Train Support Vector Machine Using Fuzzy C-means Without a Priori Knowledge for Hyperspectral Image Content Classification

Akar H. Taher

Department of Software Engineering, Faculty of Engineering, Koya University,
Koya KOY45, Kurdistan region - F.R. Iraq

***Abstract*—In this paper, a new cooperative classification method called auto-train support vector machine (SVM) is proposed. This new method converts indirectly SVM to an unsupervised classification method. The main disadvantage of conventional SVM is that it needs a priori knowledge about the data to train it. To avoid using this knowledge that is strictly required to train SVM, in this cooperative method, the data, that is, hyperspectral images (HSIs), are first clustered using Fuzzy C-means (FCM); then, the created labels are used to train SVM. At this stage, the image content is classified using the auto-trained SVM. Using FCM, clustering reveals how strongly a pixel is assigned to a class thanks to the fuzzification process. This information leads to gaining two advantages, the first one is that no prior knowledge about the data (known labels) is needed and the second one is that the training data selection is not done randomly (the training data are selected according to their degree of membership to a class). The proposed method gives very promising results. The method is tested on two HSIs, which are Indian Pines and Pavia University. The results obtained have a very high accuracy of the classification and exceed the existing manually trained methods in the literature.**

***Index Terms*—Automatic training, Clustering, Cooperative classification, Fuzzy C-means, Support Vector Machine.**

I. Introduction

Nowadays, hyperspectral image (HSI) classification attracts the attention of researchers due to the rich information they contain. Moreover, this type of image can be used in many applications for the same reason. Among the applications of HSI, mining and geology (Goetz, et al., 1985), ecology (Ryan, et al., 2014), civil or military surveillance (Lagueux, et al., 2012), agriculture (Lacar, Lewis and Grierson, 2001), medicine (Akbari, et al., 2010), food safety and quality (Feng and Sun, 2012), and teledetection (Tarabalka, et al., 2010; Cariou, Moan and Chehdi, 2020; Alameddine, Chehdi and Cariou, 2021; Cariou, Le Moan and Chehdi, 2022; Dong, et al., 2022; Sellami and Tabbone, 2022; Sun, et al., 2022) can be listed. The problem with most of the methods used for HSI classification is that they need *a priori* knowledge to train the classifier.

In the classification process, each pixel vector of the HSI must be given a distinct label. In the past two decades, a variety of pixel-wise (spectral-based) techniques has been used to solve this problem, including k-nearest neighbors (Bruzzone and Cossu, 2002), support vector machine (SVM) (Bruzzone and Cossu, 2002), and sparse representation (Chen, Nasrabadi and Tran, 2013). Among the vast number of classification methods, SVM has relatively demonstrated good performance for identifying high-dimensional data even when a small number of training samples are available (Camps-Valls and Bruzzone, 2005). In HSI classification, SVM can successfully overcome the Hughes phenomena (Hughes, 1968) and the problem of small training sample sizes. As a result, SVM and its enhanced algorithms perform better than other approaches. However, the problem with these approaches is that they strictly need previously labeled data to train the SVM. These labeled data are not available in all cases and not an easy task to obtain.

In Guo, et al., 2019, the SVM is used with a guided filter, in which two fusion methods are used to combine spectral and special features. In Shang, et al., 2022, another SVM-based method is proposed, it also contains a step of filtering. The problem with these methods is that they use a filter that may cause information loss, and it needs some parameters to be fixed for the filtering process. In Pathak and Kalita, 2019, another spectral-spatial SVM-based classification is presented, in this methods, a sliding window of fixed size is used to extract the spatial feature; however, the size of the window may affect the efficiency of the method. In Li, Li and Pan, 2019, SVM is combined with deep learning, and the results obtained using this method are very interesting. The disadvantage of using deep learning is that it dramatically increases the number of features, leading to a very high computation time. Many other SVM-based methods are proposed in the literature to classify the contents of HSI (Tarabalka, et al., 2010; Awad and Khanna, 2015; Wu, et al., 2016; Guo, et al., 2019; Li, Li and Pan, 2019; Pathak and

ARO-The Scientific Journal of Koya University
Vol. X, No. 2 (2022), Article ID: ARO.11025. 7 pages
DOI: 10.14500/aro.11025
Received: 04 July 2022; Accepted: 28 August 2022
Regular research paper: Published: 10 September 2022
Corresponding author's email: akar.taher@koyauniversity.org

Kalita, 2019; Zhao, et al., 2020; Mounika, et al., 2021; Ren, et al., 2021; Shang, et al., 2022). The problem of all the above-mentioned methods is: (1) They need labeled data *(a priori* knowledge) to train the SVM and (2) the training data are selected randomly. To this end, the main contribution of this study and the proposed method is to overcome these two disadvantages, without losing the classification rate accuracy.

The remaining of this paper is structured as follows: after this introductory section, section II explains the theory of both SVM and FCM, in section III the proposed method is introduced, then the section gives details about the datasets used in this research to validate the proposed method. The final part of this same section gives the results obtained using this newly proposed method, the method is called: Auto Train Support Vector Machine (ATSVM). Further, the results obtained using this new technique are compared to those obtained using classic SVM only and other SVM-based methods found in the literature, which are all trained with Ground Truth labels. In section IV the conclusions are given.

II. Basic Elements

To better understand this approach, its' basic elements, that is, FCM and SVM, are explained in detail below:

A. Fuzzy C-means (FCM)

FCM (Bezdek, Ehrlich and Full, 1984) is a fuzzified version of K-means clustering (MacQueen, 1967) which adds a fuzzification operation, this fuzzification gives the ability to solve more complex clustering problems (Li, et al., 2009). More clearly, this method assigns a class membership to a data point, which depends on the distance of the data point to a particular class compared to all other classes. FCM tries to minimize the objective function below:

$$J = \sum_{j=1}^{NC}\sum_{i=1}^{N} u_{ij}^{m}\left(F_i - g(c_j)\right)^2 \tag{1}$$

with the constraint:

$$\sum_{j=1}^{NC} u_{ij}^{m} = 1 \quad \forall i \tag{2}$$

Where:

- NC is the number of classes,
- N is the number of pixels in the dataset (image),
- F_i is the vector of Nf features representing the pixel x_i,
- $g(C_j)$ is the center of gravity of class C_j,
- $m \in [1,\infty]$: is the fuzzification factor,
- and u_{ij} represents the entry (i, j) of the partition matrix, with $0 \le u_{ij} \le 1$.

By giving high membership values to data points near the center of their clusters, the objective function is minimized; in contrast, low membership values are provided to data points far from the class center. The following equations update the class centers and membership functions:

$$g(c_j) = \sum_{i=1}^{N} \frac{u_{ij}^{m}}{\sum_{k=1}^{N} u_{ik}^{m}} F_i \tag{3}$$

$$u_{ij} = \frac{\left\| F_i - g(c_j) \right\|^{\frac{1}{m-1}}}{\left\| \sum_{j=1}^{NC} F_i - g(c_j) \right\|^{\frac{1}{m-1}}} \tag{4}$$

The four steps of FCM can be centered on the following phases:

Step 1: Set the membership matrix $U = \left[u_{ij}\right],\ 1 \le j \le NC,\ 1 \le i \le N$ randomly between 0 and 1 and satisfy the condition in Equation (2).

Step 2: Find cluster (class) centers $g(C_j)$ using Equation (3).

Step 3: Update the degree of membership u_{ji} using Equation (4).

Step 4: Redo *steps 2* and *3* until the algorithm converges (negligible difference between the current membership matrix and the previous membership matrix or the maximum number of iterations).

B. SVM

SVM is a powerful classification method that was initially developed by Boser, Guyon, and Vapnik (Boser, Guyon and Vapnik, 1992; Vapnik, 1995). On the other hand, SVMs are a collection of connected supervised techniques used for regression and classification (Gove and Faytong, 2012). More clearly, SVM is a classification and regression prediction method that automatically avoids over-fitting while maximizing predictive accuracy using machine learning theory. SVMs can be characterized as systems that use a high-dimensional feature space as the hypothesis space for a linear function that is trained using an optimization theory-based learning technique which incorporates a learning bias.

A classification process is composed of training and testing data that consist of some data samples (Duda and Hart, 1973). Each sample of data in the training group contains one target value and several features. SVM aims to create a model which predicts the target value of data samples in the testing group in which they contain the features only (Cristianini and Shawe-Taylor, 2000). Being a supervised approach, SVM relies on known labels to determine whether the system is operating properly. This information gives the desired response, validating the accuracy and the efficiency of the system. The first step in SVM classification involves determining whether characteristics have a close relationship to the recognized classes. This is referred to as feature selection or extraction. Feature selection and SVM classification together can be deployed to identify important elements that are involved in whatever processes recognize the classes or not (Cristianini and Shawe-Taylor, 2000).

http://dx.doi.org/10.14500/aro.11025

III. Proposed Method and Datasets Used for Validation

A. FCM and SVM in Cooperation

In this article, a cooperative approach that combines FCM and SVM is proposed in which the data are first clustered through using FCM, then the obtained class labels are used to train the SVM (see the diagram in Fig. 1).

In this proposed approach, after clustering the datasets using FCM, the obtained labels are used to train SVM instead of using known labels coming with datasets (known labels are not available in all cases). The choice of these two methods to cooperate (i.e., FCM and SVM) is not done arbitrarily. First, the reason behind choosing FCM is that the fuzzy decision gives very important information about the classification of each data point (pixel) in the image, that is, the degree of membership of each pixel to the specific clusters. Thereafter, this information is used to choose the pixels which are used to train the SVM. Second, the reason behind choosing SVM is that this method has shown very promising results in the classification of high-dimensional data and HSIs as we mentioned before in the introductory section.

B. Datasets

To validate the results of the proposed method, the Pavia University and the Indian Pines HSIs are used, as they are very well known and widely used HSIs.

Pavia University dataset

The Reflective Optics System Imaging Spectrometer (ROSIS-03) optical sensor was used to capture this image of an urban area. According to the specs, the ROSIS-03 sensor captures 115 bands with a spectral coverage of 0.43–0.86 μm. Each pixel has a spatial resolution of 1.3 m. The test site was close to the University of Pavia's Engineering School in Pavia, Italy. The pixels are 610 by 340. Due to noise, 12 channels were eliminated. Processing was done on the remaining 103 spectral channels. Nine classes of interest were considered: Tree, asphalt, bitumen, gravel, metal sheet, shadow, bricks, meadow, and soil (Fig. 2) (Tarabalka, et al., 2010).

Indian Pines dataset

This HSI was captured by the Airborne Visible/Infrared Imaging Spectrometer sensor over the agricultural Indian Pine test site in Northwest Indiana in the USA. The spatial dimension of the image is 145 by 145 pixels and has a 20 m per pixel spatial resolution. The spectral dimension is 224 components. The number of components is reduced to 200 by removing components covering the region of water

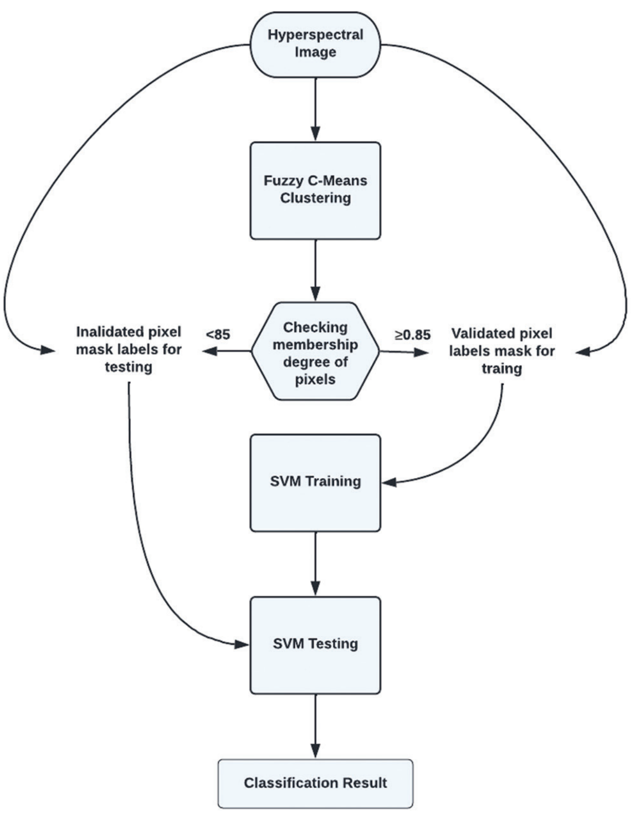

Fig. 1. Proposed method (ATSVM) flow diagram.

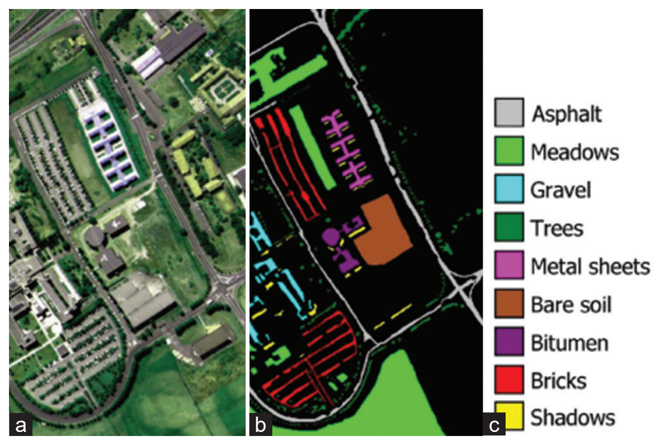

Fig. 2. Pavia University image. (a) Three-band color composites. (b) Ground truth. (c) Color code and class names.

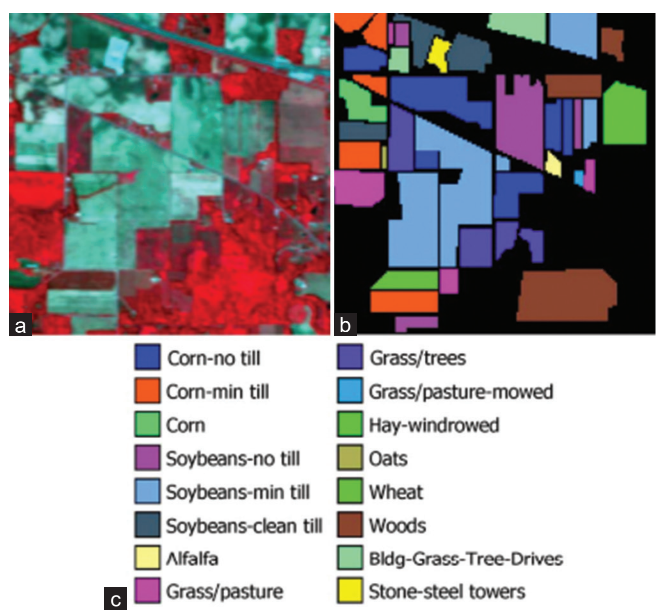

Fig. 3. Indian Pines image. (a) Three-band color composites. (b) Ground truth. (c) Color code and class names.

absorption: [104–108], [150–163], and 220. Sixteen classes of interest were considered: Alfalfa, Corn-notill, Corn-mintill, Corn, Grass-pasture, Grass-trees, Grass-pasture-mowed, Hay-windrowed, Oats, Soybean-notill, Soybean-mintill, Soybean-clean, Wheat, Woods, Buildings-Grass-Trees-Drives, and Stone-Steel-Towers (Fig. 3) (Tarabalka, et al., 2010).

C. Results

Algorithms used and fixing their parameters

To test the proposed approach, the FCM and SVM algorithms of MATLAB™ version 2021 are used. The fixed parameter for each algorithm is given in Tables I and II.

Table I
FCM Fixed Parameters

FCM
Distance: Euclidian
Number of iterations: 100
Fuzzification parameter (m): 2
Tolerance: 10e-5
Number of classes: 16 for Indian Pines and 9 for Pavia University
Validation index: 0.85 (new parameter not exiting in standard FCM, added by the proposed method)

Table II
SVM Fixed Parameters

SVM
Model type:
o Preset: Medium Gaussian SVM
o Kernel function: Gaussian
o Box constraint level: 1
o Multiclass method: One-versus-One
o Standardize data: true
Optimizer options:
o Hyperparameter options disabled

Feature selection: Disabled

Table III
FCM Classification Example (U)

	Class 1	Class 2	Class 3
Pixel 1	0.9	0.03	0.07
Pixel 2	0.8	0.09	0.11
Pixel 3	0.02	0.88	0.1

Validation process

To validate the proposed method (ATSVM), the method is applied to the previously presented datasets. First, the data are organized in a matrix of (*SxF*) format (samples in rows and features in columns). Then, data are clustered with the FCM algorithm. The output of FCM is a fuzzy decision for each data sample (pixels in the HSI). In the proposed method, a validation parameter index is introduced which is fixed to 0.85. This parameter is used to select the data points used for training the SVM algorithm (the data points with membership degree ≥0.85 are chosen for training). More clearly, this index is compared to the degree of membership of each pixel after the FCM clustering, for example, if the whole data are clustered to three classes using FCM, each pixel will have a degree of membership to the three class and the summation of all the degrees of a pixel is equal to one (Equation 2 and Table III). The reason behind using the pixels with a high degree of membership to a class is that in the case classification, confidence is high, for more clarification, the example in Table III. In this case, Pixels 1 and 3 are chosen for training SVM, but Pixel 2 is not chosen as its greatest degree of membership is smaller than the fixed threshold (0.85).

After choosing the pixels with high confidence of classification, their classification result is defuzzied. This is done by giving the label of where the membership degree is the largest, by this step, each pixel will get a label that indicates the class they belong to. In the example in Table III, Pixel 1 will get Label 1 and Pixels 3 will get Label 2, as the maximum membership degree is in Classes 1 and 2, respectively. These labels are used to train the SVM. Normally, these labels need to be known *a priori*, but in this proposed method, they are created by the FCM algorithm to train the SVM. At this point, the SVM algorithm is trained using these created labels. The pixels that are not used for training (as Pixel 2 in Table III) are used for testing.

The results of ATSVM for Pavia University and Indian Pines are shown in Table IV and it is observed that the proposed methods have the correct classification rate equal to 0% for some classes. This is because these classes contain a little number of samples and the SVM classier is not trained sufficiently by this number of samples. This problem is no unique, and it is repeated in other methods found in the literature (Tables V and VI). To show the efficiency of the

table IV
Confusion Matrix of the Class-Specific Accuracy (Csa) for Pavia University His Using ATSVM

Class	Asphalt	Meadows	Gravel	Trees	Painted metal sheets	Bare Soil	Bitumen	Self-blocking bricks	Shadows
Asphalt	8134	34	0	0	0	13	0	17	259
Meadows	123	8783	0	61	46	0	1	10	95
Gravel	0	0	981	0	0	0	0	0	0
Trees	0	0	202	6790	160	0	0	22	0
Painted metal sheets	0	0	154	0	963	0	0	0	0
Bare soil	0	0	0	0	0	949	0	0	0
Bitumen	0	0	1	0	0	0	3110	0	0
Self-blocking bricks	0	0	0	3	3	1	0	7184	0
Shadows	0	11	0	0	0	0	165	0	4501
CSA%	96.18068	96.31539	100	94.64734	86.21307	100	99.96786	99.90266	96.2369

 http://dx.doi.org/10.14500/aro.11025

Table V
Comparison of ATSVM Applied on Pavia University Hsi with Other Methods Found in the Literature

Class names	ATSVM	SVM	SVM+Watershed	SVM+PartClust	SVM-HSEG	GS-SVM
Asphalt	96.18	84.93	93.64	90.72	94.77	93.50
Meadows	96.31	70.79	75.09	92.73	89.32	95.50
Gravel	100	67.16	66.12	82.09	96.14	86.00
Trees	94.64	97.77	98.56	99.21	98.08	97.50
Painted metal sheets	86.21	99.46	99,91	100	99.82	99.50
Bare soil	100	92.83	97.35	96.78	99.76	98.08
Bitumen	99.96	90.42	96.23	92.46	100	99.00
Self-blocking bricks	99.9	92.78	97.92	97.8	99.29	93.50
Shadows	96.23	98.11	96.98	97.74	96.48	98.38
OA%	96.77	81.01	85.42	93.59	93.85	96.06

Table VI
Comparison of ATSVM Applied on Indian Pines Hsi With Other Methods Found in The Literature

Class names	ATSVM	SVM	SVM+ISODATA	SVM+EM	SVM-MSF	SVM-MSF+MV	GA-SVM
Alfalfa	0	74.4	12	0	94.9	94.9	88.96
Corn-notill	100	78.2	79.32	71.65	91	93.2	75.83
Corn-mintill	100	69.6	84.95	84.15	95.7	96.6	68.84
Corn	70.46	91.9	75.83	60.66	95.7	95.7	82.51
Grass-pasture	100	92.2	93.75	93.97	94.6	94.6	87.54
Grass-trees	100	91.7	94.8	99.11	92.4	97.3	90.68
Grass-pasture-mowed	0	100	91.67	93.97	100	100	71.43
Hay-windrowed	100	97.7	97.51	99.09	99.8	99.8	99.58
Oats	100	100	16.67	0	100	100	87.51
Soybean-notill	100	82	83.85	82.02	98	93.9	72.84
Soybean-mintill	0	58	93.16	95.05	82	82	74.52
Soybean-clean	83,14	87.9	85.17	90.05	86	97.2	79.08
Wheat	99.51	98.8	93.19	98.95	99.4	99.4	96.08
Woods	99.52	93	97.17	95.36	97.6	99.7	92.41
Buildings-grass-trees-drives	97.92	61.5	79.53	69.3	68.8	68.8	85.78
Stone-steel-towers	0	97.8	86.05	86.05	95.6	95.6	89.13
OA%	96.62	78.2	88.53	87.25	88.4	91.8	82.83

table VII
Confusion Matrix of the Class-Specific Accuracy (Csa) for Indian Pines Hsi Using ATSVM

Class names	Alfalfa	Corn-notill	Corn-mintill	Corn	Grass-pasture	Grass-trees	Grass-pasture-mowed	Hay-windrowed	Oats	Soybean-notill	Soybean-mintill	Soybean-clean	Wheat	Woods	Buildings-grass-trees-drives	Stone-steel-towers
Alfalfa	0	46	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn-notill	0	1428	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn-mintill	0	0	830	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn	0	0	0	167	70	0	0	0	0	0	0	0	0	0	0	0
Grass-pasture	0	0	0	0	483	0	0	0	0	0	0	0	0	0	0	0
Grass-trees	0	0	0	0	0	730	0	0	0	0	0	0	0	0	0	0
Grass-pasture-mowed	0	0	0	0	0	0	0	28	0	0	0	0	0	0	0	0
Hay-windrowed	0	0	0	0	0	0	0	478	0	0	0	0	0	0	0	0
Oats	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
Soybean-notill	0	0	0	0	0	0	0	0	0	972	0	0	0	0	0	0
Soybean-mintill	0	0	0	0	0	0	0	0	0	0	2455	0	0	0	0	0
Soybean-clean	0	0	0	0	0	0	0	0	0	0	94	499	0	0	0	0
Wheat	0	0	0	0	0	0	0	0	0	0	0	1	204	0	0	0
Woods	0	0	0	0	0	0	0	0	0	0	0	0	0	1259	6	0
Buildings-grass-trees-drives	0	0	0	0	0	0	0	0	0	0	6	0	0	2	378	0
Stone-steel-towers	0	0	0	0	0	0	0	0	0	0	93	0	0	0	0	0
CSA %	0	100	100	70.46	100	100	0	100	100	100	100	84.14	99.51	99.52	97.92	0

proposed method, the obtained results are compared with other SVM-based methods proposed in the literature (Tarabalka, et al., 2008; Tarabalka, Benediktsson and Chanussot, 2009; Fauvel, et al., 2013; Zhao, et al., 2020). It is important to mention, all these methods (unlike ATSVM) require previously known labels to train the SVM. The comparing results for the test images are shown in Tables V-VII, respectively, with overall accuracy (OA) of 96.77% for Pavia University and 96.62% for Indian Pines HSIs.

It is observed that the proposed methods have the correct classification rate equal to 0% for some classes. This is because these classes contain a little number of samples and the SVM classier is not trained sufficiently by this number of samples. This problem is no unique and it is repeated in other methods found in the literature (Tables V and VI). To show the efficiency of the proposed method, the obtained results are compared with other SVM-based methods proposed in the literature (Tarabalka, et al., 2008; Tarabalka, Benediktsson and Chanussot, 2009; Fauvel, et al., 2013; Zhao, et al., 2020). It is important to mention, all these methods (unlike ATSVM) require previously known labels to train the SVM. The comparing results for the test images are shown in Tables V and VI.

IV. Conclusion

A new unsupervised SVM-based clustering method is proposed. It can be concluded from the obtained results that the proposed method (ATSVM) is giving excellent results. In addition, the advantage of the proposed method is that it does not need *a priori* knowledge to train the SVM (no previously known labels are needed). Further, using FCM enables the choice of the train data instead of choosing them randomly. The method works well on big classes and it is less efficient on smaller classes. The problem of low accuracy in small classes is not unique to our proposed method (detailed results are shown in the previous section). FCM clustering needs the number of clusters to be defined. The future work will be, first improve the FCM algorithm to detect the correct number of clusters automatically and second calculate the introduced fixed threshold automatically.

References

Akbari, H., Kosugi, Y., Kojima, K. and Tanaka, N., 2010. Detection and analysis of the intestinal ischemia using visible and invisible hyperspectral imaging. *IEEE Transactions on Biomedical Engineering*, 57, pp.2011-2017.

Alameddine, J., Chehdi, K. and Cariou, C., 2021. Hierarchical unsupervised partitioning of large size data and its application to hyperspectral images. *Remote Sensing*, 13, p.4874.

Awad, M. and Khanna, R., 2015. *Support Vector Machines for Classification*. Springer, Berlin. pp.39-66.

Bezdek, J.C., Ehrlich, R. and Full, W., 1984. FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences*, 10, pp.191-203.

Boser, B.E., Guyon, I.M. and Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers. In: *Proceedings of the 5th Annual Workshop on Computational Learning Theory, COLT'92*. Association for Computing Machinery, New York, USA. pp.144-152.

Bruzzone, L. and Cossu, R., 2002. A multiple-cascade-classifier system for a robust and partially unsupervised updating of land-cover maps. *IEEE Transactions on Geoscience and Remote Sensing*, 40, pp.1984-1996.

Camps-Valls, G. and Bruzzone, L., 2005. Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43, pp.1351-1362.

Cariou, C., Le Moan, S. and Chehdi, K., 2022. A novel mean-shift algorithm for data clustering. *IEEE Access*, 10, pp.14575-14585.

Cariou, C., Moan, S. and Chehdi, K., 2020. Improving K-nearest neighbor approaches for density-based pixel clustering in hyperspectral remote sensing images. *Remote Sensing*, 12, p.3745.

Chen, Y., Nasrabadi, N.M. and Tran, T.D., 2013. Hyperspectral image classification via kernel sparse representation. *IEEE Transactions on Geoscience and Remote Sensing*, 51, pp.217-231.

Cristianini, N. and Shawe-Taylor, J., 2000. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, Cambridge.

Dong, Y., Liu, Q., Du, B. and Zhang, L., 2022. Weighted feature fusion of convolutional neural network and graph attention network for hyperspectral image classification. *IEEE Transactions on Image Processing*, 31, pp.1559-1572.

Duda, R.O. and Hart, P.E., 1973. *Pattern Classification and Scene Analysis*. 1st ed. Wiley, New York.

Fauvel, M., Tarabalka, Y., Benediktsson, J.A., Chanussot, J. and Tilton, J.C., 2013. Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 101, pp.652-675.

Feng, Y.Z. and Sun, D.W., 2012. Application of hyperspectral imaging in food safety inspection and control: A review. *Critical Reviews in Food Science and Nutrition*, 52, pp.1039-1058.

Goetz, A.F.H., Vane, G., Solomon, J.E. and Rock, B.N., 1985. Imaging spectrometry for earth remote sensing. *Science*, 228, pp.1147-1153.

Gove, R. and Faytong, J., 2012. Chapter 4 machine learning and event-based software testing: Classifiers for identifying infeasible Gui event sequences. In: Hurson, A. and Memon, A. Eds. *Advances in Computers*. Elsevier, Amsterdam, Netherlands. pp.109-135.

Guo, Y., Yin, X., Zhao, X., Yang, D. and Bai, Y., 2019. Hyperspectral image classification with SVM and guided filter. *EURASIP Journal on Wireless Communications and Networking*, 2019, p.56.

Hughes, G., 1968. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14, pp.55-63.

Lacar, F.M., Lewis, M.M. and Grierson, I.T., 2001. Use of hyperspectral imagery for mapping grape varieties in the Barossa Valley, South Australia. In: *Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International*. pp.2875-2877.

Lagueux, P., Puckrin, E., Turcotte, C.S., Gagnon, M.A., Bastedo, J., Farley, V. and Chamberland, M., 2012. Airborne Infrared Hyperspectral Imager for Intelligence, Surveillance and Reconnaissance Applications. *Proceedings of SPIE the International Society for Optical Engineering*.

Li, X., Lim, B.S., Zhou, J.H., Huang, S., Phua, S.J., Shaw, K.C. and Er, M.J., 2009. Fuzzy Neural Network Modelling for Tool Wear Estimation in Dry Milling Operation. *Annual Conference of the PHM Society*. Available from: http://www.papers.phmsociety.org/index.php/phmconf/article/view/1403 [Last accessed 2022 Jul 02].

Li, Y., Li, J. and Pan, J.S., 2019. Hyperspectral image recognition using SVM combined deep learning. *Journal of Internet Technology*, 20, pp.851-859.

MacQueen, J., 1967. Some Methods for Classification and Analysis of Multivariate Observations. Vol. 5. In: *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*.

Mounika, K., Aravind, K., Yamini, M., Navyasri, P., Dash, S. and

Suryanarayana, V., 2021. Hyperspectral Image Classification Using SVM with PCA. In: *2021 6th International Conference on Signal Processing, Computing and Control*. pp.470-475.

Pathak, D.K. and Kalita, S.K., 2019. Spectral Spatial Feature Based Classification of Hyperspectral Image Using Support Vector Machine. In: *2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, Noida, India. pp.430-435.

Ren, J., Wang, R., Liu, G., Wang, Y. and Wu, W., 2021. An SVM-based nested sliding window approach for spectral spatial classification of hyperspectral images. *Remote Sensing*, 13, p.114.

Ryan, J.P., Davis, C.O., Tufillaro, N.B., Kudela, R.M. and Gao, B.C., 2014. Application of the hyperspectral imager for the coastal ocean to phytoplankton ecology studies in Monterey Bay, CA, USA. *Remote Sens*, 6, pp.1007-1025.

Sellami, A. and Tabbone, S., 2022. Deep neural networks-based relevant latent representation learning for hyperspectral image classification. *Pattern Recognition*, 121, p.108224.

Shang, Y., Zheng, X., Li, J., Liu, D. and Wang, P., 2022. A comparative analysis of swarm intelligence and evolutionary algorithms for feature selection in SVM-based hyperspectral image classification. *Remote Sensing*, 14, p.3019.

Sun, L., Zhao, G., Zheng, Y. and Wu, Z., 2022. Spectral spatial feature tokenization transformer for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60, pp.1-14.

Tarabalka, Y., Fauvel, M., Chanussot, J. and Benediktsson, J.A., 2010. SVM and MRF-based method for accurate classification of hyperspectral images. IEEE Geoscience and Remote Sensing Letters, 7, pp.736-740.

Tarabalka, Y., Benediktsson, J. and Chanussot, J., 2009. Classification of Hyperspectral Data Using Support Vector Machines and Adaptive Neighborhoods. *International Conference on Image Processing*.

Tarabalka, Y., Chanussot, J., Benediktsson, J., Angulo, J. and Fauvel, M., 2008. Segmentation and Classification of Hyperspectral Data Using Watershed. *IGARSS 2020 Proceedings*, pp.652-655.

Vapnik, V.N., 1995. *The Nature of Statistical Learning Theory*. Springer, New York.

Wu, Y., Yang, X., Plaza, A., Qiao, F., Gao, L. and Cui, Y., 2016. Approximate computing of remotely sensed data: SVM hyperspectral image classification as a case study. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9, pp.1-13.

Zhao, C., Zhao, H., Wang, G. and Chen, H., 2020. Improvement SVM classification performance of hyperspectral image using chaotic sequences in artificial bee colony. *IEEE Access*, 8, pp.73947-73956.