# A Novel Evolutionary Swarm Fuzzy Clustering Approach for Hyperspectral Imagery

Pedram Ghamisi, *Student Member, IEEE,* Abder-Rahman Ali, *Student Member, IEEE,*
Micael S. Couceiro, *Member, IEEE,* Jon Atli Benediktsson, *Fellow, IEEE*

***Abstract*—In land cover assessment, classes often gradually change from one to another. Therefore, it is difficult to allocate sharp boundaries between different classes of interest. To overcome this issue and model such conditions, fuzzy techniques that resemble human reasoning have been proposed as alternatives. Fuzzy C-Means is the most common fuzzy clustering technique, but its concept is based on a local search mechanism and its convergence rate is rather slow, especially considering high dimensional problems (*e.g.*, in processing of hyperspectral images). Here, in order to address those shortcomings of hard approaches, a new approach is proposed, *i.e.*, Fuzzy C-Means which is optimized by Fractional Order Darwinian Particle Swarm Optimization. In addition, to speed up the clustering process, the histogram of image intensities is used during the clustering process instead of the raw image data. Furthermore, the proposed clustering approach is combined with support vector machine classification to accurately classify hyperspectral images. The new classification framework is applied on two well-known hyperspectral data sets; Indian Pines and Salinas. Experimental results confirm that the proposed swarm-based clustering approach can group hyperspectral images accurately in a time-efficient manner compared to other existing clustering techniques.**

***Index Terms*—Hyperspectral image analysis, clustering, Fractional Order Darwinian Particle Swarm Optimization, Fuzzy C-Means, Support Vector Machine classifier.**

## I. Introduction

CONVENTIONAL hard classification techniques do not consider the continuous changes of different land cover classes from one to another. For instance, the standard crisp K-Means can be considered as the most popular clustering technique in the field of pattern recognition [1]. However, this technique uses hard partitioning, in which each data point belongs to exactly one cluster. To model the gradual boundary changes, ”soft” classifiers have been used. Fuzzy classifiers are soft classification techniques that deal with vagueness in class definitions and model the gradual spatial transition between land cover classes [2]. To overcome the hard partition of K-Means, Fuzzy C-Means (FCM) was introduced in [3], which is a generalization of the standard crisp K-Means scheme, in which a data point can belong to all clusters with different degrees of membership.

Although FCM is an improvement on K-Means, it is known for being very sensitive to its initial cluster configuration and may fall into sub-optimal solutions [4]. Therefore, in the literature, researchers have tried to improve the resilience of the FCM approach by optimizing it with bio-inspired optimization techniques (*e.g.*, Particle Swarm Optimization (PSO) [5]). However, a general problem with these techniques, such as the PSO algorithm, is that they may get trapped in local optimum points, in such a way that they may be successful in some problems, but fail in others [6].

To further improve the existing techniques, we combine FCM with Fractional Order Darwinian PSO (FODPSO) previously proposed by Couceiro *et al.* and Ghamisi *et al.* and applied for different applications [6–10]. The FODPSO algorithm benefits from a cooperation paradigm in which particles within each swarm cooperate with one another, while multiple swarms compete to find the most adequate solution, *i.e.*, the optimal solution. By combining the FODPSO with the FCM technique, herein denoted as FODPSO-FCM, each particle will be represented by a given cluster configuration and the FCM objective function. The emergent collective properties of the FODPSO, together with a fractional order velocity and a set of punish-reward rules designed to simulate Darwin's natural selection mechanism, will converge to the optimal cluster configuration. In addition, in order to speed up the clustering process, the histogram of image intensities is proposed to be used as an input instead of the raw image data.

Moreover, in order to evaluate the efficiency of the clustering approaches in terms of classification accuracies, a new supervised classification framework is proposed. The classification approach integrates the classification map obtained by a Support Vector Machine (SVM) classifier and the map obtained by FODPSO-FCM through majority voting. SVM has the capability of performing well in terms of classification accuracies in classifying high dimensional data with a limited number of training samples [11, 12]. However, the SVM is a hard classifier and, therefore, it cannot model gradual changes between different classes. As a result, the combination of the proposed fuzzy clustering and SVM may lead to better classification accuracies. Then, the capability of the proposed classification framework is tested on two well-known AVIRIS hyperspectral data sets; Indian Pines and Salinas, and compared to other alternatives, including K-Means, in terms of CPU processing time, classification accuracies and statistical significant metrics.

P. Ghamisi and J. A. Benediktsson are with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland (corresponding author, e-mail: benedikt@hi.is)

Abder-Rahman Ali, abder-rahman.a.ali@ieee.org

Micael S. Couceiro is with Artificial Perception for Intelligent Systems and Robotics (AP4ISR), Institute of Systems and Robotics (ISR), University of Coimbra, 3030-290, Coimbra, Portugal, micaelcouceiro@isr.uc.pt and Ingeniarius, Lda., Rua da Vacaria, n.37, 3050-381, Mealhada, Portugal, micael@ingeniarius.pt

This research was supported in part by the Icelandic Research Fund for Graduate Students.

The obtained results confirm the capability of the proposed clustering approach in terms of improving the classification accuracies of SVM. It should also be noted that the proposed approach is time-efficient as it can converge to the most optimal solution in a few seconds.

The rest of the paper is organized as follows: Section II describes the background and Section III is devoted to the description of the proposed methodology. Experimental results are investigated in Section IV and, finally, the main concluding remarks, as well as possible future works, are discussed in Section V.

## II. Preliminaries

In the proposed approach, first, the input hyperspectral data set are grouped by the novel clustering approach (FODPSO-FCM). In parallel, the input hyperspectral data are classified by SVM. Finally, the outputs of the SVM and FODPSO-FCM are integrated through a majority voting process, from which the final classification map is created. By doing this, it is possible to take advantage of SVM which can handle high dimensional data with a limited number of training samples and the proposed clustering approach which is based on fuzzy theory and can model gradual changes between different classes.

Below, in order to introduce the new clustering FODPSO-FCM approach, different alternatives of K-Means (*e.g.*, the classical K-Means, FCM and PSO-FCM) will first be described. This will be followed by a detailed description of the proposed FODPSO-FCM clustering approach.

### A. K-Means

K-Means clustering is a well-known geometric clustering algorithm for finding clusters and cluster centers in a set of unlabeled data [13–15]. Let $X$ be a set of data points (*e.g.*, pixels). The algorithm benefits from a local search approach to partition the points in $k$ clusters [14]. The desired number of cluster centers is chosen by the user, and K-Means iteratively moves the centers in order to minimize the total within cluster variance [13].

K-Means is very widely used and has been applied in a great variety of applications [16–18]. The main attraction of this method lies in its simplicity and its observed speed [14].

K-Means aims at minimizing the sum of squared distances between all points and the closest cluster center, and proceeds as follows [19, 20]:

1) Choose *k* initial cluster centers: $z_1(1), z_2(1), z_3(1), \ . \ . \ ., z_k(1)$.

2) At the *k*-th iterative step, distribute the samples $X$ among the *k* clusters using the following: $X \in C_j(k)$ if $\|X - z_j(k)\| < \|X - z_i(k)\|$; $\forall\, i, j = 1, 2, \ \ldots, k;\ \ i \neq j$

   where $C_j(k)$ and $z_j(k)$ denote centers and set of samples, respectively, and $\|.\|$ is the distance norm.

3) Compute the new cluster centers $z_j(k+1)$; $j = 1, 2, ..., k$, such that the sum of the squared distances of all points in $C_j(k)$ to the new cluster center is minimized. The new cluster center is given by:

$$z_j(k+1) = \frac{1}{N_j} \sum_{X \in C_j(k)} X;\ j = 1, 2,\ \ldots, k \qquad (1)$$

   where $N_j$ is the number of samples in $C_j(k)$.

4) The algorithm will convergence when $|z_j(k+1) - z_j(k)| < \epsilon$, being $\epsilon \in \mathbb{R}$, and the process terminates. Otherwise, go to step 2.

K-Means is popular because of its simplicity (*i.e.*, centroids, or cluster centers, are easy to understand and interpret). Also, because K-Means can find the local minimum of the optimization criterion (the sum of squared errors), and, for this, K-Means is often used as a post-processing approach for other clustering methods [21].

The K-Means clustering approach is considered fast when compared to most clustering algorithms. This is due to the fact that only distances and means are needed to be calculated, and thus, the computational complexity is not very high. However, since it is an iterative process, part of the speed will depend on the number of iterations, which, in turn, depends on the data distribution. Speed will also depend on the dimension of the data used, because, the higher the dimension, the slower the process.

K-Means is also considered fast due to its alternative optimization scheme. In order to assign the samples to the nearest centroids, the samples have to be linearly scanned. This linear scanning procedure considers a single pass over the data set (sample). In other words, it is one reading of the full set of data points without going back to the previous point. Therefore, in order to update the centroids, only the arithmetic means are computed.

The time complexity for K-Means can be considered as:

$$T = O(I \times P \times k) \qquad (2)$$

where *I* is the number of iterations, *P* is the number of points, and *k* is the number of clusters.

The worst-case behavior (*i.e.*, when the number of iterations is high) of K-Means is slow, but that does not usually happen, since most of the K-Means convergence occurs in the very first iterations, and so the user can specify an early termination condition (*i.e.*, maximum number of iterations). But, it is important to note that K-Means converges faster when the data contains well-separated clusters. Thus, from the above, K-Means is considered a fast clustering method.

### B. Fuzzy C-Means Clustering

FCM is an unsupervised clustering algorithm that has been applied successfully to a number of problems involving feature analysis, clustering, and classifier design [22]. This algorithm was proposed by Bezdek [23], as an improvement of the early hard K-Means clustering [22]. The FCM algorithm can be described as follows:

Let $X = \{X_1, ..., X_b, ..., X_q\}$ be the set of $q$ objects, and $Z = \{z_1, ..., z_b, ..., z_k\}$ be the set of $k$ centroids in a $p$-dimensional feature space. The FCM partitions $X$ into $k$ clusters by minimizing the following objective function [24]:

$$J = \sum_{j=1}^{q}\sum_{i=1}^{k}(\mu_{ij})^m \left\|X_j - Z_i\right\|^2 \tag{3}$$

where $1 < m \leq \infty$ is the *fuzzifier*, $Z_i$ is the $i$-th centroid corresponding to cluster $\beta_i$, $\mu_{ij} \in [0,1]$ is the fuzzy membership of the pattern $X_j$ to cluster $\beta_j$, such that,

$$Z_i = \frac{1}{q_i}\sum_{j=1}^{q}(\mu_{ij})^m X_j \quad where \quad q_i = \sum_{j=1}^{q}(\mu_{ij})^m \tag{4}$$

and,

$$\mu_{ij} = \frac{1}{\sum_{c=1}^{k}(\frac{d_{ij}}{d_{cj}})^{\frac{2}{m-1}}} \quad where \quad d_{ij}^2 = \left\|X_j - Z_i\right\|^2. \tag{5}$$

FCM starts by randomly choosing $k$ objects as centroids (means) of the $k$ clusters. Memberships are calculated based on the relative distance of the object $X_j$ to the centroids using (5). After the memberships of all objects have been found, the centroids of the clusters are calculated using (4). The process terminates when the norm difference between two consecutive iterations is less than a predefined threshold [24].

## III. Methodology

### A. PSO-FCM

Researchers have been proposing several extensions of the FCM over the past few years in order to improve its convergence. The combination of FCM and the well-known PSO has been one of the most successful ones (*i.e.*, [25, 26]).

PSO is a biologically inspired technique derived from the collective behavior of birds flocks. Here, each particle presents itself as a possible solution of the problem, *e.g.*, the best cluster centers of a given hyperspectral image. These particles travel through the search space to find an optimal solution by interacting and sharing information with other particles, namely their individual best solution (local best) and computing the globally best solution [27]. In other words, PSO-based algorithms consider multiple particles, wherein each particle has its own local solution, *e.g.*, best cluster center it found, and the whole swarm has a global solution which is the best among the several local solutions of all particles, *e.g.*, best cluster center found among all best cluster centers provided by each particle.

In each step $t$ of the PSO-FCM, the fitness function, represented by (3), is used to evaluate the success of particles. To model the swarm, each particle $n$ moves in a multidimensional space according to the position $\mathbf{x}_n[t]$, and velocity $\mathbf{v}_n[t]$, which are highly dependent on the locally best $\tilde{\mathbf{x}}_n[t]$ and the globally best $\tilde{\mathbf{g}}_n[t]$ information:

$$\begin{aligned}&\mathbf{v}_n\left[t+1\right] = \\ &w\mathbf{v}_n\left[t\right] + \rho_1\mathbf{r}_1\left(\tilde{\mathbf{g}}_n[t] - \mathbf{x}_n[t]\right) + \rho_2\mathbf{r}_2\left(\tilde{\mathbf{x}}_n[t] - \mathbf{x}_n[t]\right),\end{aligned} \tag{6}$$

$$\mathbf{x}_n\left[t+1\right] = \mathbf{x}_n\left[t\right] + \mathbf{v}_n\left[t+1\right]. \tag{7}$$

Coefficients $w$, $\rho_1$ and $\rho_2$ assign weights to the inertial influence, the global best and the local best when determining the new velocity $\mathbf{v}_n[t+1]$, respectively, with $\rho_1 + \rho_2 < 2$ [7], provided that different results can be obtained by assigning different influences for each component. The parameters $r_1$ and $r_2$ are random vectors with each component generally a uniform random number between 0 and 1. The intent is to multiply a new random component per velocity dimension, rather than multiplying the same component with each particles' velocity dimension.

The velocity dimension, as well as the position dimension, correspond to the total number of desired cluster centers in the image. In other words, each particles position will be represented as a k-dimension vector. Moreover, each particle moves in a multidimensional space according to position from the discrete time system (6-7), wherein $\mathbf{v}_n[t], \mathbf{x}_n[t] \in \mathbb{R}^k$.

### B. FODPSO-FCM

Despite good results that have been achieved using the traditional PSO proposed by Kennedy and Eberhart [5], PSO also has demonstrated an important drawback: The susceptibility to local solutions. As such, under complex problems with multiple local optima, the PSO may fail. In order to overcome this problem, many authors have suggested adjustments to the traditional PSO algorithm, namely considering evolutionary properties such as the hybridization of Genetic algorithms and PSO [28] as well as the FODPSO algorithm [6–8, 29]. The FODPSO considers the use of fractional calculus which, as opposed to its integer-order derivative counterpart, endows particles with a 'memory' of all past events. In other words, by slightly increasing the memory complexity of the PSO, one can decrease the computational complexity as particles converge toward the optimal solution.

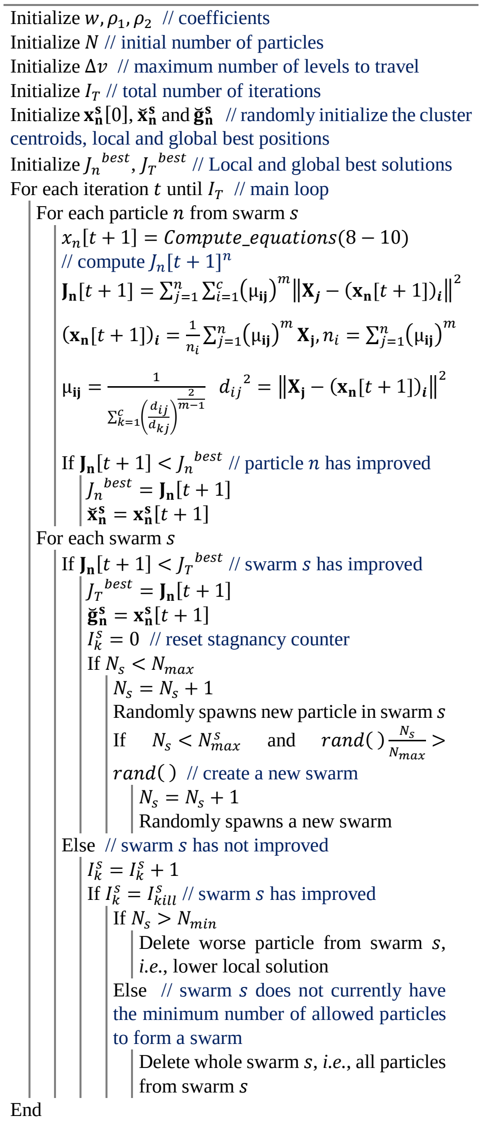
The basic idea of FODPSO-FCM is to run many simultaneous parallel PSO-FCM algorithms, each one as a different swarm, on the same test problem and then a simple selection mechanism is applied. When a search tends to a sub-optimal solution, the search in that area is simply discarded and another area is searched instead. In order to model the swarm $s$, each particle $n$ moves in a multidimensional space according to the position $\mathbf{x}_n^s[t]$, and the velocity $\mathbf{v}_n^s[t]$, that highly depend on the locally best $\tilde{\mathbf{x}}_n[t]$ and the globally best $\tilde{\mathbf{g}}_n[t]$, shown below [30] :

$$\begin{aligned}\mathbf{w}_n^s[t+1] &= \alpha\mathbf{v}_n^s[t] + \frac{1}{2}\alpha(1-\alpha)\mathbf{v}_{vn}^s[t-1] \\ &+ \frac{1}{6}\alpha(1-\alpha)(2-\alpha)\mathbf{v}_{vn}^s[t-2] \\ &+ \frac{1}{24}\alpha(1-\alpha)(2-\alpha)(3-\alpha)\mathbf{v}_{vn}^s[t-3],\end{aligned} \tag{8}$$

$$\mathbf{v}_{vn}^s[t+1] = \mathbf{w}_n^s[t+1] + \rho_1 r_1(\tilde{\mathbf{x}}_n^s[t] - x_n^s[t]) + \rho_2 r_2(\tilde{\mathbf{g}}_n^s[t] - \mathbf{x}_n^s[t]), \tag{9}$$

$$\mathbf{x}_n^s\left[t+1\right] = \mathbf{x}_n^s[t] + \mathbf{v}_n^s[t+1]. \tag{10}$$

The fractional coefficient $0 < \alpha < 1$ weights the influence of past events on determining a new velocity. With a small



```
Initialize $w, \rho_1, \rho_2$ // coefficients
Initialize $N$ // initial number of particles
Initialize $\Delta v$ // maximum number of levels to travel
Initialize $I_T$ // total number of iterations
Initialize $\mathbf{x}_\mathbf{n}^\mathbf{s}[0]$, $\breve{\mathbf{x}}_\mathbf{n}^\mathbf{s}$ and $\breve{\mathbf{g}}_\mathbf{n}^\mathbf{s}$ // randomly initialize the cluster centroids, local and global best positions
Initialize $J_n{}^{best}, J_T{}^{best}$ // Local and global best solutions
For each iteration $t$ until $I_T$ // main loop
    For each particle $n$ from swarm $s$
        $x_n[t+1] = Compute\_equations(8-10)$
        // compute $J_n[t+1]^n$
        $\mathbf{J_n}[t+1] = \sum_{j=1}^{n}\sum_{i=1}^{c}(\mu_{\mathbf{ij}})^m \|\mathbf{X}_j - (\mathbf{x_n}[t+1])_i\|^2$
        $(\mathbf{x_n}[t+1])_i = \frac{1}{n_i}\sum_{j=1}^{n}(\mu_{\mathbf{ij}})^m \mathbf{X_j}, n_i = \sum_{j=1}^{n}(\mu_{\mathbf{ij}})^m$
        $\mu_{\mathbf{ij}} = \frac{1}{\sum_{k=1}^{c}\left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}}$   $d_{ij}{}^2 = \|\mathbf{X_j} - (\mathbf{x_n}[t+1])_i\|^2$
        If $\mathbf{J_n}[t+1] < J_n{}^{best}$ // particle $n$ has improved
            $J_n{}^{best} = \mathbf{J_n}[t+1]$
            $\breve{\mathbf{x}}_\mathbf{n}^\mathbf{s} = \mathbf{x}_\mathbf{n}^\mathbf{s}[t+1]$
    For each swarm $s$
        If $\mathbf{J_n}[t+1] < J_T{}^{best}$ // swarm $s$ has improved
            $J_T{}^{best} = \mathbf{J_n}[t+1]$
            $\breve{\mathbf{g}}_\mathbf{n}^\mathbf{s} = \mathbf{x}_\mathbf{n}^\mathbf{s}[t+1]$
            $I_k^s = 0$ // reset stagnancy counter
            If $N_s < N_{max}$
                $N_s = N_s + 1$
                Randomly spawns new particle in swarm $s$
                If $N_s < N_{max}^s$ and $rand()\frac{N_s}{N_{max}} > rand()$ // create a new swarm
                    $N_s = N_s + 1$
                    Randomly spawns a new swarm
        Else // swarm $s$ has not improved
            $I_k^s = I_k^s + 1$
            If $I_k^s = I_{kill}^s$ // swarm $s$ has improved
                If $N_s > N_{min}$
                    Delete worse particle from swarm $s$, i.e., lower local solution
                Else // swarm $s$ does not currently have the minimum number of allowed particles to form a swarm
                    Delete whole swarm $s$, i.e., all particles from swarm $s$
End
```

Fig. 1. FODPSO-FCM algorithm ($s$: the number of a swarm, $t$: the number of iterations, $\alpha$: Fractional order coefficient, $\rho_1$: cognitive coefficient, $\rho_2$: social coefficient, $N^s$: current number of swarms, $N^s_{min}$: minimum number of swarms, $N^s_{max}$: maximum number of swarms, $N_{min}$: minimum number of particles in a swarm, $N_{max}$: maximum number of particles in a swarm, $I_T$: number of iterations, $I_{kill}$: maximum number of iterations for stagnation, $J$: fitness function.

$\alpha$, particles ignore their previous activities, thus ignoring the system dynamics and become susceptible to get stuck in local solutions (*i.e.*, exploitation behavior). On the other hand, with a large $\alpha$, particles will present a more diverse behavior, which allows exploration of new solutions and improves the long-term performance [30]. However, if the exploration level is too high, the algorithm may take too much time to find the global solution.

The parameters $r_1$ and $r_2$ are random vectors with each component generally a uniform random number between 0 and 1 [30].

As particles move in that multidimensional space, they share their own solution with other particles inside the same swarm by computing the fitness function presented in (3) [30]. Thus, following the insights presented for the first time by [31, 32] and more recently in [6], in order to analyze the general state of each swarm, the fitness of all particles is evaluated and the neighborhood and individual best positions of each of the particles are updated. If a new global solution is found, a new particle is spawned. Whilst if the swarm fails to find a fitter state in a defined number of steps, the particle is deleted. Fig. 1 describes how the proposed FODPSO-FCM clustering method works.

### C. Proposed classifier based on FODPSO-FCM and SVM

In order to take advantage of SVM and the novel clustering method (FODPSO-FCM), the outputs of them are integrated through majority voting. Fig. 2 depicts the general idea of integrating the outputs of FODPSO-FCM (clustering map) and SVM (classification map) with the majority voting. Clustering methods provide a few cluster centers as their outputs. Each cluster consists of several pixels that are closest to the cluster center. These samples get the same clustering label. In the same way, the output of classification is a few number of classes and each class consists of several pixels with the same classification label. To perform the majority voting jointly on the output of the clustering and classification steps, first, the number of pixels with different classification labels in each cluster is counted. Then, all pixels in each cluster get the most frequent class label in the cluster. For a better understanding of how the results of the novel fuzzy clustering approach (FODPSO-FCM) and the classification technique (SVM) are combined through majority voting, the outline of the whole procedure can be described step-by-step as follows:

1) The input data are classified by SVM;
2) The input data are clustered by FODPSO-FCM;
3) The results of step 1 and step 2 are combined using the majority voting.

It should be noted that the classification framework in this work has been only used for the evaluation of the capability of different clustering techniques and the main novelty of this paper goes to the proposition of the new clustering approach.

## IV. Experimental Results

### A. Data Description

Two data sets are used in experiments. These data sets are described below.

*1) Indian Pines:* This data set was captured by AVIRIS of Indian Pines (NW Indiana) in 1992 and has 16 classes by AVIRIS (see Fig. 3), mostly related to different land covers. The data set consists of 145 × 145 pixels with a spatial resolution of 20 m/pixel. Here, 200 data channels were considered after the removal of the spectral bands affected by atmospheric absorption. The number of training and test samples are displayed in Table I.

*2) Salinas:* This data set was captured by AVIRIS over Salinas Valley, CA, USA, and it is characterized by high spatial resolution (3.7-m pixels) consisting of 512 by 217 samples. The original data set consists of 224 data channels,

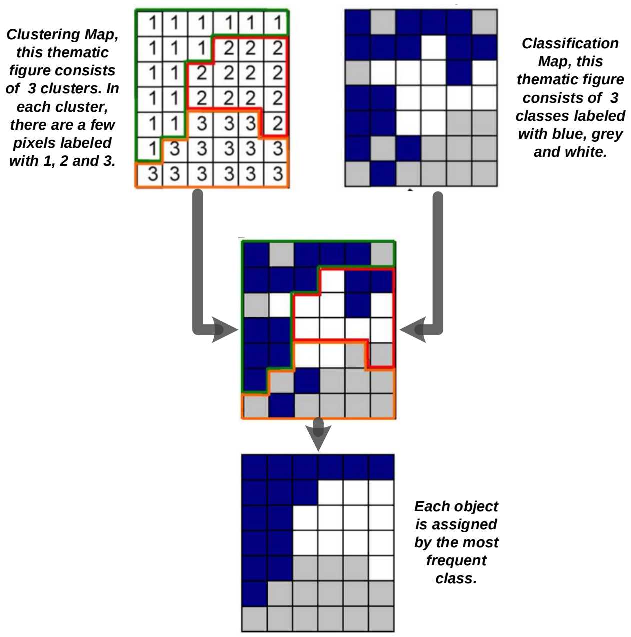

Clustering Map, this thematic figure consists of 3 clusters. In each cluster, there are a few pixels labeled with 1, 2 and 3.
1 1 1 1 1 1
1 1 1 2 2 2
1 1 2 2 2 2
1 1 2 2 2 2
1 1 3 3 3 2
1 3 3 3 3 3
3 3 3 3 3 3
Classification Map, this thematic figure consists of 3 classes labeled with blue, grey and white.
Each object is assigned by the most frequent class.

Fig. 2. Procedure of the majority voting for combining the output of the clustering and the classification techniques (based on [33]).

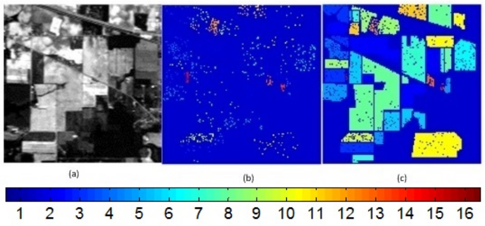


Fig. 3. The AVIRIS Indian Pines data set. a) Spectral band number 27 ($\lambda = 646.72nm$); b) training samples, c) test samples, where each color represents a specific information class. The information classes are listed in Table I.

TABLE I
INDIAN PINES: THE NUMBER OF TRAINING AND TEST SAMPLES

| | Class | Number of Samples | |
|---|---|---|---|
| Number | Name | Training | Test |
| 1 | Corn-notill | 50 | 1384 |
| 2 | Corn-mintill | 50 | 784 |
| 3 | Corn | 50 | 184 |
| 4 | Grass-pasture | 50 | 447 |
| 5 | Grass-trees | 50 | 697 |
| 6 | Hay-windrowed | 50 | 439 |
| 7 | Soybean-notill | 50 | 918 |
| 8 | Soybean-mintill | 50 | 2418 |
| 9 | Soybean-clean | 50 | 564 |
| 10 | Wheat | 50 | 162 |
| 11 | Woods | 50 | 1244 |
| 12 | Bldg-Grass-Tree-Drives | 50 | 330 |
| 13 | Stone-Steel-Towers | 50 | 45 |
| 14 | Alfalfa | 15 | 39 |
| 15 | Grass-pasture-mowed | 15 | 11 |
| 16 | Oats | 15 | 5 |

TABLE II
THE NUMBER OF TRAINING AND TEST SAMPLES.

| | Class | Number of Samples | |
|---|---|---|---|
| Number | Name | Training | Test |
| 1 | Brocoli_green_weeds_1 | 252 | 1757 |
| 2 | Brocoli_green_weeds_2 | 474 | 3252 |
| 3 | Fallow | 239 | 1737 |
| 4 | Fallow_rough_plow | 169 | 1225 |
| 5 | Fallow_smooth | 342 | 2336 |
| 6 | Stubble | 516 | 3443 |
| 7 | Celery | 442 | 3137 |
| 8 | Grapes_untrained | 1395 | 9876 |
| 9 | Soil_vinyard_develop | 775 | 5428 |
| 10 | Corn_senesced_green_weeds | 407 | 2871 |
| 11 | Lettuce_romaine_4wk | 141 | 927 |
| 12 | Lettuce_romaine_5wk | 232 | 1695 |
| 13 | Lettuce_romaine_6wk | 124 | 792 |
| 14 | Lettuce_romaine_7wk | 121 | 949 |
| 15 | Vinyard_untrained | 906 | 6362 |
| 16 | Vinyard_vertical_trellis | 231 | 1576 |

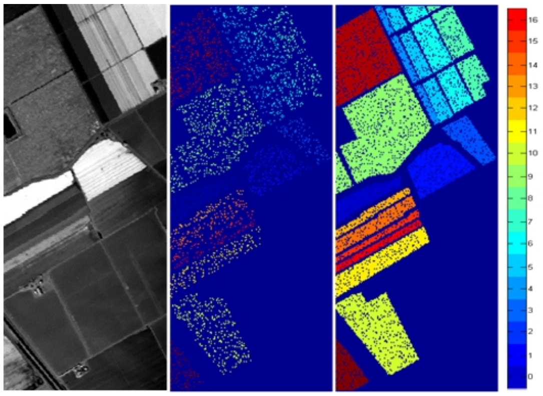


Fig. 4. Example of the Salinas test case. (a) Data channel 57. (b) Training samples. (c) Test samples. Each color represents a specific information class. The information classes are listed in Table II.

but here, 20 water absorption bands are discarded. The data set includes vegetation, bare soils, and vineyard fields. The Salinas reference data contain 16 classes. Fig. 4 shows the Salinas data set and its corresponding reference map. The number of training and test samples are displayed in Table II.

### B. Algorithm setup and general description

To compare the performance of the different K-Means alternatives used in this paper, all methods were programmed in MATLAB, running on a computer having Intel(R) Core(TM) i7 CPU 2.40 GHz and 16GB (15.9GB usable) of memory.

Table III presents the initial parameters for both the PSO- and FODPSO-FCM methods. As previously described, the PSO and FODPSO methods are parameterized algorithms. Therefore, one needs to choose parameter values that result in faster convergence. The cognitive, social and inertial weights were chosen by taking into account several works focusing on the convergence analysis of the traditional PSO (cf., [6, 34]). For instance, to guarantee the convergence of the process, Jiang *et al.* [34] presented a set of attraction domains that altogether present a relation between $\rho_1$, $\rho_2$ and $w$, wherein $0 \leq w < 1$ and $\rho_1 + \rho_2 > 0$. Based on the attraction domain in

TABLE III
DEFINITION OF PSO AND FODPSO PARAMETERS USED IN EXPERIMENTS.

| Parameter | PSO | FODPSO |
|---|---|---|
| Number of iterations $I_T$ | 100 | 100 |
| Number of particles $N$ | 150 | 20 |
| Social coefficient $\rho_1$ | 0.8 | 0.8 |
| Cognitive coefficient $\rho_2$ | 1.2 | 1.2 |
| Inertial coefficient $w$ | 0.8 | - |
| Maximum step between iterations $\Delta v$ | 2 | 2 |
| Minimum number of particles in a swarm $N_{min}$ | - | 10 |
| Maximum number of particles in a swarm $N_{max}$ | - | 30 |
| Initial number of swarms $N^s$ | - | 4 |
| Minimum number of swarms $N^s_{min}$ | - | 2 |
| Maximum number of swarms $N^s_{max}$ | - | 6 |
| Stagnancy counter $I^s_{kill}$ | - | 10 |
| Fractional order coefficient $\alpha$ | - | 0.8 |

TABLE IV
INDIAN PINES: CLASSIFICATION ACCURACIES (OA IN PERCENTAGE AND KAPPA HAS NO UNITS) ALONG WITH CPU PROCESSING TIME IN SECONDS.

| Method | OA | Kappa | Time(s) |
|---|---|---|---|
| SVM | 82.56 | 0.8019 | – |
| K-Means | 86.92 | 0.8510 | 0.0090±0.0039 |
| FCM | 87.24 | 0.8544 | 0.2737±0.0039 |
| PSO | 86.22±1.68 | 0.8428±0.0191 | 3.3471±0.2666 |
| FODPSO | 88.03±0.90 | 0.8635±0.0102 | 2.5030±0.1358 |

[34], if one would choose an inertial coefficient $w = 0.8$, the sum between the cognitive and social components would need to be less than 7, *i.e.*, $\rho_1 + \rho_2 < 7$. Note that the threshold velocities of particles and the maximum number of particles within each swarm in the FODPSO are smaller than the PSO algorithm. This was experimentally adjusted here to provide swarms of 20 particles with the same level of diversity (*i.e.*, exploration and exploitation) as swarms of 150 particles.

The same set of parameters for both data sets was chosen which infers that the proposed method is data set distribution independent and there is no need to set any parameters for it.

Since the compared methods are based on stochastic techniques and results can be different in different runs, both the PSO-FCM and FODPSO-FCM were evaluated over 50 runs for each data set, in which and the average results are reported in Tables IV and V.

In order to carry out a fair evaluation, the input is classified only once by SVM, while the output of this step is used for all different levels. By doing that, the accuracy of the classification for different methods only depends on the outcome of the clustering method.

The number of clusters for all clustering techniques studied in this work was set equal to the number of classes for both data sets.

The hyperplane parameters for the SVM classifier were chosen by 5-fold cross validation.

Below, the terms K-Means, FCM, PSO and FODPSO are referred to situations when the hyperspectral data are clustered by K-Means, FCM, PSO-FCM and FODPSO-FCM, respectively, and the outputs combined with SVM through the majority voting step.

### C. Discussion

*1) Indian Pines:* Table IV gives information regarding the classification accuracies and the CPU processing time. As can be seen, FODPSO+FCM outperforms other alternatives in terms of classification accuracies. The main reason behind this is the capability the algorithm has to distributively compare a set of ever-improving solutions provided by particles in a coopetitive fashion, *i.e.*, depicting both cooperation and competition behaviors. In other words, while particles within the PSO only cooperate with one another, the several competitive swarms within the FODPSO are able to avoid suboptimal solutions and, henceforth, achieve a higher OA. The FODPSO+FCM approach is immediately followed by the FCM, thus presenting better results than both PSO-based approach and the crisp K-Means. As can be seen, all different alternatives of K-Means (K-Means, FCM, PSO and FODPSO) improve the classification accuracies of SVM within a very short CPU processing time. In other words, the use of all the different alternatives only adds up a very short overhead time on the CPU processing time of SVM, while significantly improving the classification accuracies.

One can also observe that the non-stochastic algorithms can cluster the input data faster than the PSO- and FODPSO-based approaches. That is an expected drawback in this specific situation since particles depict an exploitation behavior near the sub-optimal and optimal solutions, thus requiring a fine evaluation of each minor step performed by each particle. Put it differently, the methods take approximately the same time to retrieve roughly the same solution. Afterwards, the biological approaches require extraordinary computation to slightly improve the result around such solution.

One-way MANOVA analysis was carried out to assess whether K-Means, FCM, PSO-based FCM and FODPSO-based FCM methods have a statistically significant effect on the classification performance. The significance of the different types of algorithms used (independent variable) on the CPU processing time, overall accuracy (OA) and the kappa coefficient (dependent variables) was analyzed using a one-way MANOVA after checking the assumptions of multivariate normality and homogeneity of variance/covariance, for a significance level of 5%.

The assumption of normality of each of the univariate dependent variables was examined using a paired-sample *Kolmogorov-Smirnov* (p-value<0.05) [35]. Although the univariate normality of each dependent variable has not been verified, since n=50, and by considering the Central Limit Theorem (CLT) [36], normality was assumed [36].

The MANOVA analysis revealed that the type of algorithm led to a statistically significant different outcome on the multivariate composite (F=1196.4569; p-value<0.0001). As the MANOVA detected significant statistical differences, we proceeded to the commonly-used ANOVA for each dependent variable. By carrying an individual test on each dependent variable, it was possible to observe that the CPU processing time presents statistically significant differences between methods (F=6046.51; p-value<0.0001), in which both the

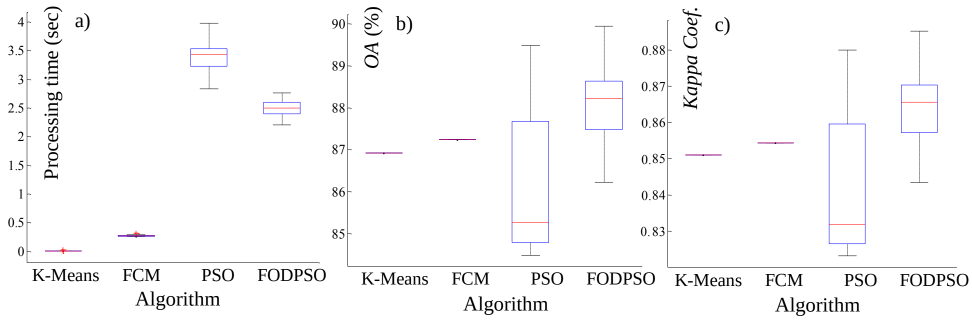


Fig. 5. Indian Pines: Performance comparison between methods. a) Processing time (sec); b) Overall accuracy (OA) (%); and c) Kappa Coefficient (has no units).

TABLE V
SALINAS: PERFORMANCE COMPARISON BETWEEN METHODS. A) PROCESSING TIME (SEC); B) OVERALL ACCURACY (OA) (%); AND C) KAPPA COEFFICIENT (HAS NO UNITS).

| Method | OA | Kappa | Time(s) |
|---|---|---|---|
| SVM | 94.02 | 0.9334 | – |
| K-Means | 95.49 | 0.9498 | 0.0211±0.0014 |
| FCM | 96.29 | 0.9587 | 0.1746±0.0096 |
| PSO | 95.92±0.49 | 0.9546±0.0054 | 3.59±0.1059 |
| FODPSO | 96.54±0.54 | 0.9615±0.0060 | 3.0196±0.5046 |

PSO- and FODPSO-based approaches take longer time than K-Means and FCM. Likewise, the OA also depicts statistically significant differences between methods (F=31.19; p-value<0.0001), as well as the KC (F=31.41; p-value<0.0001), in which the FODPSO-based approach is the one presenting an overall better performance for both metrics. In other words, the proposed FODPSO-based solution produces better solutions than the alternatives, at the cost of a higher CPU processing power. However, it does so with a lower CPU processing power than the PSO-based method.

To easily assess the differences between the algorithms and further understand the trade-off between performance and algorithmic complexity, we show the outcome of trials graphically using boxplot charts (Fig. 5). In the figure, the top/bottom of the blue boxes and the horizontal red lines in between correspond to the first and third quartiles and the median values, respectively. As one may observe, by benefiting from the FODPSO approach, the classification performance improves both in terms of OA and Kappa coef. However, there is an increase in the CPU processing time when compared to the non-stochastic methods. Even so, the FODPSO is still able to present a reduced computational complexity when compared to its original counterpart (PSO).

*2) Salinas:* Table V gives information regarding the classification accuracies (OA in percentage and Kappa) along with CPU processing time in seconds. Again, FODPSO has the best performance in terms of classification accuracies.

By comparing the CPU processing time reported in Tables V and IV, it can be observed that the new implementation of different alternatives of clustering methods (K-Means, FCM, PSO and FODPSO) can cluster different input data sets regardless of their size in almost the same CPU processing time. The two different data sets (Salinas with the size of $512 \times 217 \times 204$ and Indian Pines with the size of $145 \times 145 \times 200$) demanded almost the same CPU processing time to be clustered by using the new implementation proposed in this paper. The main reason behind this result is that computing the solution of each particle under both *PSO*- and *FODPSO*-based approaches requires about the same time, in which the influence of the image dimension is almost neglected.

As before, a one-way MANOVA analysis was carried out to assess whether K-Means, FCM, PSO-based FCM and FODPSO-based FCM methods have a statistically significant effect on the classification performance.

The MANOVA analysis revealed that the proposed algorithm led to a statistically significant outcome on the multivariate composite (F=645.9372; p-value<0.0001) when compared to the other approaches. As the MANOVA detected significant statistical differences, we proceeded to the commonly-used ANOVA for each dependent variable. It was possible to observe that the CPU processing time (F=1573.44; p-value<0.0001), the OA (F=47.81; p-value<0.0001) and the Kappa. (F=47.44; p-value<0.0001) present statistically significant differences between methods. Again, the proposed FODPSO-based solution produces better solutions in terms of accuracies than the alternatives, at the cost of higher CPU processing power.

Let us now compare the outcome of trials graphically using boxplot charts (Fig. 6). As one may observe, by benefiting from the FODPSO approach, the classification performance improves in terms of both OA and Kappa Coef. Although there is an increase in the CPU processing time as compared to the non-stochastic methods, the FODPSO is still able to present a reduced computational complexity when compared to its original counterpart (PSO), which even presents the worst results when compared to FCM under certain situations.

For both data sets, the new approach based on FODPSO-FCM leads to faster results than PSO-FCM since FODPSO with the fractional coefficient $\alpha$, is able to balance particles' behavior between the exploration and the exploitation. Finally, it should be noted that, for all the different alternatives of FCM applied in this paper, the histogram of image intensities is used instead of the raw image data used during the clustering process, which leads to the clustering of the whole data in a

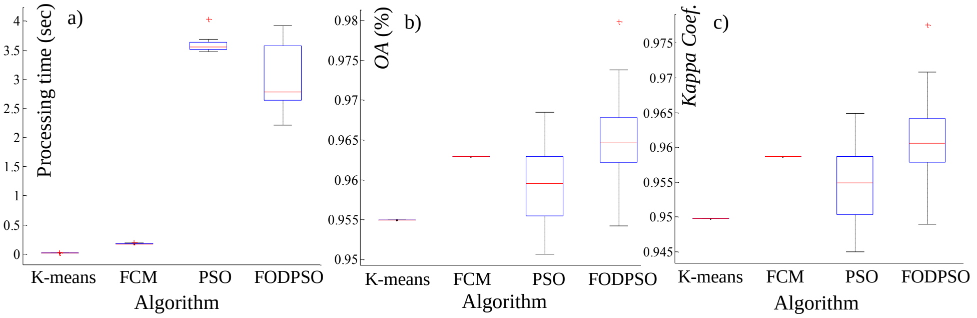


Fig. 6. Salinas: Performance comparison between methods. a) Processing time (sec); b) Overall accuracy (OA) (%); and c) Kappa Coefficient.

very fast way.

The computational complexity of the FODPSO algorithm depends, as its traditional counterpart PSO, on the number of possible solutions. In the case of PSO, the number of solutions is equal to the number of particles in the swarm, $N$, and remains the same throughout the optimization procedure. Therefore, the complexity of PSO is $O(N)$. In FODPSO, the number of solutions is not only based on the number of particles. Actually, the number of solutions is the sum of the number of particles, $N$, within all swarms, $s$, and the order of the fractional order extension, $\hat{r}$, which was defined here as four (see (8)). Therefore, the computational complexity of FODPSO is $O(\hat{r} \times N \times s)$. In the experiments, PSO was defined with 150 particles, while FODPSO was initially defined with 20 particles in each of the four starting swarms (Table III). In other words, in the very beginning, FODPSO has a higher computational complexity because of its 320 possible solutions as compared to the 150 possible solutions of the PSO algorithm. However, as the FODPSO algorithm converges, the number of particles and swarms within FODPSO starts to drop, being possible to achieve a minimum of 10 particles and two swarms, converging towards 80 possible solutions, presenting an inferior computational complexity than the traditional PSO algorithm. For more details, please refer to [37].

## V. CONCLUSION

In this paper, a novel evolutionary based fuzzy clustering approach is proposed. This approach benefits from a fractional calculus approach to improve the convergence rate of the traditional FCM, while, at the same time, it benefits from the same natural selection mechanism as the original PSO to avoid stagnation around local optima. Then, the proposed fuzzy clustering approach is used to improve the classification of hyperspectral images with SVM. It is well known that the SVM is able to efficiently classify high dimensional data with a limited samples. However, it is a hard classifier and cannot model temporal gradual changes between different classes. In contrast, the novel clustering technique is based on fuzzy concepts and can model gradual changes between different classes. In other words, the combination of the SVM and the novel fuzzy clustering technique can be considered a desirable strategy to classify hyperspectral images. The experimental evaluation for two benchmark hyperspectral images demonstrates that the performance of the proposed method, which uses a biologically inspired behavior based on natural selection and non-integer convergence, results in a statistically significant improvement in terms of overall classification accuracy and kappa coefficient. It should be noted, thanks to the efficient implementation of all clustering approaches in this paper based on image histograms, the clustering approaches here are very fast and they can lead to a conclusion in a few seconds (even less than one second for K-Means and FCM).

As a possible future work, it is envisioned to integrate the new approach with the Hidden Markov Random Field (HMRF) [38], since the new approach does not incorporate spatial and contextual information of the adjacent pixels and only considers the spectral information.

In addition, in this paper only rural areas with homogeneous regions have been taken into account since the number of clusters for different clustering alternatives here can be adjusted automatically (equal to number of classes). However, for complicated areas, this number should be increased in order to avoid under-segmentation. As a possible future work, the automatic selection of the number of clusters for complicated data sets can be of interest.

## REFERENCES

[1] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. Chichester: John Wiley Sons, 2001.

[2] A. Dutta, “Fuzzy c-means classification of multispectral data incorporating spatial contextual information by using Markov random field,” *Master thesis*, 2009.

[3] J. Bezdek, *Fuzzy Mathematics in Pattern Classification*, 2nd ed. Chichester: John Wiley Sons, 1973.

[4] W. Wang, Y. Zhang, Y. Li, and X. Zhang, “The global fuzzy c-means clustering algorithm,” *In Intelligent Control and Automation*, vol. 1, pp. 3604–3607, 2006.

[5] J. Kennedy and R. Eberhart, “A new optimizer using particle swarm theory,,” *in Proceedings of the IEEE Sixth International Symposium on Micro Machine and Human Science*, vol. 34, no. 2008, pp. 39–43, 1995.

[6] M. S. Couceiro, R. P. Rocha, N. M. F. Ferreira, and J. A. T. Machado, “Introducing the fractional order darwinian pso,” *Signal, Image and Video Processing, Springer, no. Fractional Signals and Systems Special Issue*, vol. 102, no. 1, pp. 8–16, 2007.

[7] P. Ghamisi, M. S. Couceiro, J. A. Benediktsson, and N. M. F. Ferreira, "An efficient method for segmentation of images based on fractional calculus and natural selection," *Expert Syst. Appl.*, vol. 39, no. 16, pp. 12 407–12 417, 2012.

[8] P. Ghamisi, M. S. Couceiro, F. M. Martins, and J. A. Benediktsson, "Multilevel image segmentation approach for remote sensing images based on fractional-order darwinian particle swarm optimization," *IEEE Trans. on of Remote Sens. and Geos.*, vol. 52, no. 5, pp. 2382–2394, 2014.

[9] P. Ghamisi, M. Couceiro, M. Fauvel, and J. A. Benediktsson, "Integration of segmentation techniques for classification of hyperspectral images," *IEEE Geos. and Remote Sens. Letter*, vol. 11, no. 1, pp. 342–346, 2014.

[10] P. Ghamisi, M. S. Couceiro, and J. A. Benediktsson, "Classification of hyperspectral images with binary fractional order darwinian pso and random forests," *Proc. SPIE*, vol. 8892, pp. 88 920S–88 920S–8, 2013.

[11] P. Ghamisi, J. A. Benediktsson, G. Cavallaro, and A. Plaza, "Automatic framework for spectral-spatial classification based on supervised feature extraction and morphological attribute profiles," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 2147–2160, 2014.

[12] P. Ghamisi, J. A. Benediktsson, and J. R. Sveinsson, "Automatic spectral-spatial classification framework based on attribute profiles and supervised feature extraction," *IEEE Trans. Rem. Sens. and Geos.*, vol. 52, no. 9, p. 5771 5782, 2014.

[13] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning: Data mining, inference, and prediction," *Massachusetts: Addison-Wesley*, 2009.

[14] D. Arthur and S. Vassilvitskii, "How slow is the k-means method?" *SCG'06:Proc. twenty-second annual symp. comp. geom.*, 2006.

[15] S. P. Lloyd, "Least squares quantization in pcm," *IEEE Trans. Information*, vol. 28, no. 2, p. 129136, 1982.

[16] P. Agarwal and N. Mustafa, "K-means projective clustering," *PODS '04: Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGAR symposium on Principles of database systems*, p. 155165, 2004.

[17] F. Gibou and R. Fedkiw, "A fast hybrid k-means level set algorithm for segmentation," *4th Annual Hawaii International Conference on Statisticsand Mathematics*, p. 281291, 2005.

[18] R. Herwig, A. Poustka, C. Muller, C. Bull, H. Lehrach, and J. O'Brien, "Large-scale clustering of cdna-fingerprinting data," *Genome Research*, p. 10931105, 1999.

[19] S. Ray and R. H. Turi, "Determination of number of clusters in k-means clustering and application in colour image segmentation," *Proc. 4th Int. Conf. Advances Pat. Rec. Digital Tech.*, 1999.

[20] J. Tou and R. Gonzalez, "Pattern recognition principles," *Massachusetts: Addison-Wesley*, 1974.

[21] A. K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.

[22] B. Zhou, M. Ha, and C. Wang, "An improved algorithm of unbalanced data svm," *Fuzzy Information and Engineering*, pp. 549–555, 2010.

[23] J. Bezdek and R. Ehrlich, "FCM: The fuzzy c-means clustering algorithm," *Computer and Geosciences*, vol. 10, no. 22, p. 191203, 1981.

[24] P. Maji and S. Pal, "Maximum class separability for rough-fuzzy c-means based brain mr image segmentation," *T. Rough Sets*, pp. 114–134, 2008.

[25] Z. Xian-cheng, "Image segmentation based on modified particle swarm optimization and fuzzy c-means clustering algorithm," *Second International Conference on Intelligent Computation Technology and Automation*, pp. 611–616, 2009.

[26] H. Izakian and A. Abraham, "Fuzzy c-means and fuzzy swarm for fuzzy clustering problem," *Expert Systems with Applications*, vol. 38, no. 3, pp. 1835–1838, 2011.

[27] G. V. Y.D. Valle, S. Mohagheghi, J. Hernandez, and R. Harley, "Particle swarm optimization: Basic concepts, variants and applications in power systems," *IEEE Trans. Evolutionary Computation*, pp. 171–195, 2008.

[28] P. Ghamisi and J. A. Benediktsson, "Feature selection based on hybridization of genetic algorithm and particle swarm optimization," *IEEE Geos. and Rem. Sens. Let.*, vol. 12, no. 2, pp. 309–313, 2015.

[29] P. Ghamisi, M. Couceiro, and J. Benediktsson, "A novel feature selection approach based on FODPSO and SVM," *IEEE Trans. Geosc. Remote Sens.*, vol. 53, no. 5, pp. 2935–2947, May 2015.

[30] A.-R. Ali, M. Couceiro, A. M. Anter, and A. E. Hassenian, "Evaluating an evolutionary particle swarm optimization for fast fuzzy c-means clustering on liver ct images," *IGI Glo. Comp. Vis. Image Proc. Intel. Sys. Multi. Tech.*, pp. 1–21, 2014.

[31] J. Tillett, T. M. Rao, F. Sahin, R. Rao, and S. Brockport, "Darwinian particle swarm optimization," *Proceedings of the 2nd Indian International Conference on Artificial Intelligence*, pp. 1474–1487, 2005.

[32] P. Ghamisi, M. S. Couceiro, N. M. F. Ferreira, and L. Kumar, "Use of darwinian particle swarm optimization technique for the segmentation of remote sensing images," *IGARSS*, 2012.

[33] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2013.

[34] M. Jiang, Y. Luo, and S. Yang, "Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm," *Inf. Proc. Let.*, vol. 102, no. 1, pp. 8–16, 2007.

[35] J. A. Peacock, "Two-dimensional goodness-of-fit testing in astronomy," *Monthly Notices Royal Astronomy Society*, vol. 202, pp. 615–627, 1983.

[36] J. Pallant, *SPSS Survival Manual*, 4th ed. Kindle Edition ed., Open University Press, 2011.

[37] M. S. Couceiro, F. M. Martins, R. P. Rocha, N. M. Ferreira, and S. Sivasundaram, "Introducing the fractional

order robotic darwinian pso," *AIP Conf. Proc.-American Inst. Phys.*, vol. 1493, no. 1, p. 242, 2012.

[38] P. Ghamisi, J. A. Benediktsson, and M. O. Ulfarsson, "Spectral-spatial classification of hyperspectral images based on hidden markov random fields," *IEEE Trans. Remote Sens. Geos.*, vol. 52, no. 5, pp. 2565–2574, 2014.

**Pedram Ghamisi** graduated with a B.Sc. degree in Civil (Survey) Engineering from the Tehran South Campus of Azad University. Then, he obtained the M.Sc. degree with first Class Honours in Remote Sensing at K.N.Toosi University of Technology in 2012. He received the Best Researcher Award for M.Sc. students in K. N. Toosi University of Technology in the academic year 2010-2011. Mr. Ghamisi was the recipient of the IEEE Mikio Takagi Prize which was awarded for the first place in the Student Paper Competition at the 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Melbourne, July 2013. He is currently a Ph.D. student in Electrical and Computer Engineering at the University of Iceland. He is currently a Ph.D. student in Electrical and Computer Engineering at the University of Iceland. His research interests are in Pattern Recognition, Machine Learning, remote sensing and image analysis with the current focus on spectral and spatial techniques for hyperspectral image classification and the integration of LiDAR and hyperspectral data for land cover assessment. He serves as a reviewer for a number of journals including IEEE TGRS, IEEE Trans. Image Processing, IEEE JSTARS, IEEE GRSL and Pattern Recognition Letters.

**Abder-Rahman Ali** received his BSc in Computer Science from the University of Jordan (Jordan), MSc Software Engineering from DePaul University (USA), and has also taken 2-years worth academic years of medical and health related courses in the UK, and is currently pursuing his Ph.D. degree in France. His interest in medical image analysis started during his masters degree studies, where he conducted research at the Intelligent Multimedia Processing Laboratory at DePaul University, relating to content based image retrieval for lung nodules. He is very passionate to the idea of applying computer science to medical imaging, and software engineering to medical device software systems, in an eventual goal to come up with algorithms and systems that aid in Computer Aided Diagnosis (CAD). He also likes to adopt fuzzy logic in his research.

**Micael Couceiro** obtained the BSc, Teaching Licensure and Master degrees on Electrical Engineering (Automation and Communications), at the Engineering Institute of Coimbra, Polytechnic Institute of Coimbra (ISEC-IPC). He is currently a PhD student on Electrical and Computer Engineering (Automation and Robotics) at the Faculty of Sciences and Technology, University of Coimbra (FCTUC), under a PhD Grant from the Portuguese Foundation for Science and Technology (FCT), having delivered his Thesis on September 2013. Over the past 5 years, he has been conducting scientific research on several areas besides robotics, namely computer vision, sports engineering, economics, sociology, digital media, and others, all at the Mobile Robotics Laboratory, Institute of Systems and Robotics (MRL-ISR), and at RoboCorp research group from IPC. This resulted on more than 20 scientific articles in international impact factor journals and more than 40 scientific articles at international conferences. Besides research, he has been invited for lecturing, tutoring and organization of events (*e.g.*, professional courses, national and international conferences, among others), both in the public and private domains.

**Jón Atli Benediktsson** received the Cand.Sci. degree in electrical engineering from the University of Iceland, Reykjavik, in 1984, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1987 and 1990, respectively. He is currently Pro Rector for Science and Academic Affairs and Professor of Electrical and Computer Engineering at the University of Iceland. His research interests are in remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing, and he has published extensively in those fields. Prof. Benediktsson was the 2011-2012 President of the IEEE Geoscience and and Remote Sensing Society (GRSS) and has been on the GRSS AdCom since 2000. He was Editor in Chief of the IEEE Transactions on Geoscience and Remote Sensing (TGRS) from 2003 to 2008 and has served as Associate Editor of TGRS since 1999, the IEEE Geoscience and Remote Sensing Letters since 2003 and IEEE Access since 2013. He is on the Editorial Board of the Proceedings of the IEEE, the International Editorial Board of the International Journal of Image and Data Fusion and was the Chairman of the Steering Committee of IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (J-STARS) 2007-2010. Prof. Benediktsson is a co-founder of the biomedical start up company Oxymap (www.oxymap.com). He is a Fellow of the IEEE and a Fellow of SPIE. Prof. Benediktsson is a member of the 2014 IEEE Fellow Committee. He received the Stevan J. Kristof Award from Purdue University in 1991 as outstanding graduate student in remote sensing. In 1997, Dr. Benediktsson was the recipient of the Icelandic Research Council's Outstanding Young Researcher Award, in 2000, he was granted the IEEE Third Millennium Medal, in 2004, he was a co-recipient of the University of Iceland's Technology Innovation Award, in 2006 he received the yearly research award from the Engineering Research Institute of the University of Iceland, and in 2007, he received the Outstanding Service Award from the IEEE Geoscience and Remote Sensing Society. He was co-recipient of the 2012 IEEE Transactions on Geoscience and Remote Sensing Paper Award and in 2013 he was co-recipient of the IEEE GRSS Highest Impact Paper Award. In 2013 he received the IEEE/VFI Electrical Engineer of the Year Award. In 2014 he was co-recipient of the 2012-13 Best Paper Award from the International Journal of Image and Data Fusion. He is a member of the Association of Chartered Engineers in Iceland (VFI), Societas Scientarium Islandica and Tau Beta Pi.