.4.4.1 Problem Specific Studies

This section provides an overview of the findings from our systematic literature review conducted using the MEDLINE database and the search phrases were **(Tendinopathy OR "Tendon Rupture") AND Classification**. The last time this search was verified was in August 2022. This search yielded 44 total results, but many of them were not of interest to us because they dealt with topics such as surgical management, treatment options, epidemiology, the impact of bacterial presence, turkey arthritis, the Greek shoulder pain and disability index, macroscopic vs microscopic study, and other topics.

Gillet et al. [241] sought to explain how Magnetic Resonance Imaging (MRI) might be useful for Achilles tendinitis patients receiving fluoroquinolone therapy. 14 Achilles tendons were examined by MRI (T1 and T2 or T2*weighted sequences) in 9 patients with typical tendinopathy (13 cases of tendinitis and 1 rupture) during fluoroquinolone therapy. Here, tendinous involvement was classified according to the prominence of intra or peri-tendinous changes. In this study, the tendinous involvement was classified into 5 stages (adapted from [242]):

- Stage 0: normal calcaneal tendon, appearing, due to its fibrous nature, as a dark band easily differentiated from the surrounding tissue

- Stage I: thickening of the tendon (NB: one patient underwent echography of Achilles tendon that permitted independent verification of this abnormality detected by MRI
- Stage II; thickening of the tendon with abnormal tendon signal, including structural longitudinally and centrally located images; in patients with marked signal changes predominating horizontally it was considered that there was a potential risk of rupture, especially if signal changes reached the surface
- Stage III: edema in the peri-tendon
- Stage IV: partial or complete rupture

It was commonly agreed to consider the thickening of the tendon as a focal inflammatory reaction, as sometimes observed in active sportsmen. On the other hand, the thickening of the tendon with structural alterations was considered a "degenerative" change. In all these symptomatic patients, the most frequently noticed intratendinous MR abnormalities considered to be potential rupture risks led authors to adopt a more radical treatment (use of canes, and application of below-the-knee plaster casts for 4 weeks) to prevent any complications. Such radical orthopedic treatment, the interest of which remains to be confined by other longitudinal and control studies, may account for the absence of additional ruptures. The authors came to the conclusion that in patients suffering from fluoroquinolone-induced tendinopathy, MRI appears a helpful and accurate method to identify, classify and follow the disease, of course after withdrawal of this class of antibiotics. In addition, as regards orthopedic treatment management during painful tendinopathy, MRI may also be a useful investigative tool in examining the early pathogenic changes during fluoroquinolone-induced tendinopathy. McLoughlin et al. [243] characterized the patellar tendinitis features using Magnetic Resonance Imaging (MRI) to gain a better understanding of the underlying pathophysiology that would ultimately help with further therapeutic treatment. The authors used 15 patients who were clinically diagnosed with patellar tendonitis and underwent gadolinium-enhanced MRI of the knee. In this study, an experienced musculoskeletal radiologist interpreted the images. The following features were extracted on both non-enhanced and contrast-enhanced images: tendon dimensions, tendon contour, surface change, pathologic intra-tendinous findings, including location, signal intensity characteristics, enhancement characteristics of tendon injury, the presence of cystic tendon degeneration and intra-tendinous fluid; enthesial damage; patellar apical (lower-pole) chondral-bone avulsion; and pathologic peritendinous conditions, particularly edema. In addition, the patellar margin and retinacula were evaluated. The overall appearances in the abnormal tendon were grouped on the basis of signal intensity characteristics which are as follows: (a) Grade 1 was defined by an enhancing intra-tendinous area next to the patellar apex, which had intermediate signal intensity on fl-weighted SE and T2- weighted fast SE images and high signal intensity on T2*weighted GRE images. This area was surrounded peripherally by a variable non-enhancing zone with intermediate signal intensity on images obtained with all sequences. In grade 2 damage, the characteristics of grade 1 damage were present except for perichondral-bone avulsion from the patellar apex. In grade 3 damage, a homogeneous, non-enhancing intra-tendinous area was depicted next to the patellar apex. This area had intermediate signal intensity on images obtained with all sequences, in a predominately globular on linear configuration oriented

obliquely to the axis of the don. Cases were included in grade 3 in which the abnormal area demonstrated intermediate to high (as opposed to only high) signal intensity on GRE (Gradient-recalled-echo) images on demonstrated only minimal enhancement on all images. Bauer et al. [244] proposed a comprehensive MRI grading protocol for supraspinatus tendinosis and partial thickness tears and studied its reliability. Digital MRI scans (3 Tesla) of 65 patients investigated for assessment of supraspinatus pathology or subacromial impingement were evaluated. These patients were evaluated by 3 independent and experienced musculoskeletal (MSK) radiologists. Tendonosis, partial thickness (PT) tears, and anteroposterior (AP) extent of tears were scored by each radiologist on two separate occasions (t1, t2), 2 months apart. The inter-observer and intra-observer agreement and weighted kappa values for each parameter were calculated. An instructional scoring questionnaire and scoring template were completed for each scan. They were scored (0, 1, 2) on coronal and sagittal scans. The greatest extent of tendonosis was identified by scrolling from anterior to posterior on coronal images and from lateral to medial on sagittal images. The dimensions of tendonosis were measured using the InteleViewer software. A score of 0 was given for a normal tendon on PD and T2 images. A score of 1 indicates focal tendonosis where the signal increase was the non-fluid signal intensity and was not measured greater than 10 mm in maximal extent in any dimension. A score of 2 indicated generalized tendonosis where the signal intensity was the non-fluid signal intensity and was measured as greater than 10 mm in any dimension. Tear thickness was scored from 0 to 4 on coronal T2 weighted images while scrolling from anterior to posterior. A tendon tear was identified by fluid signal intensity. The highest grade tear thickness was identified. The total tendon thickness at that location was calculated by measuring the perpendicular distance from the greater tuberosity footprint to the bursal tendon surface. The thickness of the partial tendon tear was calculated as a percentage of the total tendon thickness at that location. AP tear size was measured and scored 0 to 3 on sagittal T2 weighted images while scrolling from medial to lateral toward the supraspinatus footprint. The greatest extent of a tear was measured in the AP dimension. Total tendinopathy was calculated from the sum of tendonosis, tear thickness, and AP tear size scores. A tendinopathy total score varied from 0 (normal tendon) to a maximum score of 9 (generalized tendonosis with a full-thickness tear more than 10 mm in AP dimension). Bauer et al. [244] concluded that their protocol was simple to learn and obtained excellent intra-observer and good inter-observer reliability. Neer's classification [245] method divides rotator cuff ailments into three groups: inflammation, calcific tendinitis, and complete or partial thickness tears. [246] developed a computer-assisted diagnostic (CAD) system to aid ultrasound operators in the accurate diagnosis of rotator cuff diseases and to streamline ultrasound tests. The database consisted of 99 shoulder ultrasound images in 93 adult patients. Lesions were classified into three categories, including 43 cases of tendon inflammation, 30 cases of calcific tendinitis, and 26 cases of supraspinatus tear. In order to classify each lesion case, the lesion area and texture features were taken from the complete lesion and integrated with a multinomial logistic regression classifier. After performing feature selection, the classifier merged the relevant texture features to create a prediction model. The optimal combination of the subset feature combination was identified using Stepwise Backward elimination. The corresponding subset features were chosen for the prediction model when the lowest error rate was attained. Lesion area, local homogeneity (SD), cluster prominence (mean), cluster prominence (SD), and Haralick correlation (mean) were the features used. The performance achieved an overall accuracy of 87.9% in the classification of

rotator cuff inflammation, calcific tendinitis, and tears. The individual accuracy in diagnosing inflammation and tears groups was relatively higher (88.4% and 92.3%, respectively) and that for diagnosing calcific tendinitis was relatively lower (83.3%). The authors claim that the diagnostic suggestions generated by the proposed CAD system would be practical and promising for clinical use. Achilles tendon (AT) has a considerable impact on the foot's flexibility and mobility. AT is linked to a number of diseases, including Achilles tendinitis and rupture, much like any other organ. In this study, a software program that can identify changes in ultrasound images of Achilles tendon and categorize them as normal or abnormal is developed [247]. A total of 57 AT ultrasound images were collected from normal and abnormal cases in JPEG format from the radiology department at the University of Salford. [247] determined the ROI in AT ultrasound images automatically by dividing the image into sub-images and targeting the most vulnerable area. Determining the ROI area and size based on the medical reports states that the most vulnerable area in AT is 2 to 6 cm proximal to the calcaneal insertion point. In this study, a matrix with 68 different texture features was extracted from the whole sample. Most image features were constructed based on histogram and texture features. Statistical descriptors, co-occurrence matrices for four different angles, and run length encoding algorithm were used for extracting different texture features. Furthermore, different histogram features such as skewness and kurtosis were extracted from the samples after different segmentation methods were applied. However, using the KPCA algorithm as a feature reduction method with the Gaussian kernel function, these features were reduced to the minimum number that corresponds to the largest eigenvalues. DT, Discriminant Analysis, Quadratic SVM, k-NN, and different ensemble classifiers were used to classify the dataset images through the extracted features. The Bagged tree type of ensemble classifiers performed better than the other classifiers in terms of performance. The authors have shown that the proposed approach works best at the most vulnerable site in AT, which is 2 cm from the calcaneal insertion with a length of 4 cm. This should provide physicians with helpful diagnosis information. The purpose of this study [248] was to ascertain whether a quantitative texture-based method could accurately identify tendon abnormalities in pallapugno players. 14 elite players of Pallapugno who did not present any neuromuscular or skeletal impairment volunteered to participate in the study. Numerous texture features were extracted from within the ROI of each image, including both the dominant side and the non-dominant side of the player. Specifically, a total of 90 texture features were extracted for each ROI, for a total of 1080 texture features extracted for each subject (90 features $\times$ 6 images $\times$ 2 sides). The texture features could be divided into the following three main categories: (1) first-order statistical descriptors, (2) Haralick features, and (3) higher-order spectra, entropy features, and Hu's moments. These features on the dominant and non-dominant sides were used to perform a multivariate linear regression analysis (MANOVA) to determine tendon abnormality and, more importantly, the occurrence of subclinical tendinopathy. This study confirmed the prospect that a quantitative texture-based method could be helpful for the standardized diagnosis (as well as monitoring and prognosis) of sub-clinical tendinopathy. All that was needed was to identify the tendon ROI within the B-mode image. Tendinopathy, calcific tendinitis, tears, bursitis, and bursal responses are some of the rotator cuff disorders and among them, rotator cuff tears have a higher prevalence rate. To identify supraspinatus tears in ultrasound examinations and to lower inter-operator variability, Change et al. [249] created a computer-aided tear classification (CTC) system in this study. It included 89 ultrasound images of supraspinatus

tendinopathy and 102 supraspinatus tears from 136 patients. For each case, intensity and texture features were extracted from the entire lesion and combined in a binary logistic regression classifier for lesion classification. The intensity features are the mean, variance, skewness, and kurtosis, namely, the first-, second-, third-, and fourth-order central moments of a histogram. Grayscale co-occurrence matrices (GLCMs) were employed to represent the texture features. The averages of the GLCM-based statistical features were calculated, extracted from the lesion areas, and combined with the intensity features in the classifier to express tissue characteristics, such as brightness, contrast, heterogeneity, etc. The main objective of Golman et al. [250] was to determine the association between partial patellar tendon tears (PPTT) characteristics and treatment guidelines as well as to create a classification system for partial patellar tendon injuries based on magnetic resonance imaging (MRI). 85 of the 171 athletes who participated in the study reported symptoms suggestive of patellar tendinopathy, while 86 of the control patients had a patellar tendon that was healthy and intact. 56 individuals in total got a PPTT and further examination for tear size and placement. The relationship between tear characteristics and clinical outcome was defined with the use of statistical comparisons and univariate and logistic regression models. And the partial patellar tendon tear characteristics were defined in two folds which are, (1) signal change on fluid-sensitive sequences with signal in the tendon that was isointense to the joint fluid and (2) morphological defect with evidence of disruption or discontinuity of the fibers on axial or sagittal images. Six maximum tendon dimensions were measured with sagittal or axial (anterior to posterior) views: (1) the length (proximal to distal) of the patellar tendon, (2) the width (medial to lateral) of the patellar tendon, (3) the thickness (anterior to posterior) of the patellar tendon on the first axial slice below the inferior patellar bone, (4) the length of the tear, (5) the width of the tear, and (6) the thickness of the tear. After statistical analysis, the proposed classification system assigned different grades to the individual patellar tendons, and the result was compared to the gold standard Blazina grading system. Logistic regression and sensitivity analyses showed that patellar tendon thickness measurement predicted whether the tendon was normal or symptomatic. The most sensitive predictor for the presence of a partial tendon tear was the thickness of the tendon, in which thickness greater than 8.8 mm was strongly correlated with the presence of a tear. The authors concluded that their proposed classification system (grade 1-4) for PPTT can be used to help guide treatment. The authors of [251] tried to determine the reliability of the Soong classification, which relates the position of the implant to the watershed line of the distal radius, for predicting flexor tendinopathy in distal radius fractures treated with volar plate fixation. The final study sub-set included 659 distal radius fractures managed with volar plate fixation in 648 patients. Descriptive statistics were used to summarize data, including percentages and counts for categorical and ordinal data and means with interquartile ranges for continuous data. Power analysis was performed to determine sample sizes for two-sample proportions with a two-sided test with an alpha of 0.05 and a beta of 0.80 to compare the proportion of flexor pollicis longus rupture between Soong 0 and 1 versus Soong 2 volar plate positions for tendinopathy and rupture. To address risk factors for tendon complications or inadequate reduction, each variable was analyzed in a univariate fashion because of the limited number of postoperative complications, limiting our ability to perform multivariable analyses. Specifically, analysis was performed to determine the relationship between Soong grade and flexor tendon rupture and adequacy of fracture reduction. All analyses were carried out using the JMP statistical software package. The Soong classification

is not the sole indicator of flexor tendinopathy, according to the authors, who also suggested that it may be interpreted as an indicator of how appropriately fracture reduction was performed in general. Since the pathophysiology of subscapularis (SS) lesions is still relatively unknown, Leite et al. [252] aimed to determine the influence of the coracoid morphology and humeral version on SS tears. T1- and T2-weighted fat-saturated images of 232 shoulders, with the patient in the supine position with the arm alongside the body, elbow extended, and forearm supinated. The authors claim that they were the first to study the influence of different parameters of the coracoid process morphology and humeral version on SS tears. Coracoid morphology was evaluated, including (1) proximal segment length, (2) distal segment length, and (3) CA. The CA was measured using MRI sagittal cuts at the section corresponding to the greatest coracoid cross-sectional length. The proximal and distal coracoid segments' long axes were traced, and the angle formed under these represented the CA. Whereas the humeral version was determined using MRI axial sections, first drawing a line (L1) joining the anterior and posterior margins of the articular cartilage, at the point of maximal head diameter, previously measured and defined. Thereafter, a perpendicular line (L2) passing at the L1 line midpoint was drawn. As such, L2 represented the central axis of the humeral head. A third line (L3) parallel to the MRI scanner orientation was drawn. Then, the angle drawn between L2 and L3 corresponded to the humeral version. With an AUC of 90.0%, the CA, which was discovered to be the best predictor of SS tears, was good at predicting lesions. Based on the analysis of the CA ROC curve, a cutoff value of 105.5 degrees with a sensitivity of 82.4% and a specificity of 78.9% for SS tears was established. This study was an inaugural in using these parameters to create a classification system that allowed to divide of coracoids by shape according to the relative risk of SS tears (flat coracoid with CA superior to 120 degrees, curved coracoid with CA between 95 and 120 degrees, and hooked coracoid with CA inferior to 95 degrees). The rotator cuff's most delicate part, the supraspinatus tendon (SST), was the subject of this study. An automatic tendinopathy recognition framework based on convolutional neural networks was proposed by Jahanifar et al. [253] to assist in diagnosis. Segmentation and classification of tendons were divided into two key components in this framework. This dataset is divided into 70 healthy and 30 tendinopathy cases. Among them, 22 cases out of 70 normal cases are females, and the remaining are males. About the same statistics hold for tendinopathy cases (12 out of 30 tendinopathy cases are females). A novel CNN-based model was developed for the segmentation task, which benefited from exclusive structural blocks where three main building blocks were utilized: Normal, Reducing, and Enlarging cells. The Normal and Reducing blocks were directly borrowed from NASNet. Inspired by U-Net and NASNet models, their proposed segmentation network was called NASUNet. For the classification task, a CNN-based classification framework was introduced that can operate with different base models any of NASNet, Inception-ResNet v2 [254], Xception [255] or 169-layered DenseNet [256] as long as their original top layers are eliminated). The classification framework was equipped with positional information about the tendon in the input to boost its performance. The proposed positioning scheme was based on the polar system, in which every position inside the object was determined via two parameters: distance (radius) from the object origin (r) and angle of deviation from the object orientation (θ). These two parameter maps were built based on the resulting segmentation map. Techniques such as knowledge transfer, transfer learning, and data augmentation were used to address data shortages and make the networks well-generalized. In cross-validation experiments, the proposed tendinopathy recognition

model achieves 91% accuracy, 86.67% sensitivity, and 92.86% specificity, showing state-of-the-art performance against other models. Wang et al. [257] aimed to develop and validate an efficient ultrasound image-based radiomic model for determining Achilles tendinopathy in skiers. The dataset consists of 88 feet of skiers clinically diagnosed with unilateral chronic Achilles tendinopathy and 51 healthy feet. In the study, Pyradiomics (version 3.0.1) was used to transform RGB ultrasonic images into red, green, and blue monochromatic grayscale images, and grayscale images respectively, and extracted radiomic features successively. The extracted features included: 14 shape features, 18 first-order statistical features, 22 gray level co-occurrence matrix (GLCM) features, 16 gray level run length matrix (GLRLM) features, 16 gray level size zone (GLSZM) features, 14 gray level dependence matrix (GLDM) features, 5 neighboring gray-tone difference matrix (NGTDM) features, 364 wavelet transform features, 91 square transform features, 91 square root transform features, 91 logarithmic transform features, and 91 exponential transform features. A total of 833 quantitative features were obtained for each color channel, and 3332 radiomic features were extracted from each region of interest. In order to determine the best combination of feature selection and machine learning modeling, the study used three commonly used feature selection algorithms: least absolute shrinkage and selection operator (LASSO), Random Forest and support vector machine-based recursive feature elimination (SVM-RFE), and three machine learning modeling algorithms: Logistic Regression, Random Forest, and SVM. Based on the above experiments, the authors found that the Random Forest selection-based SVM model in G-channel was the optimal ultrasonic radiomic model. By using radiomics quality evaluation (RQS) [258], their RQS score was 20 (55.6%, the total score is 36), showing the scientific integrity and reproducibility of this study. The study proved the application value of ultrasound-based radiomics in orthopedics and athlete imaging examination and may provide reference experience for other parts of ultrasound-individualized examination.

1.4.4.2 Criticism on Existing Studies

The current literature on diagnosing tendon pathology faces several limitations, hindering progress toward developing efficient diagnostic systems. Many studies rely heavily on manually extracted features for tendon grading, introducing the potential for human error and inconsistency. This approach is both time-consuming and costly, as it demands the expertise of skilled professionals, making it difficult to scale and reproduce across diverse datasets. Other studies have explored traditional machine learning models for classification, but these models often depend on handcrafted features that fail to capture the full complexity of tendon pathology, resulting in suboptimal performance in real-world clinical applications.

Moreover, while the majority of research has focused on image-based approaches, the field of automated tendon pathology detection remains relatively underexplored, with few studies investigating advanced methodologies. As previously discussed, the challenge of working with limited datasets necessitates a robust pipeline capable of capturing data in a more representative manner. To address these limitations, it is crucial to explore alternative representations of medical data, such as graph-based methods and superpixels, which can model the relationships between anatomical structures more effectively than traditional pixel-level analysis. Graph-based techniques allow for the integration of spatial and relational information among key structures, providing a deeper understanding of tendon abnormalities. By combining image-based approaches

with graph representations, the detection of pathological changes could be enhanced, resulting in more accurate and generalizable diagnostic models. Expanding research in this area may lead to the development of more robust automated systems for tendon pathology detection, ultimately improving diagnostic precision and reliability.

1.4.4.3 Inspiring Studies Shaping Our Proposal

Superpixels

In this image classification study, superpixels are used to generate graphs to perform graph classification, enabling the prediction of whether each graph (derived from an MRI slice) represents a pathological condition or not. Each superpixel functions as a node, with edges defined based on spatial or feature-based relationships between neighboring superpixels. This approach constructs a graph that captures both local and global image structures, where the nodes represent meaningful regions of interest and the edges encode their spatial arrangement. By applying graph-based learning techniques, such as Graph Neural Networks (GNNs), this graph-based representation facilitates effective image classification while maintaining crucial spatial and contextual information. Details on the superpixel generation process are already discussed in the previous section 1.3.4.3.

Graphs

Using graphs instead of conventional image grids for image processing and analysis offers several significant advantages, including flexible representation, efficient computation, enhanced feature extraction, improved analysis, scalability, robustness, and integration with other data types. Graph-based methods can represent images with varying resolutions and irregular structures more naturally, reduce computational complexity through sparse representation, and capture topological and geometric features more effectively, which is crucial for tasks like object recognition and image segmentation [259]. Additionally, they provide a robust framework for handling noise and distortions, support adaptive filtering, and facilitate the fusion of multi-modal data, making them highly suitable for complex applications such as medical imaging and remote sensing [260]. By leveraging graph theory algorithms, It is possible to gain deeper insights into image structures and relationships, thus enhancing the overall understanding and analysis of visual data [261]. For a more comprehensive understanding of graph-based networks, please refer to the previous section 1.3.4.3.

Reservoir Computing

Reservoir computing (RC) is an innovative computational paradigm designed to simplify the training of recurrent neural networks (RNNs) for tasks involving temporal data. This approach leverages the dynamics of a fixed, randomly connected network, known as the reservoir, to project input data into a high-dimensional space. By utilizing these rich dynamics, reservoir computing enables the effective processing of complex temporal patterns while significantly reducing the training complexity typically associated with RNNs. At the core of RC lies a high-dimensional, nonlinear "reservoir" system. This reservoir, often implemented as a recurrent neural network (RNN), receives the input signal and undergoes complex internal dynamics. The intricate interplay within the reservoir allows it to capture the inherent temporal information present in the input data. Crucially, the reservoir itself remains largely untrained, with only the readout layer requiring specific adjustments for the desired

task [262]. One of the primary advantages of reservoir computing is the simplicity of its training process. Unlike traditional RNNs, which require extensive training of all network parameters, reservoir computing only involves training the output layer. This significantly reduces computational requirements and allows for faster and more efficient model development. RC has emerged in the last decade as an alternative to gradient descent methods for training RNNs [263]. ESN [264] is one of the important RC models, being practical, conceptually simple, and easy to implement. It can avoid nonconverging and computationally expensive issues by applying the least squares problem as the alternative training method. Reservoir computing can be implemented using various types of networks, with Echo State Networks (ESNs) and Liquid State Machines (LSMs) being the most prominent. ESNs utilize a randomly connected recurrent network, while LSMs are inspired by biological neural microcircuits and use spiking neurons to emphasize temporal dynamics.

Reservoir Computing on Graphs

Graph Echo State Network (GraphESN) model is a generalization of the Echo State Network (ESN) approach to graph domains. GraphESNs allow for an efficient approach to Recursive Neural Networks (RecNNs) modeling. GraphESNs consist in a hidden reservoir layer of non-linear recursive units, responsible for the encoding process, plus a readout layer of linear feed-forward units. The reservoir is initialized to implement a contractive state transition function and then is left untrained, while the readout may be trained e.g. by using efficient linear models. The condition of contractility, inherited from ESNs and TreeESNs, assumes in this case a relevant specific meaning in terms of the extension of the class of data structures supported. Indeed, contractivity of GraphESN dynamics implies stability of state computation also in case of cyclic dependencies and thus allows to extend the applicability of the efficient RC approach to a large class of cyclic/acyclic, directed/undirected, labeled graphs. Moreover, being characterized by fixed contraction state dynamics, the GraphESN may represent an architectural baseline for other RecNNs implementing adaptive contraction state transition functions, such as GNNs. In this concern, it is interesting to empirically evaluate the effective distance occurring between the performance of methods based on fixed or adaptive encoding under the contractivity constraint. On the other hand, experimental results of GraphESNs, based on the inherent ability to discriminate among graph structures in the absence of learning of recurrent connections, represent a baseline performance for the class of RecNN models with trained recurrent connections [265].

Several interesting studies served as significant motivation for adopting this graph-based approach and, more broadly, the echo state network framework. These studies highlighted the effectiveness of graph based networks in capturing complex relationships within data and demonstrated the potential of reservoir computing techniques in various applications. Micheli et al. analyzed the impact of different reservoir designs on node classification accuracy and the quality of node embeddings computed by GESN, focusing on both dense and sparse reservoir layouts [266]. As measures of embedding richness, they adopted both graph topology-dependent metrics, previously employed in the analysis of embedding smoothing, and topology-independent metrics from the areas of information theory and numerical analysis. In particular, they proposed the application of entropy measures for quantifying information in node embeddings. Li et al. proposed a novel reservoir computing-based model called the Grouped Dynamical Graph Echo State Network (GDGESN) for handling spreading

process classification (SPC) tasks [267]. In this model, they designed a novel augmentation strategy named the snapshot merging strategy to form new snapshots by merging neighboring snapshots over time, and then multiple reservoir encoders were set to capture spatiotemporal features from the merged snapshots. Following that, logistic regression was adopted to decode the sum-pooled embeddings into classification results. Experimental results on six benchmark SPC datasets showed that their proposed model achieved better classification performance than DynGESN and several kernel-based models. In another study, the authors analyzed graph neural network models that combined iterative message-passing with untrained weights and graph pooling operations [268]. They alternated randomized neural message passing with graph coarsening operations to create multiple views of the graph, which were concatenated to form graph embeddings for classification. The proposed architecture's main advantage was its speed in computing graph-level representations, further enhanced by pooling. Results from popular graph classification benchmarks, comparing various topological pooling techniques, supported their claims. Sun et al. introduced Echo State Temporal Knowledge Graph Network (ESTN), an innovative approach leveraging the Echo State Networks for Temporal Knowledge Graphs (TKG) reasoning [269]. Specifically, the ESTN contained a Graph Embedder (GE) which integrated both short-term and long-term information within the TKG, enhancing the depth and accuracy of the data representation in graph latent space (embeddings). A Time Module (TM) was designed to facilitate direct prediction over multiple time steps, significantly improved the prediction efficiency. Extensive experiments demonstrated that their method not only show robust performance, but also offers an alternative approach to TKG reasoning, emphasizing how the distinct features of ESN can be effectively harnessed in TKG reasoning tasks. In a study of Micheli et al., authors addressed the challenges of heterophilic graphs with GESN for node classification [270]. The authors showed that reservoir models were able to achieve better or comparable accuracy with respect to most fully trained deep models that implement ad hoc variations in the architectural bias or perform rewiring as a pre-processing step on the input graph, with an improvement in terms of efficiency/accuracy trade-off.

Chapter 2

Tendon, Cartilage, and Muscle Interactions

This chapter delves into the complex interplay between key anatomical structures of the knee: the cartilage, quadriceps muscles, quadriceps tendon, and patellar tendon. By examining these components in tandem, we aim to gain a deeper understanding of their biomechanical relationships, how they function together in maintaining knee stability, and their roles in both musculoskeletal health and pathology. This exploration not only highlights their individual contributions but also emphasizes how disruptions in one structure can affect the others, leading to conditions such as tendinopathy or cartilage degeneration. Understanding these interconnected roles is crucial for advancing treatment strategies and improving rehabilitation outcomes.

The findings presented in this chapter are published in the proceedings of the "IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering". The full article can be accessed at doi: 10.1109/MetroXRAINE54828.2022.9967653.

Article Title: The Role of Muscle and Tendon in Predicting Cartilage Degeneration and Tendinopathy

Keywords: CT, MRI, Cartilage, Muscle, Tendon, Degeneration, Tendinopathy, Segmentation, Classification, ML

2.1 Introduction

Knee cartilage degeneration is the progressive breakdown of the cartilage tissue in the knee joint, leading to pain, reduced mobility, and conditions such as osteoarthritis. Knee cartilage degeneration is influenced by multiple factors, with age being one of the strongest determinants. Osteoarthritis (OA), a common condition resulting from articular cartilage degeneration, is characterized by joint pain, structural changes, and functional impairment [271]. Over 25 million people suffer from OA, with prevalence rates of 42.6% in men and 62.4% in women over the age of 40 [272]. Age-related changes in articular cartilage compromise its mechanical properties and strength, increasing its susceptibility to damage from mechanical loading while diminishing the ability of cells to repair and maintain the tissue [271]. Consequently, age is a key risk factor for primary OA and plays a critical role in the onset and progression of the disease. In the early stages of OA, there is typically an increase in cartilage water content, rising from the normal range of 60–85% to over 90%. Since cartilage has limited self-repair capacity, early detection of degenerative changes is essential. Measuring water content has been proposed as a valuable diagnostic tool, with non-destructive near-infrared spectroscopy (NIRS) offering a quantitative approach to assess water levels in articular cartilage [273]. Histological studies have confirmed elevated water content as an early indicator of OA [274], suggesting that water content determination could aid in the diagnosis of early degenerative changes [275].

Gender is also a significant factor in the development of osteoarthritis, with studies indicating that women are more susceptible to the disease, suggesting that gender differences influence its progression and natural history. In particular, women are more susceptible to the development of osteoarthritis. Factors that could contribute to their greater susceptibility include thinner cartilage, joint instability, and uneven mechanical loading. In addition, the development of osteoarthritis can be triggered by the sharp drop in sex hormone levels during menopause [276]. While many other factors may exert a lesser impact on the development of osteoarthritis, it is also possible to find a correlation between cartilage degeneration and tendons or quadriceps muscles. In fact, degenerative changes in the patellofemoral joint cartilage affect quadriceps muscle dysfunction [277]. In particular, quadriceps muscle thickness assessment can be considered a practical method in the diagnosis and follow-up of knee degeneration [278], so it is an important target for symptom management and functional decline. Additionally, few studies have examined the relationship between cartilage degeneration in the patellofemoral joint [279].

It is conceivable that degeneration and inflammation are not mutually exclusive processes, but work synergistically in the pathogenesis of tendinopathy [280]. Tendon injuries are common in sports as stresses and strains are imposed on the tendinous component during physical activity. Injuries are not restricted to competitive athletes but affect recreational sports participants and many working people [281]. A number of factors such as genetics and gender could also play role [282]. The essential lesion of patellar tendinopathy is a non-inflammatory failed healing response. Clinical assessment is the key to diagnosis. Patellar tendinopathy and quadriceps tendinopathy share similar histopathological features. Macroscopic intratendinous changes in the patellar tendon (PT) and quadriceps tendon (QT) include as poorly demarcated intratendinous regions with focal loss of tendon structure. The affected parts of the tendon lose their normal shiny white appearance and become gray and amorphous. The thickening can

be diffuse, fusiform or nodular [280]. Patients with chronic symptoms may exhibit quadriceps wasting [282]. In addition, associated with tendinopathy, another important factor may be the amount of systemic adiposity [283]. For example, a review of the literature demonstrated that 153 of 800 patients (19%) had varying degrees of quadriceps fat pad edema (QFP). The analysis showed significantly higher grades of QT interval alterations in patients with intermediate and intensive QFP edema compared to controls, as well as significantly different grades of QT alterations in patients with mild and intermediate QFP edema versus those with heavy QFP edema. In particular, an increase in mean QT thickness was observed in patients with intermediate and intensive QFP edema compared to controls [284].

Various classifications exist concerning the articular surface of the joint, including both the cartilage and the subchondral bone [285]. Additionally, recently new techniques have been implemented, such as near-infrared spectroscopy (NIRS), combined with machine learning techniques, to classify cartilage integrity. In this case, the interaction of light with cartilage structure and composition allows the identification of integrity-related features in spectra of increased OH peaks from increased water content for evaluation fabrics [286]. Machine learning has been used for another classification to distinguish hemiplegic and healthy reflex pairs using three feature categories of the reflex response acceleration waveform (global parameters, temporal organization, and spectral features) [287]. Machine learning has also been used to predict bone mineral density (BMD) changes in patients undergoing total hip arthroplasty (THA) [288].

By focusing exclusively on features from the quadriceps and patellar tendons, we aimed to gain a deeper understanding of the potential correlations between the quadriceps (muscle and tendon), patellar tendons, and cartilage. To achieve this, we employed various methods, including medical image segmentation, 3D knee modeling using specialized software, and feature extraction. We extracted diverse features from the patellar tendons and quadriceps, which were then used in machine learning models for patient classification. Two distinct classification models were developed, demonstrating their ability to predict the presence of cartilage degeneration and patellar tendinopathy. Fig. 2.1 illustrates our study workflow.

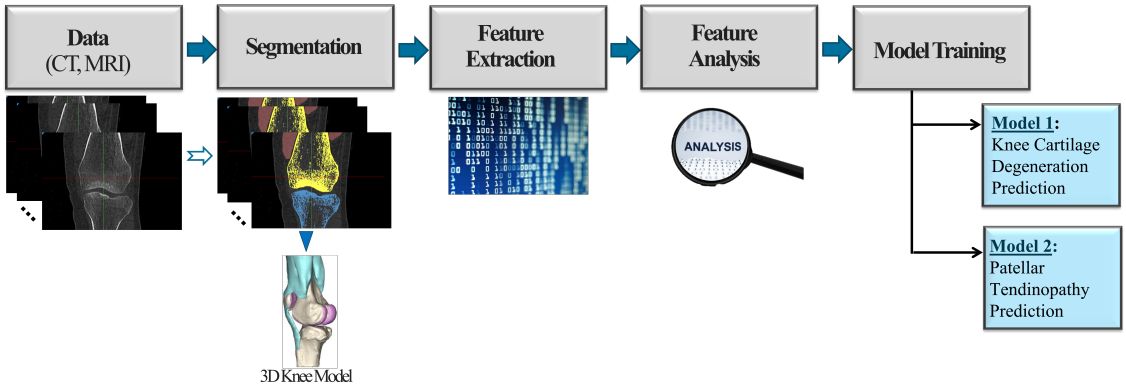


Figure 2.1: Graphical Pipeline

2.2 Materials and Methods

2.2.1 Dataset

The dataset used in this study includes 47 patients in total (24 females, 23 males, mean age = 50 ± 19 Y, mean height = 174.15 ± 9.77 cm, mean weight = 87.36 ± 17.39 kg, mean BMI = 28.946 ± 6.17). At Landspitali University Hospital in Iceland, these patients underwent plain radiography, CT, and MRI scanning of one knee. The main objective of the RESTORE project (for which this dataset was originally collected) was to implement patient-specific solutions for cartilage regeneration. Of the 47 patients, 24 (12 F, 12 M, mean age = 64 ± 12 Y, mean height = 173.292 ± 9.01 cm, mean weight = 91.71 ± 15.74 kg, mean BMI = 30.62 ± 5.41) had degenerative knee cartilage and were awaiting prosthesis replacement. 15 individuals (9 F, 6 M, mean age = 35 ± 11 Y, mean height = 173.47 ± 10.12 cm, mean weight = 87.0667 ± 16.67 kg, mean BMI = 29.27 ± 6.87) had knee trauma with probable cartilage injury, while 8 participants (3 F, 5 M, mean age = 34 ± 14 Y, mean height = 178 ± 11.63 cm, mean weight = 74.88 ± 19.38 kg, mean BMI = 23.25 ± 3.61) served as controls.

2.2.2 Data Pre-processing

Ground-truth mask generation is a critical step in data pre-processing, as these masks form the foundation for reliable analysis and model development in medical imaging studies. They ensure that the regions of interest (ROIs) are precisely identified. In this study, the masks were generated using Materialise Mimics, an advanced 3D design and modeling software developed by Materialise NV (https://www.materialise.com/en). For each individual, eight distinct ground truth masks were created.

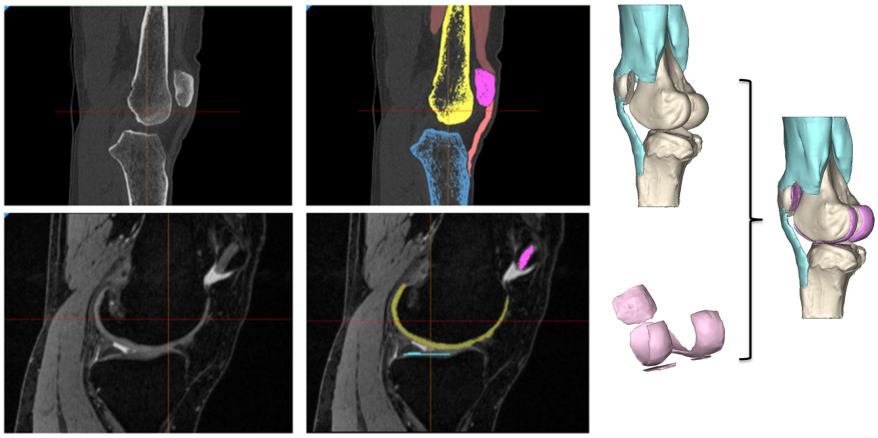


Figure 2.2: Original Vs. masked CT and MRI scans

Masks on CT: The first row of fig. 2.2 illustrates five masks derived from CT scans: three for bones (femur, tibia, patella) and two for other tissues (patellar tendon, quadriceps). These bone masks were generated using the Mimics software's default bone Hounsfield Unit (HU) threshold values (min: 226 HU, max: 3037 HU). To refine these masks, the region-growing feature was applied, resulting in more compact and accurate representations. For mask generation, 2D sagittal slices were selected for their superior visibility of the regions, and the interpolation feature was subsequently employed to create 3D models. Unlike bone masks, patellar tendon and quadriceps masks could not be generated using HU thresholds, as in Mimics, no default HU values

exist for these tissue types. Instead, these masks were manually drawn on each sagittal slice using the paint feature, following the anatomical shapes. Interpolation was again used to construct the complete 3D masks, with adjustments made for refinement as needed.

Masks on MRI: The second row of fig. 2.2 shows three cartilage masks derived from MRI scans. Similar to the patellar tendon and quadriceps, these masks were manually created using the paint feature. These bone and cartilage masks were generated as part of a separate study [289], [290].

For all masks, whether derived from CT or MRI, boundaries were carefully reviewed to ensure accuracy and to exclude surrounding tissues. The Mimics software's erase mask and add-to-mask options were used to refine the ROIs. Combining all individual masks, a comprehensive 3D knee model was constructed for each patient.

2.2.3 Data Analysis and Feature Extraction

In this phase, various features were extracted from both the patellar tendon, and quadriceps, categorized into different groups.

1. Category 1: Gray-level co-occurrence matrix (GLCM)
   This statistical approach to assess texture takes into account the spatial relationship of pixels. The features extracted from this category are homogeneity, contrast, dissimilarity, angular second moment (ASM), energy, correlation [291][292]. These features were extracted in four different pixel pair angles and also in two different pixel pair distance offsets (distance: 1 and 2), which resulted in 48 features in total.

2. Category 2: Amount of fat and water
   Another category of feature extraction was the total number of pixels, the total amount of fat and water present in the patellar tendon and quadriceps. A muscle threshold range of -400 to -51 HU was used to filter fat-containing pixels and -10 to 10 HU for water [293].

3. Category 3: Tendon thickness
   We also used tendon thickness as one of our features. The thickness measurements were derived exclusively for the patellar tendon using Materialise 3-matic, a powerful 3D data optimization software developed by Materialise. To automatically calculate the tendon thickness at various points on the model's surface, 3-Matic's built-in wall thickness analysis tool was used given the associated tendon 3D models. Later, the color mapping feature of the software was used to visualize thickness variations by color-coding the model's surface.

   A visualization of the patellar tendon thickness is provided later in fig. 2.5.

4. Category 4: Profile line analysis (PLA)
   The main goal of this analysis was to check whether there are changes in the radiodensity (HU) in the longitudinal direction of the patellar tendon, which could give us a signature. Fig. 2.3 shows the tendon formation with fibers and fiber placement direction. The direction (red arrow, same as fiber direction) shows the direction we looked for HU changes. We also wished to ascertain whether HU changes were present according to the presence and type of pathology and

whether such changes were sufficient to distinguish one patient group from another. Features extracted from PLA are AUC, slope, max HU, min HU, and avg. HU. We also extracted some additional features since our PLA provides similar to sequential data. These additional features (e.g., absolute energy, absolute sum of changes, approximate entropy, autocorrelation, skewness, kurtosis, etc.) were extracted following the study of reference number [294].

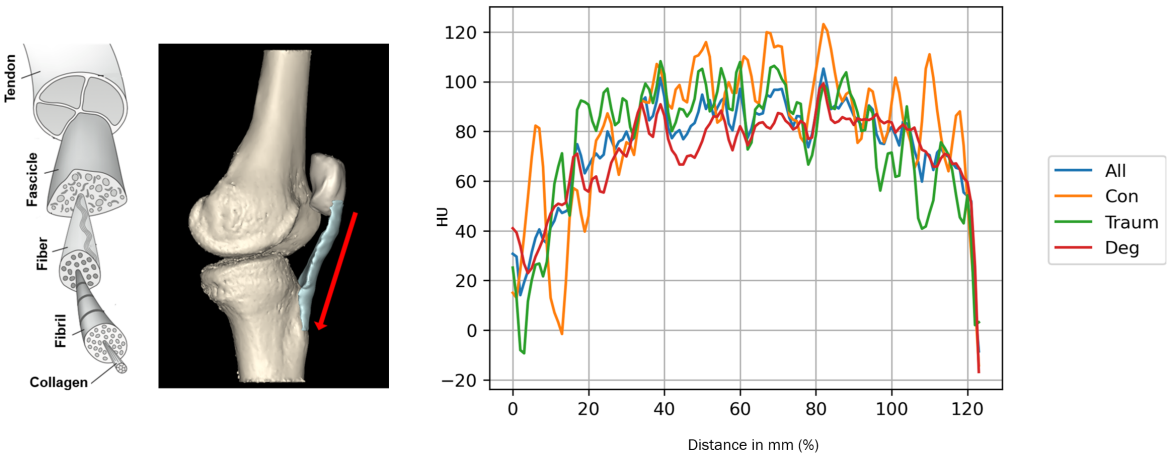


Figure 2.3: Profile Line Analysis (PLA)

5. Category 5: Radio-density related features
   Other extracted features are min HU, max HU, avg HU, std HU, skewness, kurtosis, peak value of histogram, peak value histogram position etc.

2.2.4 Feature Selection

For feature selection, an embedded method was chosen from the three available supervised strategies: filter, wrapper, and embedded. Embedded methods integrate feature selection into the learning algorithm, overcoming the limitations of both filter and wrapper methods while combining their advantages. One common approach is to use tree-based methods, such as Random Forest and Gradient Boosting, which provide feature importance scores. These scores indicate which features have the greatest impact on the target variable, helping to select the most relevant features for the model. In the case of a Random Forest classifier, feature importance is calculated by measuring how much each feature reduces impurity when used in a decision split across all trees in the forest. Features that contribute most to this reduction are considered important. The Mean Decrease Impurity (MDI) is used as a metric to evaluate feature importance. MDI averages the decrease in impurity (Gini index) across all splits using a particular feature. The Gini Index is a measure of class impurity used in decision trees, where a lower value indicates a more pure split. It is commonly used alongside MDI to evaluate the importance of features. MDI is relatively more stable than Mean Decrease in Accuracy (MDA) since MDA can fluctuate with small datasets as it depends on model performance and requires repeated evaluations. Future work could explore the use of MDA for comparison. On the other hand, Gradient Boosting builds trees sequentially and minimizes a loss function (Log Loss). At each split, the algorithm evaluates how much the split reduces the loss, and features that frequently cause large reductions are considered more important.

2.2.5 Model Training

Two distinct classification tasks were conducted: one for predicting cartilage degeneration and another for predicting patellar tendinopathy. For the prediction of cartilage degeneration, patients were divided into two groups. The patients of the traumatic and control groups were considered as a non-degenerative group and the other group as a degenerative group. Since the acquisition of this dataset was specifically focused on cartilage degeneration, it was quite difficult to identify the condition of the patellar tendons. But, after careful observation of CT and MRI data, the patellar tendons were labeled as tendinopathic or not (taking into account facts such as signal change in MRI, very minor insertional tendinopathy, tendon thickening, abnormal shape of tendon etc.). Different feature sets were used for different experiments for the two predictions. The best performing experiments using two classifiers (Random Forest and Gradient Boosting) are mentioned in the results section. Here, both of these classifiers' predictions are based on the decision tree's prediction potential. The Random Forest classifier (composed of many decision trees) uses bagging and randomization when generating each individual tree, with the goal of producing an uncorrelated forest of trees whose committee prediction is more accurate than that of any individual tree [295] [296]. Random forests are a popular machine learning method that have been increasingly used in biomedical applications [297]. It has also been shown to handle challenges arising from small sample sizes [298]. On the other hand, Gradient Boosting (GB) classifier combines several weak learning models to produce a powerful predictive model [299]. This classifier (GB) is typically optimal for small datasets [300]. In both classification models, a 10-fold cross-validation with stratified sampling was used to split our data. For model evaluation, Accuracy, Sensitivity and Specificity metrics are employed [301]. Here, to get an idea of overall performance, we use accuracy, which measures the ratio of correct predictions over the total number of instances evaluated. For more specific assessment, specificity and sensitivity are used where the sensitivity metric measures the fraction of positive cases that are correctly classified and the specificity metric measures the fraction of negative cases that are correctly classified [302].

2.3 Results

The analysis of the fat and water contents revealed significant differences between different patient groups. Fig. 2.4 illustrates these variations, highlighting their potential as predictors for patellar tendinopathy and knee cartilage degeneration. The first two plots in fig. 2.4 show the fat content in the patellar tendon and quadriceps across subjects with and without patellar tendinopathy. Subjects with patellar tendinopathy generally exhibited lower fat content in both regions. This finding may be explained by the breakdown of fat during the early stages of inflammation, where it serves as a local energy source for tissue repair. This reduction in fat content makes it a strong predictor for identifying tendinopathy, as the condition is consistently associated with decreased fat levels in both the patellar tendon and quadriceps. In contrast, the third and fourth plots in fig. 2.4 demonstrate that water content serves as a robust predictor for knee cartilage degeneration. The average water pixel count in both the patellar tendon and the quadriceps differs significantly between degenerative and non-degenerative groups, with cartilage degeneration being associated with lower water content. This

reduction likely indicates dehydration, which correlates with cartilage degeneration. Dehydration compromises cartilage health, contributing to degenerative changes in the knee. In summary, both fat and water content provide strong predictive markers for distinguishing between patient groups in the context of patellar tendinopathy and knee cartilage degeneration.

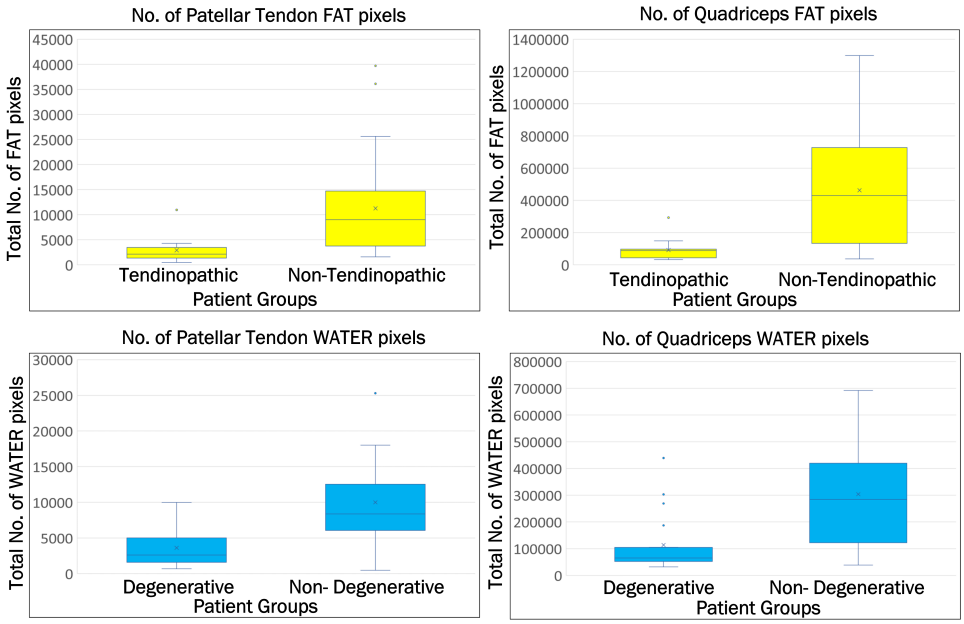


Figure 2.4: Fat and Water Contents in Patellar tendon and Qaudriceps

Fig. 2.5 demonstrates a clear distinction in patellar tendon thickness based on the presence or absence of inflammation. In this study, tendons affected by inflammation exhibited a maximum thickness of 5.75 mm and a mean thickness of 2.72 mm, while tendons without inflammation showed a maximum thickness of 5.17 mm and a mean thickness of 2.29 mm. These results suggest that inflammation significantly contributes to increased tendon thickness, which could serve as a valuable indicator for patellar tendinopathy.

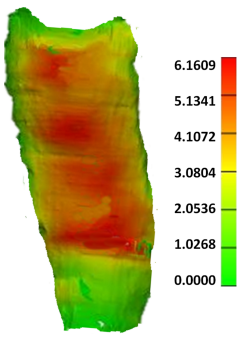


Figure 2.5: Sample Thickness

Fig. 2.3 shows that along the direction of the profile line, there is an increasing pattern in HU whereas when the tendon attaches to the patella and tibial tuberosity, HU is relatively lower. One possible explanation for the lower HU values in the insertional areas could be that these regions are covered by fat and likely contain different types of fibers as well. Furthermore, it is also noticeable that the increase in HU is

relatively less for degenerative patients compared to control (healthy). This difference clearly indicates that, for the pathological subjects, the tissue or material absorbs less X-ray radiation.

In terms of population distribution, the majority of the elderly were in the degenerative group, whereas the majority of the young subjects were in the control group. As a result, we sought to ascertain whether there was any correlation between age and degeneration using our PLA curves. There was only one older subject (67 years old) in the control group, and this person did not follow the control group's PLA, rather followed the degenerative group's. The average area under the curve (AUC) for degenerative subjects was 4841, whereas it was 5886.43 for control subjects. The AUC value for the only elderly subject in the control group was 4422.6, which is close to the AUC of degenerative subjects. Hence, there is an evidence of relationship between age and cartilage degeneration.

This association was once again observed in cartilage degeneration predictions. Table 2.1 presents the performance of two different classifiers whose main goal was to classify degenerative vs. non-degenerative patients. To perform this, several sets of features were used. The best performing experiment for predicting cartilage degeneration was the one which used all patellar tendon and quadriceps tendon features (except GLCM features and PLA features acquired as sequential data using reference number [294]) with 89.4% accuracy. Here, age was the most important feature and contributed 10% to this prediction. The other important features were the total number of patellar tendon pixels (6%) and the amount of water present in both quadriceps and patellar tendons (7% and 5% contribution respectively). The outcomes of the top two best performing experiments are shown in table 2.1.

The performances of patellar tendinopathy is shown in table 2.2. The total number of fat pixels present in quadriceps and patellar tendons, with an importance of 8.6% and 7.5% respectively, were the most importance features. The total number of pixels in patellar tendon was also important, accounting for 5% of the total importance. The peak value of the histogram (drawn with Hounsfield units of patellar tendon) (7.5% importance) and the maximum HU value of quadriceps were two other noteworthy features (6.4% contribution).

Table 2.1: Evaluation Metrics of Knee Cartilage Degeneration Prediction

Type of Features	Classifier	Accuracy	Sensitivity	Specificity
All Features (Except *)	RF	0.894	0.917	0.869
All Features(Except *)	GB	0.851	0.917	0.783

Table 2.2: Evaluation Metrics of Patellar Tendinopathy Prediction

Type of Features	Classifier	Accuracy	Sensitivity	Specificity
All Features (Except *)	RF	0.83	0.643	0.909
All Features (Except *)	GB	0.787	0.571	0.879

* GLCM features + PLA features extracted as sequential data ([294])

2.4 Conclusion

To the best of our knowledge, this is the first study to assess cartilage degeneration and patellar tendinopathy using only quadriceps and patellar tendon features. Previous studies, such as [289], analyzed a range of bone and cartilage features and employed machine learning techniques to predict cartilage conditions.

Our findings demonstrate that even solely considering the quadriceps (muscles and tendon) and patellar tendon is sufficient to predict cartilage degeneration. In particular, age, the total number of patellar tendon pixels, and the water content in both the quadriceps and patellar tendons were significant predictors of cartilage health. While other studies have established links between cartilage degeneration, age, and quadriceps water content, our work uniquely focuses on these features in the context of quadriceps and patellar tendon analysis, as reviewed in the literature.

For instance, in [303], subscapularis tendon tears were classified based on their 3D anatomical footprint. Similarly, Dakin et al. [304] characterized supraspinatus tendons across early, intermediate, and advanced stages of disease, emphasizing inflammation pathways mediated by interferon, nuclear factor κB, glucocorticoid receptor, and activator of transcription 6. Olewnik et al. [305] proposed a classification of the popliteal muscle tendon based on its proximal attachment.

Our study further highlights the pivotal role of fat content in both quadriceps and patellar tendons in predicting patellar tendinopathy. The average number of quadriceps fat pixels in the tendinopathic group versus the non-tendinopathic group was 11265.17 and 462625.08, respectively. Similarly, the average number of fat-containing pixels in the patellar tendon was 2914.69 for the tendinopathic group and 91967.61 for the non-tendinopathic group. This novel finding underscores the direct impact of fat content on patellar tendinopathy prediction and provides a foundation for future research into the causes and correlations of tendon degeneration. Together, our two distinct classification models create a framework for understanding the interactions between muscle, tendon, and cartilage.

This study, however, has limitations. The relatively small sample size (47 patients) and the limited representation of patients with tendon pathology pose challenges. Moreover, as the dataset was primarily designed to study cartilage degeneration, extracting comprehensive information about quadriceps and patellar tendons was difficult. In many cases, patellar tendon lengths were truncated, and similar issues were observed with quadriceps, making it challenging for even expert radiologists to interpret the status of the patellar tendons.

Another critical consideration is the potential for variability in radiologists' interpretations of subtle pathologies, which highlights the importance of integrating artificial intelligence and automated medical image analysis to support clinical diagnostics. Although tendinopathy is not typically evaluated using CT scans in a clinical setting, combining multiple imaging modalities offers a more comprehensive evaluation of anatomical structures. Future work should extend these methods to other tendons and joints and validate findings through surgical correlation.

Chapter 3

Medical Image Segmentation

This study presents a comprehensive, end-to-end tendon segmentation module. The proposed approach incorporates a two-step segmentation process: an initial coarse segmentation based on superpixel techniques, followed by a refined final segmentation step. The preliminary superpixel segmentation provides a coarse but computationally efficient estimate of the tendon region, while the subsequent refinement step ensures more precise boundary delineation. This modular approach aims to improve both segmentation accuracy and processing speed, contributing to more reliable analysis and diagnosis of tendon-related pathologies in clinical settings.

This work has already been published in a Q1 ranked journal named "Computer Methods and Programs in Biomedicine" (Impact Factor: 4.9, CiteScore: 12.3), doi: https://doi.org/10.1016/j.cmpb.2024.108398.

Article Title: Beyond pixel: Superpixel-based MRI segmentation through traditional machine learning and graph convolutional network

Keywords: Magnetic Resonance Imaging, Superpixel, Graph Convolutional Network, Segmentation via Node Classification, Achilles Tendon

3.1 Introduction

Tendons are strong connective tissues composed of collagen fibers that connect muscles to bones, transmitting and absorbing force. The Achilles tendon is the strongest tendon in the human body, is prone to injury, with Achilles tendinosis and tendinopathy being common, particularly among middle-aged men, and significantly affects the quality of life. Medical imaging techniques such as ultrasound and MRI play a crucial role in diagnosing tendinopathy, while machine learning can further enhance diagnosis by analyzing image data to detect and predict tendon abnormalities. Khatun et al. investigated various features of the quadriceps muscle and patellar tendon to assess their relationship with cartilage degeneration and tendinopathy [306]. Additionally, the study by [307] examined different features from medical images to explore bone and cartilage. However, research related to tendinopathy presents several challenges due to the complexity of the condition and the lack of knowledge of the exact reasons underlying the pathologies. Before categorizing the pathology, it is crucial to conduct tendon segmentation to identify the ROI. Only after this step can further studies be pursued.

In computer vision, image segmentation encompasses a large class of finely related problems. Gupta et al. [119] proposed an automated segmentation of supraspinatus tendons using ultrasound images (US) by image processing techniques. It integrates the curvelet transformation and concepts of logical and morphological operators. An adaptive texture-based Active Shape Model was suggested by Chuang et al. [120] to segment tendon and synovium sheath. Martins et al. [125] proposed a segmentation approach of finger extensor tendon in ultrasound images based on an active contour framework. One drawback of active contour, active shape, and curvelet transform-based segmentation techniques is that edge/gradient information, which is used to guide contour deformation, is not reliable in ultrasound due to the presence of speckle noise and imaging artifacts [308]. [244] performed tendon grading for tear thickness, tear size, etc. by manually extracting features such as non-fluid signal intensity-related signal changes, anteroposterior dimensions, etc., which is known to be costly. Similarly, [250] utilized MRI data, with specialists manually extracting parameters like length, width, and thickness, incurring both time and cost expenses. Recently, superpixel segmentation has attracted a lot of interest in computer vision as it provides a convenient way to compute image features and reduces the complexity of subsequent image-processing tasks. Xu et al. developed a method using a machine learning algorithm based on variable-size super pixel segmentation [168]. In [171], authors proposed a new similarity-based superpixel generation method that was integrated with texton representation to form a spatio-color-texture map of the breast histology image. Zhu et al. [172] proposed a novel lung cancer detection method for CT images based on the superpixels and the level-set segmentation methods. Also, Signoroni et al. and Wan et al. utilized superpixel-based segmentation [167] [309]. According to the literature, superpixels can be crucial from several perspectives in the case of medical image segmentation.

In traditional machine learning, unsupervised segmentation techniques analyze image intensity or gradients, performing well when boundaries are clear. Supervised methods, however, use training samples to incorporate prior knowledge. Deep learning (DL) models have significantly advanced image segmentation, often achieving superior accuracy on benchmark tasks. Kuok et al. proposed a unique finger tendon segmenta-

tion technique where a hybrid of effective convolutional neural network techniques was applied [124]. However, these types of algorithms are usually considered data-hungry. This means that these algorithms mostly require large amounts of high-quality labeled data to effectively learn and generalize patterns. The performance of many machine learning models, including deep learning neural networks, tends to improve as the volume and diversity of training data increase. Despite their potential, there is a lack of research on tendon segmentation using deep learning.

Machine learning in graphs, or graph neural networks, is a rapidly evolving field, focused on developing algorithms to extract valuable information from structured data. Aiming for automated segmentation, Cai et al. proposed a graph-based decision fusion process combined with deep convolutional neural networks (CNN) [200]. Tian et al. proposed an interactive segmentation method based on a graph convolutional network (GCN) to refine the automatically segmented results [203]. Node classification is perhaps the most popular machine-learning task in graph data, especially recently.

Given the existing literature landscape, we propose a comprehensive end-to-end system tailored for Achilles tendon segmentation, eliminating the need for expensive manual feature extraction or an extensive dataset. Our segmentation method involves two distinct approaches. As an initial step for both approaches, our data (MRI) undergoes coarse segmentation, which is based on superpixel generation. Superpixels are characterized by perceptually homogeneous regions. In comparison to pixel representation, superpixel representation decreases the number of image dependencies and offers better support to identify regions depending on image properties [152]. This phase provides a preliminary segmentation that is supposed to reduce the complexity of the final segmentation task. In our first approach, some of the generated superpixels contain tendon region, while others contain different tissues. Using traditional machine learning, the classification of each superpixel as tendon or non-tendon will result in tendon segmentation. Moving on to our second approach, these superpixels are organized into graphs rather than conventional grids. Graph neural networks, such as GCNs, excel at categorizing graph topologies and yielding a unified categorization of nodes. In our study, this type of graph neural network is employed to distinguish superpixels/nodes as tendons or non-tendons. By primarily relying on superpixels for the initial segmentation (coarse segmentation) and subsequently leveraging graph neural networks, we adopt a strategy that is less data-intensive. To sum up, our system comprises the following elements:

1. Utilization of Simple Linear Iterative Clustering (SLIC) algorithm to generate superpixels from MRI data.
2. Extraction of 94 radiomics features from each superpixel.
3. Classification of each superpixel as tendon or non-tendon, leading to tendon segmentation through two approaches:
    - Superpixel classification using Random Forest and Support Vector Machine classifiers.
    - Superpixel/Node classification based on Graph Convolution Network (GCN).

An overview of our proposed framework to perform Achilles tendon segmentation is shown in subsection 3.2.1. More details about this pipeline are highlighted in the next sections.

3.2 Materials and Methods

3.2.1 Workflow

The pipeline of our work is illustrated in fig. 3.1 which contains several steps. Each of the steps is discussed in detail in the following sections.

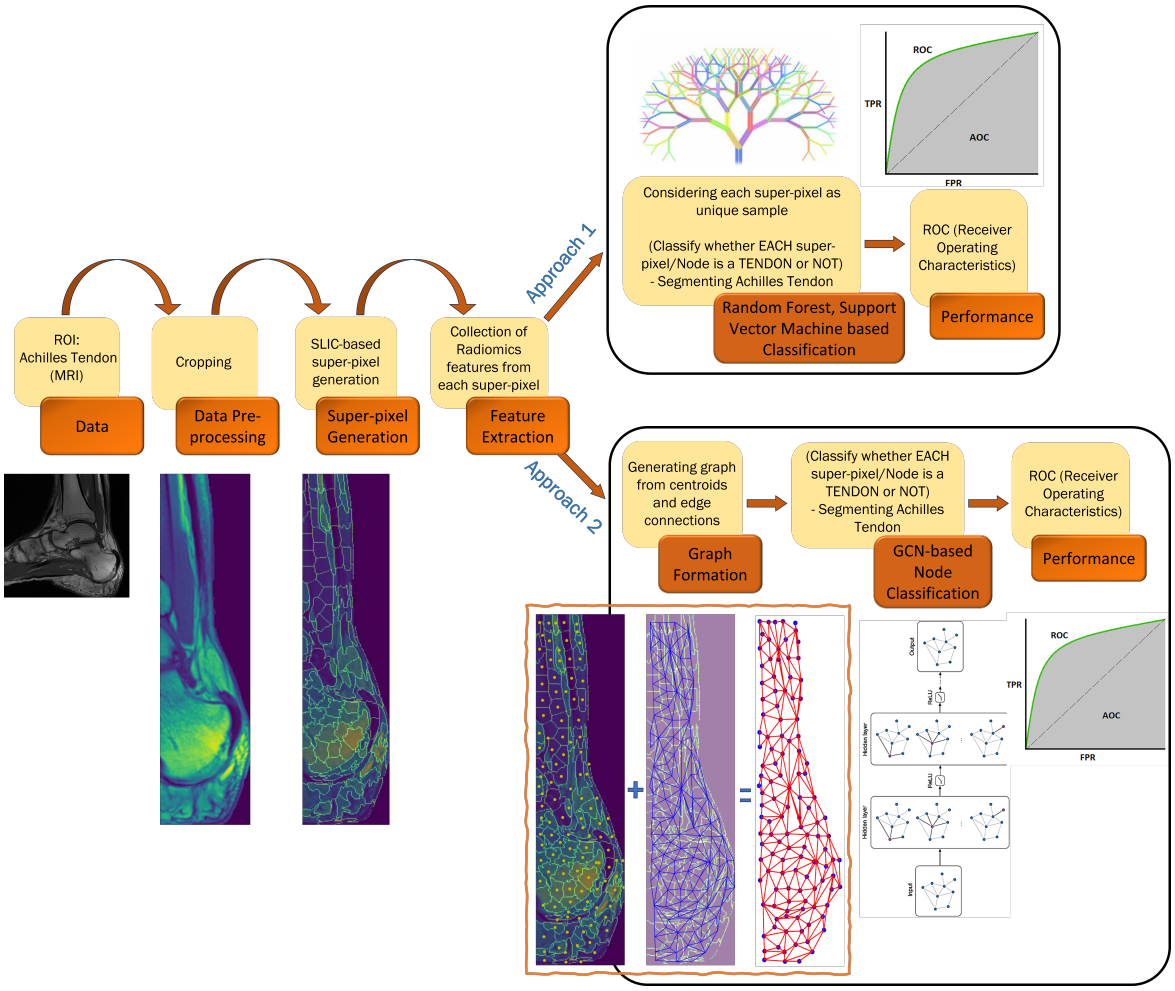


Figure 3.1: Graphical Pipeline

3.2.2 Data

The dataset employed in this study includes a diverse cohort of 76 subjects with a mean age of 45 years and a standard deviation of 19.8 years, including 42 men and 34 women. The participants went through MRI scans at Landspitali University Hospital in Iceland. The study participants are from distinct diagnostic categories. In particular, the Achilles tendon status of a total of 47 participants is listed as healthy. In 11 participants, calcific insertional tendinopathy is identified, characterized by the presence of calcium deposits at the insertion point of the tendon on the calcaneus. Six participants shows typical symptoms of Achilles tendinitis, which is an inflammation of the Achilles tendon. Two participants have a chronic tear in the distal part of the tendon. Lastly, the remaining 10 participants sustain an Achilles tendon injury.

Fig. 3.2 depicts two randomly chosen MRI slices from different subjects to show what normal Achilles tendons look like (dark band, marked with dotted red color). The slices were chosen to display sagittal views of the Achilles tendon, an imaging plane oriented parallel to the sagittal plane of the body.

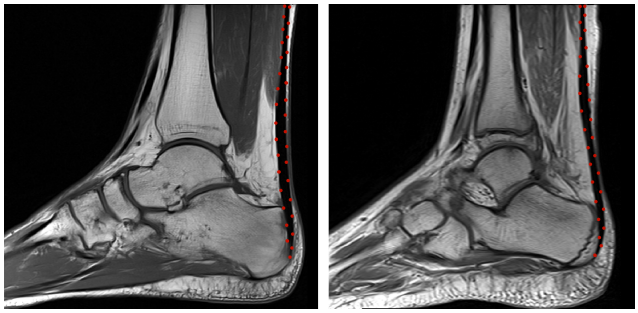


Figure 3.2: Healthy Achilles Tendons

Instead, fig. 3.3 shows two randomly selected pathological samples. These samples are crucial to our study since they offer in-depth information on deviations from normal subjects and shed light on a range of abnormalities.

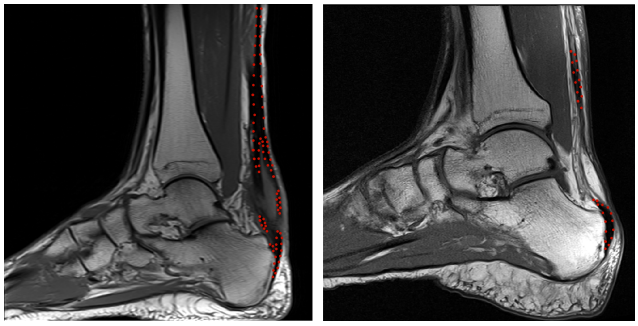


Figure 3.3: Pathological Achilles Tendons

Furthermore, the dimensions of the MRIs are variable. Some of the example dimensions are 384 x 384 x 19, 448 x 464 x 26, etc. Where the format X x Y x Z indicates the matrix size, with X and Y indicating the number of pixels or data points in the horizontal and vertical directions of each 2D image slice, and Z representing the number of slices. The slice thickness and pixel spacing also vary, with slice thickness typically ranging from approximately 3 to 5 mm, and pixel spacing from approximately 0.25\0.25 to 1.0\1.0. In this study, a T1-weighted Turbo Spin Echo (T1 TSE) MRI sequence is used. This sequence uses a specific sequence of radiofrequency pulses and magnetic fields to create images. This particular sequence is especially useful for imaging soft tissues, such as tendons, muscles, fat, etc.

3.2.2.1 Amount of Data

In this study, a careful selection process is employed from a cohort of 76 subjects, focusing exclusively on the 2D MRI slices containing the region of interest (ROI). Other slices not containing ROI are excluded. This thorough selection procedure resulted in a substantial dataset, comprising a total of 411 MRI slices. The rationale behind this choice is rooted in the observation that, on average, each subject in our study group provided a notable subset of 4 to 5 sagittal TSE MRI slices featuring

Achilles tendon. This selection is made to ensure that the dataset is complete and suitable for subsequent in-depth analysis of the Achilles tendon.

3.2.3 Data Pre-processing

3.2.3.1 Mask Generation

Ground truth masks are an essential component of data-driven work since they serve as the basic block for building, evaluating, and training machine learning models. To generate ground truth masks for our ROI, a software called Materialise Mimics is used, which is an image processing software for 3D design and modeling developed by Materialise NV (https://www.materialise.com/en).

Fig. 3.4 demonstrates the final Achilles tendon mask (in green) created by the above-mentioned software. It shows a particular example of an original Achilles tendon with pathology and the associated generated mask. Fig. 3.4, serves as a more metaphorical illustration that emphasizes the complexity of specific occurrences in our database. This specific figure is an example of a situation where creating a ground truth mask required greater attention and assistance from experts. This particular case draws attention to the range of difficulties that can be encountered when making precise masks. To reduce human bias while generating these ground truth masks, two different individuals were involved in the mask generation and correction phases. Different experts and time points provided a rigorous validation method for the consistency and quality of the created masks.

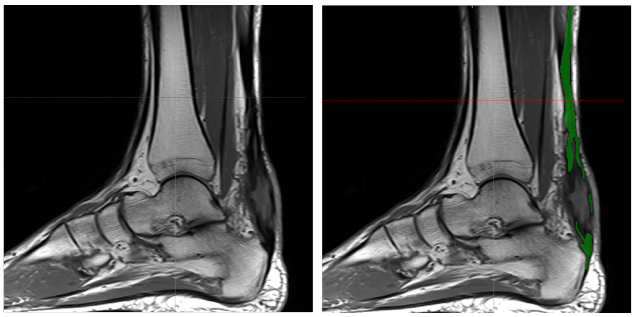


Figure 3.4: Generated Mask (Pathological Subject)

3.2.3.2 Cropping

One of the fundamental image pre-processing techniques is image cropping, which attempts to eliminate unexpected regions and unnecessary noise from an image by altering its aspect ratio or improving its composition. It is also important, from this point of view, that not all regions of the MRI image are strictly relevant or very informative for the diagnostic purposes of a particular ROI. Additionally, image cropping proves valuable in making data more manageable, especially considering the vast size of full-scale MRI images, which consume significant storage space. Moreover, from a computational standpoint, data cropping serves as a vital step in our data pre-processing.

The following steps are followed to crop both the MRI data and the associated masks:

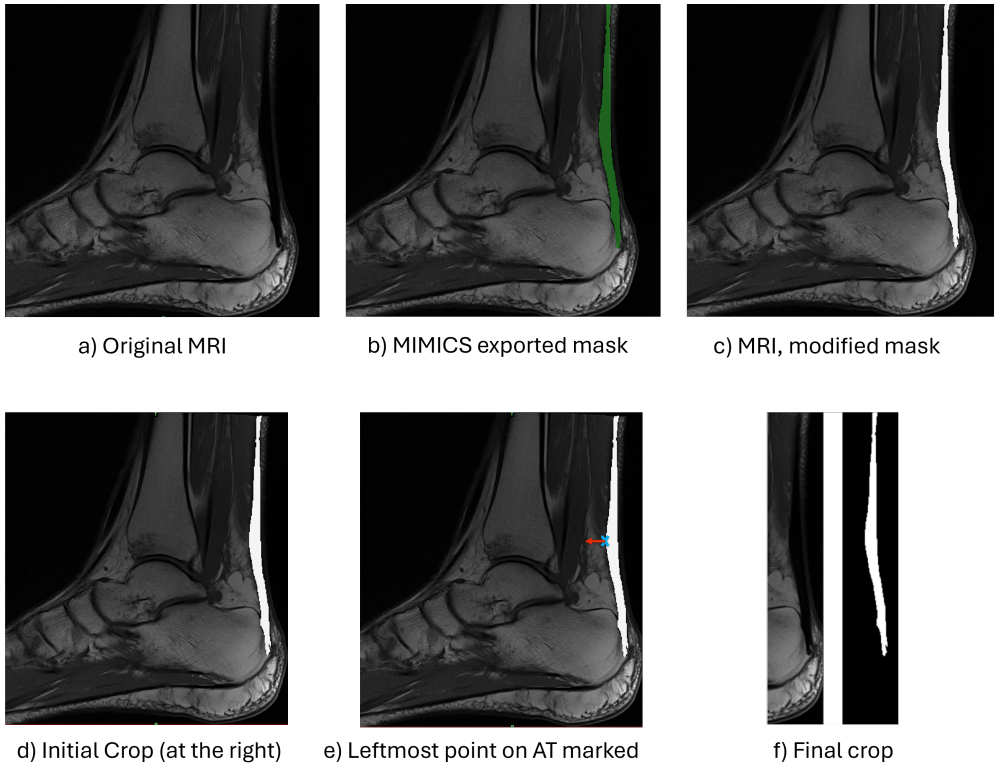


Figure 3.5: Cropping Pipeline

1. Loading individual MRI slices and corresponding ground truth masks: Each MRI slice is loaded, which typically represents a 2D image depicting a sagittal cross-section of the anatomy of interest. At the same time, the corresponding ground truth mask is loaded, which is another 2D image indicating the ROI highlighted within the MRI slice. As an example, fig. 3.5a represents a sagittal MRI slice, and fig. 3.5b represents it's ground truth mask.

2. Conversion of ground truth masks to binary masks: Our MIMICS-exported ground truth masks are initially in color but are subsequently converted to binary masks for further processing. Subsequently, for each MRI slice, the binary mask is superimposed on the MRI slice to highlight the ROI. This visual overlay is illustrated in fig. 3.5c, where all tissues are displayed in their natural appearance, with only the Achilles tendon highlighted in white.

3. Finding contour: This initial cropping phase consists of identifying the contours of the leg. This contour follows the shape of the calcaneus bone and executes the primary cropping step, as illustrated on the right side of fig. 3.5d.

4. Recording the full width of the MRI slices: For each subject, the entire width of the MRI slice image is recorded. This width serves as a reference for subsequent cropping operations.

5. Recording tendon's leftmost point for in-depth analysis: For a more detailed analysis of the ROI, the leftmost point within the Achilles tendon mask is recorded. This step is crucial for gaining a comprehensive understanding of the spatial extent of the ROI. The position of the leftmost point in a specific case is depicted in fig. 3.5e (in blue with a cross mark).

6. Final crop based on the leftmost coordinates and widths: The final cropped image, followed by a bounding box, is determined using the coordinates of the leftmost point obtained in the previous step. Each image is cropped to include the left portion of the Achilles tendon, starting from the leftmost point. This cropped region has a width equal to 7% of the total image width on the left (starting from the leftmost point) and right sides remaining as usual from step 3, 3.5d. The cropping operation produced the final representation of the ROI, illustrated in fig. 3.5f. This approach ensures precise extraction of the desired region for further analysis and examination.

Some additional examples of cropped images are shown later in fig. 3.6.

During testing, it is important to emphasize that the test ground-truth labels are not revealed. Cropping of the test data is conducted based on the leg contour, starting from the rightmost point of the calcaneus bone, with a portion of the image width cropped to the left side. More specifically, the first step is to identify the rightmost point of the calcaneus bone (fig. 3.5d), and from that point, approximately 24% of the total width of the initially cropped image (fig. 3.5d) is cropped to the left. This procedure mirrors the earlier steps, with two key differences: for testing, the reference point is the rightmost point of the calcaneus bone instead of the leftmost point of the ROI, and a larger percentage of the image width is chosen, as the reference point this time is shifted slightly to the right. This approach ensures that the test data encompasses both the region of interest (ROI) and adjacent areas, maintaining consistency in input formatting with the training data, all without relying on the test ground-truth labels for cropping.

3.2.4 Superpixel Generation

The purpose of this step is to generate superpixels on cropped data. Superpixels are perceptual groupings of pixels or over-segmented segments of an image. They capture image redundancy, provide a convenient primitive for computing image features, and significantly reduce the complexity of subsequent image processing tasks [214].

To achieve this, Simple Linear Iterative Clustering (SLIC) algorithm is used. Generally, SLIC algorithm generates superpixels by clustering pixels based on their color similarity and proximity in the image plane, using a five-dimensional [labxy] space, where [lab] represents the pixel color vector in the CIELAB color space and [xy] represents the pixel position. However, for grayscale images, intensity similarity (I) is used instead of color similarity. This is done in a three-dimensional [ixy] space, where [i] represents the pixel intensity and [xy] represents the pixel position. It is necessary to normalize the spatial distances to use the Euclidean distance in this 3D space because the maximum possible distance between two intensity values is fixed, whereas the distance in the XY plane depends on the image size [310]. Further details about this algorithm are outlined below:

1. Initialization:

   - *Grid Placement*: Divide the image into a grid with initial cluster centers placed roughly equal-sized spaced. The spacing S between these centers is determined by the desired number of superpixels K. Typically, $S = \sqrt{N/K}$, where N is the total number of pixels in the image.

- *Refinement of Centers*: Adjust the initial cluster centers to positions with the lowest gradient in a 3×3 neighborhood to avoid placing centers on edges.

2. Distance Measure:

   - *Intensity and Spatial Proximity*: Define a distance measure D that combines intensity similarity and spatial proximity. For a pixel i with intensity I_i and coordinates (x_i, y_i), and a cluster center k with intensity I_k and coordinates (x_k, y_k), the distance D is computed as:

$$D = \sqrt{(d_c)^2 + \left(m\frac{d_s}{S}\right)^2}$$

     where:

     - $d_c = |I_i - I_k|$ is the absolute difference in intensity.
     - $d_s = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}$ is the Euclidean distance in spatial coordinates.
     - m is the compactness parameter that regulates the trade-off between intensity similarity and spatial proximity.
     - S is the grid spacing.

     The term $\left(m\frac{d_s}{S}\right)^2$ adjusts the spatial distance component based on the compactness parameter m, which influences the relative importance of spatial distance compared to intensity distance.

3. Assignment of Pixels to Nearest Centers:

   - *Local Search*: For each cluster center, consider all pixels within a $2S \times 2S$ region around the center. Assign each pixel to the nearest cluster center based on the distance D.

4. Update Cluster Centers:

   - *Recompute Centers*: After assigning pixels to cluster centers, update each center to be the mean intensity and spatial position of all pixels assigned to it. This adjustment ensures that the center represents the average characteristics of its superpixel.

5. Iterate:

   - *Repeat Assignment and Update*: Repeat the assignment and update steps iteratively until convergence is reached. Convergence typically occurs when cluster centers no longer change significantly between iterations or after a predefined number of iterations.

6. Enforce Connectivity:

   - *Post-processing*: Ensure that each superpixel forms a contiguous region. This is often achieved by reassigning isolated pixels to the nearest large superpixel, ensuring spatial coherence.

The simplicity of this approach makes it extremely easy to use a single parameter that specifies the number of superpixels, and the efficiency of the algorithm makes it very practical [214]. To implement this, the cv.ximgproc.createSuperpixelSLIC() function is used, which is part of the OpenCV library, a widely used library in the computer vision community. The results of the superpixel generation step for both a healthy and pathological case are presented later in the result section 3.3. The following parameters are used in our superpixel generation approach:

1. image:

   This parameter searches for the input image of interest on which to perform superpixel generation/segmentation.

2. region_size:

   This parameter controls the compactness and size of the superpixels in the algorithm's initial clustering step. It is a purely computational parameter and is measured in pixels, not real-world units like millimeters or inches. This parameter defines the average size of the superpixels. Smaller values result in smaller superpixels, while larger values result in larger superpixels.

   In many image processing tasks, professionals often opt for a fixed region size for superpixels. However, our study faces a particular challenge due to the size variability within our input images. Therefore, we took a more adaptable approach by tying the region size to a ratio relative to the original image height. This choice is supported by the observation that the image height remains relatively stable even after cropping. Our decision to use $(13/375) \times image\_height$ as the region size is validated through rigorous experimentation. We specifically tested this ratio with both healthy and pathological cases and found that it effectively captured our ROI. One might interpret it as follows: In the context of an image with an approximate height of 375, a superpixel size of 13 is considered an ideal choice for defining the size of regions. Therefore, this dynamic approach to region size calculation offers flexibility between different image sizes, ensuring that superpixel generation remains effective and adaptable. It is worth noting that after applying this ratio-based approach to all cases, we achieved an average region size of approximately 13.8345. This level of consistency underlines the reliability of our initial approach.

3. Compactness Factor:

   It is one of the other important parameters in our study that plays a significant role in shaping the properties of the generated superpixels. When the compactness factor is set to a very high value, the resulting superpixels tend to be extremely compact, meaning they may not conform well to natural boundaries within the image. On the other hand, when this value is small, the superpixels are more closely aligned to the edges of the image, but their sizes and shapes may become less regular and uniform. In this specific study, we carefully selected a compactness factor value of 10. This choice was made after our careful experimentation and analysis. By setting the compactness factor to 10, there is a balance between these two extremes. It allows the generation of superpixels, which are not very compact, to still maintain a strong adherence to the boundaries of the underlying image. This careful selection allowed us to obtain a

segmentation result that aligns well with the natural form of the image, achieving a balance between compactness and adherence to boundaries.

3.2.4.1 Background Removal on SLIC-based Coarse Segmentation

Background removal is a widely used technique in the field of image processing. Specifically, it is a critical preprocessing step in SLIC-based medical image segmentation. It not only improves segmentation accuracy but also improves computational efficiency, reduces errors, and facilitates consistent and interpretable results. The quality of background removal can also be affected by the complexity of the background, including any patterns, textures, and overlapping elements. Complex backgrounds can be more difficult to remove and can produce an uneven or inconsistent final product. In our study, this step is a challenge due to the distinctive characteristics of our ROI, namely the Achilles tendon located close to the contour of the leg or background. The adjacency of this region to the image boundary introduced greater complexity to the task.

Furthermore, as illustrated in fig. 3.6, different ankle MRI images show a wide range of attributes at their edges. In some cases, a thin layer of fatty tissue is present. Additionally, there were scenarios where the boundary comprised predominantly fat, and in some other cases, the boundary exhibited a lack of clear delineation. These multifaceted scenarios made the simple application of conventional background removal techniques impractical. Given these distinct conditions, it was necessary to devise a more specialized and personalized approach to effectively address this issue of background removal in the context of our study. The steps involved in the background removal process of this study are shown below:

1. Gaussian Blur: This step reduces noise and enhances the overall image quality.

2. Grayscale Conversion: The image data is then converted to grayscale. The data type is changed from int16 to uint8 to ensure compatibility and appropriate representation for subsequent steps.

3. Contrast Enhancement: Next, Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to enhance the contrast of the grayscale images. This step helps in making the image details more prominent.

4. Thresholding: OpenCV's binary thresholding technique, following the OTSU method, is applied to contrast-enhanced images. This operation is used to create a binary mask, which helps to distinguish foreground from background regions.

5. Bitwise Operation: The binary mask obtained in the previous step is used for a bitwise operation with the contrast-enhanced image. This step effectively isolates the regions of interest while suppressing the background.

6. Morphological Closing: Later, a morphological closing operation is performed using a morphological ellipse with a size of 5×5. This step helps in closing small gaps or holes in the foreground regions.

7. Morphological Dilation and Erosion: To further refine the binary mask and ensure a clean separation between the foreground and background, morphological

dilation (kernel of 7×7, iteration = 1) followed by morphological erosion (kernel of 3×3, iteration = 7) is applied. These operations help to smooth the edges and eliminate any remaining artifacts.

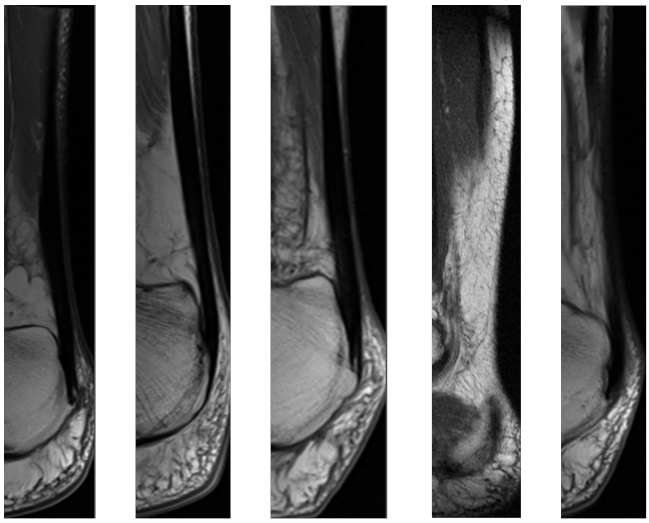


Figure 3.6: Cropped Images (Variations in boundaries)

The outcome of this step, displaying superpixels without background, for both a healthy case and a pathological case is presented in the results section 3.3, particularly in fig. 3.11a and 3.12.

3.2.5 Feature Extraction

As illustrated in our workflow, our methodology involves extracting radiomic features from the generated superpixels. Radiomics is an emerging field of research that addresses the extraction of mineable high-dimensional data from medical images. Radiomics enables the extraction of a significant number of quantitative characteristics from standard-of-care images from modalities like CT, MRI, and PE [311]. Leveraging the feature extractor provided by PyRadiomics (an open-source Python package), a total of 94 radiomics features are extracted [312] which are categorized into the following groups:

- First Order Statistics (19 Features): First-order radiomics features are statistical measures that describe the distribution of pixel intensities within a region of interest in an image. They include metrics such as mean, median, standard deviation, skewness, kurtosis, entropy, range, interquartile range etc.
- Gray Level Co-occurrence Matrix (24 Features): This is a method used to analyze the spatial relationships between pixel intensities in an image. It generates a matrix that records how frequently different pairs of pixel with specific values occur in a specified spatial relationship (e.g., horizontally, vertically) within a region of interest. From this matrix, 24 features are extracted, including measures of contrast, correlation, energy, homogeneity, etc.
- Gray Level Run Length Matrix (16 Features): These features are used to analyze texture by evaluating the lengths of consecutive runs of pixels with the same

intensity. It produces 16 features including measures like Short Run Emphasis which highlights the prevalence of short runs, and Long Run Emphasis which focuses on longer runs. Other features include Gray Level Non-Uniformity, Run Length Non-Uniformity, etc.

- Gray Level Size Zone Matrix (16 Features): This is a radiomics method used to describe the texture of an image by analyzing the size of zones with uniform gray levels. It generates 16 features that include metrics such as Zone Percentage which measures the proportion of zones of a certain size, and Gray Level Non-Uniformity, which quantifies the variability in zone sizes across different gray levels. Other features like Zone Size Non-Uniformity and Small Zone Emphasis capture the distribution and emphasis of various zone sizes.

- Neighbouring Gray Tone Difference Matrix (5 Features): This radiomics method is used to characterize texture by assessing the differences in gray tone values between neighboring pixels. It calculates five key features: Coarseness, which measures the average size of the texture patterns; Contrast, which reflects the variation in gray levels between neighboring pixels; Busyness, indicating the frequency of gray tone changes; Complexity, which captures the texture's irregularity; and Strength, which quantifies the intensity of texture patterns.

- Gray Level Dependence Matrix (14 Features): These features are used in texture analysis to describe the spatial relationship between pixels in an image. It captures how frequently certain gray levels occur at specific distances and orientations relative to each other.

3.2.5.1 Data Standardization

Prior to feeding the features into classifiers, they are standardized using *StandardScaler* from scikit-learn library [313]. The *StandardScaler* standardizes features by removing the mean and scaling to unit variance, following the formula:

$$z = \frac{x - \mu}{\sigma}$$

where x represents the original feature value, μ is the mean of the training data, and σ is the standard deviation of the training data. First, we did fit and transform the training data with *StandardScaler* class which calculated and stored the mean and standard deviation. These computed statistics are subsequently applied to the test data ensuring that it is standardized on the same scale as the training data. This approach maintains consistency and prevents data leakage.

3.2.5.2 Motivation behind Superpixel-based Feature Extraction

While convolutional neural networks (CNNs) are a common choice for such segmentation tasks, our approach takes a different approach by emphasizing the extraction of radiomic features from superpixels. This choice is in line with our hypothesis, which is:

- Robustness: Superpixels group pixels with similar features, which can make radiomic features more resistant to noise and small variations in the image. CNN may be more sensitive to fine-grained variations, which can lead to overfitting when dealing with limited training data or noisy medical images.

- Data Availability: Another big motivation was the amount of data available. With many medical imaging applications, data is limited, and acquiring labeled data to train deep learning models can be difficult. Radiomics, with its reduced dimensionality and potentially more robust features, may require fewer labeled samples for effective classification.

3.2.6 Approach 1: Superpixel-based Features and Random Forest, Support Vector Machine Classifiers

Our first approach for Achilles tendon segmentation is illustrated in fig. 3.7. After successfully performing the previous steps (creating superpixels on cropped images and extracting radiomic features from these superpixels), a classification task is performed using two different classifiers (a Random Forest classifier and a Support Vector Machine) to classify each superpixel. In this scenario, the input data for our classifiers comes from the steps mentioned above, where the inputs are radiomic features of these superpixels and their class labels. Any superpixel/node belonging to our ground truth mask is labeled as a tendon node, while those representing surrounding tissues are labeled as a non-tendon node. These details are organized as a .csv file for each image. As a result, for the total of 411 images (previously discussed in 3.2.2.1), there are a total of 411 corresponding .csv files. Therefore, by predicting whether a given superpixel or node falls into the label category of the tendon class or not, the segmentation process for the tendon region can be performed effectively.

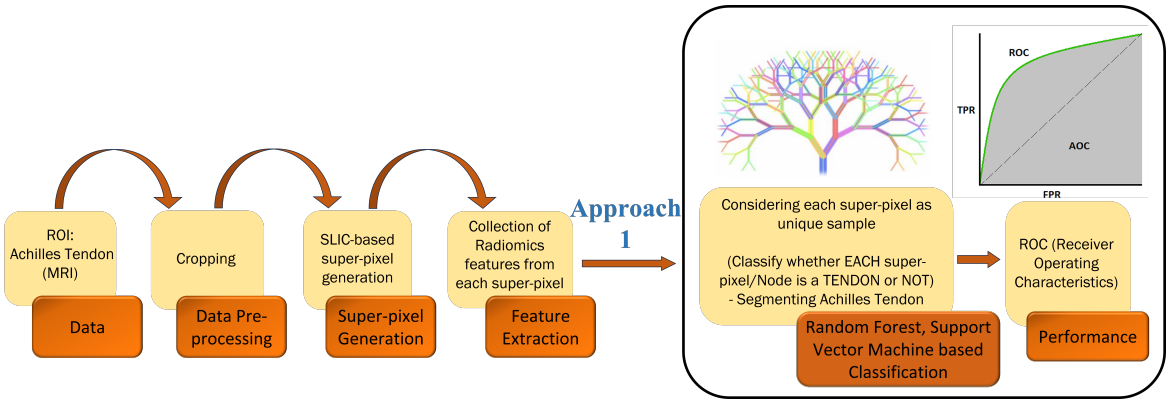


Figure 3.7: Approach 1, Superpixel-based features and RF, SVM classifiers

3.2.6.1 Model Training

During the implementation phase, the IDs of our 76 subjects are shuffled and split, resulting in an approximate 80%-20% split. As a result, there are 60 subjects in one group, called Dataset 1, and the remaining 16 subjects in another group, called Dataset 2. All associated MRI slices for each subject are then aggregated within their respective groups. It produced 320 MRI slices/images in Dataset 1 and 91 images in Dataset 2. This initial split is for training, validation, and final testing. Dataset 1 is used exclusively for training and validation processes. Dataset 1 contains a total of 30,219 superpixels/nodes while Dataset 2 contains a total of 8,423 superpixels.

Dataset 1 is then used for training and validation via a leave-one-group-out cross-validation method that employs both Random Forest (RF) and Support Vector Ma-

chine (SVM) classifiers. Cross-validation is a resampling technique utilized for the assessment of machine learning models on a limited data sample. Leave-one-group-out cross-validation (LOGOCV) is a variation of the k-fold cross-validation technique used in machine learning and statistical modeling. In LOGOCV, instead of splitting the data into k-folds as in traditional k-fold cross-validation, the data is divided into groups or clusters, and each time one group is left out as the validation set while the model is trained on the remaining groups. RF classification, as an ensemble learning method tailored for classification tasks, builds numerous decision trees during training and produces a class prediction based on the mode of the classes determined by the individual trees. On the other hand, SVM is a powerful supervised machine learning algorithm used for classification tasks. In the context of classification, SVM aims to find the optimal hyperplane that separates data points of different classes with the maximum margin. The hyperplane is defined as the decision boundary that best separates the classes in the feature space.

Next, Dataset 2 is used for final testing using the 10 best-performing models trained and validated in Dataset 1. During training, special emphasis is placed on ensuring that superpixels/nodes derived from an MRI slice and the same subject are exclusively assigned to the training or validation phase, avoiding any partial assignment to both. This means that each subject (all superpixels of all associated slices) is used either for training or validation.

To implement our RF and SVM classifiers, the Scikit-learn library is used, more specifically sklearn.ensemble.RandomForestClassifier and sklearn.svm.SVC are used respectively where in both cases a balanced class weight is used since we have data imbalance (most superpixels fall into the non-tendon category rather than tendon superpixels). The balanced mode uses the values of y to automatically adjust weights inversely proportional to class frequencies in the input data as n_samples / (n_classes * np.bincount(y)), and all other parameters are set to the default values provided by the Scikit-learn library's built-in packages mentioned earlier.

3.2.7 Approach 2: Superpixel-based Features and GCN-based Node classification

The workflow of our second approach for Achilles tendon segmentation is depicted in fig. 3.8, which is based on graph-based learning. In graph-based datasets, entities (nodes) are connected by relationships (edges), making them a natural representation of various real-world scenarios. One of the most popular baseline graph neural network models, the graph convolutional network (GCN), employs symmetric-normalized aggregation as well as the self-loop update approach. This approach was first outlined by [190] and has proved to be one of the most popular and effective baseline GNN architectures. Our approach is also GCN-based, with the primary objective of node classification ultimately leading to tendon segmentation.

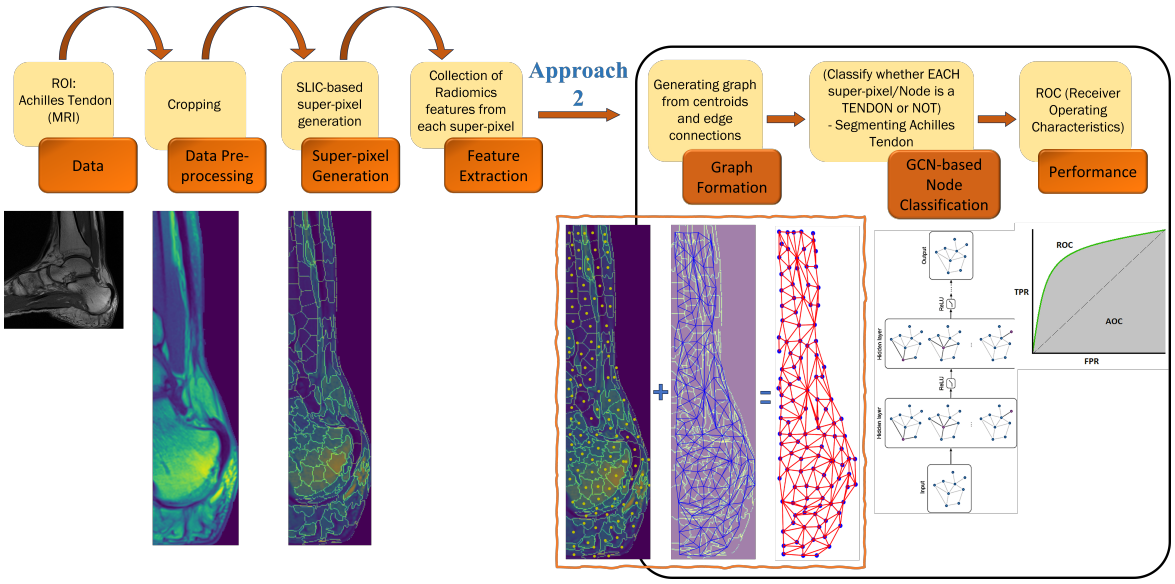


Figure 3.8: Approach 2, Superpixel-based features GCN-Based Node classifier

3.2.7.1 Graph Formation

Graph construction is the preliminary step of our second approach. A graph $G = (V, E)$ is defined by a set of nodes V and a set of edges E between these nodes. It is denoted that an edge going from node $u \in V$ to node $v \in V$ as $(u, v) \in E$. A convenient way to represent graphs is through an adjacency matrix $A \in \mathbb{R}^{|V| \times |V|}$. To represent a graph with an adjacency matrix, the nodes in the graph are ordered so that every node indexes a particular row and column in the adjacency matrix. The presence of edges is represented as entries in this matrix: $A[u, v] = 1$ if $(u, v) \in E$ and $A[u, v] = 0$ otherwise. If the graph contains only undirected edges, then A will be a symmetric matrix, but if the graph is directed (i.e., edge direction matters), then A will not necessarily be symmetric.

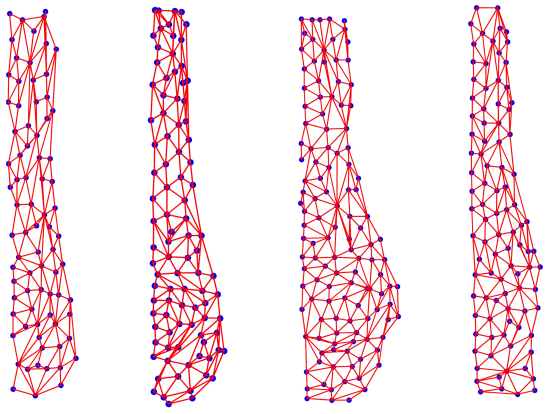


Figure 3.9: Samples of Graphs

In our methodology, the process of creating these nodes and edges is based on the coarse segmentation results obtained via the SLIC (Simple Linear Iterative Clustering) algorithm. Each distinct superpixel generated by SLIC is treated as an independent node within the graph, while edges are formed by connecting vertices associated with neighboring superpixels, forming unweighted edges. After forming a graph for each

slice, the structure of the resulting graph resembles the representation illustrated in fig. 3.9.

As highlighted in the previous section on feature extraction, these extracted radiomic features for each node within each image are systematically stored in a .csv file. Similarly, edge connections between nodes are saved in a separate .csv file. This file containing the list of edge connections is subsequently interpreted as an adjacency matrix. In this format, the first dimension keeps track of the source node, while the second dimension keeps track of the corresponding destination node. In summary, for each MRI slice, the following information is extracted and stored: a .csv file containing radiomic features per node/superpixel, another .csv file containing class labels per node, and a separate .csv file including the list of edge connections between the nodes.

3.2.7.2 Model Training

After the graphs are generated, it resulted in a total of 411 graphs from 76 subjects. They are divided in the same way as our previous approach (discussed in 3.2.6.1) which produced 60 subjects in Dataset 1 and their 320 graphs (from 320 MRI images) and the remaining 16 subjects in Dataset 2, which formed 91 graphs (from 91 MRI images).

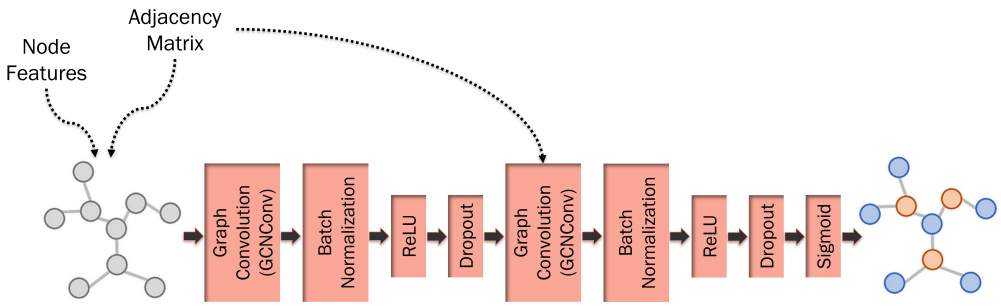


Figure 3.10: Model Architecture

In this phase, our node classification task is carried out through the utilization of a two-layer GCN architecture, as shown in fig. 3.10. The adoption of this neural network architecture is motivated by the graph-structured nature of our input data. As illustrated in fig. 3.10, a graph convolution operation is executed in the first layer, leveraging the node features and the edge connections among the nodes. This operation is implemented using the GCNConv class from PyTorch Geometric (PyG), a foundational component for Graph Convolutional Networks. GCNConv receives two primary inputs: 1) The number of input features for each node in the graph (input dim) and 2) The number of output features for each node in the graph (output dim). The input to this layer is a node feature matrix, and the output is also a node feature matrix, essentially mapping features from the input space to a new feature space. In the process of performing message passing between neighboring nodes in the graph, GCNConv aggregates information from these neighboring nodes to compute new features for each node. This aggregation is carried out by calculating a weighted sum of the features associated with the neighboring nodes, and the weights are learned during the training process. Notably, GCNConv incorporates a normalization step as part of the aggregation procedure, scaling the aggregated information in proportion to the inverse of each node's degree. This normalization aids in ensuring that nodes with varying degrees contribute equally. Subsequently, a linear transformation is applied

to the aggregated information, with GCNConv featuring learnable parameters, including weight matrices for linear transformations and normalization coefficients. These parameters are adjusted and optimized during the model's training phase.

Here, node features are the radiomics features extracted from each superpixel. A is the adjacency matrix of the graph.

$$\hat{A} = \tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}$$

is the normalized adjacency matrix, where $\widetilde{A} = A + I$, I is the identity matrix, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$ represents the degree of the node i.

To further enhance the model's stability and expedite the training process, a layer of batch normalization is employed, which acts as a standard regularization technique frequently utilized in deep neural networks. Following batch normalization, the Rectified Linear Unit (ReLU) is employed as the activation function. ReLU is a widely used activation function in deep learning models, defined as ReLU(x) = max(0, x), effectively replacing negative values with zeros while preserving positive values. This introduces non-linearity into the model, enabling it to capture complex, non-linear relationships within the data. This non-linearity aids in the model's ability to identify and propagate information relevant to the node classification task while suppressing extraneous data. Additionally, the utilization of ReLU helps in mitigating the vanishing gradient problem and contributes to faster convergence during training.

Furthermore, a dropout layer is introduced as a form of regularization within our network. The torch.nn.Dropout module is utilized, which automatically applies dropout during training, but refrains from doing so during the model's inference phase. This behavior is governed by the self.training attribute. In our training regimen, a dropout rate of 0.5 is specified.

The previously described block is reused, constituting what is referred to as the 2-layer GCN mode. Finally, a sigmoid layer is employed in the final layer, which is particularly well-suited for binary classification tasks requiring a decision between two classes, such as 0 or 1. The sigmoid layer compresses the model's output into a range between 0 and 1, representing the probability of belonging to the positive class.

Once the model is defined, it is trained and validated in a leave-one-group-out cross-validation fashion (similar to approach 1). It means that in our setup, in each run, training on all graphs except one subject's graphs is discarded for validation. Through this iterative process, the model is systematically assessed on distinct data subsets, facilitating a more dependable performance estimation and mitigating the risk of overfitting.

For the sake of the model optimization process, the Adam optimizer is used, which is a widely adopted algorithm for training neural networks. Adam is known for its adaptability in adjusting the learning rate during training, drawing from the strengths of both the AdaGrad and RMSprop optimizers. An optimizer is a function or an algorithm that adjusts the attributes of the neural network, such as weights and learning rates. Thus, it helps in reducing the overall loss and improving accuracy. Two critical parameters are provided to the optimizer: 1) learning rate, a hyperparameter, that regulates the size of the optimization steps taken throughout the training process. It influences the pace at which the optimizer refines the model's parameters based on the gradients of the loss function, 2) Another input parameter is weight decay, which serves as a regularization hyperparameter. This feature introduces L2 regularization,

commonly referred to as weight decay, into the optimization procedure. Weight decay introduces a penalty term in the loss function, depending on the magnitude of the model's parameters. This penalty encourages smaller parameter values and serves as a preventive measure against overfitting.

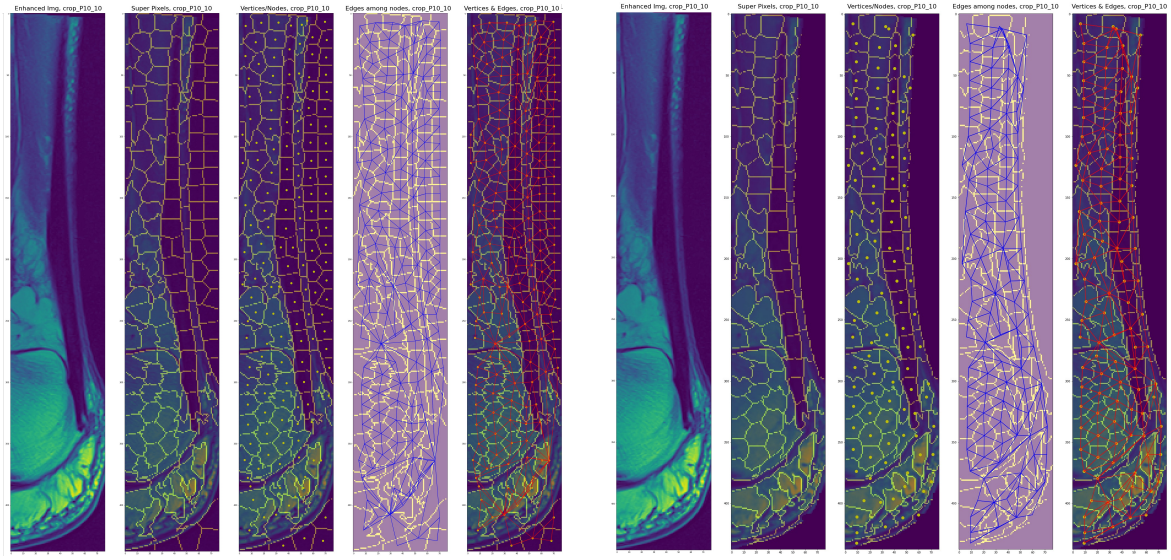
The loss function used in our approach is called BCEWithLogitsLoss (imported from torch.nn), which stands for Binary Cross-Entropy Loss with Logits, and it is used in binary classification tasks. BCEWithLogitsLoss allows different weights for positive and negative examples. This feature is especially useful when dealing with imbalanced datasets, where one class contains significantly more samples than the other. Since our problem contains node imbalance (most nodes are non-tendon nodes, while a smaller amount of tendon nodes). In our implementation, assigning distinct weights to positive and negative examples in binary classification tasks is achieved using a parameter known as pos_weight. This parameter represents a scalar value that determines the relative significance of the positive class (class 1) compared to the negative class (class 0) when calculating the loss. The following hyperparameters are used in our model training: epochs=100, batch_size=32, dropout_rate=0.5, optimier=ADAM, learning_rate=0.01, weight_decay=5e-4, loss_function= BCEWithLogitsLoss. All other parameters are set to their default values as provided by the built-in packages of the associated libraries mentioned earlier.

3.3 Results

3.3.1 Superpixel Generation Performance

Fig. 3.11a provides a visual representation of the generated superpixels, incorporating background for a comprehensive understanding. The importance of the final superpixel segmentation (after background removal) becomes clear in the later stages of our analysis, as demonstrated in fig. 3.11b and 3.12. In figs. 3.11 and 3.12, the five columns/images represent the following stages: a) the contrast-enhanced image, b) the image segmented into superpixels, c) superpixels with vertices (nodes) highlighted, d) edges connecting the nodes, and e) the final graph structure (nodes and edges) overlaid on the image. This critical phase plays a fundamental role in the overall process, serving as the basis for subsequent feature extraction. Carefully delineating superpixels containing only foreground regions sets the stage for training our model, as subsequent features are extracted from them. These extracted features are key to training our model to distinguish between tendon and non-tendon attributes.

To evaluate the performance of our tendon segmentation (via node classification), the Receiver Operating Characteristics: Area Under the Curve (ROC-AUC) is used as a performance metric. It is a valuable evaluation metric in certain situations due to its ability to provide insights into the performance of a classification model, especially in binary classification problems. AUC represents the degree or measure of separability. It tells how much the model is capable of distinguishing between classes. The higher the AUC, the better the model is at predicting 0 as 0 and 1 as 1. Another advantage is the data imbalance, which is the biggest motivation in our study for choosing this evaluation metric. When dealing with imbalanced datasets where one class significantly outperforms the other, precision can be misleading. ROC-AUC takes into account the trade-off between true positive rate and false positive rate and is less affected by class imbalance. Being a threshold-independent evaluation metric, AUC offers a more bal-



(a) Superpixels with Image Background (b) Superpixels without Image Background

Figure 3.11: Generated Superpixels

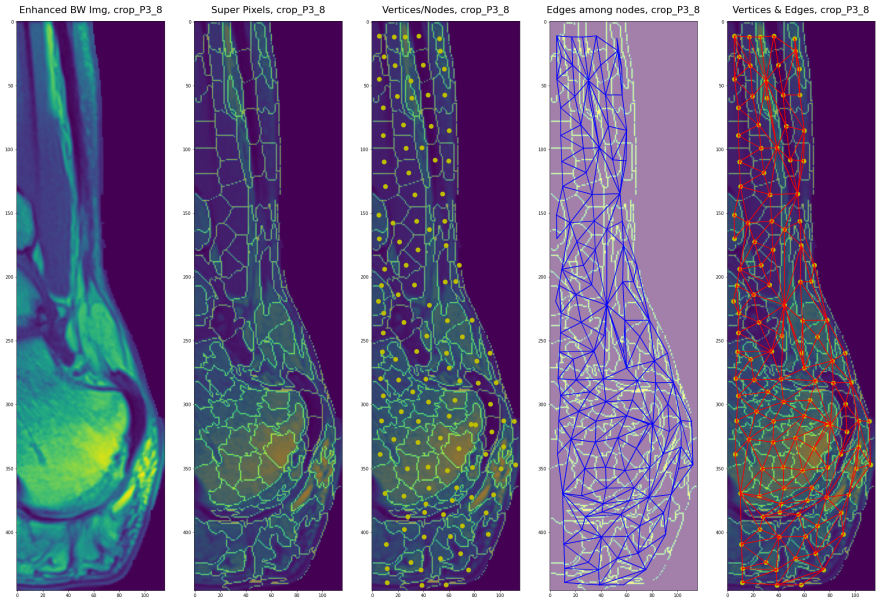


Figure 3.12: Superpixels of a pathological subject

anced assessment of model performance. On the other hand, sensitivity measures the proportion of true positive cases correctly identified by a classification model. These two metrics are also widely used as evaluation metrics for binary classification tasks with data imbalances. Additionally, we used specificity and balanced accuracy as other metrics. Specificity assesses the proportion of true negatives correctly identified by the model, while balanced accuracy is the simple average of sensitivity and specificity which offers a more comprehensive evaluation by considering both true positives and true negatives.

3.3.2 Performance of Approach 1 (Superpixel-based Features and RF, SVM Classifiers)

Since our segmentation task is achieved through superpixel classification, we have selected the Area Under the ROC Curve (AUC) as our evaluation metric to assess classification performance. This metric ultimately reflects the effectiveness of our tendon segmentation module. AUC score serves as a vital evaluation metric to evaluate the performance of our RF and SVM classifiers, which are built on our superpixel-based coarse segmentation method. This metric provides a comprehensive view of the classifier's ability to discriminate between positive (tendon) and negative (non-tendon) instances within our dataset. Additionally, sensitivity, specificity, and balanced accuracy (the average of sensitivity and specificity) are emphasized as alternative evaluation metrics.

To ensure the robustness and reliability of our evaluation, a leave-one-group-out cross-validation procedure is used to train and evaluate the model on Dataset 1. The average scores of all evaluation metrics are shown in table 3.1. In addition to the leave-one-group-out cross-validation performance, the performance of the classifiers on the test data (Dataset 2) using the top 10 models from the training phase, the average scores of evaluation metrics are reported in table 3.1. This provides insight into the model's performance on unseen and out-of-sample data, which is valuable for evaluating its real-world applicability and generalization beyond the training dataset. Combining the cross-validation results and performance of test data offers a comprehensive evaluation of the effectiveness of the RF and SVM classifier-based node classification in our Achilles tendon segmentation task.

3.3.3 Performance of Approach 2 (Superpixel-based Features and GCN-based Node Classification)

To conduct a performance analysis comparable to Approach 1, the AUC score is used as the evaluation metric for the model. Both the AUC score, sensitivity, specificity,

Table 3.1: Evaluation Metrics for Superpixel Classification/Tendon Segmentation

(a) Training and Validation (on Dataset 1)

Model Name	AUC	Sensitivity	Specificity	Balanced Acc.
Random Forest	0.984	0.871	0.983	0.927
Support Vector Machine	0.983	0.943	0.940	0.942
GCN	0.922	0.879	0.896	0.888

(b) Test (on Dataset 2)

Model Name	AUC	Sensitivity	Specificity	Balanced Acc.
Random Forest	0.992	0.904	0.975	0.939
Support Vector Machine	0.987	0.966	0.941	0.953
GCN	0.933	0.899	0.902	0.901

and balanced accurracy on Dataset 1 are shown in Table 3.1. Using a similar setup, the top 10 best models (based on AUC score) from the training phase are used for the final prediction of the test data (Dataset 2). The mean scores of all the evaluation metrics are recorded in Table 3.1.

3.4 Discussion

The results outlined in the performance evaluation underscore the effectiveness of our both approaches in tendon segmentation, especially in the area of node classification.

Several superpixel generation algorithms exist including Normalized cuts, Entropy Rate Superpixels (ERS), turbopixels method [161], mean-shift [160], watersheds in digital space [162], Simple Linear Iterative Clustering (SLIC) and g-SLIC algorithms, which address many requirements and outperform other state-of-the-art algorithms. In our study, we have chosen the SLIC algorithm for superpixel generation due to its efficiency in producing compact and uniform superpixels with linear time complexity by limiting the search space around each cluster center. SLIC balances color similarity and spatial proximity through a controllable compactness parameter, resulting in contiguous and noise-insensitive superpixels. Compared to other methods, SLIC is simpler to implement and equally robust, making it ideal for various image-processing tasks. Our choice of SLIC over other algorithms is also supported by the literature [149], [314], [178], [180], [315], [159], etc.

As demonstrated in fig. 3.11b and 3.12, for superpixel generation, our technique consistently shows good performance in accurately delineating regions of interest in both healthy and pathological cases. In particular, our configured superpixel generation parameters produce segmentation results that align perfectly with the intrinsic structure of the image. This result attests to the effectiveness of our approach in finding a harmonious balance between compactness and adherence to boundaries. This optimized balance is especially crucial when dealing with diverse datasets that include both healthy and pathological instances, demonstrating the versatility and robustness of our methodology in capturing different ROIs.

Our first approach which uses Random Forest (RF) and Support Vector Machine (SVM) classifiers for node classification while exploiting superpixel-derived radiomic features shows notable levels of performance. Both RF and SVM classifiers are well established for their effectiveness and have been widely applied in various classification tasks. Recent studies have leveraged these classifiers in several contexts, including using RF with GLRLMS feature extraction to achieve maximum classification accuracy in identifying the severity of COVID-19 [316], its application in the classification of MRI-based brain tumors with superior accuracy [317], use of RF and SVM classifiers in breast cancer detection [318], among other notable applications. Being inspired by the continued success of these methodologies, our approach incorporated both RF and SVM-based classifiers paired with superpixel generation. The results (see Table 3.1) highlight the robust capabilities of RF-based and SVM-based classifiers when coupled with superpixel generation, shedding light on the potential effectiveness of such segmentation approaches within our research domain. Notably, the AUC scores obtained from both classification configurations on our test data are 0.992 and 0.987 for RF and SVM, respectively. While the RF classifier has a slightly higher AUC, SVM demonstrates superior sensitivity performance. Notably, the model's performance on

the test data appears to be slightly better than on the training and validation sets. This could potentially be attributed to the presence of relatively fewer complex structures or images in the test dataset, making the task less challenging. Nevertheless, the overall performance across the training, validation, and test sets is consistent and demonstrates strong results. Another observation is that SVM shows the highest sensitivity across both datasets, which highlights its effectiveness in accurately identifying positive instances (the tendon class). This high sensitivity is particularly crucial for our study, as accurately detecting the tendon class is more important than achieving the highest specificity, which is the strength of the Random Forest classifier. Additionally, SVM outperforms the Random Forest classifier in terms of balanced accuracy on both datasets. This further underscores the superior performance of the SVM classifier over RF. When evaluating AUC and balanced accuracy, it's crucial to understand that they measure different aspects of model performance and are not directly comparable. Balanced accuracy is determined at a specific threshold, which may not be optimal for every model. Therefore, when the chosen threshold (the default of 0.5) is not ideal, balanced accuracy may be lower than AUC. This seems to be the case here, where the default threshold resulted in a lower balanced accuracy. In contrast, AUC assesses performance across all possible thresholds, providing a more comprehensive view of the model's ability to distinguish between tendon and non-tendon classes. This broader perspective explains why AUC shows slightly higher value and aligns better with our metrics of interest. Overall, depending on the specific point of interest, both RF and SVM performed equally well in our study.

In contrast, our second approach (graph-based approach) also demonstrates promising performance. This approach leverages a versatile graph structure, coupled with the generation of superpixels. On our test data, this approach achieved an AUC of 0.933, with a sensitivity of 0.899, a specificity of 0.902, and a balanced accuracy of 0.901. While these metrics are slightly lower compared to the performance of RF and SVM-based experiments, they still indicate potent performance. These results lead to two key observations. First, as a standalone model the graph-based approach performs well. However, when compared to non-graph-based models, its performance appears relatively lower. This discrepancy seems to be attributed to the limited amount of available data. Deep learning-based models typically require a substantial amount of data for effective training, and with only 76 unique subjects in our dataset, the potential of the graph-based approach may not have been fully realized. Initially, we hypothesized that the incorporation of coarse segmentation would enhance the performance of the graph-based final classification. However, this approach did not outperform the non-graph-based models, suggesting that further exploration is needed. This is addressed in the future work section.

To our knowledge, our study is among the first in the field of Achilles tendon segmentation which makes direct comparisons with other state-of-the-art methods challenging. Alzyadat et al. [319] conducted automatic segmentation of Achilles tendon using deep CNN. They used datasets of 3708 for training and 2472 for validation. Their approach involved ensembling different networks for the final segmentation. Due to the differences in their methodology and the amount of data used, direct performance comparison with our study is not feasible. A very recent study [320] conducted tendon segmentation on ultrasound images using grey-level co-occurrence matrix features and hidden Gaussian Markov random fields. The primary aim was to provide a quantitative and automated method for detecting potential structural changes in tendinopathy. As

with previous studies, the objectives, methodologies, and data differ from those used in our research. Within the domain of medical image segmentation, there is a very recent study that has focused on the segmentation of diverse anatomical structures in medical images. Known as MedSAM [321], it is a deep learning-powered foundation model. It is trained on a large-scale dataset of over one million image-mask pairs. The network utilized in MedSAM was built on a transformer architecture. Notably, the study underscores the significance of the training dataset size in determining final performance metrics. Moreover, considerations such as data modality, data quality, and regions of interest are also some of the crucial factors when comparing the performances of different segmentation models.

A key limitation of this study is the small dataset, which may have constrained the performance of our GCN-based approach compared to traditional machine learning models, as deep learning methods generally require larger datasets. Additionally, the diversity within the dataset made it challenging to optimize background removal, possibly necessitating further adjustments for varying data. Another limitation is that our pipeline is trained and tested on cropped data, which may not fully address the challenges of applying the approach to full-scale images, highlighting an area for future research. Moreover, since superpixels define the ROI, different images or ROIs will require adjustments to the superpixel generation parameters. Finally, being one of the pioneering efforts in Achilles tendon segmentation, our study had limited opportunities for direct comparison with state-of-the-art methods.

3.5 Conclusion

Our proposed module for tendon segmentation has demonstrated its effectiveness through the utilization of superpixel generation as a coarse segmentation step preceding the final segmentation task. This approach formulates the segmentation task as a superpixel classification problem, aiming to classify each superpixel as either tendon or non-tendon. The primary motivation behind using superpixel-based coarse segmentation was to address the fact of traditional neural networks, which must simultaneously learn both lower-level and higher-level information within the same architecture. By grouping similar pixels, this method simplifies the image, reduces complexity, preserves boundary information, and integrates higher-level features, leading to more accurate and robust segmentation results.

In terms of contribution, our proposed tendon segmentation module delivered impressive results across several key metrics. The good performance is strengthened by computationally efficient superpixel generation, which streamlines image processing without compromising precision. Additionally, the versatile nature of our module implies this through its multiple approaches (both traditional machine learning and GCN-based approaches). This flexibility empowers users to choose the method that best aligns with the computational resources, overall setups, and data properties. Finally, our generalizable framework which is built on robust principles and modular design, holds exciting potential for adaptation to various medical imaging tasks beyond tendon segmentation. Overall, this module's combination of high performance, versatility, and generalizability positions it as a valuable tool for advancing medical image analysis.

To further refine our module, future work will involve exploring its performance on larger and more diverse datasets to solidify its generalizability and robustness. Another area for exploration could include working with full-scale images instead of cropped data. Incorporation of additional features such as demographic-related or pain-level data can be another edition. Additionally, investigating more advanced graph network architecture holds promise for potentially boosting performance and capturing intricate tendon-related features. Ultimately, integrating this segmentation module into clinical workflows for tendon pathology diagnosis, treatment planning, and outcome assessment represents a crucial step toward its real-world impact on improving patient care. In conclusion, our research introduces a new methodology and valuable insights into segmentation, paving the way for improved understanding and diagnosis of tendon-related pathologies.

Chapter 4

Medical Image Classification

This study introduces an end-to-end tendon pathology detection (classification) module leveraging a Graph Echo State Network (GESN). The proposed approach is designed to effectively model the complex spatial relationships within medical images by representing them as graphs, where nodes correspond to meaningful regions (e.g., superpixels) and edges capture their interactions. The GESN framework processes these graph representations to identify pathological features, offering a novel solution for automated tendon pathology detection. This method aims to enhance diagnostic accuracy while maintaining computational efficiency, making it suitable for practical clinical applications.

This work is being prepared for submission to a Q1-ranked journal named "Medical Image Analysis" (Impact Factor: 10.7, CiteScore: 22.1).

Article Title: Graph Echo State Network for MRI-based Tendon Pathology Classification

Keywords: Magnetic Resonance Imaging, Soft Tissue Pathology, Superpixel, Graph Echo State Network, Graph Classification

4.1 Introduction

Tendons connect muscles to bones, enabling movement, but are prone to conditions like tendinosis, tendinopathy, and tears from trauma or wear. Diagnosis involves medical history, physical exams, and imaging techniques like Ultrasound and MRI, which provide critical insights but remain underutilized in research.

Gillet et al. sought to explain how MRI might be useful for Achilles tendinitis patients receiving fluoroquinolone therapy [241]. In that study, the tendinous involvement was classified into 5 stages (adapted from [242]). The authors came to the conclusion that in patients suffering from fluoroquinolone-induced tendinopathy, MRI appears a helpful and accurate method to identify, classify and follow the disease. McLoughlin et al. characterized patellar tendinitis features by using MRI to gain a better understanding of the underlying pathophysiology [243]. In that study, several features were extracted by an experienced musculoskeletal radiologist, and later, the subjects were graded accordingly. Bauer et al. proposed a comprehensive MRI grading protocol for supraspinatus tendinosis and partial thickness tears and studied its reliability [244]. Here also, the authors required experts to manually extract several features. [246] developed a computer-assisted diagnostic system to aid ultrasound operators in the accurate diagnosis of rotator cuff diseases and to streamline ultrasound tests. To classify each lesion case, the lesion area and texture features were taken from the complete lesion and integrated with a multinomial logistic regression classifier. In [247], a software program that can identify changes in ultrasound images of the Achilles tendon and categorize them as normal or abnormal is developed. In this study, a matrix with 68 different texture features was extracted from the whole sample. Discriminant Analysis, Quadratic SVM, k-NN, and different ensemble classifiers were used to classify the dataset images. The purpose of [248] was to ascertain whether a quantitative texture-based method could accurately identify tendon abnormalities in pallapugno players. This study confirmed the prospect that a quantitative texture-based method could be helpful for the standardized diagnosis of sub-clinical tendinopathy. Chang et al. created a computer-aided tear classification system [249]. In each case, intensity and texture features were extracted from the entire lesion and combined into a binary logistic regression classifier for lesion classification. Golman et al. determined the association between partial patellar tendon tear characteristics and treatment guidelines and created a classification system for partial patellar tendon injuries based on MRI [250]. Logistic regression and sensitivity analyses showed that patellar tendon thickness measurement predicted whether the tendon was normal or symptomatic. Wang et al. aimed to develop and validate an efficient ultrasound image-based radiomic model for determining Achilles tendinopathy in skiers [257]. A total of 833 quantitative features were obtained for each color channel, and 3332 radiomic features were extracted from each region of interest. To determine the best combination of feature selection and machine learning modeling, the study used three commonly used feature selection algorithms. Based on the above experiments, the authors found that the Random Forest selection-based SVM model in G-channel was the optimal ultrasonic radiomic model.

The existing literature evidences that medical images are utilized to extract features such as tendon thickness, echotexture, vascularity, and so on to identify tendon-related pathologies. Often, these features are manually extracted by experts through a meticulous process that involves visual inspection and annotation. This manual

approach is time-consuming and costly, requiring highly trained specialists and considerable resources. Additionally, manual extraction is prone to human error and can be unconsciously biased.

To address these challenges and improve the precision and efficiency of feature extraction and model training, we propose automatic feature extraction. Furthermore, our proposed module involves converting raw images into graphs and comparing the performance of graph-based methods with our baseline (non-graph-based alternatives). Using graphs instead of conventional image grids offers several significant advantages, including flexible representation, enhanced feature extraction, improved analysis, and scalability [259]. Additionally, a key concept used in our study is Reservoir Computing (RC), an innovative computational paradigm designed to simplify the training of recurrent neural networks (RNNs). This approach leverages the dynamics of a fixed, randomly connected network, known as the reservoir, to project input data into a high-dimensional space. By utilizing these rich dynamics, reservoir computing enables the effective processing of complex temporal patterns while significantly reducing the training complexity. Crucially, the reservoir itself remains largely untrained, with only the readout layer requiring specific adjustments for the desired task [262]. RC has recently emerged as an alternative to gradient descent methods for training RNNs [263]. Echo State Network (ESN) [264] is a prominent RC model, being practical, conceptually simple, and easy to implement. It can avoid nonconverging and computationally expensive issues by applying the least squares problem as an alternative training method. One of the studies that particularly inspired us considering Graph ESN is the work by Micheli et al., which explored the issue of heterophily in node classification [270].

Based on the rationale outlined above, we apply reservoir computing to graph structures derived from MRI images, specifically utilizing the Graph Echo State Network (GESN) to classify each MRI slice as pathological or non-pathological. Following slice-level predictions, a majority voting scheme aggregates these results at the subject level. Since we are working with a custom ankle MRI dataset and no publicly available benchmarks exist for direct comparison, we have also developed a baseline module. This baseline operates at the image level without incorporating graph structures and employs two distinct strategies. Our proposed module along our baseline is summarized as follows:

1. Baseline (Non-Graph based Module): Image classification based on radiomic features.
    - Strategy no. 1: Radiomic features are extracted directly from each MRI image.
    - Strategy no. 2: Radiomic features are extracted from each superpixel.
        * A Simple Linear Iterative Clustering (SLIC) algorithm is used to generate the superpixels.

2. Proposed Graph-based Module: Graph Echo State Network for Image classification.
    - Graphs are constructed from given superpixels.
    - Superpixels are generated using the SLIC algorithm (similar to our baseline).

4.2 Materials and Methods

4.2.1 Data

The dataset used in this study comprises images from 76 participants with a mean age of 45 years ± 19.8 years, including 42 men and 34 women. All of whom underwent MRI scans at Landspitali University Hospital, Iceland. The participants were divided into different diagnostic categories, with 45 having healthy Achilles tendons. The remaining 31 participants were diagnosed with conditions such as calcific insertional tendinopathy, Achilles tendinopathy, chronic tears in the distal part of the Achilles tendon, or other tendon injuries. For a comprehensive analysis, a careful selection procedure was implemented to extract only the 2D MRI slices that included the region of interest (ROI), thereby omitting any slices that did not feature it. This method resulted in a dataset comprising 411 MRI slices, with each participant contributing an average of 4 to 5 sagittal TSE MRI slices specifically targeting the Achilles tendon. Among the 411 MRI slices, 233 were labeled as healthy (reflecting 45 healthy individuals), while 178 slices were labeled as pathological (representing 31 pathological individuals).

4.2.2 Data Pre-processing

Image cropping is a crucial pre-processing step for refining MRI data by eliminating irrelevant regions. This technique reduces noise, enhances composition, and optimizes the aspect ratio, focusing on diagnostically significant areas like the region of interest.

Our image cropping pipeline, illustrated in fig. 3.5, outlines the step-by-step procedure applied to both MRI images and their corresponding masks. This technique follows a similar approach to the one described in the previous chapter 3.2.3.2.

4.2.3 Experimental Setup

We address the task of pathology identification (classification) using a graph-based module and compare its performance with a non-graph-based baseline. Both methods utilize cross-validation techniques to ensure thorough and reliable training and validation, specifically implementing a five-fold cross-validation for each.

Before diving into the details of our module, we first outline the common preliminary steps, which include superpixel generation and radiomic feature extraction. These steps are essential to prepare the data and extract meaningful features. After detailing these common steps, we present a more detailed section of our implementation, highlighting their unique strategies and results.

4.2.3.1 Superpixel Generation

Superpixels are perceived as groupings of pixels that simplify image processing by over-segmenting images into multiple meaningful regions [214]. This technique allows for a more manageable representation of images. We utilize the Simple Linear Iterative Clustering (SLIC) algorithm to generate superpixels, which has shown success in various studies [149], [314], [178], [180], [315], [159], etc. It clusters pixels based on color similarity and spatial proximity in a five-dimensional [labxy] space. For grayscale images, intensity similarity is used in a three-dimensional [ixy] space. Normalization of spatial distances is deemed essential due to the fixed maximum intensity distance

when compared to the variable distance in the XY plane [310]. More details on this algorithm and its performance on our custom dataset (though applied to a different final objective) can be found in our previous related work [2].

To implement the SLIC algorithm, the cv.ximgproc.createSuperpixelSLIC() function (from the OpenCV library) is used. The parameters for superpixel generation include the following: a) image: The input image for superpixel generation, b) region_size: This parameter significantly influences the compactness and size of the superpixels. A region size is adaptively chosen to $(13/375) \times image\_height$ and it is validated through rigorous experimentation. This choice is particularly beneficial because our study encounters variability in image sizes. By relating the region size to the original image height, we ensure that the superpixels remain appropriately scaled regardless of the input image dimensions. Specifically, in images with an approximate height of 375 pixels, a superpixel size of 13 is determined to be optimal, and resulted in an average region size of approximately 13.83, c) Compactness Factor: After careful experimentation, a factor of 10 balances compactness and natural boundary adherence.

After superpixels are generated, background removal is performed, which is a crucial preprocessing step for SLIC-based medical image segmentation. This technique is known to enhance accuracy, improve computational efficiency, and reduce errors, although its effectiveness can be influenced by background complexity. In our study, challenges are encountered as the Achilles tendon is near the leg's contour, complicating the removal process. Our ankle MRI images display various edge characteristics, including thin fatty tissue layers and poorly defined boundaries. These complexities made conventional background removal methods impractical, necessitating a tailored approach. Our detailed background removal process is presented in details in the previous chapter.

4.2.3.2 Feature Extraction

Our feature extraction process utilizes the PyRadiomics feature extractor [312], resulting in the extraction of 94 radiomic features. These features are categorized into several groups: 19 features from First Order Statistics, 24 from the Gray Level Co-occurrence Matrix, 16 from the Gray Level Run Length Matrix, 16 from the Gray Level Size Zone Matrix, 5 from the Neighboring Gray Tone Difference Matrix, and 14 from the Gray Level Dependence Matrix.

Before these radiomic features are input into different classifiers, they undergo standardization using the StandardScaler (from the scikit-learn library). Initially, the training data is fitted and transformed, enabling the StandardScaler to compute and store the mean and standard deviation. These calculated statistics are subsequently applied to the test data, ensuring it is standardized to align with the scale of the training data. This approach preserves consistency and mitigates the risk of data leakage.

4.2.4 Baseline (Non-Graph based Module)

Here, two distinct strategies are implemented, each involving a different fashion of radiomics feature extraction. The pipeline used for our baseline is shown in fig. 4.1.

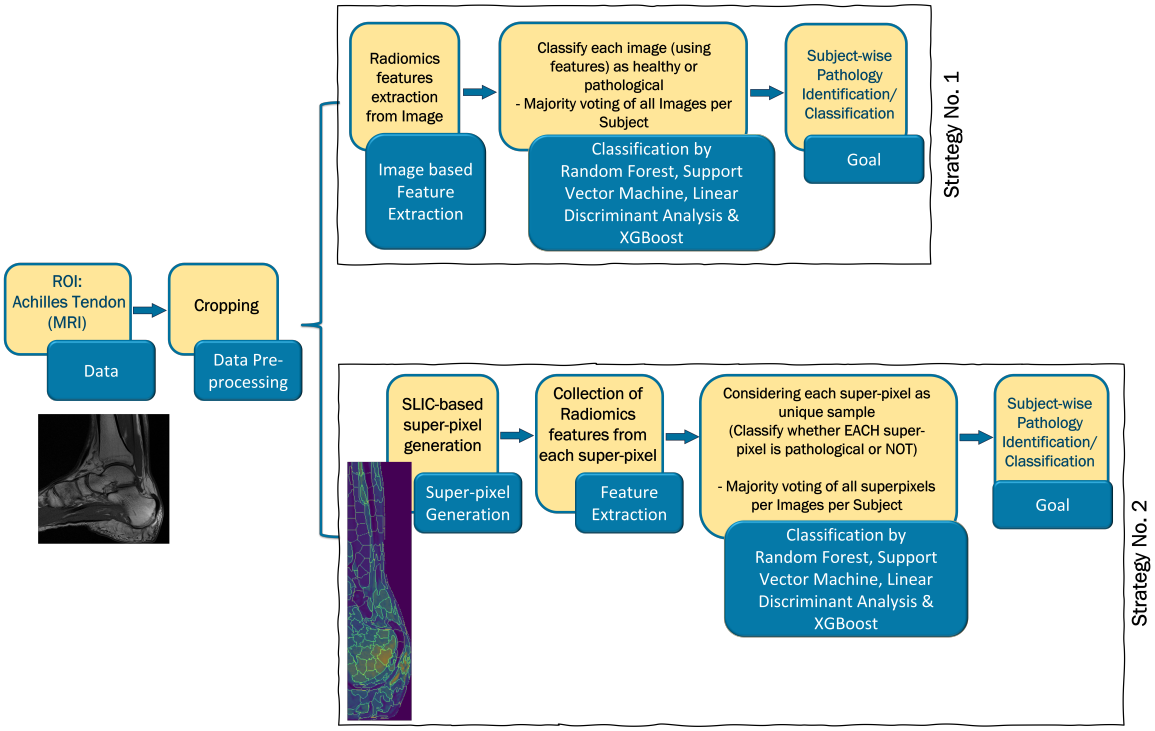


Figure 4.1: Baseline Module

4.2.4.1 Strategy No. 1, Classification based on Image-based Feature Extraction

A total of 94 radiomic features are extracted from each pre-processed MRI image, resulting in 411 sets of image-based features corresponding to 411 MRI images. A classification task is then performed to determine whether each MRI slice is pathological or not using four different classifiers. The input for these classifiers includes radiomic features, with class labels of 0 for healthy cases and 1 for pathological cases. To classify the subject, a majority vote is conducted across the performance of all slices.

Model Training using Image-based Features

Each model undergoes training and validation through a five-fold cross-validation method. Initially, the IDs of the 76 subjects are shuffled and then divided into five distinct folds. Approximately 15 subjects are allocated for validation in each fold, while the remaining 61 subjects are designated for training. Typically, the training folds comprises around 330 MRI slices, whereas the validation fold contains about 81 MRI slices. The precise number of slices in each fold may vary slightly due to differences in the number of MRI slices associated with each subject. Throughout the training process, attention is given to ensure that all MRI slices from a single subject are exclusively assigned to either the training or validation set, avoiding any split between the two. This guarantees that each subject, along with all their corresponding slices, is utilized entirely for either training or validation, thus preventing any overlap. The radiomic features derived from these MRI slices serve as input data, accompanied by their respective binary class labels.

Four different classifiers are used in this approach, namely Random Forest (RF), Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and eXtreme

Gradient Boosting (XGBoost). RF is an ensemble method which is constructed with multiple decision trees whose outputs are combined to enhance accuracy. SVM operates by finding the optimal hyperplane that separates data into distinct classes, maximizing the margin between them for effective classification. LDA is a linear classifier which projects the data onto a lower-dimensional space, aiming to maximize the separation between classes by modeling their differences. Finally, XGBoost is a powerful gradient boosting algorithm that builds decision trees sequentially, improving predictions by focusing on correcting errors made by previous trees. These diverse classifiers are employed to ensure robust performance evaluation and to explore different approaches to the pathology identification task. The Scikit-learn library is utilized to implement these classifiers, specifically sklearn.ensemble.RandomForestClassifier, sklearn.svm.SVC, sklearn.discriminant_analysis, along with the xgboost library for the XGBoost classifier. For RF and SVM classifiers, a balanced class weight is applied to address class imbalance, ensuring equal importance is given to both classes. The balanced mode uses the values of y to automatically adjust weights inversely proportional to class frequencies in the input data as n_samples / (n_classes * np.bincount(y)). In the case of the SVM classifier, a linear kernel is used to simplify the decision boundary. For all the classifiers (RF, SVM, LDA, and XGBoost), other hyperparameters are kept in their default settings, allowing performance evaluation to focus on minimal fine-tuning. This setup ensures a fair comparison across classifiers while leveraging their standard configurations.

In addition, each of the four classifiers mentioned above is trained using the features in different ways to observe their behavior in terms of performance. In the first experiment, all features are utilized, while in the subsequent three experiments, only a subset of important features are employed. The number of important features tested is 32, 16, and 8. To determine feature importance for each classifier, various built-in attributes are applied. For RF and XGBoost classifiers, the .feature_importances_ attribute from the Scikit-learn library is used to calculate feature importance. In RF, this attribute indicates how much each feature contributes to reducing the impurity of the splits, while in XGBoost, it measures feature importance based on gain or improvement in predictive performance. In the case of SVM and LDA classifiers, the .coef_[0] attribute from the Scikit-learn is applied. For SVM, .coef_ contains the coefficients (weights) that define the hyperplane, whereas in LDA, it holds the coefficients for the linear discriminant function.

4.2.4.2 Strategy No. 2, Classification based on Node-based Feature Extraction

In this strategy, the radiomic features are extracted from each node (superpixel) within each MRI slice where each superpixel is treated as a unique sample. This strategy is also tested using the four classifiers mentioned above. The input for these classifiers consists of radiomic features from each superpixel, with each superpixel labeled as pathological if the subject's true class label is pathological, or healthy otherwise. Due to the lack of precise information about the exact location of the pathology, superpixels can only be labeled based on the overall subject's class label. After training, the final performance is evaluated using a two-stage majority voting process: first, a majority vote is conducted on the predictions of all superpixels from each MRI slice to determine whether the slice is predicted as healthy or pathological. Then, a second majority vote

is performed to determine whether the subject is classified as pathological or not based on the slice level predictions.

Model Training using Node-based Features

To maintain consistency with the previous experimental setup, this method is also tested using a five-fold cross-validation approach. In this setup, subject IDs are shuffled and split into five folds, ensuring that all MRI slices from each subject are kept together, either entirely in the training set or the validation set. As mentioned above, although each superpixel is treated as a unique sample and classified individually, a majority vote is performed afterward. This is ensured that all superpixels from a given subject's MRI slices are used for either training or validation, with no overlap between the two groups.

All four of the above-mentioned classifiers are also applied in this strategy, to compare their performance and evaluate whether the models perform better when features are extracted from entire MRI images or from individual superpixels. This allows to evaluate the impact of different feature extraction techniques on classification performance. Similar to the first strategy, each of these classifiers is tested using all features, and a subset of the most important features (32, 16, and 8).

4.2.5 Proposed Graph-based Module

This module consists of several steps (shown in fig. 4.2), including transforming MRI images into graphs and then performing graph classification. Each step is explained in detail in the following sections.

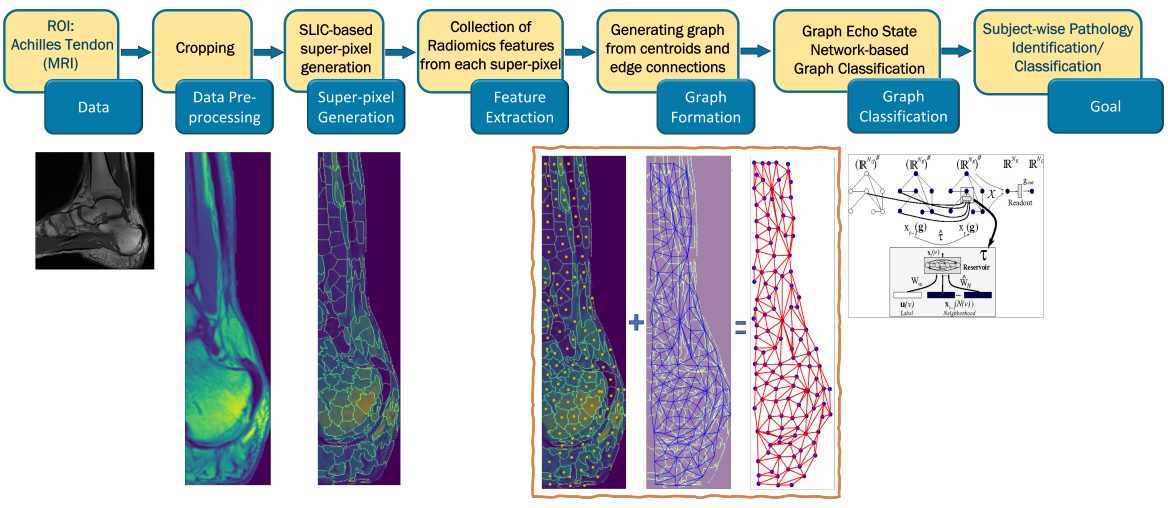


Figure 4.2: Our Proposed Module

4.2.5.1 Superpixel Generation and Feature Extraction

A crucial step in this module involves the generation of superpixels, followed by the extraction of radiomic features from each superpixel (node). The details of these two processes are outlined in the previous sections 4.2.3.1 and 4.2.3.2.

4.2.5.2 Graph Formation

The initial phase of our proposed graph-based module involves constructing graphs, denoted as $G = (V, E)$, which consists of a collection of nodes V and edges E connecting these nodes. An edge that connects node $u \in V$ to node $v \in V$ is represented as $(u, v) \in E$. A common method for graph representation is the use of an adjacency matrix $A \in \mathbb{R}^{|V|\times|V|}$. In this matrix, each node corresponds to a specific row and column, and the presence of edges is indicated by the matrix entries:

$$A[u, v] = \begin{cases} 1 & \text{if } (u, v) \in E \\ 0 & \text{otherwise} \end{cases}$$

If the graph comprises solely undirected edges, the matrix A will exhibit symmetry; however, for directed graphs, where edge direction is significant, A may lack symmetry.

In this study, nodes and edges are generated based on the coarse segmentation outcomes derived from the SLIC (Simple Linear Iterative Clustering) algorithm. Each unique superpixel produced by SLIC is regarded as a separate node within the graph, while edges are created by linking nodes that correspond to adjacent superpixels, resulting in unweighted edges.

4.2.5.3 Model Training

To classify each graph as pathological or non-pathological, we implement a Graph Echo State Network (GESN) that features a linear readout layer in its final layer, with the weights and biases trained using Ridge regression. The model utilized in our study is illustrated in fig. 4.3. The GESN is an extension of the Echo State Network (ESN) designed for graph-structured data [265]. The key innovation in GESN is the recurrent reservoir layer, which iteratively computes the states of the graph vertices while ensuring contractive dynamics for stability. The model and its training process are outlined in detail in the following steps.

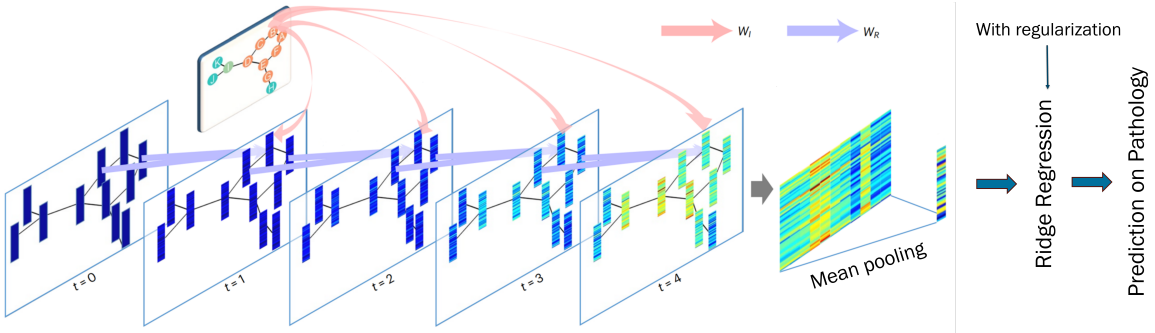


Figure 4.3: Model Architecture [322]

State Transition Function

The state transition function computes the state of each vertex v in a graph g based on the input features of the vertex and the states of its neighboring vertices. The state of vertex v at iteration t is updated using the following equation:

$$x_t(v) = f(W_{\text{in}}u(v) + \hat{W}_N x_{t-1}(N(v))) \tag{4.1}$$

Here, $x_t(v)$ represents the state of vertex v at time step t. The matrix $W_{\text{in}} \in \mathbb{R}^{N_R \times N_U}$ is the input-to-reservoir weight matrix, initialized from a uniform distribution

and scaled by a hyperparameter, the scaling factor. This scaling factor controls how much influence the input has on reservoir dynamics. A larger scale implies a stronger influence of the input on the reservoir state, while a smaller scale reduces its effect. The effect of different scaling factors is shown in the results section. The various scaling factors examined in our study include 0.01, 0.1, 0.4, 0.7, and 1.0. In this context, N_R is the number of reservoir units. Reservoir units serve as dynamic memory components that transform input data into high-dimensional representations, enabling the network to capture complex temporal patterns and relationships within the data. Various values for this unit have been tested, such as 32, 64, 128, 256, 512, and 1024. The input dimension N_U is defined as the number of nodes multiplied by the number of features per node. The input feature vector for vertex v is denoted by $u(v)$, and $x_{t-1}(N(v))$ is the concatenation of the states of all neighbors of v from the previous time step $t-1$. The recurrent weight matrix $\hat{W}_N \in \mathbb{R}^{N_R \times kN_R}$ is also initialized from a uniform distribution and scaled by the reservoir spectral radius, a hyperparameter further discussed in a later section about contractivity. Lastly, tanh is chosen as a non-linear activation function f. The number of time steps (t) continues to iterate until it reaches convergence. The convergence condition for a vertex v in the graph is defined as $|x_t(v) - x_{t-1}(v)| < \epsilon$, where $x_t(v)$ is the state of vertex v at time step t, $x_{t-1}(v)$ is the state of vertex v at the previous time step $t-1$, and ϵ is a small threshold value that determines the acceptable change in state for convergence. In our implementation, 1e-8 is chosen as ϵ. And, the states are typically initialized as $x_0(v) = 0$ for all $v \in V(g)$.

Global State Transition Function

The global state of a graph g is computed by applying the local state transition function to each vertex. The global encoding function τ is defined as:

$$x(g) = \hat{\tau}(g, x(g)) \tag{4.2}$$

where g is the entire graph, which provides context for the states and $x(g)$ is the vector that encapsulates the states of all vertices in the graph. The function $\hat{\tau}$ essentially aggregates the information from all vertices and their states to compute a new global state. This means it takes into account how the vertices depend on each other based on their connections (edges) in the graph. The function ensures that when updating the global state, it respects the relationships between vertices. This is crucial because the dynamics of a graph are not merely a collection of individual vertices; they emerge from the interactions between them.

In summary, the iterative calculation of the global state involves starting with initial states for all vertices, initially set to zeros, then using the local state transition function to update the state of each vertex. Following this, the global encoding function is applied to derive the new global state. This process is repeated for a predetermined number of iterations or until a specified convergence criterion is achieved.

Contractivity Condition

Contractivity is a property that ensures stability in a dynamic system. The goal is to ensure that, as the model processes input data, the changes in the states of the vertices do not grow uncontrollably. Instead, they should converge towards a stable configuration. The weight values of the recurrent connections are initialized according to a contractive condition on the global state transition function and then are left untrained. The contractivity condition is expressed as:

$$\sigma = (\max_{g} |V(g)|)\|\hat{W}_N\|^2 < 1 \tag{4.3}$$

Here, σ is the contraction coefficient of the network. $|V(g)|$ represents the number of vertices in graph g, and $\hat{W}_N$ is the recurrent weight matrix that governs the state transitions. The matrix $\hat{W}_N$ is rescaled to ensure that $\sigma < 1$ holds. $\hat{W}_N$ is scaled by the reservoir spectral radius, which serves as a hyperparameter tested in our pipeline. The various reservoir spectral radius tested are 0.1, 0.4, 0.7, 1.0, 5.0, and 10.0. By initializing the recurrent weights to satisfy contractivity conditions and keeping them fixed, GESNs can maintain stable state dynamics without the overhead of extensive training, allowing for faster implementation and reduced computational costs. This approach not only simplifies the model by minimizing the number of parameters that need to be adjusted but also ensures that the system remains stable during operation, preventing erratic behavior.

State Mapping Function

For structure-to-element transductions, GESN utilizes a *state mapping function* to generate a fixed-size representation of the entire graph. One method of achieving this mapping is known as Mean State Mapping, where the fixed-size state (element) is calculated as the average of the states of all vertices in that particular graph. This approach resembles average pooling, as it aggregates the information from all node embeddings to derive a comprehensive graph embedding. This can be expressed as:

$$X(x(g)) = \frac{1}{|V(g)|} \sum_{v \in V(g)} x(v) \tag{4.4}$$

By considering all vertices, this method captures the collective dynamics of the graph, potentially leading to richer representations.

Readout Function and Training

In this step, a linear transformation is executed by the output weight matrix W_{out}, which connects the fixed-size state representation/graph embedding $X(x(g))$ to the final output $y(g)$. This matrix applies a linear mapping to the fixed-size representation, effectively learning how to weight the contributions of the various features encapsulated in $X(x(g))$ to generate the prediction. The output $y(g)$ is computed as:

$$y(g) = W_{\text{out}} X(x(g)) + b,$$

where b is the bias term. The components that are learned or trained in this process are the output weight matrix W_{out} and the bias b by ridge regression. During the training phase, the model adjusts the values of W_{out} and b to minimize the error between the predicted outputs $y(g)$ and the actual target values, allowing it to better map the fixed-size representations to the desired outcomes. In this context, regularization is applied to the readout layer and is treated as a hyperparameter in our Ridge Regression model.

Once our readout layer is trained using regularized Ridge Regression, the learned weights W_{out} and biases b are used to make predictions on the new graphs.

4.3 Results

The performance metrics used to evaluate the effectiveness of both our proposed module and our baseline in this study are outlined. In particular, accuracy and sensitivity are presented. Accuracy is measured by evaluating the proportion of correctly classified instances out of all predictions made, while sensitivity focuses on the ability of the model to correctly identify positive instances. Since the goal is to identify whether pathology is present or not, and the pathology is considered the positive class, sensitivity is regarded as a very important metric in this context.

4.3.1 Baseline, Strategy No. 1

The performance of the first strategy of our baseline (non-graph based approach) is presented below, showing the mean results from 5-fold cross-validation. For each of the four classifiers, features are utilized in four different ways to observe the effects of various feature usage styles. Further comments on the performance are discussed in the discussion section.

Feature Extraction from Images					
No. of Features	Classifier	Cross-Validation	Mean Accuracy	Mean Sensitivity	Mean Specificity
All	Random Forest	Skfold = 5	0.708 ± 0.13	0.676 ± 0.35	0.733 ± 0.19
32 most important	-	-	0.696 ± 0.14	0.643 ± 0.34	0.725 ± 0.17
16 most important	-	-	0.657 ± 0.13	0.581 ± 0.32	0.711 ± 0.18
8 most important	-	-	0.696 ± 0.18	0.548 ± 0.32	0.8 ± 0.2
All	Support Vector Machine	Skfold = 5	0.683 ± 0.12	0.61 ± 0.17	0.733 ± 0.11
32 most important	-	-	0.696 ± 0.13	0.643 ± 0.22	0.733 ± 0.15
16 most important	-	-	0.618 ± 0.15	0.605 ± 0.25	0.622 ± 0.18
8 most important	-	-	0.525 ± 0.1	0.643 ± 0.27	0.444 ± 0.16
All	Linear Discriminant Analysis	Skfold = 5	0.656 ± 0.13	0.538 ± 0.2	0.733 ± 0.09
32 most important	-	-	0.711 ± 0.1	0.614 ± 0.16	0.778 ± 0.07
16 most important	-	-	0.617 ± 0.07	0.414 ± 0.23	0.756 ± 0.13
8 most important	-	-	0.63 ± 0.1	0.414 ± 0.28	0.778 ± 0.12
All	XGBoost	Skfold = 5	0.684 ± 0.11	0.619 ± 0.29	0.733 ± 0.17
32 most important	-	-	0.658 ± 0.11	0.552 ± 0.31	0.733 ± 0.17
16 most important	-	-	0.685 ± 0.11	0.652 ± 0.17	0.711 ± 0.18
8 most important	-	-	0.577 ± 0.09	0.486 ± 0.33	0.644 ± 0.2

4.3.2 Baseline, Strategy No. 2

Same as strategy no. 1, the mean performances of strategy no. 2 are presented below following the same fashion. This allows for a direct comparison of the performance between both strategies.

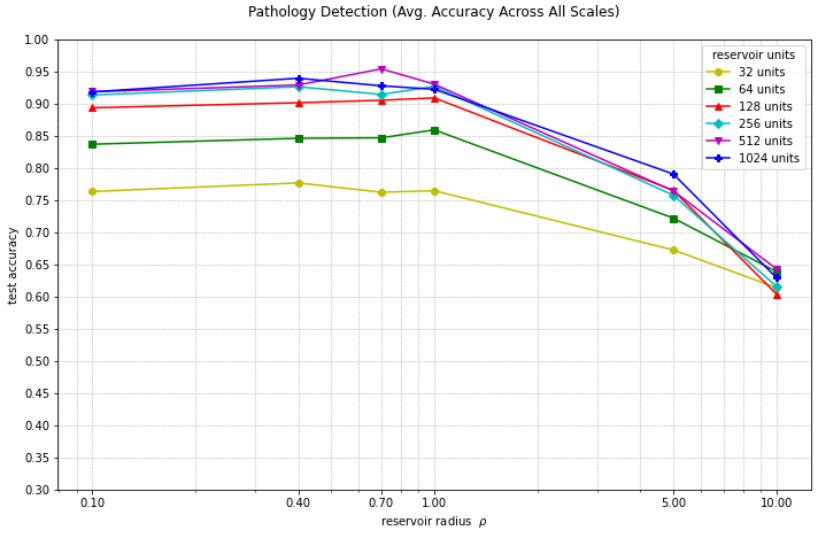
Feature Extraction from Superpixels					
No. of Features	Classifier	Cross-Validation	Mean Accuracy	Mean Sensitivity	Mean Specificity
All	Random Forest	Skfold = 5	0.658 ± 0.11	0.586 ± 0.1	0.711 ± 0.19
32 most important	-	-	0.618 ± 0.09	0.519 ± 0.13	0.689 ± 0.18
16 most important	-	-	0.645 ± 0.05	0.552 ± 0.1	0.711 ± 0.09
8 most important	-	-	0.632 ± 0.07	0.552 ± 0.1	0.689 ± 0.13
All	Support Vector Machine	Skfold = 5	0.604 ± 0.11	0.705 ± 0.08	0.533 ± 0.16
32 most important	-	-	0.577 ± 0.1	0.705 ± 0.08	0.489 ± 0.15
16 most important	-	-	0.551 ± 0.13	0.671 ± 0.11	0.467 ± 0.15
8 most important	-	-	0.564 ± 0.1	0.705 ± 0.08	0.467 ± 0.13
All	Linear Discriminant Analysis	Skfold = 5	0.564 ± 0.1	0.448 ± 0.1	0.644 ± 0.16
32 most important	-	-	0.565 ± 0.09	0.419 ± 0.07	0.667 ± 0.19
16 most important	-	-	0.578 ± 0.1	0.357 ± 0.13	0.733 ± 0.23
8 most important	-	-	0.565 ± 0.12	0.257 ± 0.22	0.778 ± 0.29
All	XGBoost	Skfold = 5	0.645 ± 0.08	0.519 ± 0.23	0.733 ± 0.11
32 most important	-	-	0.658 ± 0.06	0.519 ± 0.23	0.756 ± 0.08
16 most important	-	-	0.658 ± 0.06	0.519 ± 0.23	0.756 ± 0.08
8 most important	-	-	0.658 ± 0.06	0.519 ± 0.23	0.756 ± 0.08

4.3.3 Proposed Graph-based Module

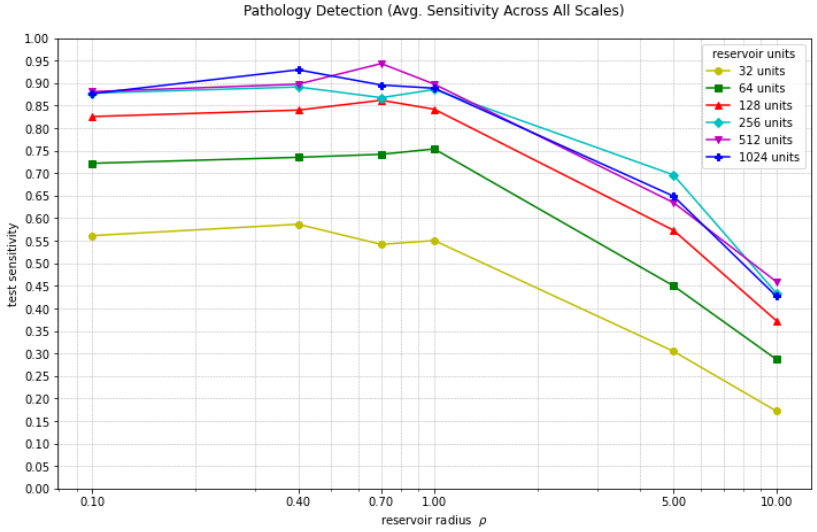
The performance of our graph-based classification approach using the Graph Echo State Network is presented here. As in baseline, the model's effectiveness is evaluated using two key metrics which are mean accuracy and mean sensitivity. These metrics are measured across various values of the reservoir spectral radius and scaling factors. Both of these parameters have a significant impact on the model's performance.

In fig. 4.4a, the model's mean accuracy across various scaling factors is shown, while fig. 4.4b presents the corresponding sensitivities. As illustrated in fig. 4.4, the highest mean accuracy achieved is 0.953 ± 0.013, which occurs when the reservoir

units used are 512 and the reservoir spectral radius is 0.7. This highest mean accuracy represents the mean performance using 512 reservoir units and a spectral radius of 0.7, the mean across all scaling factors, with the experiment repeated for two trials. Running experiments across multiple trials ensures that the model's performance is not dependent on a single set of initial parameters, providing a more robust and reliable evaluation by averaging results over different initialization. Associated with this highest mean accuracy, the highest mean sensitivity obtained is 0.943 ± 0.035.



(a) Avg. Accuracy (across all Scales)

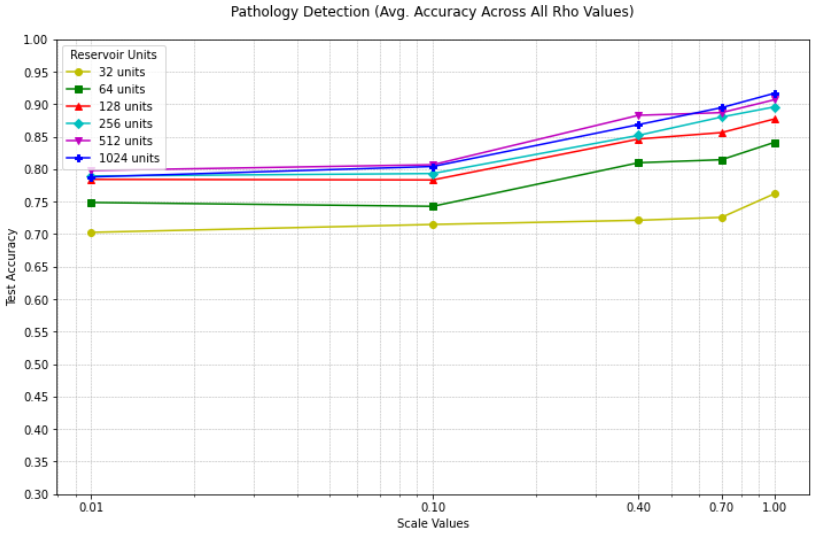


(b) Avg. Sensitivity (across all Scales)

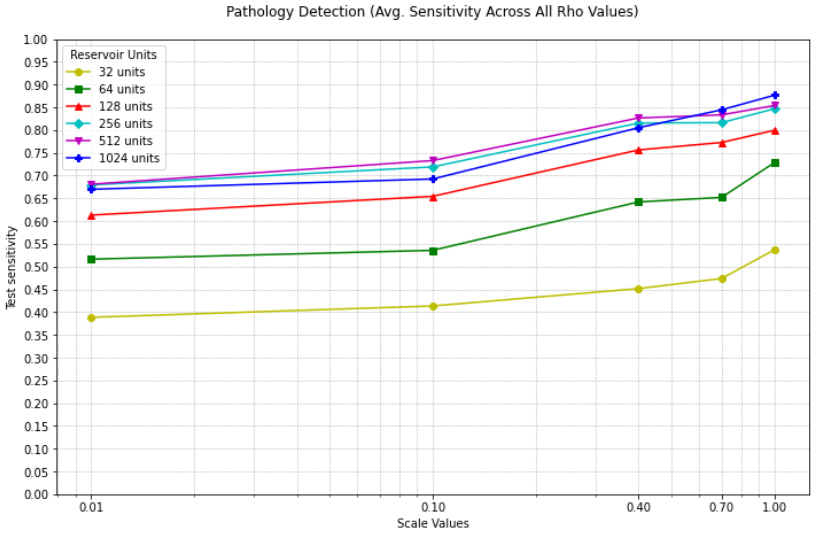
Figure 4.4: Avg. Accuracy and Avg. Sensitivity (across all scales)

Alternatively, fig. 4.5 presents the mean performance for various reservoir spectral radius values across each scaling factor. The highest mean accuracy is 0.916 ± 0.039 which is achieved using 1024 reservoir units, a scaling factor of 1.0, and considering all reservoir radius values (fig. 4.5a). This data point can be read as the mean accuracy for all spectral radius values with a scaling factor of 1.0 and 1024 reservoir units over two trials. The corresponding highest mean sensitivity is 0.876 ± 0.055 (fig. 4.5b).

A detailed comparison of the results is covered in the discussion section.



(a) Avg. Accuracy (across all Reservoir radius)



(b) Avg. Sensitivity (across all Reservoir radius)

Figure 4.5: Avg. Accuracy and Avg. Sensitivity (across all Reservoir Spectral radius)

4.4 Discussion

This section explores the performance of our proposed graph-based module alongside the baseline, emphasizing how different parameters impact their respective outcomes. By analyzing these modules in detail, we aim to understand the strengths and weaknesses inherent in each module.

In baseline strategy no. 1, the radiomic features are extracted from the entire cropped image. In the first experiment for each of the four classifiers, all features are used, while in the remaining three experiments for each classifier, only a certain number of the most important features are used. Based on the given performance metrics, it is clear that these important features do not outperform the performance when all features are used. An improvement in mean accuracy is noted only with the Linear Discriminant Analysis classifier when using the 32 most important features compared to using all features. However, considering overall performance, including sensitivity metrics, the Random Forest classifier shows the best performance when all features

are used. This indicates that Random Forest when using all features, remains a powerful classifier compared to Support Vector Machine, Linear Discriminant Analysis, and XGBoost classifiers. In the strategy no. 2 of our baseline, rather than extracting radiomic features from the entire cropped image, features are extracted from each superpixel individually. Each superpixel is treated as a unique sample, and if a subject's class label is pathological, all generated superpixels from all associated slices of that subject are labeled as pathological superpixels. Each superpixel is then trained as a unique sample, and the final prediction is determined by majority voting among all the superpixels of each subject. This is considered a more naive approach as labeling a subject and its associated slices as pathological does not necessarily imply that all superpixels are pathological since pathology is not necessarily present in the entire area. However, since the exact location of the pathology is unknown, and only the subject's class label is present, precise labeling of each superpixel is not possible. As a result, each superpixel's label follows the subject's class label. This limitation is reflected in the performance, as the mean sensitivities are generally poor, regardless of the classifier or the number of features used. Consequently, the overall performance, including mean accuracy, is also poor. Although the sensitivity of SVM is comparatively higher in strategy no. 2 than in strategy no. 1, the overall performance of SVM is better with strategy no. 1 when considering both accuracy and specificity. Strategy no. 1 provides a more balanced approach to detecting both positive and negative classes, with higher overall accuracy compared to strategy no. 2.

Overall, between the two strategies of our baseline, it is evident that strategy no. 1 performs better. This outcome is expected, as extracting features from the entire image likely captures more comprehensive global information, providing a more accurate representation of the subject. In contrast, strategy no. 2, which labels and treats each superpixel as an individual pathological or healthy sample couldn't perform as well as the first strategy. The class labels used in strategy no. 1 are inherently more reliable, reflecting the subject's overall condition rather than making the assumption that every superpixel shares the same label. By capturing more meaningful features, and by avoiding the noise introduced by naively assigning labels to each superpixel, strategy no. 1 results in more consistent and accurate classification performance. Thus, the superior performance of strategy no. 1 reinforces the importance of leveraging global information in such classification tasks.

The performance of our proposed module demonstrates the effectiveness of our graph-based classification system that transforms each MRI image into a graph for echo state network (ESN)-based classification. Here the classification problem is treated in terms of regression problem where the inputs of the regression model are graph embeddings and associated labels. These graph embeddings are derived from the echo state properties of each graph. The performance evaluation reveals that increasing the number of reservoir units—specifically, 32, 64, 128, 256, 512, and 1024 improves performance, with configurations of 512 to 1024 units yielding the best performance. While only 32 units lead to significantly lower performance. The improvement with larger reservoirs can be attributed to several factors. First, a larger reservoir serves as a dynamic, high-dimensional state space, enabling the GESN to capture complex patterns and interactions within the input graphs; with only 32 units, the limited capacity fails to effectively represent underlying data patterns. Second, the dynamics of the reservoir, driven by non-linear activation functions, benefit from more units, enhancing the network's ability to model non-linear relationships. Third, larger reser-

voirs can approximate a wider range of functions, which is crucial for complex tasks like graph classification and pathology identification, where intricate relationships exist among data points. Additionally, having more units facilitates a diverse set of activations, allowing each unit to respond to different features of the input data and resulting in a richer representation. Finally, as more reservoir units are incorporated, the network can learn hierarchical feature representations, with lower layers capturing basic structures and higher layers identifying abstract patterns critical for tasks like pathology identification. Overall, while the 32-unit configuration proves inadequate, the enhancements seen with 512 and 1024 units underscore the GESN's capability to effectively model complex relationships in the data. Regarding the scaling factor, our results indicate that a scaling factor of 1.0 is the most effective choice for our problem. This scaling factor directly influences the uniform weight initialization, allowing the weights to fully utilize the variance of the uniform distribution. Such an approach can significantly enhance the model's ability to learn complex relationships within the data. Additionally, for the reservoir spectral radius, values between 0.4 and 0.7 appear to yield the best performance, as these values are less than 1, aligning with theoretical expectations. Notably, when the spectral radius exceeds 1, there is a marked decline in performance, highlighting the importance of maintaining an optimal range for this parameter.

As a discussion on the excellence of our GESN-based graph classification task from a more global perspective, the performance of our model resonates with the power of GESN over traditional Graph Convolutional Networks (GCNs) and other deep learning models, particularly in terms of computational efficiency and training simplicity. Our proposed graph-based module also achieved significantly better performance than the baseline module. GESN utilizes a reservoir of fixed, randomly initialized weights, eliminating the need for iterative backpropagation-based weight updates and reducing the computational cost associated with training deep models. This approach allows GESN to handle dynamic graphs and temporal data effectively without requiring extensive parameter tuning.

4.5 Conclusion

This study aimed to develop an end-to-end system for tendon pathology detection (classification) using a Graph Echo State Network (GESN) based on the reservoir computing paradigm. Unlike traditional neural networks, the GESN produces rich graph embeddings without the need for iterative weight training. By transforming medical images into graphs, the system captures both local and global patterns, providing a more expressive representation compared to standard image grids. These GESN-generated graph embeddings are then processed through a regularized Ridge regression model, resulting in significant improvements in pathology detection performance compared to our non-graph-based baseline approach.

In the baseline approach, radiomic features extracted from entire cropped images outperformed those derived from individual superpixels, underscoring the benefit of using a more global view of the image rather than focusing on isolated regions. In contrast, our GESN-based module advances this concept by employing graph-based representations and projecting the data into much higher-dimensional spaces, which

enhances the quality of the representations without relying on backpropagation-based weight updates, setting it apart from traditional neural networks.

The study also emphasizes the importance of hyperparameters, such as the spectral reservoir radius and scale factor, in determining the performance of the GESN module. Fine-tuning these parameters is shown to have a substantial impact on the effectiveness of graph classification. However, the small dataset used in this study is a limitation, suggesting that future research could focus on testing this on larger datasets for more robust evaluation. Additionally, optimizing superpixel sizes and incorporating full, uncropped images could improve the system's overall performance.

In summary, this study highlights the promise of graph-based approaches in medical imaging, particularly for diagnosing tendon pathology, and showcases the GESN's ability to offer more expressive and efficient data representations. Future work could further explore these insights by addressing current limitations and refining the proposed methods.

Chapter 5

Tendon Reflex Responses

This study explores how age, height, weight, and gender influence reflex response times in healthy individuals, aiming to establish a baseline for understanding normal neuromuscular function. By examining these factors, the research provides valuable insights into how demographic variables affect tendon reflexes, offering a reference point for future studies on neuromuscular conditions and their potential variations across different populations.

This study has been accepted for publication in the "Frontiers in Neurology" journal (Impact Factor: 2.7, CiteScore: 4.9), under the "Neuromuscular Disorders and Peripheral Neuropathies" section.

Article Title: Assessing Neuromuscular System via Patellar Tendon Reflex Analysis Using EMG in Healthy Individuals

Keywords: Electromyography, Nerve Conduction Velocity, Patellar Tendon Reflex, Response Time Measurement, Reference Database

5.1 Introduction

Tendon reflex is a standard component of neurological evaluation. The characteristics of the patellar tendon reflex (PTR) provide fundamental insight regarding the diagnosis of neurological status. The PTR is a deep tendon reflex that assesses the femoral nerve and spinal cord segments L2-L4 function. The PTR is elicited by tapping the patella tendon, causing the quadriceps to stretch. The stretch stimulates the muscle spindles, which leads to impulses being sent to the spinal cord via sensory afferents. The sensory fibers branch as they reach the spinal cord; some enter the grey matter of the cord and make monosynaptic contact with the lower motor neuron (LMNs), which stimulates the muscle to contract. As a result of this, the leg extends at the knee joint as the quadriceps muscle contract suddenly and the hamstring relaxes [323] [324] [325]. Due to the muscle contraction, the muscle spindles shorten again and their afferent activity is reduced. Furthermore, sensory neurons act indirectly with interneurons to inhibit flexor motor neurons that would otherwise contract the opposite muscle, the hamstring.

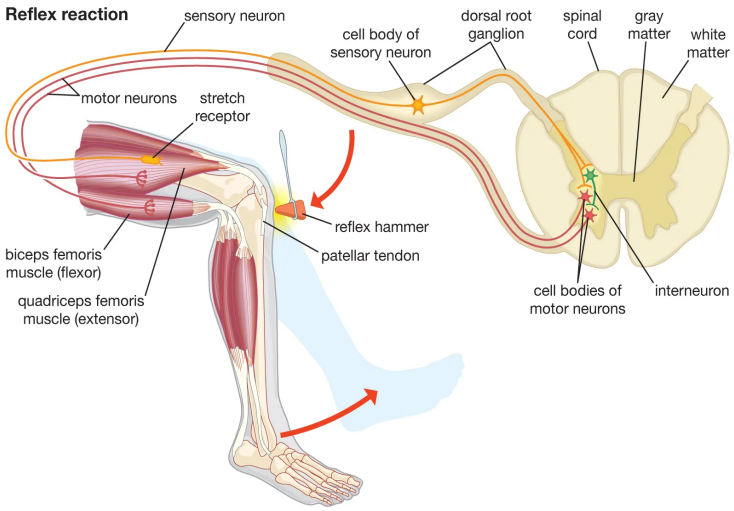


Figure 5.1: Patellar Tendon Reflex [326]

This reflex is an essential tool for clinicians in evaluating the health and functionality of the nervous system. By observing and measuring the reflex responses, healthcare professionals can identify any abnormalities or changes that may indicate underlying neurological conditions. These reflexes are involuntary responses to stimuli, and alterations in their normal patterns can serve as early warning signs of potential issues such as nerve damage, neurological disorders, or systemic diseases. Therefore, careful analysis of reflex responses not only aids in diagnosing existing conditions, but also plays a crucial role in monitoring the progression of neurological health over time [325].

The patellar tendon reflex is chosen for this study because it is commonly used to assess the integrity of the spinal cord and peripheral nervous system. This choice aligns with the broader goals of the two large-scale projects involved, which aimed to study both the spinal cord and tendons. The simplicity, reliability, and clear measurement of neural and muscular response times make the patellar tendon reflex ideal for examining how different demographic factors influence reflex timing.

Research has demonstrated that spinal reflexes vary significantly between younger and older individuals. Chandrasekhar et al. found that patellar reflex amplitude decreases with age, highlighting substantial differences in reflex magnitude across three

age groups [327]. Dudgeon et al. further concluded that aging reduces reflex excitability, particularly during the Jendrassik maneuver, which involves clenching teeth and locking fingers while the patellar tendon is struck [328]. Additionally, Koceja et al. revealed that older adults exhibit reduced monosynaptic reflex responses, particularly in standing positions, suggesting that age-related changes in the nervous system, such as motoneuron loss and decreased muscle spindle sensitivity, contribute to impaired balance and increased fall risk among the elderly [329]. Carroll et al. investigated the in vivo mechanical properties of the patellar tendon, noting that differences related to aging are more associated with force output than direct age effects; however, decreased signal intensity indicated alterations in the tendon's internal milieu, potentially affecting muscle spindle excitation during reflex hammer testing [330]. Although Chandrasekhar et al. reported a significant age effect on the magnitude of patellar reflex responses, they found no influence of gender, concluding that neurologically normal individuals experience an age-dependent decline in patellar reflex. Notably, this study did not address whether the delay between the reflex hammer strike and the onset of EMG responses changes with age [331]. Kamen et al. examined force-time characteristics of the patellar tendon reflex and found that older individuals produced greater overall reflex force; they also noted that a contralateral conditioning stimulus resulted in short-latency inhibition followed by longer-latency facilitation, both of which were more pronounced in aged subjects, indicating that age-related changes may occur at the spinal level and could influence the reaction delay of the quadriceps muscle following a reflex hammer strike [332]. Hwang et al. investigated the impact of body weight on the soleus H-reflex and found that the amplitudes of the reflex significantly decreased as body weight load was reduced [333]. Although this study did not examine response times, it highlighted the relationship between body weight and reflex amplitude. In the context of gender differences, one study utilizing surface electromyography indicated that males exhibit a slower patellar reflex compared to females [334]. Supporting this, other studies have reported higher reflex responses in females [335]. Conversely, some studies found no significant differences between the genders [336]. As a result, the existence of sex differences in deep tendon reflexes remains a contentious topic in the literature [337]. Campbell et al. examined the relationship between patient height and sex with patellar tendon length (PTL). They found no significant differences in mean age and body mass index between male and female patients, concluding that PTL is more strongly correlated with patient sex than with height. This suggests that the initiation of the patellar tendon reflex may differ between genders, while height appears to have a lesser impact [338]. Pazzinatto et al. investigated women with and without patellofemoral pain (PFP) to compare the amplitudes of the patellar tendon reflex and the vastus medialis Hoffmann reflex (VM H-reflex) between the two groups. They found that women with PFP exhibited significantly lower amplitudes of the patellar T-reflex compared to pain-free controls, and that the VM H-reflex was strongly correlated with the patellar T-reflex in both groups [339]. While there are numerous studies addressing tendon reflex responses, none specifically focus on how response times fluctuate when accounting for various demographic variables. This gap in the literature underscores the primary objective of our study.

This study investigates the neuromuscular system using the patellar tendon reflex test, specifically examining how age, height, weight, and gender influence reflex response times. We analyze three key aspects of reflex responses: the onset of muscular

activity (the moment the reflex is triggered), the endpoint of the reflex, and the total duration of the reflex. By comparing these response times in elderly and young individuals, we comprehend how aging affects reflex timing. Additionally, we assess the impact of height by comparing reflex responses between taller and shorter individuals and evaluate how weight influences reflex responses by examining differences between heavier and lighter individuals. The study also includes analyses that normalize for height and weight to isolate the pure effects of aging on reflex responses. Furthermore, we explore potential gender-related differences in reflex reactions. For our experiments, EMG data is collected on the patellar tendon reflex, and after obtaining the response times for different groups, statistical tests (independent t-test) are conducted to identify significantly different samples.

5.2 Materials and Methods

5.2.1 Hardware and Software

KineLive is a wireless surface EMG biofeedback system (KISO) that has been shown to be an effective tool for diagnosing and treating problems with the human muscular-skeletal system [340]. The wireless measurement units are small and light. The system consists of both a Kine measurements system and KinePro software. In this study, KISO's 4 and 8-channel EMG systems are utilized. The Kine unit amplifies the signal by 800, has a sampling frequency of 1600 Hz, and a signal bandwidth of 16-500 Hz. The analog-to-digital converter (ADC) is 10-bit and the saturation level is ±2 mV. Each unit transfers the recorded EMG signal wirelessly to the measuring system, which is connected to a computer. The KinePro software then allows the recorded signal to be analyzed visually as well as transferred to other software applications for further research. The sEMG units are used with $Triode^{\text{TM}}$ (T3402M) electrodes that have a 2 cm spacing of silver-silver chloride electrodes and nickel-plated brass snaps [341]. To gain quantitative data on deep tendon reflexes, like the PTR test, an EMG system can be used. The electrically synchronized reflex hammer used in this study includes an accelerometer to accurately measure the force and timing administered by the reflex hammer on the subject's tendon [342]. The accelerometer and its surrounding electronics are capable of accurately synchronizing with the EMG measurement with an error of less than 0.6 ms.

5.2.2 Participants

The dataset used in this study comprises 40 participants, categorized into two distinct age groups. The younger cohort comprises 25 individuals, with an average age of 23.76 ± 1.75 years, an average height of 177.48 ± 12.52 cm, and an average weight of 75.28 ± 16.17 kg. This group includes 14 males and 11 females. In contrast, the elderly cohort comprises 15 individuals, with a mean age of 59.8 ± 5.66 years, an average height of 174.33 ± 6.51 cm, and an average weight of 82.33 ± 13.68 kg, including 8 males and 7 females. Participants are selected based on predefined health criteria, ensuring that all individuals are in good health at the time of data collection. This selection process ensured that the dataset would reflect typical physiological characteristics without the influence of significant underlying health conditions, providing a more accurate

baseline for the study's focus. This database serves as a reference collection of healthy cases and can be used as a baseline for studying neurological disorders.

The physiological measurements were approved by the Medical Director of Landspítali, the Science Ethics Committee, and the Icelandic Data Protection Act. The Icelandic name of the ethics committee is Vísindasiðanefnd (The national bioethics committee, in English), under permission no. 14-150-S1, with an approval date on November 18, 2014. The study title in Icelandic is *Raförvun mænu með yfirborðsrafskautum, færnibætandi áhrif á síspennu* (translated to English: Transcutaneous spinal cord stimulation for skill-improving effects on spasticity). Fig.5.2 illustrates the distribution of subjects across various categories and age groups.

Age Group / Demographics	Young (No. of Subjects)	Elderly (No. of Subjects)
Age	25 (Mean Age: 23.76 ± 1.75 years)	15 (Mean Age: 59.8 ± 5.66 years)
Taller	14 (>= mean height 177 cm)	7 (>= mean height 174 cm)
Shorter	11 (< mean height 177 cm)	8 (< mean height 174 cm)
Heavier	13 (>= mean weight 75 Kg)	6 (>= mean weight 82 Kg)
Lighter	12 (< mean weight 75 Kg)	9 (< mean weight 82 Kg)
Male	14	8
Female	11	7

Figure 5.2: Study Population

5.2.3 Data Acquisition Protocol

The following data acquisition protocol is used in our study:

1. Administer a questionnaire to collect the subject's age, height, and weight. The questionnaire also records any known underlying illnesses or injuries, the frequency of physical activity per week, and consumption of alcohol and caffeine.

2. Have the subject sit on a bench with their legs hanging freely without touching the ground.

3. Verify that the subject is not wearing any electronic devices to minimize background noise.

4. Mark the appropriate locations for surface electrode placement. The first EMG electrode (anode) is positioned approximately one-third of the distance from the subject's hip joint to the upper edge of the patellar bone. The second EMG electrode (cathode) is placed 6 cm below this point.

5. Adjust the final electrode placement to ensure they are positioned over the rectus femoris for accurate reflex response measurement.

6. Cleanse the skin with alcohol and apply conductive gel to enhance electrical conductivity before placing the EMG electrodes.

7. Attach the EMG electrodes to the skin using KineLive sEMG units to collect EMG data.

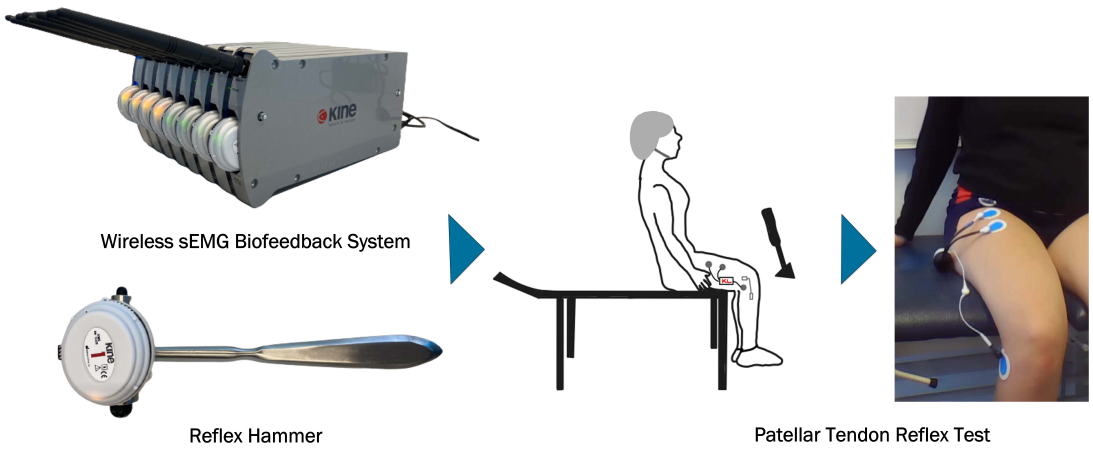


Figure 5.3: Data Acquisition

8. Place a grounding electrode on the head of the tibia bone.
9. Use a second KineLive sEMG unit integrated into a medical hammer to measure the force applied to the patellar tendon and synchronize this data with the reflex EMG data.
10. Connect the KineLive sEMG units to the KinePro software for synchronization.
11. Initiate the recording by clicking the record measurements button in the software, with the recording lasting a total of 5 minutes.
12. During the 5-minute recording, strike the patellar tendon with the reflex hammer in about every 15 to 20 seconds while recording the EMG data in KinePro software. Ensure the subject remains still, calm, and blindfolded to prevent anticipation of the strikes and ensure unbiased reflex data.

To reduce variability in strike localization, the midpoint of the patellar tendon is pre-marked to ensure consistent and accurate targeting. This pre-marking helps to eliminate guesswork and enhances precision during the procedure. Additionally, to standardize the force and direction of the strike, the examiner's elbow is supported, providing stability and enabling controlled, precise wrist movements.

5.2.4 Data Analysis

The pipeline of this study, illustrated in fig. 5.4, is designed to select only impactful signals for the final data analysis. This pipeline is implemented using Matlab R2023b. A portion of this pipeline was developed and used in a previous study with different objectives [343].

Initially, the recorded data (.d3d format) are imported into the Matlab environment. The first step involves identifying and extracting the hammer impact or impulse peaks to synchronize them with the corresponding electromyography (EMG) signals. This synchronization enabled the segmentation of the EMG data into individual epochs, where each epoch represented the EMG response to a specific hammer stimulus. These epochs are plotted for visual inspection to confirm accurate segmentation. Next, each EMG response to the hammer impact is carefully reviewed to assess

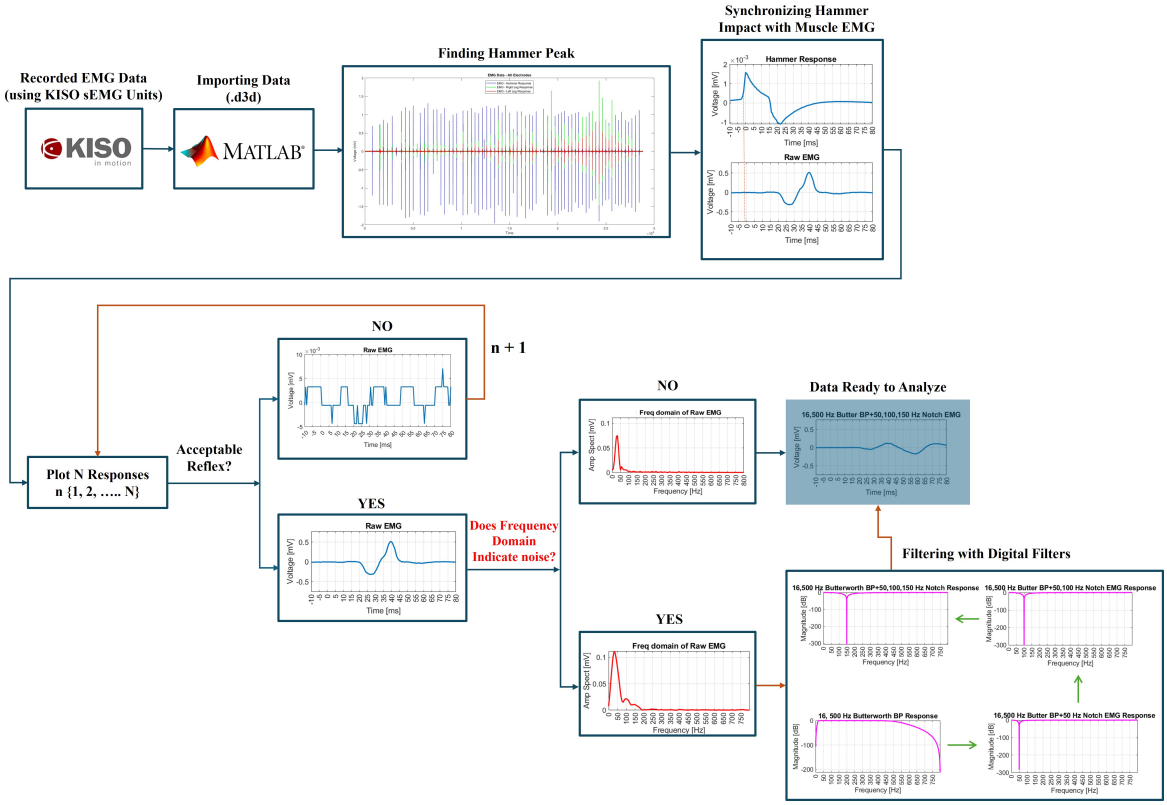


Figure 5.4: Pipeline to Select Impactful Signals Only

whether it contains an acceptable reflex or not. Responses that don't meet this criteria are discarded. The remaining acceptable reflexes are then checked for digital noise to ensure their reliability for further analysis. Any responses exhibiting excessive noise or deemed insufficient for detecting critical time points, such as latency, are excluded. Valid, noise-free responses are retained in a new dataset for further processing, including frequency domain analysis.

During the frequency domain analysis, any interference at 50 Hz, commonly associated with power line noise, is addressed by applying a 50 Hz notch filter. This filter effectively eliminated the 50 Hz interference. Prior to this step, a 4th-order Butterworth bandpass filter, with a lower cutoff frequency of 16 Hz and an upper cutoff frequency of 500 Hz, is applied to the EMG data. This bandpass filter helps to isolate the relevant signal components by removing unwanted low-frequency and high-frequency noise. In cases where additional noise at 100 Hz or 150 Hz is detected, corresponding notch filters are applied to eliminate this specific interference. The effectiveness of each digital filter is validated by examining the frequency and magnitude responses in the frequency domain. Once filtering is complete, the data are deemed suitable for the final analysis.

5.2.5 Response Time Analysis Across Subject Groups

Following the data processing, the response times of different subject groups are analyzed. Response times are measured in terms of the start of the response, the end of the response, and the total duration. The start of the response is defined as the precise moment when the reflex is triggered, indicating the onset of muscular activity following the stimulus. The end of the response marks the point when the muscle

activity ends. The total duration represents the time from the initial reflex trigger to the complete relaxation of the muscle. These metrics provide valuable information about the efficiency and speed of the neuromuscular response. After processing, these response times are analyzed across various subject groups to identify patterns or differences in reflex behavior based on factors such as age, height, weight, and gender. The specific groups studied in this study are outlined in the following sections.

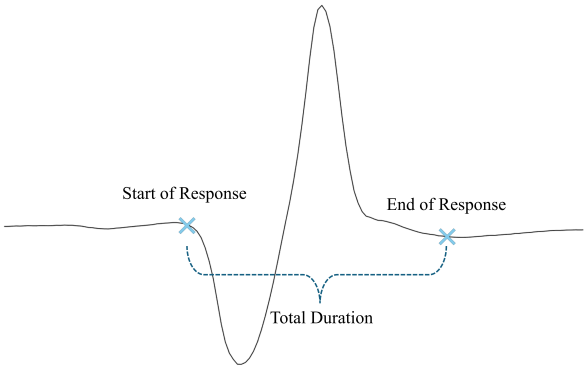


Figure 5.5: Measured Response Times

1. **Age Group Only**: The response times of elderly versus young individuals are analyzed. Elderly individuals are defined as those over 50 years of age, while young individuals are defined as those younger than 30 years of age. These age thresholds are selected based on the characteristics of our study population, ensuring clear distinctions between younger and older participants. The aim here is to compare the response times between these two distinct age groups.

2. **Age + Height Group**: In this experiment, the effect of height on response times within different age groups is examined. For elderly individuals, the mean height is calculated to be 174 cm. Elderly individuals whose height is greater than or equal to this value are labeled as taller, while those shorter than the mean are labeled as shorter. For the young group, the mean height is 177 cm. Young individuals taller than or equal to this height are considered taller, while those shorter than the mean are considered shorter. The response times are then compared between taller elderly versus taller young, and shorter elderly versus shorter young, to determine whether height significantly influences response times across age groups or not.

3. **Age Group Only (Height Normalized)**: Since individual heights vary, the response times are normalized by dividing the recorded times by each subject's height to account for height differences. This normalization allows for a more consistent comparison of response times across individuals, independent of height.

4. **Age + Weight Group**: This experiment examines the impact of weight on response times in different age groups. For elderly individuals, the mean weight is 82 kg. Elderly individuals who weigh more than or equal to the mean are labeled as heavier, others as lighter. Similarly, for younger individuals, the mean weight is 75 kg. Young individuals weighing more than or equal to the mean are marked as heavier, others as lighter. The response times of heavier versus lighter individuals within each age group are then compared.

5. **Age Group Only (Weight Normalized)**: To eliminate bias caused by varying weights, the response times are normalized by dividing each subject's response time by their respective weight. This normalization removes the influence of weight, allowing for an unbiased analysis of response time differences across individuals.

6. **Age + Gender Group**: In this setup, the effect of gender on response times within each age group is analyzed. Elderly males are compared to young males, and elderly females are compared to young females, to investigate potential gender-based differences in response times across age groups.

7. **Gender Group Only**: In addition to age-related comparisons, this experiment focuses exclusively on the effect of gender. Response times are analyzed by comparing all male individuals versus all female individuals, regardless of age, to assess whether gender alone has a significant impact on response times.

5.3 Results

5.3.1 Age Group Only

The comparison of response times between elderly and young individuals reveals a significant difference in the start of response ($p = 0.007$). The elderly individuals exhibits a slower onset time, with the response starting at 20.92 ± 3.013 ms, compared to 17.68 ± 3.708 ms for the young group. The end of response shows a marginal difference ($p = 0.068$), with the elderly group completing the response slightly later (78.47 ± 1.995 ms) compared to the young individuals (77.08 ± 2.397 ms), although this difference is not statistically significant. Additionally, the total duration of the reflex shows no significant difference between the two groups, with elderly individuals recording a duration of 57.55 ± 2.901 ms and young individuals 59.40 ± 4.013 ms, reflected by a p-value of 0.127.

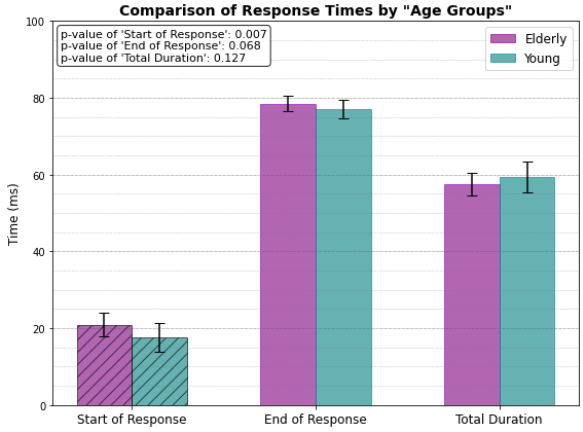
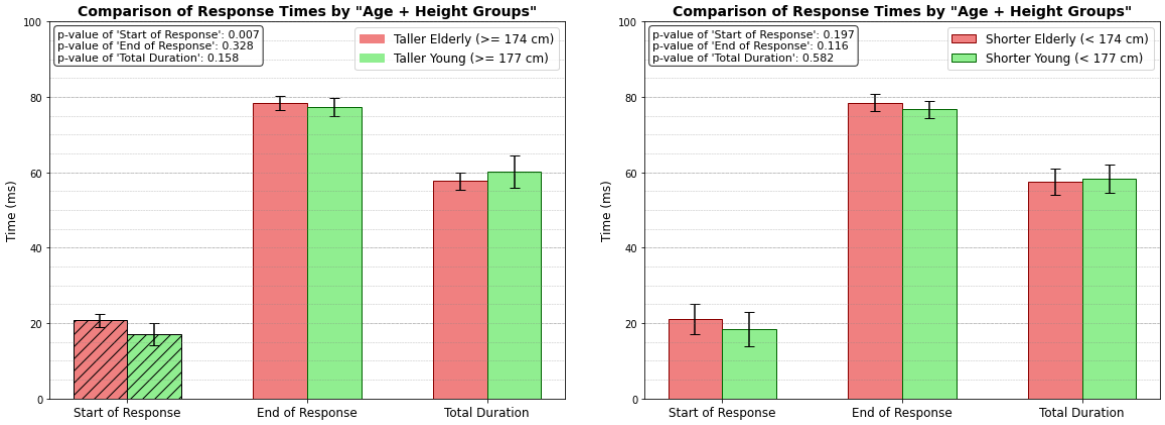


Figure 5.6: Only Age related Response

Fig. 5.6 presents the response times of the two age groups, with dashed bars used to highlight the response times that show a statistically significant difference (specifically, the start of response in this case).

5.3.2 Age + Height Group

The response times between the two groups, Taller Elderly ($\geq$ 174 cm) and Taller Young ($\geq$ 177 cm) (fig. 5.7a), shows that the taller young group gets reflex triggered significantly much faster (p = 0.007). The mean value of start of response for the taller young group is 17.14 $\pm$ 2.898 ms, compared to 20.74 $\pm$ 1.656 ms for the elderly group. However, both groups exhibit similar times for the end of the response, with no significant difference (p = 0.328). The total duration also shows no significant variation between the groups (p = 0.158).



(a) Taller Elderly Vs. Taller Young (b) Shorter Elderly Vs. Shorter Young

Figure 5.7: Age + Height related Response

Fig. 5.7b presents the measured response times for the Shorter Elderly (height $<$ 174 cm) and the Shorter Young (height $<$ 177 cm) groups. The results indicate that there is no significant difference between the groups in terms of the start of response, end of response, or total duration of the patellar tendon reflex. This lack of significant variation suggests that height does not have a considerable impact on reflex response times among shorter individuals, regardless of age.

5.3.3 Height Normalized, Age Group Only

Fig. 5.8 illustrates the height-normalized response times (unitless, since divided by heights) for both elderly and young individuals. The bar plots clearly demonstrate that there is a significant difference in the start of response between the two groups, with a p-value of 0.008. However, the analysis reveals no significant differences in the end response or total duration categories between the elderly and young individuals. This suggests that height normalization influences the timing of the reflex initiation but does not affect the other aspects of the response.

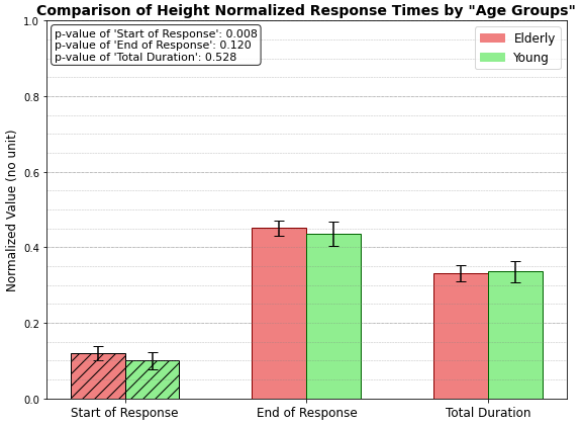
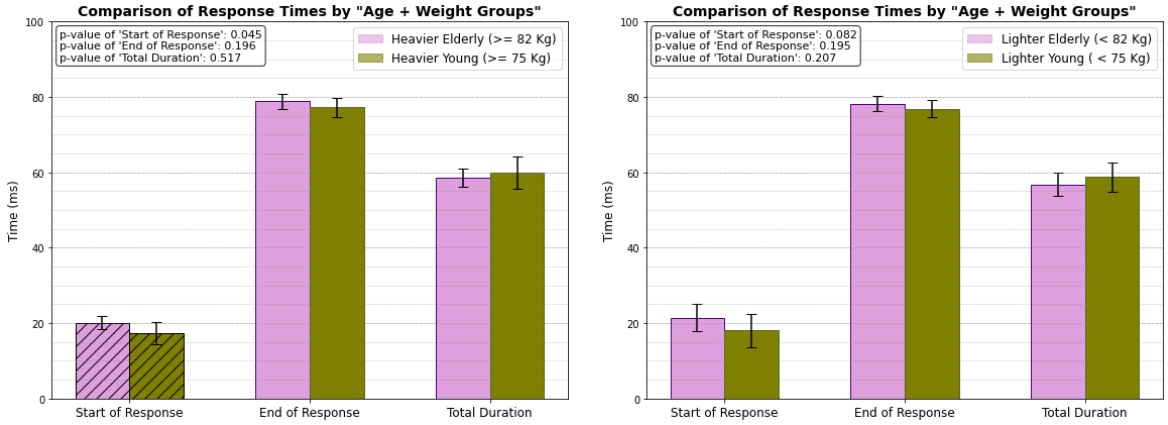


Figure 5.8: Height Normalized, Age related Response

5.3.4 Age + Weight Group

Fig. 5.9 illustrates how aging influences patellar tendon reflex response times, particularly in relation to the individuals' weight. The results indicate a significant difference in the start of response times, highlighting that weight plays a role in the initiation of the reflex response. Specifically, the heavier individuals exhibit a quicker start to their reflex action. The start of response occurs at 17.3 ± 2.947 ms for the heavier young individuals, while it is measured at 20.14 ± 1.742 ms for the heavier elderly individuals. However, for the lighter individuals, the difference in response times is less pronounced, suggesting that while weight impacts reflex initiation, the effect is not as strong in this group.



(a) Heavier Elderly Vs. Heavier Young (b) Lighter Elderly Vs. Lighter Young

Figure 5.9: Age + Weight related Response

In contrast, when evaluating the end of response and total duration, there are no significant differences observed between either the heavier or lighter groups. This finding suggests that, while weight may influence the speed of reflex initiation, it does not significantly affect the overall duration or the termination of the patellar tendon reflex response in either age group.

5.3.5 Weight Normalized, Age Group Only

When analyzing the weight-normalized response times, a significant difference emerges between the elderly and young subject groups in the total duration of muscle activation. The total duration for the young individuals, after normalizing for weight, is noticeably longer compared to that of the elderly individuals, with a p-value of 0.041 indicating statistical significance. However, no significant differences are observed in the start of response and end of response times between the two groups. This suggests that while the overall duration of muscle activation is influenced by age in the context of weight normalization, the initiation and ending of the patellar tendon reflex response remain similar across both young and elderly individuals.

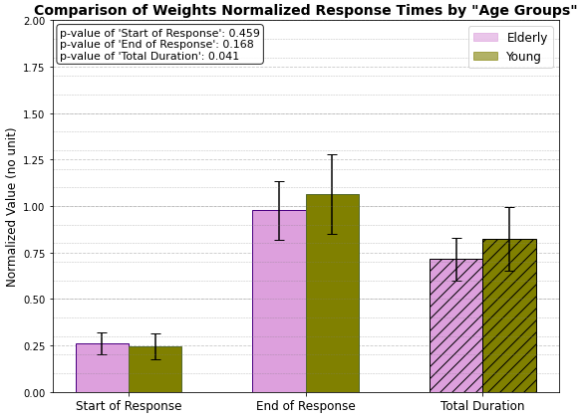
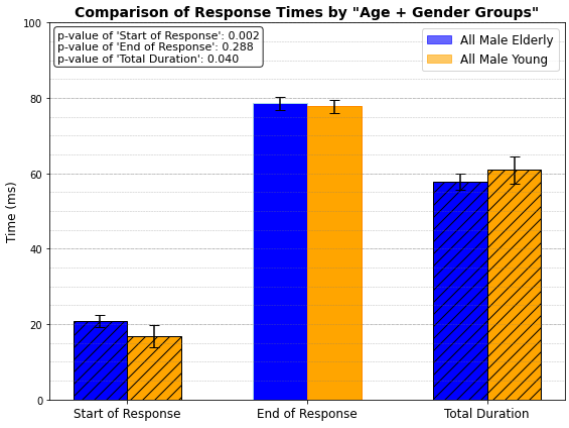
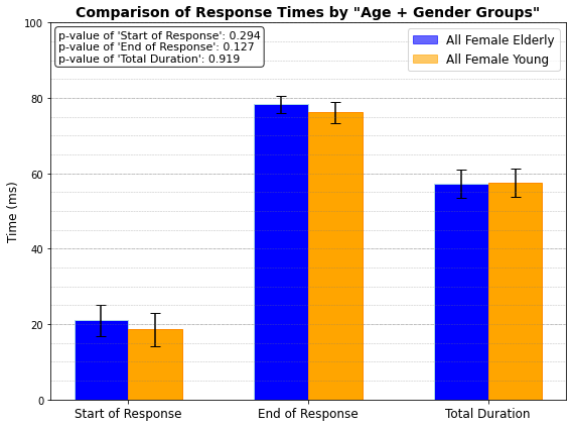


Figure 5.10: Weight Normalized, Age related Response

5.3.6 Age + Gender Group



(a) All Male Elderly Vs. All Male Young



(b) All Female Elderly Vs. All Female Young

Figure 5.11: Age + Gender related Response

In this case, the response timings are analyzed for both elderly and young individuals, with gender being a key factor. One experimental category focuses exclusively on males from both age groups, as shown in fig 5.11a, while another category examines all females from the same age groups, illustrated in fig 5.11b. The results reveal a significant difference in response timings only among males, with p-values of 0.002 for

the initial activation and 0.04 for the overall total duration. This suggests that male individuals exhibit notable variations in both the speed of reflex initiation and the length of the reflex response compared to their female group. However, when examining the female individuals, the analysis shows no significant differences in response timings between the elderly and young groups.

5.3.7 Gender Group Only

Lastly, the study focuses exclusively on gender groups, regardless of age, to investigate if there is any significant differences in reflex response timing. The results indicate that males, regardless of whether they are elderly or young, exhibit longer response times, with a statistical significance of p = 0.04. This finding aligns with the data presented in fig. 5.11a, which also considers age as a factor.

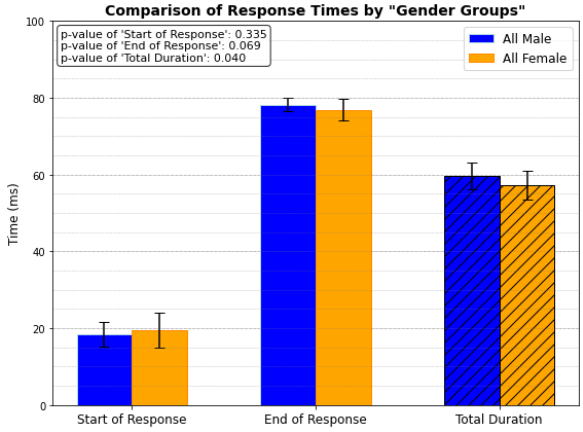


Figure 5.12: Gender related Response

5.4 Discussion

Age Group Only: The PTR test reveals notable differences in response times between elderly and young individuals, particularly at the onset of the reflex response. A significant delay in onset time is observed in the elderly group, suggesting that aging slows neural processing and motor response initiation. This delay is likely due to age-related physiological changes, such as decreased nerve conduction velocity, reduced muscle mass, and declined neuromuscular efficiency. Burke et al. indicated that aging reduces reflex excitability, particularly during the Jendrassik maneuver [328]. Further research has shown that older adults exhibit diminished monosynaptic reflex responses, especially in standing positions [329]. Overall, our results are consistent with the literature on age-related changes in reflex responses. In contrast, younger individuals generally have a higher proportion of fast-twitch muscle fibers, which may contribute to a quicker onset time. Interestingly, still the total duration of the reflex remains somewhat similar between this two groups. This suggests that while the initiation of the reflex is slower in the elderly, the overall reflex duration stays comparable. Elderly individuals may rely on compensatory mechanisms, such as greater reliance on experience or adaptations to physiological changes. According to Ward, the recruitment of additional motor regions in elderly adults provides a compensatory mechanism to counteract age-related brain changes [344]. Similarly, studies by Takeuchi et al. and Goble et al. highlighted how

overactivation of specific brain regions, such as the prefrontal cortex and supplementary motor areas, may compensate for declines in motor control and cognitive tasks in older adults [345] [346]. Furthermore, Reuter-Lorenz et al. suggested that the decreased lateralization in older adults could be compensatory, reflecting the recruitment of additional brain regions to support task performance [347]. For instance, although nerve conduction is slower, the efficiency of muscle contractions and movement coordination remains intact, allowing the reflex to occur within a similar timeframe as in younger individuals.

Age + Height Group: When categorizing individuals by height in addition to age, a similar pattern emerges where the elderly group shows a delayed onset of reflex response compared to the younger group, but there is no significant difference in the end of the reflex or the total duration. However, this pattern is significant only among taller individuals. In contrast, for shorter individuals, no significant differences are observed between the elderly and young groups in any of the reflex measures. The lack of significant variation in reflex response times between the shorter elderly and shorter young groups suggests that height does not play a major role in reflex timing for shorter individuals, regardless of age. This may be because the neural and musculoskeletal distances involved in the patellar tendon reflex arc (i.e., the pathway from the muscle spindles to the spinal cord and back) are relatively consistent among shorter individuals. Since the speed of the reflex depends on how quickly signals travel through nerves and how fast muscles respond, the shorter distances in these individuals likely lead to less variability in reflex times. In contrast, taller individuals have longer neural pathways, which may introduce more variability in reflex timing between age groups. This is consistent with the findings of Péréon et al., which demonstrated that there is a strong correlation between height and latency - specifically, as height increases, so does latency [348]. However, for shorter individuals, these distances are smaller and less variable.

Height Normalized, Age Group Only: Given the previous experimental outcomes, the objective of this step was to determine whether height impacts the performance differences between age groups. The results show that regardless of whether response times are analyzed based on height categories (taller vs. shorter) or height-normalized, the onset of the reflex is consistently delayed in elderly individuals ($p = 0.008$). However, no significant differences are observed in the end response or total reflex duration between the two age groups. This suggests that while the initiation of the reflex is slower in the elderly, the overall duration remains comparable. This is likely due to compensatory mechanisms and adaptations to physiological changes, as discussed earlier. These adaptations may allow elderly individuals to maintain a similar total reflex time, despite the delayed onset.

Age + Weight Group: In studying the combined effect of age and weight, we observe the familiar pattern that elderly individuals have slower reflex onset times compared to younger ones, though there is no significant difference at the end of response or total duration. However, two interesting observations emerge from this weight-related analysis. First, both elderly and young individuals in the heavier group demonstrate quicker reflex onset times than their lighter counterparts. This difference may be attributed to the greater muscle mass and increased mechanical load in heavier individuals, which can enhance motor responses and sensitivity in reflex pathways. Second, the expected age-related delay in reflex onset occurs only in the heavier individuals. Among lighter individuals, there is no significant difference in response times between

elderly and young individuals. This could be because lighter individuals experience less mechanical stress, which minimizes the impact of weight on their reflex performance. Additionally, lighter individuals typically experience less strain on their muscles and joints due to their lower body weight. This reduced mechanical load during reflex actions contributes to more consistent reflex performance across different age groups, as lighter individuals are less affected by the declines in muscle strength and coordination associated with aging.

Weight Normalized, Age Group Only: The findings indicate a significant difference in the total duration of muscle activation during the weight-normalized patellar tendon reflex between young and elderly individuals, with young individuals demonstrating longer activation times ($p = 0.04$). This difference may be attributed to age-related changes in muscle fiber composition and recruitment patterns. Additionally, the higher height of younger individuals could influence the time required for neural signals to travel between the muscle spindles to the spinal cord and back, potentially contributing to the longer activation duration. Although, there are no significant differences observed in the onset and end times of the reflex between the two groups.

Age + Gender Group: The results from the patellar tendon reflex test suggest that aging significantly affects reflex response times in males but not in females, highlighting a gender-specific difference in neuromuscular aging. In contrast, the absence of significant differences in all response times for females (p-values > 0.1) indicates that their reflex response times remain stable with age. This gender disparity could be attributed to biological differences, including hormonal factors, or potentially different patterns of physical activity and muscle health across the lifespan. These findings imply that age-related neuromuscular decline may be more pronounced in males.

Gender Group Only: The analysis of reflex response timings across gender groups, excluding age considerations, indicates that males exhibit significantly longer total response times than females ($p = 0.04$). This finding implies that, independent of aging effects, males may possess longer neuromuscular reflexes. One potential explanation for this difference is the average height of males compared to the mean height of females. Taller individuals, in this case, males, typically have longer nerve pathways, potentially extending the time required for signals to travel between the muscle spindles to the spinal cord and back which is involved in reflex actions. Additionally, greater limb length and increased muscle mass in males may necessitate more time for the neuromuscular system to fully activate, further contributing to the observed differences in reflex response times. While the differences in the start and end of responses are not statistically significant, the prolonged total duration in males suggests a more extended reflex action overall.

5.5 Conclusion

This study has established a quantitative baseline for interpreting patellar tendon reflex (PTR) responses by analyzing factors such as age, height, weight, and gender in healthy individuals. While physicians typically rely on qualitative assessments of PTR for general diagnostic purposes, our approach provides an objective and quantitative method for evaluating neuromuscular reflexes. The results offer a benchmark for assessing typical functional profiles in healthy individuals, which can serve as a reference point for future studies, enabling comparisons with pathological cases and aiding in the

detection of subtle abnormalities. By distinguishing normal age-related changes from pathological variations, our findings pave the way for improved diagnostic accuracy and the development of advanced therapeutic tools in neuromuscular medicine.

Our study demonstrates that aging is linked to delayed reflex response onset in elderly individuals, which supports established theories about age-related physiological changes. Supporting our findings, Chandrasekhar et al. observed a decrease in patellar reflex amplitude with age [327]. Additionally, another study indicated that aging reduces reflex excitability, particularly during the Jendrassik maneuver [328]. Further research has shown that older adults exhibit diminished monosynaptic reflex responses, especially in standing positions [329]. Overall, our results are consistent with the literature on age-related changes in reflex responses. While elderly individuals exhibit slower reflex initiation, the overall duration of the reflex remains comparable to that of younger individuals. Additionally, the interaction between age and height shows that elderly individuals experience pronounced delays in reflex onset, particularly among taller individuals, with shorter individuals not demonstrating this delay. Even when the data is height normalized, the pattern of delayed onset persists, indicating that, regardless of height considerations, younger individuals consistently show faster onset times than their elderly counterparts. In terms of the impact of weight, our study reveals that elderly individuals generally respond more slowly to reflex tests, particularly among heavier individuals. Additionally, when weight normalization is applied to assess the aging effect, younger individuals demonstrate longer reflex completion times, which correlates with their height, as taller individuals would naturally require more time to complete the reflex response. The study also reveals gender-specific differences in reflex response times, highlighting that aging impacts males more significantly than females. This suggests that biological and physiological factors may play a role in neuromuscular aging, leading to the observed differences in reflex responses. Furthermore, when considering only gender groups, males exhibit a longer total reflex duration, likely due to their generally greater height, which necessitates more time for signals to travel the longer nerve pathways. One study suggested that males exhibit slower patellar reflexes compared to females [334], which aligns with our finding that delayed reflex onset in the elderly is significantly observed only in males, not females.

Despite providing valuable insights, this research has a few limitations as well. The small sample size may limit the generalizability of the findings, and the cross-sectional data collection hinders the establishment of causal relationships. Future studies could include larger, longitudinal samples to track neuromuscular changes over time. Expanding research to include more diverse populations could deepen our understanding of age-related neuromuscular decline and the influence of gender and body mass index. This study serves as an exploratory analysis, generating preliminary insights. To address its limitations, including the use of independent t-tests with increased Type I error risk and challenges with small sample sizes, future work will involve re-analyzing the data using more robust methods, such as ANOVA test, on a larger cohort. This will allow us to perform experiments on more than two independent groups to study the effect of multiple demographic factors simultaneously. Another limitation of this study is that the mean age, height, and weight were computed based on our study population to maintain balance across all classes. With a larger dataset, standard reference values for age, height, and weight could be used to categorize different classes more effectively. The use of these mean values may influence the results, but as noted, future work should involve testing this approach on a larger cohort to validate the

findings. This may pave the way for improved diagnostic tools and more effective therapeutic strategies.

The authors would like to thank the Landspitali Science Fund (987013, 974260, 961204, 935838) and the Student Innovation Fund (2210641-1101, 2210642-1101, 2210643-1101) for their past support, which facilitated part of the data acquisition and contributed to the development of a portion of the pipeline used in this research. Additionally, we extend our thanks to the Department of Rehabilitation Medicine at Grensás, Landspitali University Hospital, for their prior support in facilitating part of the data acquisition.

Chapter 6

Conclusion

This thesis presents an in-depth exploration of tendon-related pathologies through four distinct studies. It delves into the relationship between tendinopathy and cartilage degeneration, advances techniques for tendon segmentation, assesses tendon pathologies, and explores tendon reflex responses for neuromuscular assessment. The findings not only advance our understanding of these complex relationships but also pave the way for innovative diagnostic methodologies that could significantly enhance patient care.

In the first study, we explored the relationship between patellar tendinopathy and cartilage degeneration. To our knowledge, this is the first study that focused solely on the features of the quadriceps muscles and patellar tendons to predict knee cartilage degeneration. Our findings indicate that these muscle and tendon-related features are powerful predictors of cartilage health, challenging the conventional reliance on broader anatomical features. Notably, the quantity of water content in both the quadriceps and patellar tendons emerged as a significant predictor of cartilage degeneration, along with age and total pixel counts of the patellar tendons. This novel insight suggests that the structural integrity of tendons may directly influence the health of adjacent cartilage. Furthermore, our study identified a critical link between fat content in the quadriceps and patellar tendons and the development of patellar tendinopathy. These findings highlight the potential of using targeted imaging metrics to identify at risk populations and enhance our understanding of tendon pathology. Despite limitations, including a small patient cohort, our study highlights the necessity of integrating artificial intelligence and automated assessment tools into clinical practice to improve diagnostic accuracy.

The second study presented a comprehensive tendon segmentation module that employs superpixel generation as a preliminary coarse segmentation step to improve overall segmentation accuracy. By framing the segmentation task as a superpixel classification problem, we simplified the process and achieved impressive results across various key performance metrics. This approach also preserved critical boundary information, enabling more robust segmentation outcomes. The versatility of our module, capable of accommodating both traditional machine learning and graph-based methods, allows for flexibility in application according to available resources and data characteristics. The promising performance and generalizability of our framework signify its potential for adaptation to diverse medical imaging tasks beyond tendon segmentation. Future research should focus on testing this module on larger datasets and integrating additional patient-related variables, thereby enhancing its clinical applicability and effectiveness in diagnosing tendon pathologies.

The third study of this thesis presented an end-to-end system designed for diagnosing tendon-related pathologies using a Graph Echo State Network (GESN) rooted in reservoir computing principles. Unlike traditional neural networks, the GESN generates rich graph embeddings without the need for iterative weight training. These embeddings are subsequently processed through a linear readout layer, where weights and biases are optimized using ridge regression with regularization. Our results demonstrated that radiomic features extracted from entire cropped images outperformed those derived from isolated superpixels, emphasizing the significance of anatomical relationships in improving classification accuracy. The GESN's capacity to function in higher-dimensional graph spaces enhances its ability to differentiate between healthy and pathological tendon states. Key hyperparameters, including the spectral reservoir radius and scale factor, proved crucial for optimizing the GESN's performance. Future research should focus on refining these parameters and validating the model with larger, more diverse datasets. Additionally, employing full uncropped images may yield deeper insights into tendon pathology. This study highlights the potential of graph-based methods in medical imaging for diagnosing tendon pathologies, with the GESN offering a robust framework for enhanced data representation, diagnosis, and treatment planning.

The final study took an exploratory approach, focusing on how tendon reflex responses vary across different demographics such as age, BMI, and gender in healthy individuals. By establishing this baseline, the study provides a foundation for understanding the influence of these factors on neuromuscular behavior. The results hold potential for developing a functional profile for healthy individuals, which could be crucial in identifying abnormalities in those with pathological conditions. This, in turn, could offer deeper insights into how neuromuscular responses are altered in affected individuals, supporting more accurate diagnostics and better clinical assessments.

Collectively, the findings from these studies contribute to a deeper understanding of the complex relationships between tendon health, cartilage degeneration, the complexities of segmentation and classification in medical imaging, and neuromuscular assessment. The integration of advanced imaging techniques and computational models holds promise for revolutionizing how we diagnose and treat tendon pathologies. As we move forward, it is crucial to validate these findings in larger, diverse populations and consider multi-modal imaging approaches that combine various imaging modalities to provide a comprehensive assessment of tendon health.

In conclusion, this thesis underscores the importance of interdisciplinary research in advancing the understanding of tendon-related pathologies. By integrating medical image analysis, neuromuscular measurements, and computational models, this work has contributed to the development of innovative diagnostic tools aimed at improving patient outcomes. Although the studies primarily focus on tendon pathology, the methodologies and models developed—from tendon segmentation to pathology detection—demonstrate broader applicability. The computational techniques, particularly in medical imaging and artificial intelligence, have the potential to be adapted for a variety of other medical conditions and imaging tasks. Future research should focus on validating these models with larger, more diverse datasets and exploring multimodal approaches for more comprehensive diagnostics. Ultimately, this thesis lays the groundwork for future advancements not only in tendon pathology but also in broader medical imaging fields, offering new opportunities for early diagnosis, personalized treatment, and expanded applications in healthcare.

Bibliography

[1] Z. Khatun, M. Tsirilaki, A. Lindemann, *et al.*, "The role of muscle and tendon in predicting cartilage degeneration and tendinopathy", in *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, 2022, pp. 289–294. DOI: 10.1109/MetroXRAINE54828.2022.9967653.

[2] Z. Khatun, H. Jónsson, M. Tsirilaki, *et al.*, "Beyond pixel: Superpixel-based mri segmentation through traditional machine learning and graph convolutional network", *Computer Methods and Programs in Biomedicine*, vol. 256, p. 108 398, 2024, ISSN: 0169-2607. DOI: https://doi.org/10.1016/j.cmpb.2024.108398. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260724003912.

[3] H. L. Birch, "Tendon matrix composition and turnover in relation to functional requirements: Tendon matrix composition and turnover", *International Journal of Experimental Pathology*, vol. 88, no. 4, pp. 241–248, Aug. 2007. DOI: 10.1111/j.1365-2613.2007.00552.x.

[4] K. L. Moore, A. F. Dalley, and A. M. R. Agur, *Clinically Oriented Anatomy*, 7th. Lippincott Williams & Wilkins, 2014.

[5] M. F. Mazzone and T. McCue, "Common conditions of the achilles tendon", *American Family Physician*, vol. 65, no. 9, pp. 1805–1811, 2002.

[6] M. Wong, A. H. Jardaly, and J. Kiel, *Anatomy, bony pelvis and lower limb: Achilles tendon*, August 8, Author Information and Affiliations, 2023.

[7] T. Finni, P. V. Komi, V. Lepola, and J. Avela, "In vivo behavior of the human achilles tendon during walking and running", *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 267, no. 1459, pp. 1395–1400, 2000.

[8] G. P. Williams, "Achilles tendon lesions in sport", *Sports Medicine*, 1993, 5.

[9] "Location and distribution of non-collagenous matrix proteins in musculoskeletal tissues of rat", 12, n.d.

[10] R. F. Zernicke and B. Loitz-Ramage, "Exercise-related adaptations in connective tissue", in *Strength and Power in Sport*, P. V. Komi, Ed., Blackwell Science Ltd, 2003, pp. 96–113. DOI: 10.1002/9780470757215.ch6.

[11] G. P. Riley, V. Curry, J. DeGroot, *et al.*, "Matrix metalloproteinase activities and their relationship with collagen remodelling in tendon pathology", *Matrix Biology*, vol. 21, no. 2, pp. 185–195, Mar. 2002. DOI: 10.1016/S0945-053X(01)00196-2.

[12] R. Aicale, D. Tarantino, and N. Maffulli, "Basic science of tendons", in *Bio-Orthopaedics*, A. Gobbi, J. Espregueira-Mendes, J. G. Lane, and M. Karahan, Eds., Springer Berlin Heidelberg, 2017, pp. 249–273. DOI: 10.1007/978-3-662-54181-4-21.

[13] D. Morrissey, "Guidelines and pathways for clinical practice in tendinopathy: Their role and development", *J Orthop Sports Phys Ther*, vol. 45, no. 11, pp. 819–822, Nov. 2015. DOI: 10.2519/jospt.2015.0111.

[14] K. Lipman, C. Wang, K. Ting, C. Soo, and Z. Zheng, "Tendinopathy: Injury, repair, and current exploration", *DDDT*, vol. Volume 12, pp. 591–603, Mar. 2018. DOI: 10.2147/DDDT.S154660.

[15] N. Maffulli, P. Sharma, and K. L. Luscombe, "Achilles tendinopathy: Aetiology and management", vol. 97, p. 5, 2004.

[16] S. P. Magnusson, H. Langberg, and M. Kjaer, "The pathogenesis of tendinopathy: Balancing the response to loading", *Nature Reviews Rheumatology*, vol. 6, no. 5, pp. 262–268, May 2010. DOI: 10.1038/nrrheum.2010.43.

[17] N. Maffulli, "Overuse tendon conditions: Time to change a confusing terminology", *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, vol. 14, no. 8, pp. 840–843, Nov. 1998. DOI: 10.1016/S0749-8063(98)70021-0.

[18] R. Syha *et al.*, "Tendinopathy of the achilles tendon: Volume assessed by automated contour detection in submillimeter isotropic 3-dimensional magnetic resonance imaging data sets recorded at a field strength of 3 t", *J. Comput. Assist. Tomogr.*, vol. 39, no. 2, pp. 250–256, 2015.

[19] P. Szaro, K. Nilsson-Helander, and M. Carmont, "Mri of the achilles tendon – a comprehensive pictorial review. part two", *European Journal of Radiology Open*, vol. 8, p. 100 343, 2021. DOI: 10.1016/j.ejro.2021.100343.

[20] P. Sharma and N. Maffulli, "Tendon injury and tendinopathy: Healing and repair", vol. VOLUME, p. 16, n.d.

[21] Y. Ochen *et al.*, "Operative treatment versus nonoperative treatment of achilles tendon ruptures: Systematic review and meta-analysis", *BMJ*, k5120, Jan. 2019. DOI: 10.1136/bmj.k5120.

[22] S. J. Sorrenti, "Achilles tendon rupture: Effect of early mobilization in rehabilitation after surgical repair", *Foot Ankle Int.*, vol. 27, no. 6, pp. 407–410, Jun. 2006. DOI: 10.1177/107110070602700603.

[23] M. E. Gomes, R. L. Reis, and M. T. Rodrigues, *Tendon Regeneration: Understanding Tissue Physiology and Development to Engineer Functional Substitutes*. Amsterdam: Elsevier/Academic Press, 2015.

[24] M. M. Åström and A. M. Rausing, "Chronic achilles tendinopathy", *Clinical Orthopaedics and Related Research*, vol. 316, pp. 151–164, Jul. 1995.

[25] M. Stiskal and et al., "Magnetic resonance imaging of achilles tendon in patients with rheumatoid arthritis", *Invest. Radiol.*, vol. 32, no. 10, pp. 602–608, 1997.

[26] M. Kvist, "Achilles tendon injuries in athletes", *Sports Medicine*, vol. 18, no. 3, pp. 173–201, 1994.

[27] K. M. Khan and N. Maffulli, "Tendinopathy: An achilles' heel for athletes and clinicians", *Clin J Sport Med*, pp. 151–4, 1998.

[28] C.-F. Åström, P. Gentz, A. Nilsson, S. Rausing, S. Sjöberg, and N. Westlin, "Imaging in chronic achilles tendinopathy: A comparison of ultrasonography, magnetic resonance imaging and surgical findings in 27 histologically verified cases", *Skeletal Radiology*, vol. 25, no. 7, pp. 615–620, Oct. 1996. DOI: 10.1007/s002560050146.

[29] K. M. Khan, "Are ultrasound and magnetic resonance imaging of value in assessment of achilles tendon disorders? a two year prospective study", *British Journal of Sports Medicine*, vol. 37, no. 2, pp. 149–153, Apr. 2003. DOI: 10.1136/bjsm.37.2.149.

[30] N. M. A. Ibrahim and H. H. Elsaeed, "Lesions of the achilles tendon: Evaluation with ultrasonography and magnetic resonance imaging", *The Egyptian Journal of Radiology and Nuclear Medicine*, vol. 44, no. 3, pp. 581–587, Sep. 2013. DOI: 10.1016/j.ejrnm.2013.05.006.

[31] K. W. et al., "Diffusion tensor imaging of human achilles tendon by stimulated echo readout-segmented epi (ste-rs-epi): Wengler et al.", *Magn. Reson. Med*, vol. 80, no. 6, pp. 2464–2474, Dec. 2018. DOI: 10.1002/mrm.27220.

[32] M. T. N. et al., "T 2 relaxation reveals spatial collagen architecture in articular cartilage: A comparative quantitative mri and polarized light microscopic study: T 2 imaging of articular cartilage", *Magn. Reson. Med.*, vol. 46, no. 3, pp. 487–493, Sep. 2001. DOI: 10.1002/mrm.1218.

[33] P. Szaro, K. Nilsson-Helander, and M. Carmont, "Mri of the achilles tendon—a comprehensive pictorial review. part one", *European Journal of Radiology Open*, vol. 8, p. 100 342, 2021. DOI: 10.1016/j.ejro.2021.100342.

[34] A. Shalabi, "Magnetic resonance imaging in chronic achilles tendinopathy", *Acta Radiol*, vol. 45, no. suppl-432, pp. 1–45, Sep. 2004. DOI: 10.1080/03655950410009742.

[35] W. M. V. Chan and E. Kavanagh, "Postoperative infection in the foot and ankle", *Semin Musculoskelet Radiol*, vol. 16, no. 03, pp. 241–253, Jul. 2012. DOI: 10.1055/s-0032-1320124.

[36] DeepAI. "Feature extraction". (), [Online]. Available: https://deepai.org/machine-learning-glossary-and-terms/feature-extraction.

[37] J. van Timmeren, D. Cester, and S. Tanadini-Lang, "Radiomics in medical imaging—"how-to" guide and critical reflection", *Insights Imaging*, vol. 11, p. 91, 2020. DOI: 10.1186/s13244-020-00887-2.

[38] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images are more than pictures, they are data", *Radiology*, vol. 278, pp. 563–577, 2016.

[39] H. Völzke, C. O. Schmidt, and K. Hegenscheid, "Population imaging as a valuable tool for personalized medicine", *Clinical Pharmacology & Therapeutics*, vol. 92, pp. 422–424, 2012.

[40] C. L. Schlett, T. Hendel, and S. Weckbach, "Population-based imaging and radiomics: Rationale and perspective of the german national cohort mri study", *Rofo*, vol. 188, pp. 652–661, 2016.

[41] M. Hajek, M. Dezortova, A. Materka, and R. Lerski, *Texture Analysis for Magnetic Resonance Imaging*. 2006.

[42] A. Zwanenburg, S. Leger, M. Vallieres, and S. Lock, "Image biomarker standardisation initiative", 2016. [Online]. Available: https://arxiv.org/abs/1612.07003.

[43] R. M. Haralick and K. Shanmugam, "Textural features for image classification", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 610–621, 1973.

[44] A. Zwanenburg, S. Leger, M. Vallieres, and S. Lock. "Image biomarker standardisation initiative". (2016), [Online]. Available: https://arxiv.org/abs/1612.07003.

[45] M. M. Galloway, "Texture analysis using gray-level run lengths", *Computer Graphics and Image Processing*, vol. 4, no. 2, pp. 172–179, 1975.

[46] G. Thibault, J. Angulo, and F. Meyer, "Advanced statistical matrices for texture characterization: Application to cell classification", *IEEE Transactions on Biomedical Engineering*, vol. 61, pp. 630–637, 2014.

[47] M. Amadasun and R. King, "Textural features corresponding to textural properties", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, pp. 1264–1274, 1989.

[48] O. Cosido, A. Iglesias, A. Galvez, *et al.*, "Hybridization of convergent photogrammetry, computer vision, and artificial intelligence for digital documentation of cultural heritage - a case study: The magdalena palace", in *2014 International Conference on Cyberworlds (CW)*, IEEE, 2014, pp. 369–376. DOI: 10.1109/CW.2014.58.

[49] S. Ghosh, N. Das, I. Das, and U. Maulik, "Understanding deep learning techniques for image segmentation", Year. DOI: 10.1145/3329784.

[50] K. Balaji and K. Lavanya, *Chapter 5 - Medical Image Analysis With Deep Neural Networks*. 2019. [Online]. Available: https://doi.org/10.1016/B978-0-12-816718-2.00012-9.

[51] A. Scott and et al., "Sports and exercise-related tendinopathies: A review of selected topical issues by participants of the second international scientific tendinopathy symposium (ists) vancouver 2012", *Br J Sports Med*, vol. 47, no. 9, pp. 536–544, Jun. 2013. DOI: 10.1136/bjsports-2013-092329.

[52] B. Bley and W. Abid, "Imaging of tendinopathy: A physician's perspective", *J Orthop Sports Phys Ther*, vol. 45, no. 11, pp. 826–828, Nov. 2015. DOI: 10.2519/jospt.2015.0113.

[53] J. L. Cook, K. M. Khan, P. R. Harcourt, M. Grant, D. A. Young, and S. F. Bonar, "A cross sectional study of 100 athletes with jumper's knee managed conservatively and surgically. the victorian institute of sport tendon study group.", *British Journal of Sports Medicine*, vol. 31, no. 4, pp. 332–336, Dec. 1997. DOI: 10.1136/bjsm.31.4.332.

[54] P. Szaro, G. Witkowski, R. Smigielski, P. Krajewski, and B. Ciszek, "Fascicles of the adult human achilles tendon - an anatomical study", *Annals of anatomy = Anatomischer Anzeiger : official organ of the Anatomische Gesellschaft*, vol. 191, no. 6, pp. 586–593, 2009. DOI: 10.1016/j.aanat.2009.07.006.

[55] P. Kannus and L. Józsa, "Histopathological changes preceding spontaneous rupture of a tendon. a controlled study of 891 patients", *J. Bone Joint Surg. Am.*, vol. 73, no. 10, pp. 1507–1525, 1991.

[56] G. P. Riley, V. Curry, J. DeGroot, *et al.*, "Matrix metalloproteinase activities and their relationship with collagen remodeling in tendon pathology", *Matrix Biology*, vol. 21, no. 2, pp. 185–95, 2002. DOI: 10.1016/S0945-053X(01)00196-2.

[57] J. L. Cook and C. R. Purdam, "Is tendon pathology a continuum? a pathology model to explain the clinical presentation of load-induced tendinopathy", *British Journal of Sports Medicine*, vol. 43, no. 6, pp. 409–416, Jun. 2009. DOI: 10.1136/bjsm.2008.051193.

[58] X. Wei and R. Sui, "A review of machine learning algorithms for retinal cyst segmentation on optical coherence tomography", *Sensors*, vol. 23, no. 6, p. 3144, 2023. DOI: 10.3390/s23063144. [Online]. Available: https://doi.org/10.3390/s23063144.

[59] T. Cover and P. Hart, "Nearest neighbor pattern classification", *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967. DOI: 10.1109/TIT.1967.1053964.

[60] S. Bernhard and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, 1st. MIT Press, 2001.

[61] T. K. Ho, "A data complexity analysis of comparative advantages of decision forest constructors", *Pattern Analysis & Applications*, vol. 5, no. 2, pp. 102–112, Jun. 2002. DOI: 10.1007/s100440200008.

[62] J. Neter, W. Wasserman, and M. H. Kutner, *Applied linear regression models*. Irwin, 1989.

[63] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004. DOI: 10.1109/TPAMI.2004.60.

[64] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using k-means clustering algorithm and subtractive clustering algorithm", *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.

[65] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images", in *Readings in computer vision*, M. A. Fischler and O. Firschein, Eds., Elsevier, 1987, pp. 564–584.

[66] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, R. Kikinis, and H. W. Muller-Gartner, "Markov random field segmentation of brain mr images", *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 878–886, 1997, [PubMed: 9533587]. DOI: 10.1109/42.650886.

[67] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biol. Cybern.*, vol. 36, no. 4, pp. 193–202, 1980.

[68] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang, "Phoneme recognition using time-delay neural networks", *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 3, pp. 328–339, Mar. 1989.

[69] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning", *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. DOI: 10.1038/nature14539. [Online]. Available: https://doi.org/10.1038/nature14539.

[70] UpGrad, *Basic cnn architecture*, Accessed: 2024-09-16, n.d. [Online]. Available: https://www.upgrad.com/blog/basic-cnn-architecture/.

[71] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", in *Proc. 25th Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[72] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", *arXiv preprint arXiv:1409.1556*, 2014.

[73] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[74] D. E. R. et al., "Learning representations by back-propagating errors", *Cogn. Model.*, vol. 5, no. 3, 1988.

[75] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[76] J. Zhu, Z. Yang, M. Mourshed, *et al.*, "Electric vehicle charging load forecasting: A comparative study of deep learning approaches", *Energies*, vol. 12, p. 2692, Jul. 2019. DOI: 10.3390/en12142692.

[77] H. Jiang, Y. Li, C. Zhou, H. Hong, T. Glade, and K. Yin, "Landslide displacement prediction combining lstm and svr algorithms: A case study of shengjibao landslide from the three gorges reservoir area", *Applied Sciences*, vol. 10, p. 7830, Nov. 2020. DOI: 10.3390/app10217830.

[78] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, 2017.

[79] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, *Image segmentation using deep learning: A survey*, Jan. 2020. DOI: 10.48550/arXiv.2001.05566.

[80] L. Cai, Y. Chen, N. Cai, W. Cheng, and H. Wang, "Utilizing amari-alpha divergence to stabilize the training of generative adversarial networks", *Entropy*, vol. 22, p. 410, Apr. 2020. DOI: 10.3390/e22040410.

[81] I. G. et al., "Generative adversarial nets", in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.

[82] Docsumo, *Image segmentation: An introduction and guide*, Accessed: 2024-09-16, 2023. [Online]. Available: https://www.docsumo.com/blogs/data-extraction/image-segmentation.

[83] H. Seo, M. Badiei Khuzani, V. Vasudevan, *et al.*, "Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications", *Medical Physics*, vol. 47, no. 5, e148–e167, 2020. DOI: 10.1002/mp.13649.

[84] K. Z. Mao, P. Zhao, and P.-H. Tan, "Supervised learning-based cell image segmentation for p53 immunohistochemistry", *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1153–1163, 2006, [PubMed: 16761842]. DOI: 10.1109/TBME.2006.873706.

[85] C. Wachinger and P. Golland, "Atlas-based under-segmentation", in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2014.

[86] D. Li, L. Liu, J. Chen, *et al.*, "Augmenting atlas-based liver segmentation for radiotherapy treatment planning by incorporating image features proximal to the atlas contours", *Physics in Medicine & Biology*, vol. 62, no. 1, p. 272, 2016, [PubMed: 27991439]. DOI: 10.1088/1361-6560/62/1/272.

[87] R. Nock and F. Nielsen, "Statistical region merging", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1452–1458, 2004.

[88] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-9, no. 1, pp. 62–66, 1979. DOI: 10.1109/TSMC.1979.4310076.

[89] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models", *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[90] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[91] N. Plath, M. Toussaint, and S. Nakajima, "Multi-class image segmentation using conditional random fields and global classification", in *Proceedings of the 26th International Conference on Machine Learning*, ACM, 2009, pp. 817–824.

[92] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, R. Kikinis, and H. W. Müller-Gärtner, "Markov random field segmentation of brain mr images", *IEEE Transactions on Medical Imaging*, vol. 16, no. 6, pp. 878–886, 1997, [PubMed: 9533587].

[93] J. E. van Engelen and H. H. Hoos, "A survey on semi-supervised learning", English, *Machine Learning*, vol. 109, no. 2, pp. 373–440, Feb. 2020. [Online]. Available: https://www.proquest.com/scholarly-journals/survey-on-semi-supervised-learning/docview/2367633899/se-2.

[94] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523–3542, 2022. DOI: 10.1109/TPAMI.2021.3059968.

[95] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.

[96] A. G. Schwing and R. Urtasun, "Fully connected deep structured networks", *arXiv preprint*, 2015. eprint: arXiv:1503.02351.

[97] S. Zheng and et al., "Conditional random fields as recurrent neural networks", in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1529–1537.

[98] G. Lin, C. Shen, A. Van Den Hengel, and I. Reid, "Efficient piecewise training of deep structured models for semantic segmentation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 3194–3203.

[99] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation", in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1520–1528.

[100] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[101] J. Fu, J. Liu, Y. Wang, J. Zhou, C. Wang, and H. Lu, "Stacked deconvolutional network for semantic segmentation", *IEEE Transactions on Image Processing*, 2019. DOI: 10.1109/TIP.2019.2895460.

[102] A. Chaurasia and E. Culurciello, "Linknet: Exploiting encoder representations for efficient semantic segmentation", in *Proceedings of the IEEE International Conference on Visual Communications and Image Processing*, pp. 1–4.

[103] X. Xia and B. Kulis, "W-net: A deep model for fully unsupervised image segmentation", *arXiv preprint arXiv:1711.08506*, 2017.

[104] Y. Cheng, R. Cai, Z. Li, X. Zhao, and K. Huang, "Locality-sensitive deconvolution networks with gated fusion for rgb-d indoor semantic segmentation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3029–3037.

[105] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234–241.

[106] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation", in *Proceedings of the 4th International Conference on 3D Vision (3DV)*, 2016, pp. 565–571.

[107] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2117–2125.

[108] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2881–2890.

[109] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks", in *Proc. 28th Int. Conf. Neural Inf. Process. Syst.*, 2015, pp. 91–99.

[110] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs", *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 40, no. 4, pp. 834–848, 2018.

[111] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions", 2015, arXiv:1511.07122.

[112] P. Wang and et al., "Understanding convolution for semantic segmentation", in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2018, pp. 1451–1460.

[113] M. Yang, K. Yu, C. Zhang, Z. Li, and K. Yang, "Denseaspp for semantic segmentation in street scenes", in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3684–3692.

[114] A. Paszke, A. Chaurasia, S. Kim, and E. Culurciello, "Enet: A deep neural network architecture for real-time semantic segmentation", arXiv:1606.02147.

[115] F. Visin, K. Kastner, K. Cho, A. Courville, and Y. Bengio, "Reseg: A recurrent neural network-based model for semantic segmentation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2016, pp. 41–48.

[116] L.-C. Chen, Y. Yang, J. Wang, W. Xu, and A. L. Yuille, "Attention to scale: Scale-aware semantic image segmentation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 3640–3649.

[117] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, "Semantic segmentation using adversarial networks", *arXiv preprint arXiv:1611.08408*, 2016.

[118] N. Souly, C. Spampinato, and M. Shah, "Semi-supervised semantic segmentation using generative adversarial network", in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 5688–5696.

[119] R. Gupta and et al., "Curvelet based automatic segmentation of supraspinatus tendon from ultrasound image: A focused assistive diagnostic method", *BioMed Eng OnLine*, vol. 13, no. 1, p. 157, Dec. 2014. DOI: 10.1186/1475-925X-13-157.

[120] B.-I. Chuang and et al., "A medical imaging analysis system for trigger finger using an adaptive texture-based active shape model (atasm) in ultrasound images", p. 21, DOI: 10.1371/journal.pone.0187042.

[121] Z. Zhou, G. Zhao, R. Kijowski, and F. Liu, "Deep convolutional neural network for segmentation of knee joint anatomy: Zhou et al.", *Magn. Reson. Med*, vol. 80, no. 6, pp. 2759–2770, Dec. 2018. DOI: 10.1002/mrm.27229.

[122] K. Kamnitsas and et al., "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation", *Medical Image Analysis*, vol. 36, pp. 61–78, Feb. 2017. DOI: 10.1016/j.media.2016.10.004.

[123] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation", 2016, arXiv, Jun. 15, 2016. Accessed: Sep. 19, 2022. [Online]. Available: http://arxiv.org/abs/1606.04797.

[124] C.-P. Kuok and et al., "Segmentation of finger tendon and synovial sheath in ultrasound image using deep convolutional neural network", *BioMed Eng OnLine*, vol. 19, no. 1, p. 24, Dec. 2020. DOI: 10.1186/s12938-020-00768-1.

[125] N. Martins, S. Sultan, D. Veiga, M. Ferreira, F. Teixeira, and M. Coimbra, "A new active contours approach for finger extensor tendon segmentation in ultrasound images using prior knowledge and phase symmetry", *IEEE J. Biomed. Health Inform.*, vol. 22, no. 4, pp. 1261–1268, Jul. 2018. DOI: 10.1109/JBHI.2017.2723819.

[126] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions", 2016, arXiv, Apr. 30, 2016. Accessed: Sep. 14, 2022. [Online]. Available: http://arxiv.org/abs/1511.07122.

[127] C.-Y. Lee, S. Xie, P. Gallagher, Z. Zhang, and Z. Tu, "Deeply-supervised nets", 2014, arXiv, Sep. 25, 2014. Accessed: Sep. 14, 2022. [Online]. Available: http://arxiv.org/abs/1409.5185.

[128] G.-Q. Zhou and et al., "A single-shot region-adaptive network for myotendinous junction segmentation in muscular ultrasound images", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 67, no. 12, pp. 2531–2542, Dec. 2020. DOI: 10.1109/TUFFC.2020.2979481.

[129] V. M. Zatsiorsky and B. I. Prilutsky, *Biomechanics of skeletal muscles*. Champaign, IL: Human Kinetics, 2012.

[130] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn", 2018, arXiv, Jan. 24, 2018. Accessed: Sep. 14, 2022. [Online]. Available: http://arxiv.org/abs/1703.06870.

[131] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation", 2016, arXiv, Jun. 15, 2016. Accessed: Sep. 14, 2022. [Online]. Available: http://arxiv.org/abs/1606.04797.

[132] J. Yao and et al., "Evaluation of a deep learning method for the automated detection of supraspinatus tears on mri", *Skeletal Radiol*, vol. 51, no. 9, pp. 1765–1775, Sep. 2022. DOI: 10.1007/s00256-022-04008-6.

[133] A. H. Ramakrishnan, M. Rajappa, K. Krithivasan, P. E. Chatzistergos, N. Chockalingam, and M. R. Nalluri, "A concept for movement-based computerized segmentation of connective tissue in ultrasound imaging", *Multimed Tools Appl*, Apr. 2022. DOI: 10.1007/s11042-022-12932-x.

[134] R. J. Hodgson, P. J. O'Connor, and A. J. Grainger, "Tendon and ligament imaging", *The British Journal of Radiology*, p. 16, 2012.

[135] S. Liu, "Deep learning in medical ultrasound analysis: A review", p. 15, 2019.

[136] X. Liao, J. Yin, M. Chen, and Z. Qin, "Adaptive payload distribution in multiple images steganography based on image texture features", *IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING*, p. 14, 2020.

[137] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape", p. 21,

[138] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[139] Z. Li and J. Chen, "Superpixel segmentation using linear spectral clustering", in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1356–1363.

[140] L. Xu, L. Zeng, and Z. Wang, "Saliency-based superpixels", *Signal, Image and Video Processing*, vol. 8, no. 1, pp. 181–190, 2014.

[141] H. Fu, X. Cao, D. Tang, Y. Han, and D. Xu, "Regularity preserved superpixels and supervoxels", *IEEE Transactions on Multimedia*, vol. 16, no. 4, pp. 1165–1175, 2014.

[142] J. Shen, Y. Du, W. Wang, and X. Li, "Lazy random walks for superpixel segmentation", *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1451–1462, 2014.

[143] C. Conrad, M. Mertz, and R. Mester, "Contour-relaxed superpixels", in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, 2013, pp. 280–293.

[144] P. Wang, G. Zeng, R. Gan, J. Wang, and H. Zha, "Structure-sensitive superpixels via geodesic distance", *International Journal of Computer Vision*, vol. 103, no. 1, pp. 1–21, 2013.

[145] D. Giordano, F. Murabito, S. Palazzo, and C. Spampinato, "Superpixel-based video object segmentation using perceptual organization and location prior", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4814–4822.

[146] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 898–916, 2011. DOI: 10.1109/TPAMI.2010.161.

[147] F. Yang, H. Lu, and M.-H. Yang, "Robust superpixel tracking", *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1639–1651, 2014. DOI: 10.1109/TIP.2014.2300823.

[148] Z. Tian, N. Zheng, J. Xue, X. Lan, C. Li, and G. Zhou, "Video object segmentation with shape cue based on spatiotemporal superpixel neighborhood", *IET Computer Vision*, vol. 8, no. 1, pp. 16–25, 2014. DOI: 10.1049/iet-cvi.2012.0189.

[149] J. Cheng, J. Liu, Y. Xu, *et al.*, "Superpixel classification for initialization in model-based optic disc segmentation", in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. DOI: 10.1111/j.1365-2613.2007.00552.x.

[150] J. Cheng, J. Liu, Y. Xu, *et al.*, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening", *IEEE Transactions on Medical Imaging*, vol. 32, no. 6, pp. 1019–1032, 2013. DOI: 10.1111/j.1365-2613.2007.00552.x.

[151] Ren and Malik, "Learning a classification model for segmentation", in *Proceedings Ninth IEEE International Conference on Computer Vision*, 2003, 10–17 vol.1. DOI: 10.1109/ICCV.2003.1238308.

[152] T. Gaber, A. Tharwat, A. Ibrahim, V. Snael, and A. E. Hassanien, "Human thermal face recognition based on random linear oracle (rlo) ensembles", in *2015 International Conference on Intelligent Networking and Collaborative Systems*, 2015, pp. 91–98. DOI: 10.1109/INCoS.2015.67.

[153] S. Kaur and R. K. Bansal, "Comparative analysis of superpixel segmentation methods", *International Journal of Engineering, Technology, Management, and Research*, vol. 5, no. 3, pp. 1–9, 2020. DOI: 10.29121/ijetmr.v5.i3.2018.172.

[154] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation", *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004. DOI: 10.1023/B:VISI.0000022288.19776.77.

[155] M.-Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation", in *CVPR 2011*, 2011, pp. 2097–2104. DOI: 10.1109/CVPR.2011.5995323.

[156] D. Hutchison *et al.*, "Superpixels and supervoxels in an energy optimization framework", in *Computer Vision – ECCV 2010*, vol. 6315, 2010, pp. 211–224. DOI: 10.1007/978-3-642-15555-0-16.

[157] P. Carr and R. Hartley, "Minimizing energy functions on 4-connected lattices using elimination", in *2009 IEEE 12th International Conference on Computer Vision*, 2009, pp. 2042–2049. DOI: 10.1109/ICCV.2009.5459450.

[158] J. Shen, Y. Du, W. Wang, and X. Li, "Lazy random walks for superpixel segmentation", *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1451–1462, 2014. DOI: 10.1109/TIP.2014.2302892.

[159] L. Cong, S. Ding, L. Wang, A. Zhang, and W. Jia, "Image segmentation algorithm based on superpixel clustering", *IET Image Processing*, vol. 12, no. 11, pp. 2030–2035, 2018. DOI: 10.1049/iet-ipr.2018.5439.

[160] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking", in *Computer Vision – ECCV 2008*, vol. 5305, 2008, pp. 705–718. DOI: 10.1007/978-3-540-88693-8-52.

[161] A. Ibrahim, S. Tominaga, and T. Horiuchi, "A spectral invariant representation of spectral reflectance", *Optical Review*, vol. 18, no. 2, pp. 231–236, 2011. DOI: 10.1007/s10043-011-0047-6.

[162] F. Meyer, *The watershed concept and its use in segmentation: A brief history*, arXiv, Accessed: Oct. 04, 2022. [Online]. Available: http://arxiv.org/abs/1202.0216, Feb. 2012.

[163] P. Neubert and P. Protzel, "Compact watershed and preemptive slic: On improving trade-offs of superpixel segmentation algorithms", in *2014 22nd International Conference on Pattern Recognition*, 2014, pp. 996–1001. DOI: 10.1109/ICPR.2014.181.

[164] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, 1991. DOI: 10.1109/34.87344.

[165] B. Gaonkar *et al.*, "Multi-parameter ensemble learning for automated vertebral body segmentation in heterogeneously acquired clinical mr images", *IEEE J. Transl. Eng. Health Med.*, vol. 5, pp. 1–12, 2017. DOI: 10.1109/JTEHM.2017.2717982.

[166] P.-L. Bazin *et al.*, "Direct segmentation of the major white matter tracts in diffusion tensor images", *NeuroImage*, vol. 58, no. 2, pp. 458–468, Sep. 2011. DOI: 10.1016/j.neuroimage.2011.06.020.

[167] A. Signoroni *et al.*, "Bs-net: Learning covid-19 pneumonia severity on a large chest x-ray dataset", *Medical Image Analysis*, vol. 71, p. 102 046, Jul. 2021. DOI: 10.1016/j.media.2021.102046.

[168] J. Xu *et al.*, "Three-dimensional spectral-domain optical coherence tomography data analysis for glaucoma detection", *PLoS ONE*, vol. 8, no. 2, e55476, 2013. DOI: 10.1371/journal.pone.0055476.

[169] J. Shi and J. Malik, "Normalized cuts and image segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000. DOI: 10.1109/34.868688.

[170] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: A statistical view of boosting", p. 71,

[171] A. D. Belsare, M. M. Mushrif, M. A. Pangarkar, and N. Meshram, "Breast histopathology image segmentation using spatio-colour-texture based graph partition method: Breast histopathology image segmentation", *Journal of Microscopy*, 2016. DOI: 10.1111/jmi.12361.

[172] H. Zhu *et al.*, "A novel lung cancer detection algorithm for cads based on ssp and level set", *THC*, vol. 25, pp. 345–355, Jul. 2017. DOI: 10.3233/THC-171338.

[173] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation", *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, Sep. 2004. DOI: 10.1023/B:VISI.0000022288.19776.77.

[174] C. Chevrefils *et al.*, "Texture analysis for automatic segmentation of intervertebral disks of scoliotic spines from mr images", *IEEE Trans. Inform. Technol. Biomed.*, vol. 13, no. 4, pp. 608–620, Jul. 2009. DOI: 10.1109/TITB.2009.2018286.

[175] C. Chevrefils *et al.*, "Watershed segmentation of intervertebral disk and spinal canal from mri images", in *Image Analysis and Recognition*, vol. 4633, 2007, pp. 1017–1027. DOI: 10.1007/978-3-540-74260-9-90.

[176] L. Grady, "Random walks for image segmentation", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006. DOI: 10.1109/TPAMI.2006.233.

[177] G. Yang *et al.*, "Fully automatic segmentation and objective assessment of atrial scars for long-standing persistent atrial fibrillation patients using late gadolinium-enhanced mri", *Med. Phys.*, vol. 45, no. 4, pp. 1562–1576, Apr. 2018. DOI: 10.1002/mp.12832.

[178] S. N. Kumar *et al.*, "Suspicious lesion segmentation on brain, mammograms and breast mr images using new optimized spatial feature based superpixel fuzzy c-means clustering", *J Digit Imaging*, vol. 32, no. 2, pp. 322–335, Apr. 2019. DOI: 10.1007/s10278-018-0149-9.

[179] Y. Mu *et al.*, "White matter segmentation algorithm for dti images based on superpixel full convolutional network", *J Med Syst*, vol. 43, no. 9, p. 303, Sep. 2019. DOI: 10.1007/s10916-019-1431-1.

[180] N. B. Prakash *et al.*, "Deep transfer learning for covid-19 detection and infection localization with superpixel based segmentation", *Sustainable Cities and Society*, vol. 75, p. 103 252, Dec. 2021. DOI: 10.1016/j.scs.2021.103252.

[181] S. G. Sundaram *et al.*, "Deep transfer learning based unified framework for covid19 classification and infection detection from chest x-ray images", *Arab J Sci Eng*, vol. 47, no. 2, pp. 1675–1692, Feb. 2022. DOI: 10.1007/s13369-021-05958-0.

[182] B. Li *et al.*, "Fast segmentation of vertebrae ct image based on the snic algorithm", *Tomography*, vol. 8, no. 1, pp. 59–76, Jan. 2022. DOI: 10.3390/tomography8010006.

[183] S. Li *et al.*, "Automatic dental plaque segmentation based on local-to-global features fused self-attention network", *IEEE J. Biomed. Health Inform.*, vol. 26, no. 5, pp. 2240–2251, May 2022. DOI: 10.1109/JBHI.2022.3141773.

[184] L. Backstrom and J. Leskovec, "Supervised random walks: Predicting and recommending links in social networks", in *Proceedings of the fourth ACM international conference on Web search and data mining - WSDM '11*, Hong Kong, China, 2011, p. 635. DOI: 10.1145/1935826.1935914.

[185] L. Akoglu, H. Tong, and D. Koutra, "Graph based anomaly detection and description: A survey", *Data Mining and Knowledge Discovery*, vol. 29, no. 3, pp. 626–688, May 2015. DOI: 10.1007/s10618-014-0365-y.

[186] "Proceedings of the 2017 siam international conference on data mining", N. Chawla and W. Wang, Eds., Philadelphia, PA: Society for Industrial and Applied Mathematics, 2017. DOI: 10.1137/1.9781611974973.

[187] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein, "Geometric deep learning on graphs and manifolds using mixture model cnns", in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, Jul. 2017, pp. 5425–5434. DOI: 10.1109/CVPR.2017.576.

[188] S. Zhang, H. Tong, J. Xu, and R. Maciejewski, "Graph convolutional networks: A comprehensive review", *Computational and Structural Biotechnology Journal*, vol. 6, no. 1, p. 11, Dec. 2019. DOI: 10.1186/s40649-019-0069-y.

[189] W. L. Hamilton, *Graph representation learning* (Synthesis Lectures on Artificial Intelligence and Machine Learning 3). 2020, vol. 14, pp. 1–159.

[190] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks", in *ICLR*, 2016.

[191] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks", in *arXiv*, Accessed: Oct. 09, 2022, Feb. 2017. [Online]. Available: http://arxiv.org/abs/1609.02907.

[192] X. Zhu, Z. Ghahramani, and J. Lafferty, "Semi-supervised learning using gaussian fields and harmonic functions", p. 8.

[193] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory", *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, Mar. 2011. DOI: 10.1016/j.acha.2010.04.005.

[194] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs", in *arXiv*, Accessed: Oct. 10, 2022, May 2014. [Online]. Available: http://arxiv.org/abs/1312.6203.

[195] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering", p. 9.

[196] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein, "Geometric deep learning on graphs and manifolds using mixture model cnns", in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, Jul. 2017, pp. 5425–5434. DOI: 10.1109/CVPR.2017.576.

[197] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs", in *NeurIPS*, 2017.

[198] M. McPherson, L. Smith-Lovin, and J. Cook, "Birds of a feather: Homophily in social networks", *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.

[199] D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency", in *NeurIPS*, 2004.

[200] S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds., *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II*. Springer International Publishing, 2016, vol. 9901. DOI: 10.1007/978-3-319-46723-8.

[201] S. Xie and Z. Tu, "Holistically-nested edge detection", 2015.

[202] D. Močnik *et al.*, "Segmentation of parotid glands from registered ct and mr images", *Physica Medica*, vol. 52, pp. 33–41, Aug. 2018. DOI: 10.1016/j.ejmp.2018.06.012.

[203] Z. Tian *et al.*, "Graph-convolutional-network-based interactive prostate segmentation in mr images", *Med. Phys.*, vol. 47, no. 9, pp. 4164–4176, Sep. 2020. DOI: 10.1002/mp.14327.

[204] S. Pang *et al.*, "Spineparsenet: Spine parsing for volumetric mr image by a two-stage segmentation framework with semantic image representation", *IEEE Trans. Med. Imaging*, vol. 40, no. 1, pp. 262–273, Jan. 2021. DOI: 10.1109/TMI.2020.3025087.

[205] K. M. Eschenburg, T. J. Grabowski, and D. R. Haynor, "Learning cortical parcellations using graph neural networks", *Front. Neurosci.*, vol. 15, p. 797 500, Dec. 2021. DOI: 10.3389/fnins.2021.797500.

[206] M. F. Glasser *et al.*, "A multi-modal parcellation of human cerebral cortex", *Nature*, vol. 536, no. 7615, pp. 171–178, Aug. 2016. DOI: 10.1038/nature18933.

[207] E. M. Gordon *et al.*, "Precision functional mapping of individual human brains", *Neuron*, vol. 95, no. 4, 791–807.e7, Aug. 2017. DOI: 10.1016/j.neuron.2017.07.011.

[208] G. Wang *et al.*, "Improving graph attention networks with large margin-based constraints", [Online]. Available: http://arxiv.org/abs/1910.11945.

[209] A. Suinesiaputra *et al.*, "Statistical shape modeling of the left ventricle: Myocardial infarct classification challenge", *IEEE J. Biomed. Health Inform.*, vol. 22, no. 2, pp. 503–515, Mar. 2018. DOI: 10.1109/JBHI.2017.2652449.

[210] X. Chen *et al.*, "Shape registration with learned deformations for 3d shape reconstruction from sparse and incomplete point clouds", *Medical Image Analysis*, vol. 74, p. 102 228, Dec. 2021. DOI: 10.1016/j.media.2021.102228.

[211] J. Sykes, "Reflections on the current status of commercial automated segmentation systems in clinical practice", *J Med Radiat Sci*, vol. 61, no. 3, pp. 131–134, Sep. 2014. DOI: 10.1002/jmrs.65.

[212] Q. Ma *et al.*, "Dgrunit: Dual graph reasoning unit for brain tumor segmentation", *Computers in Biology and Medicine*, vol. 149, p. 106 079, Oct. 2022. DOI: 10.1016/j.compbiomed.2022.106079.

[213] S. Wan *et al.*, "Multi-scale dynamic graph convolutional network for hyperspectral image classification", [Online]. Available: http://arxiv.org/abs/1905.06133.

[214] R. Achanta *et al.*, "Slic superpixels compared to state-of-the-art superpixel methods", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012. DOI: 10.1109/TPAMI.2012.120.

[215] Docsumo, *Image classification: A comprehensive guide*, Accessed: 2024-09-16, 2024. [Online]. Available: https://www.docsumo.com/blogs/data-extraction/image-classification.

[216] Y. Su, Y. Wang, J. Jiao, and Y. Guo, "Automatic detection and classification of breast tumors in ultrasonic images using texture and morphological features", *The open medical informatics journal*, vol. 5, no. 1, 2011.

[217] J. Xu, Y. Wei, A. Wang, H. Zhao, and D. Lefloch, "Analysis of clothing image classification models: A comparison study between traditional machine learning and deep learning models", *Fibres & Textiles in Eastern Europe*, vol. 30, no. 5, pp. 66–78, 2022. DOI: doi:10.2478/ftee-2022-0046. [Online]. Available: https://doi.org/10.2478/ftee-2022-0046.

[218] A. Şengür, Y. Akbulut, Ü. Budak, and Z. Cömert, "White blood cell classification based on shape and deep features", in *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)*, 2019, pp. 1–4. DOI: 10.1109/IDAP.2019.8875945.

[219] M. Sajjad, S. Khan, Z. Jan, *et al.*, "Leukocytes classification and segmentation in microscopic blood smear: A resource-aware healthcare service in smart cities", *IEEE Access*, vol. 5, pp. 3475–3489, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:11853371.

[220] S. Tantikitti, S. Tumswadi, and W. Premchaiswadi, "Image processing for detection of dengue virus based on wbc classification and decision tree", in *2015 13th International Conference on ICT and Knowledge Engineering (ICT & Knowledge Engineering 2015)*, 2015, pp. 84–89. DOI: 10.1109/ICTKE.2015.7368476.

[221] K. W. Wan, C. H. Wong, H. F. Ip, *et al.*, "Evaluation of the performance of traditional machine learning algorithms, convolutional neural network and automl vision in ultrasound breast lesions classification: A comparative study", *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 4, pp. 1381–1393, 2021. DOI: 10.21037/qims-20-922.

[222] V. K. Shrivastava and M. K. Pradhan, "Rice plant disease classification using color features: A machine learning paradigm", *Journal of Plant Pathology*, vol. 103, no. 1, pp. 17–26, 2021.

[223] S. B. Jadhav, V. R. Udupi, and S. B. Patil, "Identification of plant diseases using convolutional neural networks", *International Journal of Information Technology*, vol. 13, no. 6, pp. 2461–2470, 2021.

[224] K. Aurangzeb, F. Akmal, M. A. Khan, M. Sharif, and M. Y. Javed, "Advanced machine learning algorithm based system for crops leaf diseases recognition", in *2020 6th Conference on Data Science and Machine Learning Applications (CDMA)*, IEEE, 2020, pp. 146–151.

[225] E. Hossain, M. F. Hossain, and M. A. Rahaman, "A color and texture based approach for the detection and classification of plant leaf disease using knn classifier", in *2019 international conference on electrical, computer and communication engineering (ECCE)*, IEEE, 2019, pp. 1–6.

[226] M. Sharif, M. A. Khan, Z. Iqbal, M. F. Azam, M. I. U. Lali, and M. Y. Javed, "Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection", *Computers and electronics in agriculture*, vol. 150, pp. 220–234, 2018.

[227] N. Sharma, V. Jain, and A. Mishra, "An analysis of convolutional neural networks for image classification", *Procedia Computer Science*, vol. 132, pp. 377–384, 2018, International Conference on Computational Intelligence and Data Science, ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2018.05.198. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050918309335.

[228] D. Han, Q. Liu, and W. Fan, "A new image classification method using cnn transfer learning and web data augmentation", *Expert Systems with Applications*, vol. 95, pp. 43–56, 2018, ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2017.11.028. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417417307844.

[229] M. Hussain, J. Bird, and D. Faria, "A study on cnn transfer learning for image classification", Jun. 2018.

[230] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "High-resolution image classification with convolutional networks", in *Geoscience and Remote Sensing Symposium (IGARSS), 2017 IEEE International*, IEEE, 2017, pp. 5157–5160.

[231] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network", in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*, 2014, pp. 844–848. DOI: 10.1109/ICARCV.2014.7064414.

[232] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, "Joint optic disc and cup segmentation based on multi-label deep network and polar transformation", *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1597–1605, 2018. DOI: 10.1109/TMI.2018.2791488.

[233] S. Chen, Y. Liu, D. Liu, and C. Su, "Alexnet model and adaptive contrast enhancement based ultrasound imaging classification", *Computer Science*, vol. 46, no. 6A, pp. 146–152, 2019.

[234] S. Shalev-Shwartz and T. Zhang, "Stochastic dual coordinate ascent methods for regularized loss minimization.", *Journal of Machine Learning Research*, vol. 14, no. 1, 2013.

[235] C. F. Baumgartner, K. Kamnitsas, J. Matthew, S. Smith, B. Kainz, and D. Rueckert, "Real-time standard scan plane detection and localisation in fetal ultrasound using fully convolutional neural networks", in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, Springer, 2016, pp. 203–211.

[236] Y. Zhang, Y. Ye, and D. Wang, "Clinical application of image processing and neural network in cytopathological diagnosis of lung cancer", *Chinese Journal of Thoracic and Cardiovascular Surgery*, no. 04, 2005.

[237] T. Qaiser, A. Mukherjee, C. P. Reddy, *et al.*, "Her2 challenge contest: A detailed assessment of automated her2 scoring algorithms in whole slide images of breast cancer tissues", *Histopathology*, vol. 72, no. 2, pp. 227–238, 2018. DOI: 10.1111/his.13333.

[238] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields", in *2016 IEEE 13th international symposium on biomedical imaging (ISBI)*, IEEE, 2016, pp. 698–701.

[239] B. E. Bejnordi, M. Balkenhol, G. Litjens, *et al.*, "Automated detection of dcis in whole-slide h&e stained breast histopathology images", *IEEE transactions on medical imaging*, vol. 35, no. 9, pp. 2141–2150, 2016.

[240] H. Yoshida, Y. Yamashita, T. Shimazu, *et al.*, "Automated histological classification of whole slide images of colorectal biopsy specimens", *Oncotarget*, vol. 8, no. 53, p. 90 719, 2017.

[241] P. Gillet and et al., "Magnetic resonance imaging may be an asset to diagnose and classify fluoroquinolone-associated achilles tendinitis", *Fundamental & Clinical Pharmacology*, vol. 9, no. 1, pp. 52–56, Jan. 1995. DOI: 10.1111/j.1472-8206.1995.tb00265.x.

[242] R. Weinstabl, M. Stiskal, A. Neuhold, B. Aamlid, and H. Hertz, "Classifying calcaneal tendon injury according to mri findings", *J. Bone Joint Surg. Br.*, vol. 73, no. 4, pp. 683–685, 1991.

[243] R. F. McLoughlin, E. L. Raber, A. D. Vellet, J. P. Wiley, and R. C. Bray, "Patellar tendinitis: Mr imaging features, with suggested pathogenesis and proposed classification", *Radiology*, vol. 197, no. 3, pp. 843–848, 1995.

[244] S. Bauer and et al., "Reliability of a 3 t mri protocol for objective grading of supraspinatus tendonosis and partial thickness tears", *J Orthop Surg Res*, vol. 9, no. 1, p. 128, Dec. 2014. DOI: 10.1186/s13018-014-0128-x.

[245] B. C. Carofino and S. S. Leopold, "Classifications in brief: The neer classification for proximal humerus fractures", *Clinical Orthopaedics & Related Research*, vol. 471, no. 1, pp. 39–43, Jan. 2013. DOI: 10.1007/s11999-012-2454-9.

[246] R.-F. Chang, C.-C. Lee, and C.-M. Lo, "Computer-aided diagnosis of different rotator cuff lesions using shoulder musculoskeletal ultrasound", *Ultrasound in Medicine & Biology*, vol. 42, no. 9, pp. 2315–2322, Sep. 2016. DOI: 10.1016/j.ultrasmedbio.2016.05.016.

[247] J. Benrabha and F. Meziane, "Automatic roi detection and classification of the achilles tendon ultrasound images", pp. 1–7, Oct. 2017. DOI: 10.1145/3109761.3158381.

[248] K. Meiburger and et al., "Quantitative analysis of patellar tendon abnormality in asymptomatic professional 'pallapugno' players: A texture-based ultrasound approach", *Applied Sciences*, vol. 8, no. 5, p. 660, Apr. 2018. DOI: 10.3390/app8050660.

[249] R.-F. Chang, C.-C. Lee, and C.-M. Lo, "Quantitative diagnosis of rotator cuff tears based on sonographic pattern recognition", *PLoS ONE*, vol. 14, no. 2, e0212741, Feb. 2019. DOI: 10.1371/journal.pone.0212741.

[250] M. Golman and et al., "Rethinking patellar tendinopathy and partial patellar tendon tears: A novel classification system", *Am J Sports Med*, vol. 48, no. 2, pp. 359–369, Feb. 2020. DOI: 10.1177/0363546519894333.

[251] B. R. DeGeorge, D. M. Brogan, and A. Y. Shin, "The relationship of volar plate position and flexor tendon rupture: Should we question the validity of the soong classification?", *Plastic & Reconstructive Surgery*, vol. 146, no. 3, pp. 581–588, Sep. 2020. DOI: 10.1097/PRS.0000000000007080.

[252] M. J. Leite and et al., "Coracoid morphology and humeral version as risk factors for subscapularis tears", *Journal of Shoulder and Elbow Surgery*, vol. 29, no. 9, pp. 1804–1810, Sep. 2020. DOI: 10.1016/j.jse.2020.01.074.

[253] M. Jahanifar and et al., "Automatic recognition of the supraspinatus tendinopathy from ultrasound images using convolutional neural networks", 2020, arXiv, Nov. 23, 2020. Accessed: Sep. 14, 2022. [Online]. Available: http://arxiv.org/abs/2011.11208.

[254] M. Ali, I. Mohamed, S. Suliman, and Y. M. Mustafah, "Detection of anterior cruciate ligament (acl) injury in ultrasound imaging: A review", pp. 54–58, Feb. 2021. DOI: 10.1109/ICMRE51356.2021.9378354.

[255] S. Orzada and et al., "Mr imaging of the infraspinatus tendon: Thickness measurements using a radial ultrashort echo time (ute) sequence in volunteers and patients with rotator cuff tears", *RoFo: Fortschritte auf dem Gebiete der Röntgenstrahlen und der Nuklearmedizin*, vol. 193, no. 11, pp. 1215–1222, Nov. 2021. DOI: 10.1055/a-1504-6993.

[256] R.-F. Chang, C.-C. Lee, and C.-M. Lo, "Ultrasound-based computer-aided diagnosis of rotator cuff tears: Quantification of the supraspinatus tendon thickness", *Diagnostics*, vol. 11, no. 7, p. 1232, Jul. 2021. DOI: 10.3390/diagnostics11071232.

[257] L. Wang *et al.*, "Musculoskeletal ultrasound image-based radiomics for the diagnosis of achilles tendinopathy in skiers", *Journal of Ultrasound Medicine*, jum.16059, Jul. 2022. DOI: 10.1002/jum.16059.

[258] P. Lambin *et al.*, "Radiomics: The bridge between medical imaging and personalized medicine", *Nature Reviews Clinical Oncology*, vol. 14, no. 12, pp. 749–762, Dec. 2017. DOI: 10.1038/nrclinonc.2017.141.

[259] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: Going beyond euclidean data", *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 18–42, 2017. DOI: 10.1109/MSP.2017.2693418.

[260] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains", *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[261] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency", in *Advances in neural information processing systems*, 2005, pp. 321–328.

[262] M. Lukoševičius, H. Jaeger, and B. Schrauwen, "Reservoir computing: Methods for multi-step dynamic prediction", *Journal of Machine Learning Research*, vol. 13, no. Feb, pp. 2931–2981, 2012.

[263] I. Yildiz, H. Jaeger, and S. Kiebel, "Re-visiting the echo state property", *Neural Networks*, vol. 35, pp. 1–9, 2012. DOI: https://doi.org/10.1016/j.neunet.2012.07.005.

[264] H. Jaeger, "Adaptive nonlinear system identification with echo state networks", in *Proceedings of the International Conference on Neural Information Processing Systems (NIPS)*, 2002, pp. 593–600. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2002/file/426f990b332ef8193a61cc90516c1245-Paper.pdf.

[265] C. Gallicchio and A. Micheli, "Graph echo state networks", in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, 2010, pp. 1–8. DOI: 10.1109/IJCNN.2010.5596796.

[266] A. Micheli and D. Tortorella, "Designs of graph echo state networks for node classification", *Neurocomputing*, vol. 597, p. 127 965, 2024, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2024.127965. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231224007367.

[267] Z. Li, K. Fujiwara, and G. Tanaka, "Dynamical graph echo state networks with snapshot merging for spreading process classification", in *Neural Information Processing*, B. Luo, L. Cheng, Z.-G. Wu, H. Li, and C. Li, Eds., Singapore: Springer Nature Singapore, 2024, pp. 523–534, ISBN: 978-981-99-8141-0.

[268] F. Bianchi, C. Gallicchio, and A. Micheli, "Pyramidal reservoir graph neural network", *Neurocomputing*, vol. 470, pp. 389–404, 2022, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2021.04.131. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231221011152.

[269] Y. Sun, Z. Wang, and K. Wang, "Leveraging the power of echo state network for enhanced temporal knowledge graph reasoning", in *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024, pp. 1–8. DOI: 10.1109/IJCNN60899.2024.10650708.

[270] A. Micheli and D. Tortorella, "Addressing heterophily in node classification with graph echo state networks", *Neurocomputing*, vol. 550, p. 126 506, 2023, ISSN: 0925-2312. DOI: https://doi.org/10.1016/j.neucom.2023.126506. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092523122300629X.

[271] J. A. Martin and J. A. Buckwalter, "Aging, articular cartilage chondrocyte senescence and osteoarthritis", p. 8,

[272] K. Abula *et al.*, "Elimination of bmp7 from the developing limb mesenchyme leads to articular cartilage degeneration and synovial inflammation with increased age", *FEBS Letters*, vol. 589, no. 11, pp. 1240–1248, May 2015. DOI: 10.1016/j.febslet.2015.04.004.

[273] M. V. Padalkar, R. G. Spencer, and N. Pleshko, "Near infrared spectroscopic evaluation of water in hyaline cartilage", *Ann Biomed Eng*, vol. 41, no. 11, pp. 2426–2436, Nov. 2013. DOI: 10.1007/s10439-013-0844-0.

[274] T. Nishii, K. Kuroda, Y. Matsuoka, T. Sahara, and H. Yoshikawa, "Change in knee cartilage t2 in response to mechanical loading", *J. Magn. Reson. Imaging*, vol. 28, no. 1, pp. 175–180, Jul. 2008. DOI: 10.1002/jmri.21418.

[275] S. Lüssea *et al.*, "Evaluation of water content by spatially resolved transverse relaxation times of human articular cartilage", *Magnetic Resonance Imaging*, vol. 18, no. 4, pp. 423–430, May 2000. DOI: 10.1016/S0730-725X(99)00144-7.

[276] M. Peshkova *et al.*, "Gender-related aspects in osteoarthritis development and progression: A review", *IJMS*, vol. 23, no. 5, p. 2767, Mar. 2022. DOI: 10.3390/ijms23052767.

[277] B. Olesiak, A. Przedborska, and J. W. Raczkowski, "Degeneration grade of a femoropatellar joint and functional status of a knee joint", *PPS*, vol. 5, no. 2, p. 44, Feb. 2019. DOI: 10.12775/PPS.2015.05.02.003.

[278] I. Koca, A. Boyaci, A. Tutoglu, N. Boyaci, and A. Ozkur, "The relationship between quadriceps thickness, radiological staging, and clinical parameters in knee osteoarthritis", *J Phys Ther Sci*, vol. 26, no. 6, pp. 931–936, 2014. DOI: 10.1589/jpts.26.93.

[279] S. Ota, "Relationship between patellar mobility and patellofemoral joint cartilage degeneration after anterior cruciate ligament reconstruction", Nagoya University Graduate School of Medicine, School of Medicine, Nov. 2017.

[280] M. Abate *et al.*, "Pathogenesis of tendinopathies: Inflammation or degeneration?", *Arthritis Res Ther*, vol. 11, no. 3, p. 235, 2009. DOI: 10.1186/ar2723.

[281] K. M. Khan, J. L. Cook, J. E. Taunton, and F. Bonar, "Overuse tendinosis, not tendinitis: Part 1: A new paradigm for a difficult clinical problem", *The Physician and Sportsmedicine*, vol. 28, no. 5, pp. 38–48, May 2000. DOI: 10.3810/psm.2000.05.890.

[282] N. Maffulli, P. Renström, and W. B. Leadbetter, Eds., *Tendon injuries: basic science and clinical medicine*. London: Springer, 2005.

[283] A. G. Culvenor, J. L. Cook, S. J. Warden, and K. M. Crossley, "Infrapatellar fat pad size, but not patellar alignment, is associated with patellar tendinopathy", *Scandinavian Journal of Medicine & Science in Sports*, vol. 21, no. 6, e405–e411, Dec. 2011. DOI: 10.1111/j.1600-0838.2011.01334.x.

[284] B. Erber, A. Baur-Melnyk, C. Glaser, S. Goller, J. Ricke, and A. Heuck, "Quadriceps fat pad edema in mr imaging: Association with quadriceps tendon alterations in a retrospective analysis", *European Journal of Radiology*, vol. 142, p. 109 858, Sep. 2021. DOI: 10.1016/j.ejrad.2021.109858.

[285] J. M. Walker, "Pathomechanics and classification of cartilage lesions, facilitation of repair", *J Orthop Sports Phys Ther*, vol. 28, no. 4, pp. 216–231, Oct. 1998. DOI: 10.2519/jospt.1998.28.4.216.

[286] I. O. Afara *et al.*, "Machine learning classification of articular cartilage integrity using near infrared spectroscopy", *Cel. Mol. Bioeng.*, vol. 13, no. 3, pp. 219–228, Jun. 2020. DOI: 10.1007/s12195-020-00612-5.

[287] R. Lemoyne, W. Kerr, K. Zanjani, and T. Mastroianni, "Implementation of an ipod wireless accelerometer application using machine learning to classify disparity of hemiplegic and healthy patellar tendon reflex pair", *j med imaging hlth inform*, vol. 4, no. 1, pp. 21–28, Mar. 2014. DOI: 10.1166/jmihi.2014.1219.

[288] C. Ricciardi *et al.*, "Improving prosthetic selection and predicting bmd from biometric measurements in patients receiving total hip arthroplasty", *Diagnostics*, vol. 10, no. 10, p. 815, Oct. 2020. DOI: 10.3390/diagnostics10100815.

[289] F. K. Ciliberti *et al.*, "Ct- and mri-based 3d reconstruction of knee joint to assess cartilage and bone", *Diagnostics*, vol. 12, no. 2, p. 279, Jan. 2022. DOI: 10.3390/diagnostics12020279.

[290] M. K. Gislason *et al.*, "Three dimensional bone mineral density changes in the femur over 1 year in primary total hip arthroplasty patients", *Clinical Biomechanics*, vol. 78, p. 105 092, Aug. 2020. DOI: 10.1016/j.clinbiomech.2020.105092.

[291] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification", *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973. DOI: 10.1109/TSMC.1973.4309314.

[292] M. Hall-Beyer, *GLCM Texture: A Tutorial v. 3.0 March 2017*, p. 76.

[293] V. N. Chougule, A. V. Mulay, and B. B. Ahuja, "Clinical case study: Spine modeling for minimum invasive spine surgeries (miss) using rapid prototyping", p. 8,

[294] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time series feature extraction on basis of scalable hypothesis tests (tsfresh – a python package)", *Neurocomputing*, vol. 307, pp. 72–77, Sep. 2018. DOI: 10.1016/j.neucom.2018.03.067.

[295] T. Yiu, "Understanding random forest", *Understanding Random Forest*, Jun. 2019.

[296] T. K. Ho, "Random decision forests", in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, vol. 1, 1995, pp. 278–282. DOI: 10.1109/ICDAR.1995.598994.

[297] S. Han, B. D. Williamson, and Y. Fong, "Improving random forest predictions in small datasets from two-phase sampling designs", *BMC Med Inform Decis Mak*, vol. 21, no. 1, p. 322, Dec. 2021. DOI: 10.1186/s12911-021-01688-3.

[298] Y. Qi, "Random forest for bioinformatics", p. 18,

[299] J. H. Friedman, "Greedy function approximation: A gradient boosting machine", *Ann. Statist.*, vol. 29, no. 5, Oct. 2001. DOI: 10.1214/aos/1013203451.

[300] J. Jiang, R. Wang, M. Wang, K. Gao, D. D. Nguyen, and G.-W. Wei, "Boosting tree-assisted multitask deep learning for small scientific datasets", *J. Chem. Inf. Model.*, vol. 60, no. 3, pp. 1235–1244, Mar. 2020. DOI: 10.1021/acs.jcim.9b01184.

[301] A. R. Webb, G. Cawley, and K. D. Copsey, *Statistical Pattern Recognition*. 2011, p. 668.

[302] R. Parikh, A. Mathai, S. Parikh, G. C. Sekhar, and R. Thomas, "Understanding and using sensitivity, specificity and predictive values", *Indian Journal of Ophthalmology*, vol. 56, no. 1, p. 6, 2008.

[303] J. C. Yoo *et al.*, "Subscapularis tendon tear classification based on 3-dimensional anatomic footprint: A cadaveric and prospective clinical observational study", *Arthroscopy: The Journal of Arthroscopic & Related Surgery*, vol. 31, no. 1, pp. 19–28, Jan. 2015. DOI: 10.1016/j.arthro.2014.08.015.

[304] S. G. Dakin *et al.*, "Inflammation activation and resolution in human tendon disease", *Sci. Transl. Med.*, vol. 7, no. 311, Oct. 2015. DOI: 10.1126/scitranslmed.aac4269.

[305] Ł. Olewnik *et al.*, "A proposal for a new morphological classification of the popliteus muscle tendon with potential clinical and biomechanical significance", *Sci Rep*, vol. 11, no. 1, p. 14 434, Dec. 2021. DOI: 10.1038/s41598-021-93778-5.

[306] Z. Khatun, M. Tsirilaki, A. Lindemann, *et al.*, "The role of muscle and tendon in predicting cartilage degeneration and tendinopathy", in *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, Rome, Italy, 2022, pp. 289–294. DOI: 10.1109/MetroXRAINE54828.2022.9967653.

[307] R. Aubonnet, J. Ramos, M. Recenti, *et al.*, "Toward new assessment of knee cartilage degeneration", *CARTILAGE*, vol. 14, no. 3, pp. 351–374, 2023. DOI: 10.1177/19476035221144746.

[308] G. Tsechpenakis, "Deformable model-based medical image segmentation", in *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, Springer US, 2011, pp. 33–67. DOI: https://doi.org/10.1007/978-1-4419-8195-0_2.

[309] S. Wan, C. Gong, P. Zhong, B. Du, L. Zhang, and J. Yang, "Multiscale dynamic graph convolutional network for hyperspectral image classification", *IEEE Transactions on Geoscience and Remote Sensing*, 2020. DOI: 10.1109/TGRS.2019.2949180.

[310] D. Jain. "Superpixels and slic". (2019), [Online]. Available: https://darshita1405.medium.com/superpixels-and-slic-6b2d8a6e4f08 (visited on 10/04/2022).

[311] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images are more than pictures, they are data", *Radiology*, vol. 278, no. 2, pp. 563–577, Feb. 2016. DOI: 10.1148/radiol.2015151169.

[312] "Pyradiomics documentation". (2023), [Online]. Available: https://pyradiomics.readthedocs.io/en/latest/features.html.

[313] S.-l. developers. "Sklearn.preprocessing.standardscaler". Accessed: 2024-07-23. (2023), [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html.

[314] M. S. Siddiquee and N. S. Pathan, "Optic disc segmentation using superpixel based features and random forest classifier", in *2019 4th Intl. Conference on Electrical Information and Communication Technology (EICT)*. DOI: 10.1109/EICT48899.2019.9068827.

[315] A. Ibrahim and E.-S. M. El-kenawy, "Image segmentation methods based on superpixel techniques: A survey", Tech. Rep., 2020, p. 11.

[316] N. Amini and A. Shalbaf, "Automatic classification of severity of COVID-19 patients using texture feature and random forest based on computed tomography images", *International Journal of Imaging Systems and Technology*, 2021. DOI: 10.1002/ima.22679.

[317] Unknown, "Learning texture features from glcm for classification of brain tumor MRI images using random forest classifier", *WSEAS Transactions on Signal Processing*, Apr. 2022. DOI: 10.37394/232014.2022.18.8.

[318] V. D. P. Jasti, A. S. Zamani, K. Arumugam, *et al.*, "Machine learning and image processing for medical image analysis of breast cancer diagnosis", *Journal of Medical Imaging and Health Informatics*, 2021. DOI: https://doi.org/10.1155/2022/1918379.

[319] T. Alzyadat, S. Praet, G. Chetty, *et al.*, "Automatic segmentation of achilles tendon tissues using deep convolutional neural network", in *Machine Learning in Medical Imaging: 11th Intl. Workshop, MLMI 2020, Held in Conjunction with MICCAI 2020*, Springer Intl. Publishing.

[320] I. Scott, D. Connell, D. Moulton, *et al.*, "An automated method for tendon image segmentation on ultrasound using grey-level co-occurrence matrix features and hidden gaussian markov random fields", DOI: 10.1016/j.compbiomed.2023.107872.

[321] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images", *Nature Communications*, vol. 15, Jan. 2024, Open access. DOI: 10.1038/s41598-023-45678-9.

[322] S. Wang, Y. Li, D. Wang, *et al.*, "Echo state graph neural networks with analogue random resistive memory arrays", *Nature Machine Intelligence*, vol. 5, no. 2, pp. 134–143, Feb. 2023, Open access. DOI: https://doi.org/10.1038/s42256-023-00609-5.

[323] G. Pocock and C. D. Richards, *The human body: An introduction for the biomedical and health sciences*, 1st. Oxford University Press, 2009.

[324] C. R. Noback, *The human nervous system: Structure and function*, 6th. Humana Press, ISBN: 978-1-59259-730-7.

[325] K. E. R., S. J.H., J. T.M., S. S.A., H. A.J., and M. S(Eds.), *Principles of Neural Science*, 5th. McGraw-Hill Education, 2013.

[326] Encyclopedia Britannica. "Knee-jerk reflex". (n.d.).

[327] A. Chandrasekhar, N. A. Abu Osman, L. Tham, K.-S. Lim, and W. Abas, "Influence of age on patellar tendon reflex response", *PloS one*, vol. 8, e80799, 2013. DOI: 10.1371/journal.pone.0080799. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0080799.

[328] J. R. Burke, M. C. Schutten, D. M. Koceja, and G. Kamen, "Age-dependent effects of muscle vibration and the jendrassik maneuver on the patellar tendon reflex response", *Archives of Physical Medicine and Rehabilitation*, vol. 77, no. 6, pp. 600–604, 1996. DOI: 10.1016/s0003-9993(96)90302-0. [Online]. Available: https://doi.org/10.1016/s0003-9993(96)90302-0.

[329] D. M. Koceja and R. G. Mynark, "Comparison of heteronymous monosynaptic ia facilitation in young and elderly subjects in supine and standing positions", *The International Journal of Neuroscience*, vol. 103, no. 1-4, pp. 1–17, 2000. DOI: 10.3109/00207450009035005. [Online]. Available: https://doi.org/10.3109/00207450009035005.

[330] C. C. Carroll, J. M. Dickinson, J. M. Haus, *et al.*, "Influence of aging on the in vivo properties of human patellar tendon", *Journal of Applied Physiology (Bethesda, Md. : 1985)*, vol. 105, no. 6, pp. 1907–1915, 2008. DOI: 10.1152/japplphysiol.00059.2008. [Online]. Available: https://doi.org/10.1152/japplphysiol.00059.2008.

[331] A. Chandrasekhar, N. A. Abu Osman, L. K. Tham, K. S. Lim, and W. A. B. Wan Abas, "Influence of age on patellar tendon reflex response", *PLoS ONE*, vol. 8, no. 11, e80799, 2013. DOI: 10.1371/journal.pone.0080799. [Online]. Available: https://doi.org/10.1371/journal.pone.0080799.

[332] G. Kamen and D. M. Koceja, "Contralateral influences on patellar tendon reflexes in young and old adults", *Neurobiology of Aging*, vol. 10, no. 4, pp. 489–495, 1989. DOI: 10.1016/0197-4580(89)90005-6.

[333] S. Hwang, H.-S. Jeon, O.-Y. Kwon, and C.-H. Yi, "The effects of body weight on the soleus h-reflex modulation during standing", *Journal of Electromyography and Kinesiology*, vol. 21, no. 3, pp. 445–449, 2011, ISSN: 1050-6411. DOI: https://doi.org/10.1016/j.jelekin.2010.11.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1050641110001756.

[334] O. Lin-Wei, L. L. S. Xian, V. T. W. Shen, *et al.*, "Deep tendon reflex: The tools and techniques. what surgical neurology residents should know", *The Malaysian Journal of Medical Sciences : MJMS*, vol. 28, no. 2, pp. 48–62, 2021. DOI: 10.21315/mjms2021.28.2.5.

[335] L. K. Tham, N. A. Abu Osman, W. A. Wan Abas, and K. S. Lim, "The validity and reliability of motion analysis in patellar tendon reflex assessment", *PloS One*, vol. 8, no. 2, e55702, 2013. DOI: 10.1371/journal.pone.0055702.

[336] B. D. Moore, J. Drouin, B. M. Gansneder, and S. J. Shultz, "The differential effects of fatigue on reflex response timing and amplitude in males and females", *Journal of Electromyography and Kinesiology*, vol. 12, no. 5, pp. 351–360, 2002. DOI: 10.1016/s1050-6411(02)00032-9.

[337] T. Morimoto, H. Hirata, K. Watanabe, *et al.*, "The usefulness of deep tendon reflexes in the diagnosis of lumbar spine diseases: A narrative review", *Cureus*, vol. 16, no. 3, e55772, 2024. DOI: 10.7759/cureus.55772.

[338] A. Campbell, J. Caldwell, D. Yalamanchili, *et al.*, "Effect of patient height and sex on the patellar tendon and anterior cruciate ligament", *Orthopaedic Journal of Sports Medicine*, vol. 9, no. 5, 2021, ISSN: 2325-9671. DOI: https://doi.org/10.1177/23259671211003. [Online]. Available: https://journals.sagepub.com/doi/full/10.1177/23259671211003244.

[339] M. F. Pazzinatto, D. de Oliveira Silva, A. S. Ferreira, *et al.*, "Patellar tendon reflex and vastus medialis hoffmann reflex are down regulated and correlated in women with patellofemoral pain", *Archives of Physical Medicine and Rehabilitation*, vol. 100, no. 3, pp. 514–519, 2019. DOI: 10.1016/j.apmr.2018.06.024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30059658/.

[340] Kiso, Inc., *Home*, Retrieved October 2, 2023, from https://www.kisoinc.com/, n.d.

[341] Thought Technology. "Triode electrodes - t3402m". [Online]. (2022), [Online]. Available: https://thoughttechnology.com/triode-electrodes-t3402m/ (visited on 03/20/2022).

[342] Ó. Pilkington, A. Þórisdóttir, D. Weckels, *et al.*, "Electrically synchronized reflex hammer for more precise patella reflex test", 2021.

[343] S. Kristinsdóttir, A. Þ. Þórisdóttir, L. B. Halldórsdóttir, *et al.*, "A novel reflex analysis of healthy and spinal cord-injured individuals", *Current Directions in Biomedical Engineering*, vol. 8, no. 2, pp. 745–748, 2022. DOI: doi:10.1515/cdbme-2022-1190.

[344] N. S. Ward, "Compensatory mechanisms in the aging motor system", *Ageing Research Reviews*, vol. 5, no. 3, pp. 239–254, 2006.

[345] N. Takeuchi, Y. Oouchida, and S.-I. Izumi, "Motor control and neural plasticity through interhemispheric interactions", *Neural Plasticity*, vol. 2012, p. 823 285, 2012. DOI: 10.1155/2012/823285.

[346] D. J. Goble, J. P. Coxon, A. Van Impe, J. De Vos, N. Wenderoth, and S. P. Swinnen, "The neural control of bimanual movements in the elderly: Brain regions exhibiting age-related increases in activity, frequency-induced neural modulation, and task-specific compensatory recruitment", *Human Brain Mapping*, vol. 31, no. 8, pp. 1281–1295, 2010. DOI: 10.1002/hbm.20943.

[347] P. A. Reuter-Lorenz, J. Jonides, E. E. Smith, *et al.*, "Age differences in the frontal lateralization of verbal and spatial working memory revealed by pet", *Journal of Cognitive Neuroscience*, vol. 12, no. 1, pp. 174–187, 2000. DOI: 10.1162/089892900561814.

[348] Y. Péréon, S. Nguyen The Tich, E. Fournier, R. Genet, and P. Guihéneuc, "Electrophysiological recording of deep tendon reflexes: Normative data in children and in adults", *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 34, no. 3, pp. 131–139, 2004, ISSN: 0987-7053. DOI: https://doi.org/10.1016/j.neucli.2004.04.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0987705304000528.

Department of Engineering, School of Technology
Reykjavík University
Menntavegur 1
101 Reykjavík, Iceland
Tel. +354 599 6200
Fax +354 599 6201
www.ru.is
ISBN 978-9935-539-47-2 Electronic version
ISBN 978-9935-539-46-5 Print version

ORCID Zakia Khatun 0009-0007-7303-338X
https://orcid.org/0009-0007-7303-338X