IFERP connecting engineers... developing research

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 12, December 2017

PSO inspired Hyperspectral Image Classification

[1] Dr.K.Kavitha [2] Dr.S.Arivazhagan, [3] W.Jenifa
[1] Associate Professor, [2] Principal, [3] PG Student
Mepco Schlenk Engineering College, Sivakasi

Abstract: - **Hyperspectral Remote Sensing technology is used for identification and detection of objects on the earth. Hyperspectral images provide accurate classification than multispectral images but it suffers from over dimensionality problem. In order to overcome this drawback, Daubechies wavelet with Four taps (DB4) and Eight taps are used for extracting the features and to improve the classification performance Particle Swarm Optimization (PSO) technique is used for feature selection. Support Vector Machine (SVM) classifier is used for efficient classification. In this paper image acquired from AVIRIS sensor, Indian pines dataset is used. The overall accuracy obtained for DB4 and DB8 is 92% and 90% respectively.**

Keywords: **Hyperspectral Image Classification,Particle Swarm Optimization (PSO), Support Vector Machine.**

I. INTRODUCTION

Images produced from Hyperspectral sensors contain much more data than images from multispectral sensors and have a greater potential to detect differences among land and water features. For example, multispectral imagery can be used to map forested areas, while Hyperspectral imagery can be used to map tree species within the forest. For instance, the AVIRIS hyperspectral sensor [1] has 224 spectral bands ranging from 0.4μm to 2.5μm. Such a large number of bands implies high dimensionality data, presenting several significant challenges to image classification The dimensionality of input space strongly affects the performance of many supervised classification methods [2]. There is likely to be redundancy between bands and some bands may contain less discriminatory information than others. Finally high dimensional data impose requirements for storage space. Feature extraction is a special form of dimensionality Reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously then the input data will be transformed into a reduced representation set of features. Moreover, the features have to be selected [3] in the way to tackle the problem of recognizing objects of 3 images even if they have different scales, orientations, poses and also to categorize objects from backgrounds. Feature selection is a process of choosing a subset of relevant features from a large number of original features. The selected feature subset should be sufficient to describe the target concepts. By eliminating irrelevant and redundant features, feature selection could improve classification performance, make learning and executing processes faster and simplify the structure of the learned classifiers [4] Chengming et al proposed an ensemble learning framework which applies classification technique to learn multiple kernal classifier for classification problem but this ensemble framework is faster than the mixture kernal but slower than the single kernel[5].Xiangtao Zheng and Xiaoqiang Lu, adapted the Determinantal point process(DPP),Multiple Laplacian Eigen Maps(MLE) to extract an useful information regarding from Hyperspectral image yet speed of this method is less and computational complexity is related to number of bands[6] For a genetic algorithm time taken for convergence is high and also it does not provide a guarantee for providing global maxima. Patra et al analysed band selection based on Rough set, this Rough set theory is a paradigm to deal with uncertainty, vagueness, and incompleteness of data Rough based set is for small numbers of selected bands this method always provided significantly higher accuracies compared to all of the reference band selection methods this is an supervised method to select informative bands from hyperspectral image but major disadvantage of this paper is it is not suitable for large number of bands it does not provide high accuracy [7]. Jiali Zhu et al explained the feature effective selection based on genetic algorithm Band grouping provide high correlation between the bands and Searching procedure is high even though it provide a less classification accuracy[8]. Benediktsson et al proposed the use of probability estimates obtained by the support vector machine (SVM) classification, in order to determine the most reliable classified pixels as seeds of spatial regions [9]. The hyperplane used by the SVM maximize the margin between the two classes and it is very efficient for the classification [10] the multi resolution transforms are efficiently used for extracting feature. The extracted feature contain spectral and spatial information and this extracted feature has redundancy inorder to avoid that best features are selected ,this can be done by using Particle Swarm

Optimization(PSO).After Selecting the optimized feature it is efficiently classified using SVM classifiers. From the literature it is observed that Feature extraction is computationally complex and Cost for storing the data is high. Previously Genetic algorithm is used but it has no guarantee of finding global maxima and Time taken for convergence is high.For GA it undergo Mutation,Cross over so computation time is high. In order to over this drawback Particle Swarm Optimatization technique is used for feature Selection. Rest of the paper is organized as follows in section2

II. PROPOSED METHODOLOGY

A. Input Image

Indian Pines dataset consists of 16 classes like alfalfa, corn notril, corn mintril, etc, and this sensor covers the wavelength region from 0.4 to 2.5 µm using more than two hundred spectral channels, at nominal spectral resolution of 10 nm. This dataset consists of 220 bands with 145*145 pixels in each spectral band this dataset is acquired from AVIRIS Sensor.

B. Feature Extraction

ransforming the input data into the set of features is called Feature extraction. Feature extraction involves reducing amount of resources required to describe large set of features. Features are extracted from transformed image bands. This project makes use of Discrete Wavelet Transform (DWT) for getting transformed image. DWT with single level decomposition is used to divide the images into approximation and detailed coefficients. Statistical and Co-occurrence features are extracted from the approximation coefficients. The resultant features are having the property to distinguish one class from other.In DWT Daubechies wavelet with four tap (DB4) and eight tap (DB8) is used.

C. Feature Selection

The feature selection is preprocessing which aims to minimize the number of features amount, and simultaneously keep the maximum amount of discriminate information as possible for each combination of selected features a seperability criterion should be used to find the best subset. The computational load employed to test all possible combination thus, PSO is a suitable tool that one can use to lead a search that optimizes a certain seperability criterion. Although the high dimensionality of hyperspectral images provides great discriminative power, its classification is still a challenging task due to the large amount of spectral information and its small set of referenced data This is also known as Hughes phenomena or the "curse of dimensionality". Another constraint mentioned in the literature when data is in high-dimensional space is the density estimation. It is more difficult to compute than when in a lower dimensional space, since the space is quite sparse. In order to overcome such difficulties, some approaches apply feature selection technique.

D. SVM Classifier

SVM is the supervised learning algorithm that is being used as a regression and a classifier but it is mostly used as a classifier for solving the classification problems. Classification in general can be viewed as separating the different classes present in the feature space. SVM can be used to classify or able to discriminate between the different classes whether they are linearly separable or non-linearly separable. The SVM algorithm finds the maximum margin separating hyperplane. The points from each class that determine the margin planes are called the support vectors.

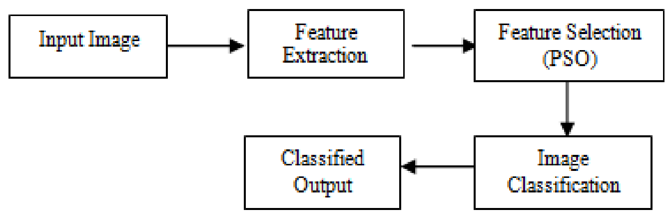


Fig. 1. Work Flow

III. PSO ALGORITHM

Feature selection is a process of selecting the best feature among the extracted features. Selected features can be optimized by a technique such as Particle Swarm Optimization (PSO).This can provide dimensionality reduction .PSO is an swarm intelligence based approximate, non-deterministic optimization technique. This technique will find the parameters that provide the maximum or minimum value of a target value. Unlike GA, PSO has no evolution operators such as crossover and mutation.

Steps Involved in PSO algorithm:

Step1. Set parameter $w_{\max}$ $w_{\min}$, c_1 and c_2 of PSO

Step2. Initialize population of particles having positions X and velocities V

Step 3. Set iteration k = 1

Step4. Calculate fitness of particles $F_i^k = f\left(X_i^k\right), \forall_i$ and find the index of the best particle b

Step 5. Select $pbest_i^k = X_i^k, \forall_i$ and $gbest = X_b^k$

Step6. W = $w_{\max} - k \times \left(w_{\max} - w_{\min}\right) \big/ (\max ite)$

Step7. Update velocity and position of particles

$$v_{ij}^{k+1} = w \times v_{ij}^k + c_1 \times rand() \times \left(pbest_i^k - X_{ij}^k\right) + c_2 \times rand() \times \left(gbest_j^k - X_{ij}^k\right); \quad \forall_i \text{ and } \forall_j$$

IFERP connecting engineers... developing research

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 12, December 2017

Step 8. Evaluate fitness $F_i^{k+1} = f\left(X_i^{k+1}\right)$, $\forall_i$ and find the index of the best particle b_1

Step 9. Update Pbest of population $\forall_i$ If $F_i^{k+1} < F_i^k$ then $pbest_i^{k+1} = X_i^{k+1}$ else $pbest_i^{k+1} = pbest_i^k$

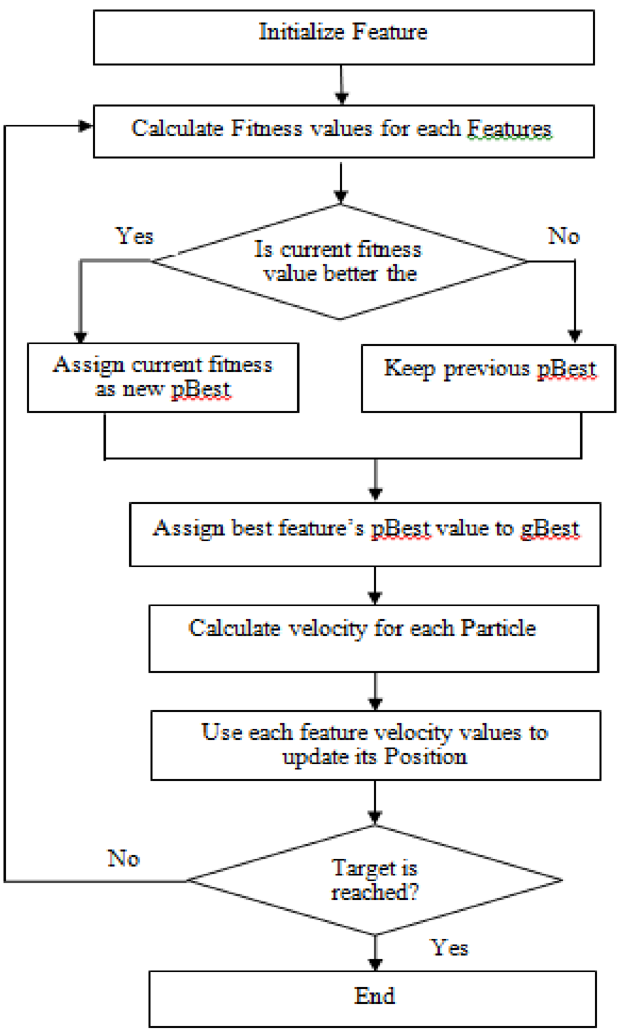
Step 10. Update gbest of population If $F_{b+1}^{k+!} < F_b^k$ then $gbest^{k+1} = pbest_{b1}^{k+1}$ and set b = b_1

else $gbest^{k+1} = gbest^k$

Step 11. If k < Maxite then k = k + 1 and goto step 6
else goto step 12

Step 12. Set optimum solution as

B. Flow Diagram of PSO algorithm



By using the PSO, the best features are selected from the derived features. In PSO a proposed solution is defined by a set of parameters known as pBest. In this paper, 8 features namely Skewness, Kurtosis, Mean, Standard Deviation, Contrast, Energy, Entropy,Homogeneity are considered as particle for the Indian Pines Dataset. After initialization, the selection process is done using fitness function. The entropy is chosen as fitness function. The entropy is used to measure the randomness of gray level distribution. The bands in each feature are analyzed using the fitness function, which features having large number of bands showing the fitness value or above the fitness value are considered as the best features. To find best feature PSO pBest value of each particle is compared with fitness value if it is greater than fitness value assign current fitness as new Pbest else keep previous as pBest. Assign best particles pBest value to gBest. This process is repeated until it reaches the maximum iterations

Entropy in any system represents disorder, where in the case of texture analysis is a measure of its spatial disorder. A completely random distribution would have very high entropy because it represents chaos. This feature can be useful to tell us if entropy is bigger for heavy textures or for the smooth textures giving us information about which type of texture can be considered statistically more chaotic.

Entropy = $-\sum_i \sum_j p(i,j)\log_2 p(i,j)$ the features which are going to be evaluated through the fitness function is given to the PSO. The features are concatenated and processed using the Optimization technique. The features will be represented in the number of rows and column.

By using this procedure above the fitness value is evaluated and the best features are selected for the AVIRIS of Indian Pines dataset the 8 features are compared namely skew, kurtosis, mean Variance, energy, contrast, homogeneity, entropy are given for the fitness evaluation from that contrast and homogeneity is selected as best feature. Contrast feature gives difference in the color and brightness of the object whereas homogeneity provides the closeness of the distribution of the elements. Features are extracted by using Daubechies tap four wavelet (db4) and Daubechies tap eight wavelet (db8) among the extracted feature best features are selected by using Particle Swarm Optimization (PSO) and SVM average and overall accuracy is calculated. Accuracy of classes is obtained as the ratio of difference between total number of samples in a class and misclassified pixels to the total number of samples in the class

$$\text{Accuracy} = \frac{\text{Total Number of Samples} - \text{Misclassified Pixels}}{\text{Total Number of Samples}} \quad (1)$$

Overall accuracy as well as classwise accuracy for each class before PSO and after PSO was calculated and the results were tabulated.

ISSN (Online) 2394-6849

International Journal of Engineering Research in Electronics and Communication Engineering (IJERECE)
Vol 4, Issue 12, December 2017

IV. RESULTS AND DISCUSSION

Results obtained after classification by using SVM classifier for Indian Pines dataset using DB4 and DB8 wavelet is discussed in this section and the classwise accuracy and Overall average accuracy were calculated in this paper. In this dataset consists of sixteen classes each class obtained different accuracy for DB4 and DB8.

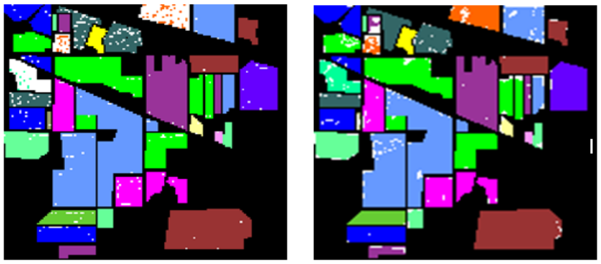


Fig. 2. Output For DB4 Without PSO, (b) For DB4 (With PSO)

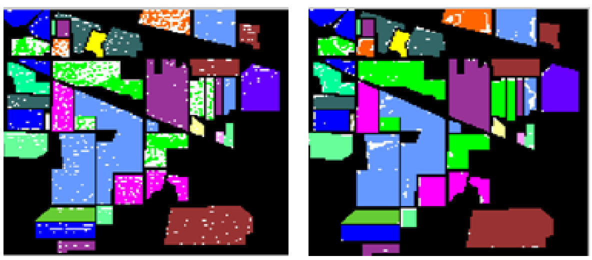


Fig. 3. Output For DB8 Without PSO, (b) For DB8 (With PSO)

Table 1 Accuracy table for various classes (DB4)

S.No	CLASS NAME	Average Accuracy(%)	
		Before PSO	After PSO
1	Alfalfa	95.46	96.5
2	Corn Notill	94.26	88.9
3	Corn Mintill	87.45	92.25
4	Corn	91.88	93.34
5	Grass Pasture	91.34	90.53
6	Grass Trees	92.37	92.06
7	Grass Pasture Mowed	73.08	99
8	Hay Windrowed	85.68	40.87
9	Oats	81	99.84
10	Soybean Notill	83.75	97.67
11	Soybean Mintill	94.53	94.78
12	Soybean Clean	95.00	90.67
13	Wheat	96.28	96.74
14	Woods	93.56	94.78
15	BGTD	85.97	92.45
16	Stone –Steel-Towers	92.65	82.98
	Overall accuracy (%)	**89.64**	**92.68**

Table 2 Accuracy table for Various classes (DB8)

The results tabulated in the Table 1 shows the accuracy for the classes in the AVIRIS Dataset has achieved very less before the selection of the best features. The classification of classes has achieved the accuracy level of 89.64% due to the 8 features namely energy,entropy, contrast, homogeneity, skew, kurtosis, mean, Variance and due to the high value and bands. But after applying PSO the classification of classes has achieved the accuracy level of 92.68%. The classes such as Grass Pasture Mowed and Oats had been achieved average of 98% the Overall accuracy as well as Average accuracy has been increased.

S.NO	CLASS NAME	Average Accuracy(%)	
		Before PSO	After PSO
1	Alfalfa	75	99
2	Corn Notill	89.9	88.8
3	Corn Mintill	89.3	98.5
4	Corn	70.6	96.7
5	Grass Pasture	76.4	90.5
6	Grass Trees	79.9	98.85
7	GrassPasture Mowed	72.0	99.7
8	Hay Windrowed	93.2	98.3
9	Oats	46.0	39.0
10	Soybean Notill	80.5	99.2
11	Soybean Mintill	71.6	95.3
12	Soybean Clean	50.3	94.9
13	Wheat	45.2	99.1
14	Woods	82.0	93.2
15	BGTD	32.4	99.7
16	Stone –Steel-Towers	55.84	98.36
	Overall accuracy (%)	**87.07**	**90.85**

The results tabulated in the Table 1 shows the accuracy for the classes in the AVIRIS Dataset has achieved very less before the selection of the best features. The classification of classes has achieved the accuracy level of 89.64% due to the 8 features namely energy,entropy, contrast, homogeneity, skew, kurtosis, mean, Variance and due to the high value and bands. But after applying PSO the classification of classes has achieved the accuracy level of 92.68%. The classes such as Grass Pasture Mowed and Oats had been achieved average of 98% the Overall accuracy as well as Average accuracy has been increased. Table 2 indicates the accuracy for the classes before PSO it achieved the overall accuracy of 87.07% and after applying PSO overall accuracy has been increased to 90.85%

V. CONCLUSION

In Hyperspectral image Classification using SVM classifier is experimented we infer that after selecting best feature overall accuracy for dataset has been increased.DB4 provide better accuracy when compared DB8 Without using the any Optimization technique the highest overall accuracy achieved for the Indian Pines dataset is nearly 89.64%, but by using the PSO and by using its fitness function evaluation the
overall accuracy for the Indian Pines dataset has been increased to 92.68% for DB4 and 90.85% for DB8 DWT.

REFERENCES

[1] Guo, Baofeng, Steve R. Gunn, Robert I. Damper, and James DB Nelson. "Band selection for hyperspectral image classification using mutual information." IEEE Geoscience and Remote Sensing Letters 3, no. 4 (2006): 522-526.

[2] Shwetank, Jain Kamal, and K. J. Bhatia. "Review of rice crop identification and classification using hyper-spectral image processing system." International Journal of Computer Science & Communication 1, no. 1 (2010): 253-258.

[3] Hughes, Gordon. "On the mean accuracy of statistical pattern recognizers." IEEE transactions on information theory 14, no. 1 (1968): 55-63.

[4] Dash, Manoranjan,and Huan Liu. "Feature selection for classification." Intelligent data analysis 1, no. 1-4 (1997): 131-156.

[5] Qi,Chengming, Zhangbing Zhou, Yunchuan Sun, Houbing Song, Lishuan Hu, and Qun Wang. "Feature selection and multiple kernels boosting framework based on PSO with mutation mechanism for hyperspectral classification." Neurocomputing 220 (2017): 181-190.

[6] Zheng, Xiangtao, Yuan Yuan, and Xiaoqiang Lu. "Dimensionality Reduction by Spatial–Spectral Preservation in Selected Bands." IEEE Transactions on Geoscience and Remote Sensing 55, no. 9 (2017): 5185-5197.

[7] Patra, Swarnajyoti, Prahlad Modi, and Lorenzo Bruzzone. "Hyperspectral band selection based on rough set." IEEE Transactions on Geoscience andRemoteSensing 53, no. 10 (2015): 5495-5503

[8] Li, Shijin, Hao Wu, Dingsheng Wan, and Jiali Zhu. "An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine." Knowledge-Based Systems 24, no. 1 (2011): 40-48

[9] Waske, Björn, and Jon Atli Benediktsson. "Fusion of support vector machines for classification of multisensor data." IEEE Transactions on Geoscience and Remote Sensing 45, no. 12 (2007): 3858-3866.

[10] Melgani, Farid, and Lorenzo Bruzzone. "Classification of hyperspectral remote sensing images with support vector machines." IEEE Transactions on geoscience and remote sensing 42, no. 8 (2004): 1778-1790

[11] Xu,Yan, Qian Du, and Nicolas H. Younan. "Particle swarm optimization-based band selection for hyperspectral target detection." IEEE Geoscience and Remote Sensing Letters 14, no. 4 (2017): 554-558

[12] Sun, Zhuo, Cheng Wang, Peng Li, Hanyun Wang, and Jonathan Li. "Hyperspectral image classification with SVM-based domain adaption classifiers." In Computer Vision in Remote Sensing (CVRS), 2012 International Conference on, pp. 268-272. IEEE, 2012.

[13] Samadzadegan, Farhad, and F. Tabib Mahmoudi. "Optimum band selection in hyperspectral imagery using swarm intelligence optimization algorithms." In Image Information Processing (ICIIP), 2011 International Conference on, pp. 1-6. IEEE, 2011.

[14] Ruimin Li, Liaoying Zhao, Xiaorun Li " Band Selection for Hyperspecral Imagery Based on Particle Swarm Optimization ".In Image and Signal Processing (CISP), 2015 8th International Congress on, pp. 346-350. IEEE, 2015.

[15] Mercier, Grégoire, and Marc Lennon. "Support vector machines for hyperspectral image classification with spectral-based kernels."In Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International, vol. 1, pp. 288-290. IEEE, 200

[16] Mercier, Grégoire, and Marc Lennon. "Support vector machines for hyperspectral image classification with spectral-based kernels." In Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International, vol. 1, pp. 288-290. IEEE, 200

[17] Sabale, Savita P,and Chhaya R. Jadhav. "Hyperspectral Image Classification Methodsin Remote Sensing-A Review." In Computing Communication Control and

Automation (ICCUBEA), 2015 International Conference on, pp. 679-683. IEEE, 2015.

[18] Ghamisi, Pedram, Micael S. Couceiro, and Jon A. Benediktsson. "Classification of hyperspectral images with binary fractional order Darwinian PSO and random forests." In Proc. SPIE, vol. 8892, pp. 1-8. 2013

[19] Melgani, Farid, and Lorenzo Bruzzone. "Classification of hyperspectral remote sensing images with support vector machines." IEEE Transactions on geoscience and remote sensing 42, no. 8 (2004): 1778-1790. Genshe Chen. "Particle swarm optimization-based hyperspectral dimensionality

[20] Yang, He, Qian Du, and reduction for urban land cover classification." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5, no. 2 (2012): 544-554

[21] M. Chi and L. Bruzzone, "Semisupervised classification of hyperspectral images by SVMs optimized in the primal", IEEE Trans. Geosci. Remote Sens., vol. 45, no. 6, pp. 1870–1880, Jun. 2007.

[22] Xia, Junshi, Naoto Yokoya, and Akira Iwasaki. "Hyperspectral image classification with canonical correlation forests." IEEE Transactions on Geoscience and Remote Sensing 55, no. 1 (2017): 421-431

[23] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines,"IEEE Trans. Geosci. Remote Sens., vol. 42, no. 8, pp. 1778–1790, Jan. 2004

[24] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new Bayesian approach with active learning,"IEEE Trans.Geosci. Remote Sens., vol. 49, no. 10, pp. 3947–3960, Oct. 2011

[25] X. Guo, X. Huang, L. Zhang, L. Zhang, A. Plaza, andJ. A. Benediktsson, "Support tensor machines for classification of hyperspectral remote sensing imagery,"IEEE Trans. Geosci. RemoteSens., vol. 54, no. 6, pp. 3248–3264, Jun.

