

Integration of Segmentation Techniques for Classification of Hyperspectral Images

Pedram Ghamisi, *Student Member, IEEE*, Micael S. Couceiro, *Student Member, IEEE*, Mathieu Fauvel and Jon Atli Benediktsson, *Fellow, IEEE*,

Abstract—A new spectral-spatial method for the classification of hyperspectral images is introduced. The proposed approach is based on two segmentation methods, Fractional-Order Darwinian Particle Swarm Optimization and Mean Shift Segmentation. The output of these two methods is classified by Support Vector Machines. Experimental results indicate that the integration of the two segmentation methods can overcome the drawbacks of each other and increases the overall accuracy in classification.

Keywords—hyperspectral image analysis, multilevel segmentation, mean shift segmentation.

I. INTRODUCTION

With recent advances in remote sensing technology, the spatial resolution of satellite images has become less than one meter [15]. The accurate classification of remote sensing images play a key role in many applications, including crop monitoring, forest applications, urban development, mapping and tracking or risk management. One way for achieving this goal would be to use the spectral and the spatial information sequentially [15]. The goal of considering spatial context in the classification step can be partially achieved by using some specific methods such as morphological filters [15] and Markov random fields [4]. The above-mentioned methods significantly increase the accuracy of the classification by incorporating spatial and spectral information. Another way for considering the spatial structures would be to perform image segmentation.

Image segmentation is a procedure which can be used to modify the accuracy of classification maps. To make such an approach effective, an accurate segmentation of the image is needed. A few methods for segmentation of multispectral and hyperspectral images have been introduced in the literature. Some of these methods are based on region merging methods, where neighboring segmented regions are merged with each other according to their homogeneity criterion, for instance *Multiresolution Segmentation Method* in the *eCognition* software is used this type of approach [5]. In [6], hierarchical segmentation algorithm is proposed, which performs region growing and spectral clustering alternately.

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

P. Ghamisi and J. A. Benediktsson are with the Faculty of Electrical and Computer Engineering, University of Iceland, Iceland. (e-mail: peg6@hi.is, benedikt@hi.is)

M. S. Couceiro is with the Engineering Institute of Coimbra, RoboCorp, 3030-199, Coimbra, Portugal (e-mail: micael@isec.pt).

M. Fauvel is with the DYNFOR lab, INRA and the University of Toulouse, France (e-mail: mathieu.fauvel@ensat.fr).

One of the best known methods for image segmentation is thresholding. Different types of optimal thresholding methods have been proposed in the literature (e.g., [16]). One strategy to find the optimal set of thresholds is to take into account an exhaustive search. A commonly used exhaustive search method is based on the Otsu criterion [1]. However, exhaustive search to find $n - 1$ optimal thresholds involves evaluation of the fitness for $n(L-n+1)^{n-1}$ combinations of thresholds [9]. Therefore, this method is not desirable from a computational point of view. Alternatively, the issue of determining $n - 1$ optimal thresholds for n -level image thresholding can be formulated as a multidimensional optimization problem. To solve the aforementioned issue, several biologically inspired algorithms have been explored in image segmentation [9].

One of the most commonly used methods based on split and merging segmentation is Mean Shift Segmentation (*MSS*) which is widely used in image processing. *MSS* is a nonparametric clustering technique which does not need embedded assumptions on the shape of the distribution and the number of clusters compared with the classic K-means clustering. *MSS* is a powerful method for segmentation of images with high redundancy [10] such as remote sensing images.

Fractional Order Darwinian Particle Swarm Optimization (*FODPSO*) segmentation (as all thresholding based methods in general) suffers from the following disadvantages: *i*) It cannot handle inhomogeneity; *ii*) it fails when the intensity of object of interest does not appear as a peak in the histogram; and *iii*) the traditional *FODPSO* based segmentation takes into account only the between class variance, thus disregarding any feedback from the within-class variance. In the *MSS* method, a kernel size needs to be tuned by the user. The tuning may be a difficult task and the final results may be affected by that dramatically.

In this letter, a new spectral-spatial classification approach is introduced for accurate classification of hyperspectral images. First, an input image will be segmented by Fractional Order Darwinian Particle Swarm Optimization (*FODPSO*). Then, the output of this step will be segmented again by *MSS*. At the end, the segmented image will be classified by Support Vector Machine (*SVM*). The paper is organized as follows: Methodology is discussed in Section II. Then, Section III is devoted to experimental results. Finally, in Section IV are outlined the main conclusions.

II. METHODOLOGY

The flowchart of the proposed method is illustrated in Fig. 1. The segmentation part consists of two different approaches: *i*) Multilevel thresholding method based on *FODPSO*; and *ii*) *MSS*. Then, the output of the segmenta-

tion methods will be classified by *SVM*. The following sub sections present a brief description of both segmentation methods.

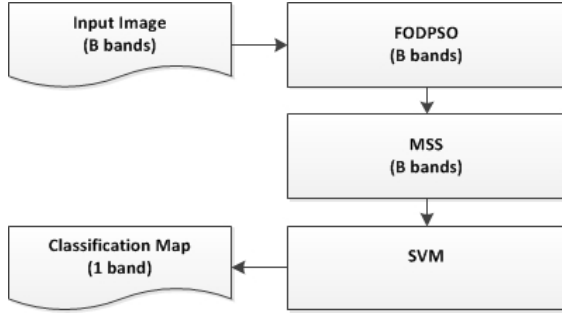


Fig. 1: Flowchart of the proposed methodology.

A. Multilevel thresholding method based on Fractional Order Particle Swarm Optimization (FODPSO)

Multilevel segmentation techniques provide an efficient way to perform image analysis. However, the automatic selection of a robust optimum n -level threshold has remained a challenge in remote sensing image segmentation. Let L represents the intensity levels in each component of a given image where a component is defined in the range $\{0, 1, 2, \dots, L-1\}$. Then, one can calculate the probability distribution p_i^C as:

$$p_i^C = \frac{h_i^C}{N}, \sum_{i=1}^N p_i^C = 1, \quad (1)$$

where i represents a specific intensity level, *i.e.*, $0 \leq i \leq L-1$, C represents the component of the pixel, *e.g.*, $C = \{R, G, B\}$ for *RGB* images, N represents the total number of pixels in the image and h_i^C denotes the number of pixels for the corresponding intensity level i in the component C . In other words, h_i^C represents an image histogram for each component C , which can be normalized and regarded as the probability distribution p_i^C .

Hence, the n -level thresholding presents $n - 1$ threshold levels t_j^C , $j = 1, \dots, n - 1$, and the operation is performed as:

$$F^C(a, b) = \begin{cases} 0, & f^C(a, b) \leq t_1^C \\ \frac{1}{2}(t_1^C + t_2^C), & t_1^C < f^C(a, b) \leq t_2^C \\ \vdots & \\ \frac{1}{2}(t_{n-2}^C + t_{n-1}^C), & t_{n-2}^C < f^C(a, b) \leq t_{n-1}^C \\ L, & f^C(a, b) > t_{n-1}^C \end{cases} \quad (2)$$

where a and b are the width (W) and height (H) pixel of the image of size $H \times W$ represented by $f^C(a, b)$ with L intensity levels for each component. The pixels of a given image will be divided into n classes D_1^C, \dots, D_n^C , which may represent multiple objects or even specific features on such objects (*e.g.*, topological features).

The simplest method of obtaining the optimal threshold is the

one that maximizes the between-class variance of each component which can be generally defined by:

$$\sigma_B^{c2} = \sum_{j=1}^n w_j^C (\mu_j^C - \mu_T^C)^2, \quad (3)$$

where j represents a specific class in such a way that w_j^C and μ_j^C are the probability of occurrence and mean of class j , respectively. The total mean value of a component is represented by μ_T^C .

For classes D_1^C, \dots, D_n^C , the probabilities of occurrence w_j^C and the means μ_j^C can be defined by (4) and (5), respectively:

$$w_j^C = \begin{cases} \sum_{i=1}^{t_j^C} p_i^C, & j = 1 \\ \sum_{i=t_{j-1}^C+1}^{t_j^C} p_i^C, & 1 < j < n, \\ \sum_{i=t_{j-1}^C+1}^L p_i^C, & j = n \end{cases} \quad (4)$$

$$\mu_j^C = \begin{cases} \sum_{i=1}^{t_j^C} \frac{ip_i^C}{w_j^C}, & j = 1 \\ \sum_{i=t_{j-1}^C+1}^{t_j^C} \frac{ip_i^C}{w_j^C}, & 1 < j < n, \\ \sum_{i=t_{j-1}^C+1}^L \frac{ip_i^C}{w_j^C}, & j = n \end{cases} \quad (5)$$

The problem of n -level thresholding is reduced to an optimization problem to search for the thresholds t_j^C that maximize the objective functions of each image component C , generally defined as:

$$\varphi^C = \max_{1 < t_1^C < \dots < t_{n-1}^C < L} \sigma_B^{c2}(t_j^C). \quad (6)$$

Computing this optimization problem involves a huge computational effort because the number of threshold levels and image components increase. Recently, biologically inspired methods, such as the well-known Particle Swarm Optimization (*PSO*), have been used as computationally efficient alternatives to analytical methods to solve optimization problems [13].

An example of such methods is the Fractional Order Particle Swarm Optimization (*FODPSO*) recently presented in [16]. This method is a natural extension of the Darwinian Particle Swarm Optimization (*DPSO*) presented by Tillet *et al.* in 2005 [14] using fractional calculus to control the convergence rate and was extended for the classification of remote sensing images in [8].

As in the classical *PSO*, particles within the *FODPSO* travels through the search space to find an optimal solution by interacting and sharing information with other particles. In each step of the algorithm, t , a fitness function is used to evaluate the success for a particle. To model the swarm, s , each particle, n , moves in a multidimensional space according to a position, $x_n^s[t]$, $0 \leq x_n^s[t] \leq L - 1$, and velocity, $v_n^s[t]$. The position and velocity values are highly dependent on the

individually best, $(\tilde{x}_n^s[t]$, and the globally best, $\tilde{g}_n^s[t]$, information:

$$\begin{aligned} v_n^s[t+1] &= \alpha v_n^s[t] + \frac{1}{2} \alpha v_n^s[t-1] + \frac{1}{6} \alpha (1 - \\ &\alpha) v_n^s[t-2] + \frac{1}{24} \alpha (1 - \alpha) (2 - \alpha) v_n^s[t-3] + \\ &\rho_1 r_1 (\tilde{g}_n^s - x_n^s[t]) + \rho_2 r_2 (\tilde{x}_n^s - x_n^s[t]), \end{aligned} \quad (7)$$

$$x_n^s[t+1] = x_n^s[t] + v_n^s[t+1]. \quad (8)$$

The coefficients ρ_1 and ρ_2 are weights which control the global and individual performance, respectively. Within the *FODPSO* algorithm, the inertial influence of particles depends on the fractional coefficient α . The parameters r_1 and r_2 are random vectors with each component is generally a uniform random number between 0 and 1.

When applying the *FODPSO* to multilevel thresholding of images, the particles' velocities are initially set to zero and their position is randomly set within the boundaries of the search space, *i.e.*, $v_n^s[0] = 0$ and $0 \leq x_n^s[0] \leq L - 1$. In other words, the search space depends on the number of intensity levels L , *i.e.*, if one wishes to perform a segmentation of a 8-bit image, then particles will be deployed between 0 and 255. Hence, associated to each particle, a possible solution φ^C will be found and compared between all particles of the same swarm. The particle that has found the higher between-class variance φ^C so far will be the best performing one (*i.e.*, $\tilde{g}_n^s[t]$), thus luring other particles toward it. It is also noteworthy that when a particle improves, *i.e.*, when a particle is able to find a higher between-class variance φ^C from one step to another, the fractional extension of the algorithm outputs a higher exploitation behavior. This allows achieving an improved collective convergence of the algorithm, thus allowing a good short-term performance.

FODPSO is a promising method to specify a predefined number of clusters with a higher between class variance. In [9], the authors demonstrated that the *FODPSO* based segmentation method performs considerably better in terms of accuracies than *Genetic Algorithm (GA)*, *Bacterial Algorithm (BA)*, *PSO* and *DPSO*, thus finding different number of clusters with a higher between-class variance and more stability in less computational processing time. For further information on the *FODPSO* algorithm please refer [9, 16].

B. Mean Shift Segmentation

MSS is a nonparametric clustering technique which requires neither embedded assumptions on the shape of the distribution nor the number of clusters in comparison to the classic K-means clustering approach. Mean shift was firstly introduced in [11]. This approach has been more recently developed for different purposes of low-level vision problems, including, adaptive smoothing and segmentation [10].

In the *MSS*, each pixel is associated with the most significant mode of the joint domain density located in its neighborhood, after nearby modes were pruned as in the generic feature space analysis technique [12]. Let us assume that we have N data points $X_i, i = 1, \dots, N$ in a d -dimensional space R^d . Then, the multivariate kernel density estimate obtained with a radially symmetric kernel, $K(X)$, and window radius, h , is:

$$f(X) = \frac{1}{Nh^d} \sum_{i=1}^N K\left(\frac{X-X_i}{h}\right), \quad (10)$$

where h is the kernels size parameter, presents the radius of kernel. The radially symmetric kernel $K(X)$ can be defined as:

$$K(X) = c_k k(\|X\|^2), \quad (11)$$

where c_k is a normalization constant. The modes of the density function are located at the zeros of the gradient function, *i.e.*, where $\nabla f(x) = 0$. After some algebraic manipulation, the gradient of the density estimator can be written as follows:

$$\nabla \hat{f}(X) = \frac{2c_k d}{Nh^{d+2}} \sum_{i=1}^N g\left(\left\|\frac{X-X_i}{h}\right\|^2\right) \left[\underbrace{\sum_{i=1}^N X_i g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)}_{\text{term 1}} - \underbrace{\sum_{i=1}^N g\left(\left\|\frac{X-X_i}{h}\right\|^2\right)}_{\text{term 2}} X \right], \quad (12)$$

where $g(X) = -k'(X)$ which gives the derivative of the selected kernel profile. The first term of (12) gives information regarding the density estimate at x computed with kernel $G(X) = c_g g(\|X\|^2)$. The second term is regarded as the mean shift vector, m , which points toward the direction of the maximum increase in density and is related to the density gradient estimate at point x obtained with kernel K . In follows, the way of calculating the mean shift for a given point X_i will be shown (*cf.*, Fig. 1):

- i. Computation of the mean shift vector $m(X_i^t)$.
- ii. Translation of the window: $X_i^{t+1} = X_i^t + m(X_i^t)$.
- iii. Iterate steps i and ii, until convergence, *i.e.*, til $\nabla f(X) = 0$.

The most important limitation of the standard *MSS* is that the value of the kernel size is unspecified. More information regarding the *MSS* can be found in [10].

III. EXPERIMENTAL RESULTS

A. Description of data sets

Pavia data: The first test case is a hyperspectral data set captured on the city of Pavia, Italy, by airborne data from the *ROSIS-03*. In our experiments, 12 noisy bands were eliminated and 103 bands were processed. The spatial resolution is 1.3 m per pixel. The original data set is 610 by 340 pixels. This data set consisted of different classes, including trees, asphalt, bitumen, gravel, metal sheet, shadow, bricks, meadow and soil. Fig. 2a and b depict Pavia data set and its reference map.

Salinas data: This scene was captured by AVIRIS sensor over Salinas Valley, California, and is characterized by high spatial resolution (3.7-meter pixels) consisting 512 lines and 217 samples. It includes vegetables, bare soils, and vineyard fields. The Salinas reference data contains 16 classes. Fig. 2c and d show the Salinas data set and its corresponding reference map.

The data sets have been classified with *SVM* and a Gaussian kernel. The hyper parameters have been selected using 5-fold cross validation. The training set was randomly composed of 12.5% of the referenced set, the experiments have been repeated 20 times, and the mean accuracy and the standard deviation have been reported in Table II and IV.

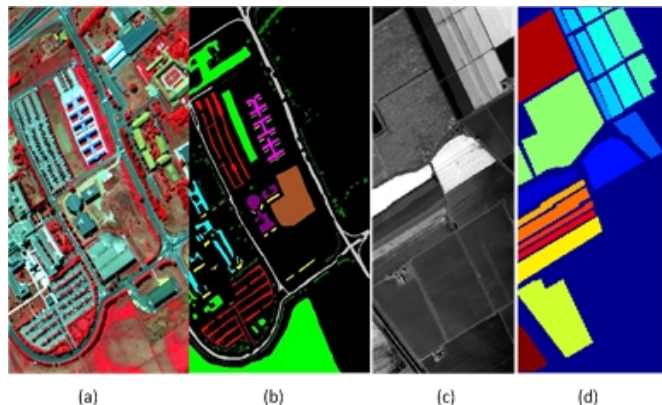


Fig. 2: An example of our test cases. a) False color composition of Pavia data set; and b) reference map where each color represents a specific class c) Salinas data set d) reference map where each color represents a specific class.

The proposed multilevel thresholding techniques based on *FODPSO* were implemented with the specific parameters shown in Table I for our test cases. These parameters are chosen based on some studies from [3], [16].

Table I: Initial parameters of the *FODPSO* for our data sets

I_T	N	ρ_i, ρ	Δv	N_{min}	N_{max}	N^s	N_{min}^s	N_{max}^s	N_{kill}	α
10	3	0.8	5	10	50	4	2	6	10	0.
0	0									6

Since the *MSS* approach is very dependent on the kernel size, two different kernel sizes were selected (5 and 20) in the experiments. The experimental evaluation will demonstrate if the proposed method is highly dependent on the size of kernel or not.

B. Results and Discussion

1) Pavia data set

Table II illustrates the kappa coefficient and overall accuracy for different methods for the Pavia data set.

As can be observed from Table II, *FODPSO+SVM* gave comparatively the worst performance in terms of accuracies. In histogram based methods, the spatial information of data such as size and shape are not taken into consideration and the final result is spatially independent and can be determined by considering only the histogram of the data. On the contrary, the *MSS+SVM* outperforms the *FODPSO+SVM* in terms of accuracies since it does not suffers from the above-mentioned disadvantages and can handle images with more complexity such as remote sensing images in a significant way. As can be seen in the table, *FODPSO+MSS+SVM* gave comparatively the best accuracies. Fig. 3 shows the output of classification for different methods.

To further improve the comparison between *MSS+SVM* and *FODPSO+MSS+SVM*, the significance of the method on the

overall accuracy (OA) and the kappa coefficient (dependent variables) was analyzed using the Multivariate Analysis of Variance (*MANOVA*) technique after checking the assumptions of multivariate normality and homogeneity of variance/covariance. This is a statistical test procedure that allows comparing multivariate means of several groups. In other words, it allows comparing different methods (as it is the case) with more than one dependent variable (i.e., overall accuracy and the kappa coefficient). In other words, the *MANOVA* merges the multiple dependent variables, thus creating a single dependent variable. For more information regarding *MANOVA*, it is referred to [2].

Table II. The kappa coefficient and overall test accuracy of different methods for the Pavia data set

Methods	Kernel size	Mean(Overall accuracy)	Mean(kappa coefficient)
<i>FODPSO+SVM</i>		90.8 ± 0.192	0.887 ± 0.003
<i>MSS(R=5)+SVM</i>	R=5	98.84 ± 0.079	0.985 ± 0.001
<i>FODPSO+MSS(R=5)+SVM</i>		98.92 ± 0.106	0.986 ± 0.001
<i>MSS(R=20)+SVM</i>	R=20	97.72 ± 0.120	0.970 ± 0.002
<i>FODPSO+MSS(R=20)+SVM</i>		98.04 ± 0.128	0.974 ± 0.002
<i>SVM</i>		94.32 ± 0.174	0.925 ± 0.002

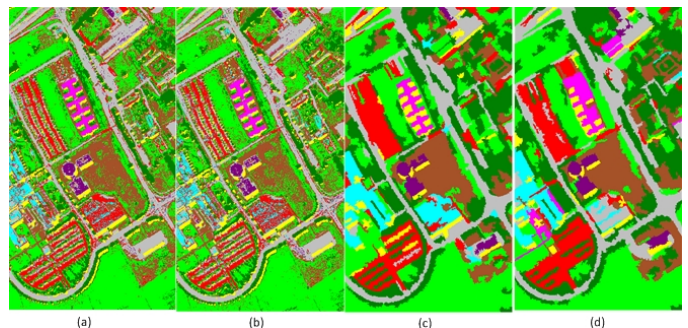


Fig. 3: Pavia classification result for a) Original image b) *FODPSO* c) *MSS* d) *FODPSO+MSS*.

The assumption of normality for each of the univariate dependent variables was examined using univariate tests of *Kolmogorov-Smirnov* (p -value < 0.05). When the *MANOVA* detected significant statistical differences, we proceeded to the commonly-used *ANOVA* for each dependent variable followed by the *Tukey's HSD Post Hoc*. The classification of the size effect (i.e., measure of the proportion of the total variation in the dependent variable explained by the dependent variable) was done according to Maroco [7] and Pallant [2]. This analysis was performed using *IBM SPSS Statistics* for a significance level of 5%.

Table III. *MANOVA* results for the Pavia data set.

Method	kappa	OA
<i>FODPSO+MSS(R=20)+SVM</i> vs <i>MSS(R=20)+SVM</i>	.004*	.003*

* The corresponding mean difference is significant at the 0.05 level

All p -values corresponding to the mean differences are equal to 0.001

A two-way *MANOVA* analysis was carried out to assess whether the algorithms used on this study have statistically significant differences with respect to the classification process. The *MANOVA* analysis revealed that the dependent variable *kappa coefficient* presents statistically significant

differences with large effect ($F_{(1,38)} = 66.656$; p -value = 0.001; $\eta_p^2 = 0.637$; $Power = 1.0$), as well as the dependent variable overall accuracy ($F_{(1,38)} = 66.491$; p -value = 0.001; $\eta_p^2 = 0.636$; $Power = 1.0$) (See Table III).

2) Salinas data set

For the Salinas data set, *FODPSO+SVM* gave the worst accuracies (Table IV). Furthermore, the overall classification accuracy by *MSS+SVM* dropped from 99.14% to 94.76% when the kernel size was increased from 5 to 20. This dramatic decrease in accuracy shows that the result of the classification by using *MSS* is highly dependent on the kernel size. The kernel size must still be tuned by user who might find the task difficult since the size can dramatically influence the final result. A larger or smaller kernel may influence the result of the segmentation and considerably reduce the efficiency of the *MSS* method. Mode candidates with a distance that is less than the kernel size are merged and may cause to lose information on an image. In contrast, a small kernel size may cause a high increase on the *CPU* processing time. As the *FODPSO* is able to find modes with maximum between-class distance, the influence of tuning the size of the kernel size is significantly reduced. In other words, the two methods can solve each other's problems and, thus, complement each other. Fig. 4 illustrates the result of the classification for the different methods. Compared to the *MSS+SVM*, the *FODPSO+MSS+SVM* increased the accuracy from 94.76% to 96.27% with the kernel size of 20. Considering both kernel sizes ($R=5$ and $R=20$), it can be stated that the *FODPSO+MSS+SVM* shows more stability. *FODPSO+MSS+SVM* improved the result of the traditional *SVM* by almost 5 and 2 percent by considering kernels with the size of 5 and 20, respectively (Table IV).

Table IV. The kappa coefficient and overall test accuracy of different methods for the Salinas data set

Methods	Kernel size	Mean(Overall accuracy)	Mean(kappa coefficient)
<i>FODPSO+SVM</i>		91.97 ± 0.17	0.91 ± 0.0020
<i>MSS(R=5)+SVM</i>	R=5	99.14 ± 0.05	0.99 ± 0.0005
<i>FODPSO+MSS(R=5)+SVM</i>		99.13 ± 0.03	0.99 ± 0.0004
<i>MSS(R=20)+SVM</i>	R=20	94.76 ± 0.19	0.94 ± 0.0021
<i>FODPSO+MSS(R=20)+SVM</i>		96.27 ± 0.11	0.958 ± 0.0013
<i>SVM</i>		94.06 ± 0.13	0.93 ± 0.0014

For the Salinas data set, the *MANOVA* was separately performed on *MSS (R=5)+SVM*, *FODPSO+MSS(R=5)+SVM* and *MSS (R=20)+SVM*, *FODPSO+MSS(R=20)+SVM* (Table V). The *MANOVA* analysis revealed that the *kappa coefficient* does not present statistically significant differences ($F_{(1,38)} = 0.415$; p -value = 0.523; $\eta_p^2 = 0.011$; $Power = 0.96$). A similar result was observed for the overall accuracy ($F_{(1,38)} = 0.386$; p -value = 0.538; $\eta_p^2 = 0.010$; $Power = 0.93$).

For $R=20$, the *MANOVA* analysis depicted that the dependent variable *kappa coefficient* presents statistically significant differences with large effect ($F_{(1,38)} = 1111.827$; p -value = 1.00; $\eta_p^2 = 0.967$; $Power = 1.00$), as well as the dependent variable overall accuracy ($F_{(1,38)} = 1112.876$; p -value = 0.001; $\eta_p^2 = 0.967$; $Power = 1.00$). Results show that by increasing the size of the kernel, the proposed method works better than others in terms of accuracies.

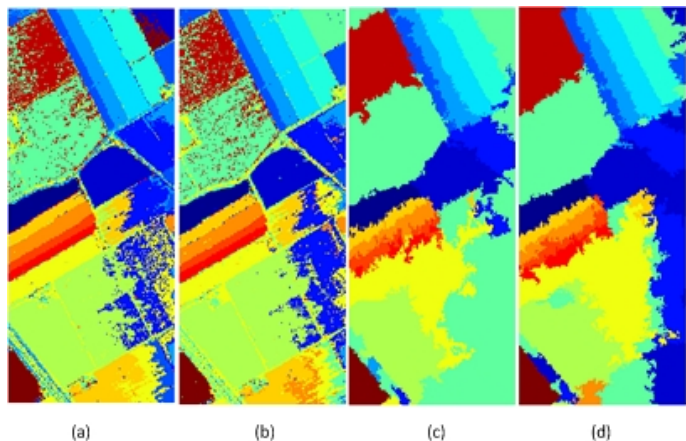


Fig. 4: Salinas classification result for a) Original image b) FODPSO c) MSS d) FODPSO+MSS.

Table V. MANOVA results for the Salinas data set.

Method	kappa	OA
<i>FODPSO+MSS(R=5)+SVM</i> vs <i>MSS(R=5)+SVM</i>	$-9.258E-5$	$-8.023E-5$
<i>FODPSO+MSS(R=20)+SVM</i> vs <i>MSS(R=20)+SVM</i>	0.017*	0.015*

* The corresponding mean difference is significant at the 0.05 level

All p -values corresponding to the mean differences are equal to 0.001

IV. CONCLUSION

In this paper, a new spectral-spatial classification approach is introduced for the accurate classification of hyperspectral images. The approach is based on the combination of Fractional Order Darwinian Particle Swarm Optimization (*FODPSO*) and Mean Shift Segmentation (*MSS*). *FODPSO* is a very powerful approach for finding the predefined number of clusters with the highest between-class value. In the proposed approach, the result of *FODPSO* is used as the input to *MSS* to develop a pre-processing method for classification. Tuning the size of the kernel can be considered as the main difficulty of *MSS* and the obtained result may considerably be affected by the kernel size. The *SVM* is used for classification on the outcome of these two segmentation methods. Results indicate that the use of both segmentation methods can overcome the shortcomings of each other and the combination can improve the result of classification significantly.

ACKNOWLEDGMENT

The ROSIS Pavia data on were obtained from Prof. Paolo Gamba of the University of Pavia, Italy. This work was in part supported by the Icelandic Research Fund for Graduate Students and PhD scholarship SFRH/BD/73382/2010 granted by the Portuguese Foundation for Science and Technology (FCT).

REFERENCES

- [1] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. on Sys., Man, Cyb.*, SMC-9, 62–66, 1979.
- [2] J. Pallant, *SPSS Survival Manual*, Open University Press; 4 edition, 2011.
- [3] M. S. Couceiro, F. M. L. Martins, R. P. Rocha & N. M. F. Ferreira.

- “Analysis and Parameter Adjustment of the RDPSO - Towards an Understanding of Robotic Network Dynamic Partitioning based on Darwin's Theory”, *International Mathematical Forum*, Hikari, Ltd., Vol. 7, Issue 32, pp. 1587-1601, 2012
- [4] A. Farag, R. Mohamed and A. El-Baz, “A Unified Framework for Map Estimation in Remote Sensing Image Segmentation,” *IEEE Trans. Geosci. Remote Sens.* Vol. 43, no. 7, pp. 1617–1634, 2005.
- [5] A. Darwish and K. Leukert, W. Reinhardt, “Image Segmentation for the Purpose of Object-based Classification,” *Proceedings of IGARSS'03*, vol. 3, pp. 2039–2041, 2003.
- [6] J. Tilton, “Analysis of Hierarchically Related Image Segmentations,” *Proceedings of IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pp. 60–69, 2003.
- [7] J. Maroco, *Análise Estatística com utilização do SPSS*. Lisboa: *Edições Silabo*, 2010.
- [8] 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,” *Proceedings IGARSS*, 2012, Germany.
- [9] 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 Systems with Applications*, vol. 39, pp. 12407–12417, 2012.. DOI: 10.1016/j.eswa.2012.04.078
- [10] D. Comaniciu and P. Meer, “Mean Shift: A Robust Approach Toward Feature Space Analysis,” *IEEE Trans. Pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603-619, 2002.
- [11] K. Fukunaga, K. and L. Hostetler, “The Estimation of The Gradient of a Density Function, with Applications in Pattern Recognition,” *IEEE Transactions on Information Theory*, vol. 21, pp. 32–40, 1975
- [12] D. Comaniciu, V. Ramesh, and P. Meer, “Kernel-based object tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 564–577, 2003
- [13] M. S. Couceiro, J. M. A. Luz, C. Figueiredo .and N. M. F. Ferreira, “Modeling and Control of Biologically Inspired Flying Robots,” *Robotica*, vol. 30, Issue 1, pp. 107-121, *Camb. University Press*, 2012.
- [14] 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.
- [15] 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, March 2013, doi: 10.1109/JPROC.2012.2197589
- [16] P. Ghamisi, M. S. Couceiro, and J. A. Benediktsson, “Extending the Fractional Order Darwinian Particle Swarm Optimization to Segmentation of Hyperspectral Images,” *SPIE Remote Sensing*, Edinburgh, UK, September 2012.