Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **6**, November **2018**

Optimization and Analysis on Fuzzy SVM for Objects Classification

Xiao-Kang Ding[1], Xiao-Jun Yang[1,*], Jin-Yun Jiang[1],
Xiao-Lei Deng[1], Jian-Chen Cai[1] and Ying-Yu Ji[1]

[1]College of Mechanical Engineering, Quzhou University,
Quzhou, Zhejiang 324000, China
[*]Corresponding Author:*yxj_qzu@126.com*

;

Received April, 2018; Revised July, 2018

Abstract. *In order to improve the classification performance of fuzzy support vector machine (FSVM), a new fuzzy membership function was proposed in this paper to solve the two-class problems. Three FSVM models were established and a normal SVM was selected as comparison reference. The performance of each model was evaluated by 6 datasets including 1 home-brewed dataset and 5 benchmark real-world datasets from the UCI machine learning repository. The experimental results showed that the proposed method could effectively eliminate the noise impact and improve the classification accuracy.*
Keywords: SVM, Fuzzy membership, Binary classification

1. **Introduction.** Support vector machine (SVM) is a machine learning algorithm first proposed by Cortes and Vapnik in 1995[1]. Compared with traditional artificial neural networks, SVM can not only simplify the learning algorithm, but also improve the technical performance, especially the generalization. Consequently, SVM has become a hot spot in the field of machine learning in recent years.

Currently, SVM algorithm has been widely used in pattern recognition [1], regression estimation [2], probability density function estimation [3], etc. S. Ikram proposed a hybrid intrusion detection model by integrating the principal component analysis (PCA) and support vector machine(SVM) to improve accuracy and reduce training time for intrusion detection [4]. Y. Wang and H. Duan proposed introduced a novel classification framework for hyperspectral images (HSIs). Spectral, spatial, and hierarchical structure information are integrated into the SVM classifier in a way of multiple kernels. The proposed classification framework can achieve 13.46-15.61% in average higher than the traditional methods [5]. S. Asra applied SVM on human behavior recognition based on hand written cursives. The paper proposed a novel method of impulse noise filter construction, based on the switching scheme with two cascaded detectors and two corresponding estimators [6]. G. Kaur et al. put forward a method to classify text by SVM and KNN, and SVM gives importance great exactness, accuracy, review than KNN, SVC [7].

SVM is becoming more and more widely used. However, in practical applications, many noisy points may make the sample data not ideal as expected, which have a significant weakening impact on the generalization performance of SVM. In view of this, the fuzzy vector machine algorithm (FSVM) was developed to calculate the fuzzy membership for each sample, by which the effects of noise on SVM could be eliminated effectively [8-10]. Based on FSVM, many kinds of estimating methods for fuzzy membership functions were presented to improve the performance of SVM [11-13]. For example, Q. Yan, S. Xia, Meng F presented a novel cost-sensitive SVM method whose penalty parameter C optimized on the basis of cluster probability density function(PDF) and the cluster PDF is estimated only according to similarity matrix and some predefined hyper-parameters. Experimental results on various standard benchmark data sets and real-world data with different ratios of imbalance show that the proposed method is effective in comparison with commonly used cost-sensitive techniques [14]. O. Almasi et al. adopted Adaptive Particle Swarm Optimization (APSO) method to minimize the generalization error by changing the attributes values of positive and negative class centers so that it can make them free of attribute-noise. As the APSO converged, the fuzzy memberships were assigned for each training data points based on their distance to the corresponding purified class centers with the same class-label. The results demonstrated that the proposed FSVM-4 had a considerable better generalization performance in comparison with the other FSVM methods [15].

According to the above researches, in this paper, a new fuzzy membership estimating method is proposed for FSVM to solve binary classification problems. Three FSVM models and a normal SVM model are established, and the performance of each model is evaluated by 6 datasets. The experimental results show that the proposed method can effectively eliminate the noise impact and improve the classification accuracy.

2. **Support Vector Machine.**

As a new machine learning algorithm, the core idea of SVM is to map non-linear problem in the low-dimensional space to high-dimensional space so that the non-linear problem can be transformed into linearly separable problem. SVM only concerns the points nearest to the hyperplane (called as support vectors), and the hyperplane can be defined by maximizing the distances from support vectors to itself. Thus, the objective function for solving the hyperplane can be expressed as follows:

$$\begin{aligned} min_{\omega,b}\frac{1}{2}||\omega||^2 \\ s.t. y^{(i)}(\omega^T x^{(i)} + b) \geq 1, i = 1, 2, ..., m \end{aligned} \tag{1}$$

However, noisy points may make the hyperplane move easily. Considering the high sensitivity of SVM to noisy points, slack variables ξ is introduced to the quadratic programming problems to permit the existence of noisy points. Then, the objective function can be converted into the following form:

$$\begin{aligned} min_{\gamma,\omega,b}\frac{1}{2}||\omega||^2 + C\sum_{i=1}^{m}\xi_i \\ s.t. y^{(i)}(\omega^T x^{(i)} + b) \geq 1 - \xi_i, i = 1, 2, ..., m \\ \xi_i \geq 0, i = 1, 2, ..., m \end{aligned} \tag{2}$$

where penalty parameter C is used to limit the impact of noisy points on the objective function. This modified model is called as soft interval classifier.

To solve the objective function, Lagrangian operator is introduced into the formula. By simplification and conversion, the hyperplane can be expressed as follows:

$$f(x) = sgn(\sum_{i=1}^{m} \alpha_i y_i K(x_i, x) + b) \tag{3}$$

where $K(x_i, x)$is the kernel function for solving non-linear separable problems. In this paper, radial basis function (RBF) is selected as the kernel function, which can be expressed as follows:

$$K(x, y) = exp(-\frac{||x - y||^2}{2\sigma^2}) \tag{4}$$

3. FSVM and A New Fuzzy Membership Function.

3.1. Fuzzy SVM.

In soft interval classifier, the value of penalty parameter C cannot be too large or too small in order to ensure the effect of classifier. Hence, fuzzy membership s_i is introduced to SVM. For a training dataset $\{x_1, x_2, x_3, ...\}$, there is a label y_i for each x_i, denoted as $\{(x_1, y_1), (x_2, y_2), ...\}$. As s_i is introduced, the dataset is denoted as $\{(x_1, y_1, s_1), (x_2, y_2, s_2), ...\}$, and the objective function can be rewritten as:

$$\begin{aligned} min_{\gamma,\omega,b}\frac{1}{2}||\omega||^2 + C\sum_{i=1}^{m} s_i\xi_i \\ s.t. y^{(i)}(\omega^T x^{(i)} + b) \geq 1 - \xi_i, i = 1, 2, ..., m \\ \xi_i \geq 0, i = 1, 2, ..., m \end{aligned} \tag{5}$$

where s_i takes a value between 0 and 1 ($0 < s_i \leq 1$). s_i represents the probability of each sample belonging to its label y_i. Meanwhile, it counteracts the impact of penalty parameter C on classifiers.

Therefore, an appropriate membership function is very important for FSVM model. Firstly, the lower bound of the membership should be defined. Secondly, the membership function should be constructed according to the characteristic datasets. At present, the commonly used membership calculation methods are mainly based on the distances from sample points to the class center.

3.2. A New Fuzzy Membership Function for FSVM.

In this paper, three FSVM models are established, denoted as FSVM-1, FSVM-2 and FSVM-3 respectively. FSVM-1 uses a conventional method as comparison reference, and the other two determine the fuzzy membership by comparing the distances from each sample to the positive and negative class centers. These three algorithms all select linear and exponential decay formulas to calculate the membership.

3.2.1. *FSVM-1.*

This model adopts a conventional membership calculation method, which determines the class center by averaging all samples. Marking the Euclidean distance from each sample point to the class center as di, then the membership function can be expressed as:

$$s_i = \begin{cases} 1 - \frac{d_i}{max(d_i)+\beta} \\ \frac{2}{1+exp(\beta d_i)} \end{cases} \tag{6}$$

where

$$d_i = \begin{cases} ||x_i^+ - x_{cen}^+|| \\ ||x_i^- - x_{cen}^-|| \end{cases} \tag{7}$$

In the membership function, β is a small constant to avoid $s_i = 0$; x_i^+ is the sample with label $y_i = 1$, and x_i^- is the sample with label $y_i = -1$; x_{cen}^+ is the positive class center and x_{cen}^- is the negative one. This function indicates that the closer the distance from sample to the class center, the greater the membership value, and the smaller the contrary; vice versa.

3.2.2. *FSVM-2.* In FSVM-2, the distances from each sample point to the positive and the negative class centers are calculated respectively. The membership function can be constructed as follows:

$$s_i = \begin{cases} f(d_i^+), if||x_i^+ - x_{cen}^+|| \geq ||x_i^+ - x_{cen}^-|| \\ 1, if||x_i^+ - x_{cen}^+|| < ||x_i^+ - x_{cen}^-|| \\ 1, if||x_i^- - x_{cen}^+|| > ||x_i^- - x_{cen}^-|| \\ f(d_i^-), if||x_i^- - x_{cen}^+|| \leq ||x_i^- - x_{cen}^-|| \end{cases} \tag{8}$$

where

$$d_i = \begin{cases} ||x_i^+ - x_{cen}^+|| \\ ||x_i^- - x_{cen}^-|| \end{cases} \tag{9}$$

When the distance of the positive sample to the positive class is less than that to the negative class, the point is considered as a "useful point", and its membership is assigned as 1. When the distance of the positive sample to the positive class is greater than that to the negative class, the point is considered as a "noisy point", and its membership value is calculated according to the decay formula; vice versa.

3.2.3. *FSVM-3.* In order to better estimate fuzzy membership for nonlinear datasets, in FSVM-3, the sample points are firstly mapped into a high-dimensional space by function $\Phi(x_i)$, and then the membership is calculated by the same method as that in FSVM-2. The membership function can be constructed as follows:

$$s_i = \begin{cases} f(d_i^+), if||\Phi(x_i^+) - \Phi_{cen}^+|| \geq ||\Phi(x_i^+) - \Phi_{cen}^-|| \\ 1, if||\Phi(x_i^+) - \Phi_{cen}^+|| < ||\Phi(x_i^+) - \Phi_{cen}^-|| \\ 1, if||\Phi(x_i^-) - \Phi_{cen}^+|| > ||\Phi(x_i^-) - \Phi_{cen}^-|| \\ f(d_i^-), if||\Phi(x_i^-) - \Phi_{cen}^+|| \leq ||\Phi(x_i^-) - \Phi_{cen}^-|| \end{cases} \tag{10}$$

where

$$d_i = \begin{cases} \sqrt{||\Phi(x_i^+) - \Phi_{cen}^+||} \\ \sqrt{||\Phi(x_i^-) - \Phi_{cen}^-||} \end{cases} \tag{11}$$

In the membership function, $\Phi(x_i^+)$ is the sample mapped to the feature space with label $y_i = 1$, and $\Phi(x_i^-)$ is the one with label $y_i = -1$, Φ_{cen}^+ is the positive class center in feature space and Φ_{cen}^- is the negative one. In the feature space, the distances from the samples to the positive and negative class centers are calculated as follows:

$$\begin{aligned} ||\Phi(x_i^+) - \Phi_{cen}^+|| &= \sqrt{||\Phi(x_i^+) - \Phi_{cen}^+||^2} = \sqrt{\Phi(x_i^+)^2 - 2\Phi(x_i^+)\Phi_{cen}^+ + (\Phi_{cen}^+)^2} \\ &= \sqrt{K(x_i^+, x_i^+) - \frac{2}{n^+}\sum_{j=1}^{n^+} K(x_i^+, x_j^+) + \frac{1}{n^{+2}}\sum_{i=1}^{n^+}\sum_{j=1}^{n^+} K(x_i^+, x_j^+)} \end{aligned} \tag{12}$$

$$\begin{aligned} ||\Phi(x_i^+) - \Phi_{cen}^-|| &= \sqrt{||\Phi(x_i^+) - \Phi_{cen}^-||^2} = \sqrt{\Phi(x_i^+)^2 - 2\Phi(x_i^+)\Phi_{cen}^- + (\Phi_{cen}^-)^2} \\ &= \sqrt{K(x_i^+, x_i^+) - \frac{2}{n^-}\sum_{j=1}^{n^-} K(x_i^+, x_j^-) + \frac{1}{n^{-2}}\sum_{i=1}^{n^-}\sum_{j=1}^{n^-} K(x_i^-, x_j^-)} \end{aligned} \tag{13}$$

$$||\Phi(x_i^-) - \Phi_{cen}^+|| = \sqrt{||\Phi(x_i^-) - \Phi_{cen}^+||^2} = \sqrt{\Phi(x_i^-)^2 - 2\Phi(x_i^-)\Phi_{cen}^+ + (\Phi_{cen}^+)^2}$$
$$= \sqrt{K(x_i^-, x_i^-) - \frac{2}{n^+}\sum_{j=1}^{n^+} K(x_i^-, x_j^+) + \frac{1}{n^{+2}}\sum_{i=1}^{n^+}\sum_{j=1}^{n^+} K(x_i^+, x_j^+)} \quad (14)$$

$$||\Phi(x_i^-) - \Phi_{cen}^-|| = \sqrt{||\Phi(x_i^-) - \Phi_{cen}^-||^2} = \sqrt{\Phi(x_i^-)^2 - 2\Phi(x_i^-)\Phi_{cen}^- + (\Phi_{cen}^-)^2}$$
$$= \sqrt{K(x_i^-, x_i^-) - \frac{2}{n^-}\sum_{j=1}^{n^-} K(x_i^-, x_j^-) + \frac{1}{n^{-2}}\sum_{i=1}^{n^-}\sum_{j=1}^{n^-} K(x_i^-, x_j^-)} \quad (15)$$

where

$$k(x, y) = \langle\Phi(x), \Phi(y)\rangle = exp(-\frac{||x - y||^2}{2\sigma^2}) \quad (16)$$

Thus, the membership value $s_i$is defined in FSVM-3.

4. **Experiments.** The experiments are achieved by the widely used ”libsvm” toolbox under Matlab language environment. Classifier models are tested by 6 different datasets, and the geometric mean of sensitivity is used to evaluate models in order to demonstrate the comprehensive classification performance. In the results, SE represents the positive class classification accuracy, SP is the negative class classification accuracy, and $GM = \sqrt{SE \times SP}$ denotes the overall performance.

4.1. **Description of Datasets.** In this study, the Tree dataset is constructed by our research, and the other five are the benchmark real-world datasets obtained from UCI machine learning repository. The details of the six datasets are listed as Table 1. These datasets all consist of two classes and contain different sample data and attributes, covering all kinds of areas. Hence, these datasets can be considered as representatives that effectively verify the proposed FSVM models.

Table 1. Information of Datasets.

Datasets	Attributes	Samples	Positive Samples	Negative Samples
Tree	6	431	381	50
Wine	13	130	59	71
Transfusion	4	748	570	178
Haberman	3	306	225	81
Pima	8	768	500	268
Quality	11	3655	1457	2198

4.2. **Experimental results of the Tree Dataset.** With different penalty parameter C, SVM and the three FSVM models are applied to the Tree dataset using linear decay formula and exponential decay formula respectively, and the classification results are shown in Tables 2 and 3.

In Table 2, it can be seen that when $C \leq 10$, FSVM-3 has the best results; while as C increases, the results of FSVM-3 become consistent with those of other models. From Table 3, the classification accuracy of FSVM-3 is equal to that of SVM, but better than those of FSVM-1 and FSVM-2. In general, the optimization effect of the proposed model on SVM classification is not as significant as assumed. This discrepancy may be related with the dataset itself. Firstly, the relevance between the positive and the negative samples is small, and there are few noisy points; secondly, the increment of C can weaken the effect of fuzzy membership function so as to maintain the results of FSVM-3 consistent with other models when C is too large.

Table 2. Classification Results of the Tree Dataset with Different C using Linear Formula.

Penalty Parameter C	Classification Results(%)	SVM	FSVM-1	FSVM-2	FSVM-3
	GM	80.28	80.16	80.16	82.7
C=2	SE	99.14	98.85	98.85	97.7
	SP	65	65	65	70
	GM	83.18	82.9	82.94	83.33
C=10	SE	98.85	98.28	98.28	99.14
	SP	70	70	70	70
	GM	86.47	86.47	86.47	86.47
C=50	SE	99.71	99.71	99.71	99.71
	SP	75	75	75	75
	GM	91.79	91.79	91.79	91.79
C=200	SE	99.14	99.14	99.14	99.14
	SP	85	85	85	85
	GM	86.35	86.35	86.35	86.35
C=500	SE	99.42	99.42	99.42	99.42
	SP	75	75	75	75
	GM	83.42	83.42	83.42	83.42
C=1000	SE	99.42	99.42	99.42	99.42
	SP	70	70	70	70

Table 3. Classification Results of the Tree Dataset with Different C using Exponential Formula.

Penalty Parameter C	Classification Results(%)	SVM	FSVM-1	FSVM-2	FSVM-3
	GM	80.28	38.91	78.88	82.8
C=2	SE	99.14	15.14	95.71	99.14
	SP	65	100	65	65
	GM	83.18	21.38	51.54	83.18
C=10	SE	98.85	4.5	26.57	98.85
	SP	70	100	100	70
	GM	86.47	54.77	85.85	86.47
C=50	SE	99.71	30	98.28	99.71
	SP	75	100	75	75
	GM	91.79	73.87	63.47	91.79
C=200	SE	99.14	54.57	40.28	99.14
	SP	85	100	100	85
	GM	86.35	85.98	51.57	86.35
C=500	SE	99.42	98.57	28	99.42
	SP	75	75	95	75
	GM	83.42	79.84	51.82	83.42
C=1000	SE	99.42	85	26.85	99.42
	SP	70	75	100	70

4.3. **Experimental results of Different Datasets.** In order to better verify the classification performance, the proposed FSVM models are tested by another five different benchmark datasets, and the memberships are also calculated using linear and exponential functions respectively, as shown in Tables 4 and 5.

From Tables 4 and 5, it is obvious that the proposed FSVM-3 has the best classification performance. According to the analysis, there is a close correlation among the data, and the class centers may overlap or approach each other in the low-dimensional space. Therefore, a wholly positive or negative sampling may occur according to the number of training samples in FSVM-1 and FSVM-2. In addition, each sample is multidimensional and has many attributes, so that it has the best result when mapped to the high-dimensional space of FSVM-3.

TABLE 4. Classification Results of Different Datasets using Linear Formula.

Dataset	Classification Results(%)	SVM	FSVM-1	FSVM-2	FSVM-3
Wine	GM	68.92	44.72	54.77	77.45
	SE	50	20	30	60
	SP	95	100	100	100
Transfusion	GM	49.62	48.67	48.32	52.51
	SE	83.5	84	86.75	82.75
	SP	29.48	28.2	26.92	33.33
Haberman	GM	58.03	57.7	60	63.09
	SE	59.42	43.42	51.42	54.28
	SP	56.67	76.67	70	73.33
Pima	GM	37.93	36.3	22.67	37.93
	SE	16	92.67	5.33	16
	SP	89.91	14.22	96.33	89.91
Quality	GM	53.56	54.31	53.65	55.47
	SE	33.3	34.3	33.9	39.5
	SP	86.17	86	84.91	77.91

TABLE 5. Classification Results of Different Datasets using Exponential Formula.

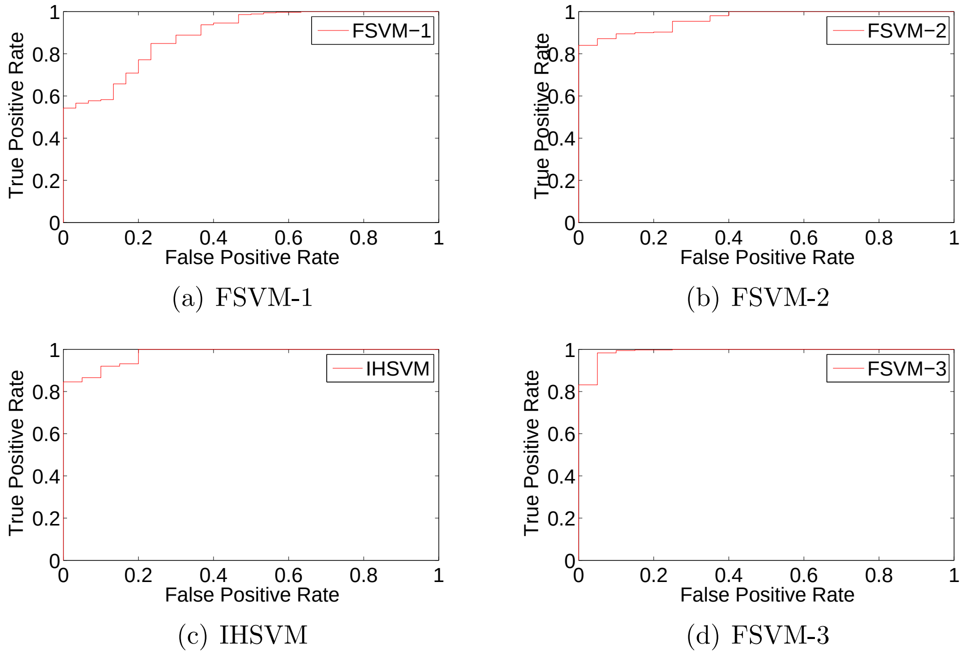
Dataset	Classification Results(%)	SVM	FSVM-1	FSVM-2	FSVM-3
Wine	GM	68.92	0	54.77	72.28
	SE	50	100	30	55
	SP	95	0	100	95
Transfusion	GM	49.62	0	22.44	51.65
	SE	83.5	0	98.25	83.25
	SP	29.48	100	5.12	32.05
Haberman	GM	58.03	62.94	63.57	64.14
	SE	59.42	59.42	58.85	61.71
	SP	56.67	66.67	68.67	66.67
Pima	GM	37.93	0	0	37.93
	SE	16	100	100	16
	SP	89.91	0	0	89.91
Quality	GM	53.56	27.27	53.7	56.25
	SE	33.3	7.7	33.6	47.7
	SP	86.17	96.58	85.83	66.3

4.4. **Performance of Different SVMs.** In order to better analyze the performance of each SVM model, the ROC (Receiver operating characteristic) curves are drawn, and compared with Improved Hypersphere Support Vector Machine (IHSVM) in reference [16].

As shown in Figure 1, the vertical axis TPR defines how many correct positive results occur among all positive samples available during the test, and the horizontal axis FPR defines how many incorrect positive results occur among all negative samples available during the test. The curve closest to the upper left corner of the ROC curve represents the highest accuracy. ROC curves can visually discriminate the merits of each model. On the other hand, by calculating the AUC (area under the ROC curve), the estimate values can also be compared, and the curve with the maximum AUC has the best estimate value. From Figure 1, it is obvious that the proposed FSVM-3 model has better performance than the others.

5. Conclusions.

In summary, a new membership function was proposed to improve classification accuracy of FSVM in this paper. In this method, distances from sample points to the positive and negative class centers were calculated respectively, and the fuzzy memberships were calculated by linear and exponential decay formula respectively through comparison



(a) FSVM-1 (b) FSVM-2

(c) IHSVM (d) FSVM-3

Figure 1. ROC curve of each model

of the distances. This method was applied in the low-dimensional space FSVM-2 and high-dimensional space FSVM-3, with normal SVM and normal FSVM-1 as comparison references. After model construction, the performance of each model was evaluated by 6 datasets. The results showed that the proposed FSVM-3 could effectively eliminate the noise impact and improve the classification accuracy for the data with large association and many noisy points.

Acknowledgment.
The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve this paper. Besides, this work is supported by Zhejiang Provincial Natural Science Foundation(No.LQ17C160001, LQ16E030008, LY16E050011), National Natural Science Foundation(No.51605253), Quzhou Science and Technology Plan Project (No.2016Y002), Provincial Public-benefit Technology Application Research of Zhejiang(No.2016C31127, 2017C37060).

REFERENCES

[1] H. Huanrui, New Mixed Kernel Functions of SVM Used in Pattern Recognition, *Cybernetics and Information Technologies,* vol. 16, no. 5, pp. 5-14, 2016.

[2] M. Premalatha, C. Vijayalakshmi, SVM approach for classification and regression with absolute value combination method for controlling complexity, *International Journal of Pure and Applied Mathematics,* vol. 101, no. 5, pp. 811-820, 2015.

[3] K. Liao, H. Hsiung, Reliability-Based Safety Analysis of a Seawall with Incomplete Information on Tsunamis, *Coastal Engineering Journal*, vol. 59, no. 3, pp. 1-25, 2017.

[4] S. Ikram, Improving Accuracy of Intrusion Detection Model Using PCA and optimized SVM, J*ournal of Computing and Information Technology*, vol. 24, no. 2, pp. 133-148, 2016.

[5] Y. Wang, H. Duan, Classification of Hyperspectral Images by SVM Using a Composite Kernel by Employing Spectral, Spatial and Hierarchical Structure Information, *Remote Sensing,* vol. 10, no. 3, pp. 441-467, 2018.

[6] S. Asra, D. Shubhangi, Human behavior recognition based on hand written cursives by SVM classifier, International Conference on Electrical, Electronics, Communication, *Computer, and Optimization Techniques*, pp. 260-268, 2017.

[7] G. Kaur, E. Kaur, Novel Approach of Text Classification by SVM-RBF Kernel and Linear SVC, International Journal of Advance research , Ideas and Innovations in Technology, vol. 3, no. 3, pp. 1014-1017, 2017.

[8] J. Zhang, P. Zhang, L. Zhuo, Fuzzy Support Vector Machine Based on Color Modeling for Facial Complexion Recognition in Traditional Chinese Medicine, Chinese Journal of Electronics, vol. 25, no. 3, pp. 474-480, 2016.

[9] Y. Wei, J. Watada, W. Pedrycz, Design of a qualitative classification model through fuzzy support vector machine with type-2 fuzzy expected regression classifier preset, *Ieej Transactions on Electrical and Electronic Engineering*, vol. 11, no. 3, pp. 348-356, 2016.

[10] S. Yang, Z. Zhang, A novel fuzzy SVM based on fuzzy c-means for credit scoring, *IEEE International Conference on Computer and Communications,* pp. 1349-1353, 2017.

[11] P.Sridevi, Identification of suitable membership and kernel function for FCM based FSVM classifier model, *Cluster Computing*, vol. , no. 6, pp. 1-10, 2018.

[12] K. Liu, X. Wu, A Fuzzy Support Vector Machine Based on New Membership Function, Computer Engineering, vol. 42, no. 4, pp. 155-159, 2016.

[13] W. Tian, F. Randi, J. Wei, Adaptive Fuzzy Support Vector Machine for Classification of Clouds in Satellite Imagery, *Geomatics and Information Science of Wuhan University*, vol. 42, no. 4, pp. 488-495, 2017.

[14] Q. Yan, S. Xia, Meng F. Optimizing Cost-Sensitive SVM for Imbalanced Data: *Connecting Cluster to Classification*, arXiv:1702.01504, 2017.

[15] O. Almasi, H. Gooqeri, B. Asl, W. Tang, A New Fuzzy Membership Assignment Approach for Fuzzy SVM Based on Adaptive PSO in Classification Problems, *Journal of Mathematics and Computer Science,* vol. 14, no. 2, pp. 171-182, 2015.

[16] C. Wu, J. Yang, An Improved Hypersphere Support Vector Machine Method for Vibration Fault Diagnosis of Wind Turbine Gearbox, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 8, no. 6, pp. 1237-1245, 2017.