International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

Clustering and classification of Images Using ABC-FCM and Naive Bayes Classifier

S. Praveena

ECE Dept, M.G.I.T, Hyderabad, India

Abstract: ***This paper presents a hybrid clustering algorithm and Naive Bayes classifier for trees, shade, building and road. It starts with the single step preprocessing procedure to make the image suitable for clustering. The pre-processed image is clustered using the ABC-FCM algorithm that is developed by hybridizing the ABC and FCM algorithm to obtain the effective clustering in satellite image and classified using Naive Bayes classifier. Classification accuracy and Mean square Error of hybrid algorithm is found better than ABC.***

Keywords: Clustering, classification, features extraction, Naïve Bayes classifier, ABC-FCM algorithm

1. Introduction

The main objective of image segmentation is to partition the image into parts of strong correlation with objects or areas of the real world image

At present, in the Literature, an extensive variety of satellite image categorization methods are: Cluster, Statistic, Bayesian Net, Artificial Neural Networks (ANN), etc [4]. By visual understanding of satellite imagery and aerial photography, now, operators can manually remove cartographic features, such as buildings, roads, and trees. Frequently consigned to as ''soft'' classifiers, as opposed to the non-fuzzy ''hard'' classifiers, with their applicability being of still larger significance under the survival of varied pixels.

2. Review of Literature

In the image processing domain, Satellite image categorization is an extremely taxing task. In order to progress precision, researchers have applied different kinds of classification techniques. According to the database of skilled knowledge for a more focused satellite image classification, a hybrid biologically inspired method was adapted that was presented by Lavika Goel [1].

As a competent land cover classifier for satellite image, a hybrid FPAB/BBO based algorithm has been presented by Navdeep Kaur Johal *et al.* [2].

A FPAB/BFO based algorithm for the categorization of satellite image has been presented by Parminder Singh *et al.* [3]. To choose the optimal and minimum set of fuzzy rules, the use of a genetic algorithm (GA) has been presented by O. Gordo *et al*. [4] and to categorize remotely sensed images using a fuzzy classifier.

A generalized multiple-kernel fuzzy C-means clustering (MKFCM) methodology for satellite image segmentation has been presented by M. Ganesh and V. Palanisamy [5]..

This paper, proposes a significant satellite image classification technique using ABC-GA clustering algorithm and Naïve Bayes classifier. The proposed approach consists of three steps

i) Pre-processing
ii) Clustering ABC-FCM algorithm
iii) Classification using Naïve Bayes Classifier.

Initially pre-processing is performed to make the image suitable for segmentation. In segmentation, the pre-processed image is segmented using hybrid ABC-FCM algorithm that is developed by hybridizing the ABC algorithm [6] and FCM algorithm to obtain the effective segmentation in satellite images. Then, feature is extracted and the classification of satellite image into four different labels (tree, shade, road and building) is done using Naïve Bayes Classifier.

3. Hybrid Clustering Algorithm and Naïve Bayes Classifier for Satellite Image Classification

3.1 Segmentation using Genetic-ABC Algorithm

In this work different colors of trees, roads , shades and buildings regions are taken in the database. Extract H,S,L,T,U,V layers. For each layer calculate features like mean and maximum value of histogram. Train these features using Naïve Bayes classifier.
During testing phase, a test image is given. For this apply the ABC-FCM clustering algorithm.

Step 1: Generate the new solutions using Eq.(3)

Step 2: ***Fitness Computation***
The fitness of each chromosome in the solution is evaluated with objective function presented by Eq(1). The objective function of the clustering process taken for this paper is intra cluster distance that is computed by the following equation

$$J_1(Y)=\frac{1}{\sum_{i=1}^{N}\min\left\{\left\|X_i-X_j\right\|\right\}};\ j=1,2,\ldots,k \quad \text{---- (1)}$$

Where, $\|X_i - X_j\|$ is a chosen distance measure between a data point X_i and the cluster centre X_j, 'N' and 'k' are the number of input data and the number of cluster centres, respectively. The finding of objective function of clustering process is illustrated in Fig(4).

Step 3: *Employed bee phase*

After initialization, in the employed bees' phase, each employed bee is sent to the food source in its memory and finds a neighboring food source. Here, 20% chromosomes are randomly generated for every new chromosome generation process. Subsequently, we have determined best fitness from the employed bee phase.

Step 4: *Onlooker bee phase*

In the onlooker bees' phase, the onlookers receive the information of the food sources shared by employed bees. Then each chromosome has chosen a solution to exploit depending on a probability related to the nectar amount of the solution (fitness values of the solution). That is to say, there may be more than one onlooker bees choosing a same solution if the source has a higher fitness. The probability P_i is calculated according to Eq(2) as followed:

$$P_i = \left[\frac{0.25}{Max\ fitness}\right] * fit + 1 \qquad (2)$$

After solutions have been chosen, each onlooker bee finds a new solution in its neighbourhood following Eq(3),

$$x_{i,j} = x_j^{\min} + rand(0,1)(x_j^{\max} - x_j^{\min}) \qquad (3)$$

Step 5: Scout bee phase

In scout bees' phase, if the value of trials counter of a solution is greater than a parameter, known as 'limit', the solution is abandoned and the bee becomes a scout bee. A new food solution has produced randomly in the search space using Eq(3), as in the case of initialization phase. The employed, onlooker, mutation and scout bees' phases will recycle until the termination condition is met. The best food source which presents the best solution.

Step 6: FCM Operation

The best solution obtained from the ABC process is taken for the FCM operation. The FCM is a clustering technique that allows a single pixel to belong to two or more clusters

3.3 Classification using Naïve Bayes Classifier

Classification step is to identify road, building, tree and shadow regions from original satellite image.

3.4 Evaluation Metrics

Accuracy

The evaluation of proposed technique in different satellite images are carried out using the following metrics as suggested by below equation.

$$\text{Accuracy} = \frac{\text{number of true positives} + \text{number of true negatives}}{\text{number of true positives} + \text{false negatives} + \text{true negatives} + \text{false positives}}$$

Mean Square Error

The Mean Square Error (MSE) for a clustering is done by summing the least distance value between the centroid and data of each cluster. The best cluster would have the nearest data of the centroid in it. It is defined as follows:

$$MSE = \sum_{n=1}^{N} \min dis(X_b, C_n) \qquad where,\ b = 1,2,....,T$$

In the above equation N is the total number of clusters, X_b is the b^{th} data in n^{th} cluster, C_n is the centroid of n^{th} cluster and T is the total number of data in the n^{th} cluster.

4. Experimental Results

The proposed technique is designed for identify the four different labels (tree, shade, road and building) of satellite images. The obtained experimental results from the proposed technique are given in figure 7 (a)-(e).These figures shows the extracted regions (tree, shade, road and building) from the Input satellite images using Naïve Bayes Classifier with kernel –normal.

Figure 7(a): Input image

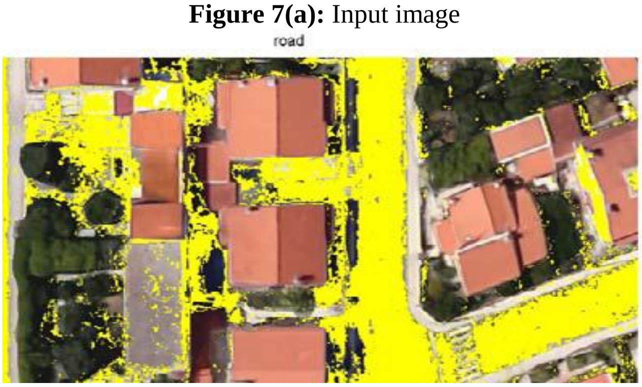

Figure 7(b): Road

International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

Figure 7(C): Shade

Figuere 7(d): Tree

Figure 7(e): Building

5. Comparative Analysis

In this paper, the proposed algorithm results are compared with against ABC. The performance of this is analyzed with parameter accuracy for five images. The accuracy value is computed by dividing the total number of similar pixels identified as land use to the number of pixels in the tree, shade, building and road region. The detailed results obtained for proposed technique is shown in Table(1). As shown in Table(1) the proposed technique is achieved the maximum accuracy value compared to ABC algorithm for all images Hence, it is clear from result analysis that the proposed algorithm shows better results as compared with existing techniques.

Table 1: Comparison of Accuracy

Image index	*Techniques*	*Tree*	*Shades*	*Roads*	*Building*
I_1	**ABC-FCM**	**0.84**	**0.7034**	**0.79**	**0.8045**
	ABC	0.77	0.6502	0.765	0.74619
I2	**ABC-FCM**	**0.80**	**0.67**	**0.77**	**0.71625**
	ABC	0.78	0.650	0.7	0.68436
I3	**ABC-FCM**	**0.80**	**0.723**	**0.884**	**0.7581**
	ABC	0.78	0.714	0.803	0.73980
I4	**ABC-FCM**	**0.85**	**0.825**	**0.949**	**0.91891**
	ABC	0.80	0.765	0.866	0.86733
I5	**ABC-FCM**	**0.83**	**0.87**	**0.881**	**0.91346**
	ABC	0.80	0.82	0.861	0.92116

Table 2 indicates the comparison of MSE values for two algorithms. The less value of MSE indicates the better clustering algorithm. The two algorithms are applied on 5 input images. For all images it is found that proposed algorithm has less MSE value which indicates the better performance.

Table 2: Comparison of MSE

Image index	*ABC-FCM*	*ABC*
I_1	**0.063293**	0.088031
I2	**0.045774**	0.078312
I3	**0.061098**	0.090004
I4	**0.063293**	0.088031
I5	**0.045774**	0.078312

6. Conclusion

In this paper, optimization algorithms for segmentation with the intention of improving the segmentation in satellite images using Naïve Bayes classifier is proposed. genetic-ABC algorithm, and iii) classification using Naïve Bayes classifier. In segmentation, the image is clustered using hybrid ABC algorithm that is developed by hybridizing the ABC algorithm and FCM algorithm to obtain the effective segmentation in satellite images. Then, feature is extracted and the classification of satellite image into four different labels (tree, shade, road and building) is done using Naïve Bayes Classifier. Finally, classification accuracy and MSE of the proposed algorithm in satellite image classification is calculated and the performance is compared with ABC.

References

[1] Lavika Goel,"Land Cover Feature Extraction using Hybrid Swarm Intelligence Techniques - A Remote Sensing Perspective," ACEEE Int. J. on Signal & Image Processing, Vol. 01, No. 03, Dec 2010

[2] Navdeep Kaur Johal, Samandeep Singh, Harish Kundra, “Hybrid FPAB/BBO Algorithm for Satellite Image

Classification," International Journal of Computer Applications, pp.0975- 8887, Vol.6, No.5, September 2010

[3] Parminder Singh, Navdeep Kaur, Loveleen Kaur,"Satellite Image Classification by Hybridization of FPAB Algorithm and Bacterial Chemotaxis,"International Journal of Computer Technology and Electronics Engineering (IJCTEE), Vol.1, No.3, 2011.

[4] O. Gordo, E. Martínez, C. Gonzalo and A. Arquero,"Classification of Satellite Images by means of Fuzzy Rules generated by a Genetic Algorithm,"IEEE Latin America Transactions, Vol. 9, No. 1, March 2011

[5] M. Ganesh, V. Palanisamy,"Multiple-Kernel Fuzzy C-Means Algorithm for Satellite Image Segmentation, "European Journal of Scientific Research, ISSN 1450-216X,Vol.83 No.2, pp.255-263,2012.

[6] Changsheng Zhang, Dantong Ouyang, Jiaxu Ning, "An artificial bee colony approach for clustering", Expert Systems with Applications, vol. 37, pp. 4761–4767, 2010.

[7] Lillesand, T.M., Kiefer, R.W. and Chipman J.W., "Remote Sensing and Image Interpretation", Fifth Edition, Wiley & Sons Ltd., England, pp.586-592, 2003.