International Journal of Fuzzy System Applications
Volume 8 • Issue 2 • April-June 2019

Classification of Remote Sensing Imagery Based on Density and Fuzzy c-Means Algorithm

Trinh Le Hung, Le Quy Don Technical University, Hanoi, Viet Nam
Mai Dinh Sinh, Le Quy Don Technical University, Hanoi, Viet Nam

ABSTRACT

The goal of data clustering is to divide a set of data into different clusters, so that the data in the same cluster show some similar characteristics. There are many clustering methods for satellite image segmentation, such as k-means, c-means, iso-data, minimum distance algorithms. Each method has certain advantages and disadvantages, but generally they are based on brightness value to divide the pixels of the image in to clusters. Actually, the probability of occurrence of frequency of appearance of pixel has certain effects on clustering results. In this article, the authors propose a method for clustering satellite imagery based on density. It consists of two main steps: find cluster centroid using density and data clustering using fuzzy c-Means algorithm (DFCM). The results obtained in this study can be used to potentially improve classification accuracy of satellite image.

KEYWORDS

Density, Fuzzy Clustering, Fuzzy C-Means, Multi-Spectral Image, Remote Sensing

INTRODUCTION

Remote sensing data clustering is an extremely important part of satellite image processing (Torahi & Chai, 2011; Mai, Trinh & Ngo, 2016; Mai & Ngo, 2015; Mai & Ngo, 2018). The results of satellite image classification can be used for a variety of purposes, such as natural resource research and environmental monitoring, urban planning and ensure national defense and security. Meanwhile, optical remote sensing data sources are often affected by weather conditions and the accuracy of the receiver, this make the image classification more complicate (Ngo, Mai & Pedrycz, 2015). In fact, uncertainty is inherently present in decision making. As such, it is increasingly imperative to research and develop new theories and methods based on fuzzy clustering (Li, 2017).

There are many satellite image classification methods (Han, Chi & Yeon, 2005; Gordo, Martinez, Gonzalo & Arquero, 2013), such as manual thresholds methods (Yang et al., 2016), unsupervised classification methods (Genitha & Vani, 2013), supervised classification methods (Jog and Dixit, 2016), fuzzy clustering method (Rauf, Valentin & Leonid, 2009) and method based on intuitionistic

DOI: 10.4018/IJFSA.2019040101

fuzzy sets (Li, 2004; Li & Cheng, 2002). These methods often use some common algorithms, such as K-means, c-Means, Iso-data, minimum distance and Fuzzy c-means. These clustering algorithms are essentially using the same strategy based on brightness to split into clusters (Jog & Dixit, 2016; Rauf et al., 2009) without regard to the density of the pixels, while high density pixels are usually located near the centroid of the cluster (Peherstorfer, Pflüger & Bungartz, 2012; Chen, Yan & Wang, 2014; Benmouiza & Cheknane, 2016).

Many scientists in the field of remote sensing data processing have proposed clustering methods based on density of pixels, in which density based spatial clustering of applications with noise (DBSCAN) is commonly used for satellite image classification (Khan, Rehman, Aziz, Fong & Saravady, 2014; Benmouiza & Cheknane, 2016). This algorithm requires only one input parameter and supports the users in determining an appropriate value for it. It discovers clusters of arbitrary shape and divides high density areas into cluster without depend on the size of data. In terms of implementation, this algorithm is also difficult to find the optimal radius of the density function around each pixel. In addition, the execution time of this algorithm is quite slow, especially when tested on large datasets, such as satellite imagery (Ngo, Mai & Nguyen, 2012). To overcome these limitations, many scientists are interested in improving this algorithm. Peherstorfer et al. (2012) presented a grid-based density estimation method to improve the speed of clustering. Chen et al. (2014) improved the DBSCAN algorithm by expanding the clusters which uses the margins of the objects, such as a pixel, to reduce the computation time. These improvements significantly reduce clustering time; however, affect the accuracy of clustering results.

To solve the above problem, this study proposed a method for approximating the centroid of cluster based on the density of pixels. Next step, the authors use approximation centroids for classification satellite imagery using the fuzzy c-means algorithm.

PROPOSAL METHODOLOGY

Scientific Basic

Density

The concept of density can be understood as the quantity representing the amount of matter in unit of measure (length, area, volume). So, the pixel density is the frequency of the pixel per unit of measure. Usually, the centroid of cluster is the average value of the pixels, so if the pixel has high frequency of appearance, that pixel is closer to the centroid of cluster (Ngo et al., 2012).

Fuzzy C-Means Algorithm

In general, fuzzy memberships in fuzzy c-means clustering algorithm (FCM) achieved by computing the relative distance among the patterns and cluster centroids (Bezdek & Ehrlich, 1984). Hence, to define the primary membership for a pattern, FCM algorithm defines the membership using value of *m*. The use of fuzzifier gives different objective function as follows:

$$J_m(U,v) = \sum_{k=1}^{N}\sum_{i=1}^{C}(u_{ik})^m d_{ik}^2 \tag{1}$$

where $d_{ik} = \left\| x_k - v_i \right\|$ is Euclidean distance between the pattern x_k and the centroid v_i, *C* is number of clusters and *N* is number of patterns, *m* is any real number greater than 1.

Degree of membership u_{ik} is determined as follow:

$$u_{ik} = \frac{1}{\sum_{j=1}^{C} \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}}, \; i = 1,\dots,C \;;\; k = 1,\dots,N \tag{2}$$

With $d_{ik} = \left\| x_k - v_i \right\|$ or $d_{jk} = \left\| x_k - v_j \right\|$ is Euclidean distance between the pattern x_k and the centroid v_i or v_j, C is number of clusters and N is number of patterns. u_{ik} is value of membership function between the pattern x_k and the centroid v_i.

Cluster centroid is computed as follows:

$$v_i = \frac{\sum_{k=1}^{N} (u_{ik})^m x_k}{\sum_{k=1}^{N} (u_{ik})^m}, \; i = 1,\dots,C \tag{3}$$

Next, defuzzification for FCM is made as if $u_i(x_k) > u_j(x_k)$ for j=1, ..., C and $i \neq j$ then x_k is assigned to cluster i.

Proposal Methodology

One of the difficulties of clustering algorithms is the initialization of the initial cluster centroid (Jog & Dixit, 2016; Rauf et al., 2009), this affects the steps taken and results clustering, if the centroid of the initiator cluster is too close together or too far apart, it will easily lead to local convergence, which makes the clustering algorithm is low accuracy or is unstable. There should be an approach to initializing the centroid of clusters that makes clustering algorithms stable and efficient. In this study, initialization of cluster center was proposed based on the density of pixels and the fuzzy c-Means algorithm applied to the land cover classification on satellite imagery.

In fact, the image information is stored as numeric values so the problem of image partitions is usually based on the degree of similarity among these values to decide whether an object belongs to any region in the image. Hence, the key to determine a pixel will belong to certain area is based on the similarity of brightness, which is calculated through a function of the distance in the color space d_{ik} between the pattern x_k and the centroid v_i.

In that, the centroid will be in the samples that the density surrounding the sample data are large. The concept of statistical variance mathematical model is used to solve the problem of selecting a surrounding data points. For the first step, the expected pattern $\bar{z}_i$ is computed by the following equation:

$$\overline{z}_i = \frac{1}{d} \sum_{i=1}^{d} x_{ij}, \; j = \overline{1, N} \tag{4}$$

And standard deviation s_i:

$$s_i = \sqrt{\frac{1}{d} \sum_{i=1}^{d} \left(x_{ij} - \overline{z}_i \right)^2}, \; j = \overline{1, N} \tag{5}$$

with *i*=1,2, ..., *d*; $X = (x_1, x_2, ..., x_N)$, $x_k \in R^d, k = \overline{1, N}$.

Considering the surround of each data point is *m*-dimensional box with radius defined by the standard deviation is $r = \min_{1<i<d} s_i$. Compute density D_i of pattern x_i:

$$D_i = \sum_{j=1}^{N} T(r - \| x_j - x_i \|) \text{ with } T(z) = \begin{cases} 1 & if\ z \geq 0 \\ 0 & if\ z < 0 \end{cases} \tag{6}$$

In which $T = 1$ if $z \geq 0$ otherwise T=0.

Call V_c is a set of pixels in order of density from high to low. Find pixel satisfying the condition:

$$D_i = \max_{1 \leq j \leq N} D_j \tag{7}$$

Put x_i into the result set V_c according to the following formula:

$$V_c = V_c \cup x_i \tag{8}$$

and $\mathrm{X}=\mathrm{X} \backslash \mathrm{x}_i$. If *X*= Θ given a set of candidate points V_c, else back to finding D_i.

If V_c is large then we can proceed with this algorithm to reduce the number of candidate clusters. The calculations can be speeded up by dividing the input data set into subsets, then the algorithm can be applied for that subset to finding candidates set V_i. Call V is the set of all candidates, then $\cup V_i = V$, apply this algorithm with set V to finding V_c set. The centroid matrix *V* can be initialized by choosing the patterns in V_c according to the density of candidates.

The above approach can approximate centroid of clusters, starting the FCM algorithm with these approximate centroid centers will reduce the number of iterations, computational time of the algorithm and improve the accuracy of clustering results. The detailed DFCM algorithm consists of the following nine main steps:

Input: Data set *X* with *N* data sample: $X = (x_1, ..., x_N)$ and $x \in R^d$, the number of cluster is C, stop condition ε.

Output: Data clusters.

Step 1: Calculate sample expectations and standard deviations by formula (4) and (5), radius of the sphere $r = \min_{1<i<d} s_i$ in the *m*-dimensional space.

Step 2: Density calculation D_i by formula (6).

Step 3: Find x_i by formula (7), and assign x_i to result set by formula (8).

Step 4: Calculated $Y = \{x_j, r_i - \| x_i - x_j \| \geq 0\}$ and set $X = X \setminus Y$, if $X = \phi$ then go to *Step* 5, else go to *Step* 1.

Step 5: Given set of centroids $V_c = \{v_j\}$.

Step 6: Calculate the value of the membership function according to the formula (2).

Step 7: Update centroid by formula (3).

Step 8: Stop condition: $\left\| J^{(t+1)} - J^{(t)} \right\| \leq \varepsilon$, if true go to *Step* 9, else go to *Step* 6.

Step 9: Assign the pixels to the cluster according to the formula (2) and given the clustering result.

The proposed algorithm consists of 9 steps, where steps 1 to 5 are performed to approximate centroid of clusters; steps 6 to 9 apply the fuzzy c-means algorithm to classify the land cover. This algorithm can be applied to different types of multispectral images, where the number of channels is the dimension of each pixel.

Indicator

To assess the effectiveness of algorithms and the quality of clusters, the authors use a number of indicators that are widely used in clustering problems such as Mean Squared Error (MSE) index (Wang & Bovik, 2009), Image Quality Index (IQI) (Wang & Bovik, 2002).

- MSE index:

$$MSE(x,y) = \frac{1}{N}\sum_{i=1}^{N}(\mathrm{x}_i - \mathrm{y}_i) \tag{9}$$

where, $X = \left\{x_i\right\} = \left\{x_1, x_2, \ldots, x_N\right\}$ and $Y = \left\{y_i\right\} = \left\{y_1, y_2, \ldots, y_N\right\}$ corresponding to the original image and segment results image. The smaller MSE value is the better quality cluster.

- IQI index:

$$IQI = \frac{4\sigma_{xy}\overline{x}\overline{y}}{(\sigma_x^2 + \sigma_y^2)(\overline{x}^2 + \overline{y}^2)} \tag{10}$$

with:

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad \overline{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\mathrm{x}_i - \overline{x})$$

$$\sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\mathrm{y}_i - \overline{y})$$

and:

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathrm{x}_i - \overline{x})(\mathrm{y}_i - \overline{y})$$

where:

$X = \left\{x_i\right\} = \left\{x_1, x_2, \ldots, x_N\right\}$ and $Y = \left\{y_i\right\} = \left\{y_1, y_2, \ldots, y_N\right\}$

corresponding to the original image and segment results image. The best value 1 is achieved if and only if $y_i = x_i$, the lowest value of -1 occurs when $y_i = 2\bar{x} - x_i$ with $i = \overline{1, N}$.

The one of simplest and most widely used criterion measure for clustering is Sum of Squared Error (SSE). It is defined as:

$$\mathrm{SSE} = \sum_{i=1}^{C} \frac{1}{N_i} \sum_{\forall x_k \in C_i} || x_k - v_i ||^2 \tag{11}$$

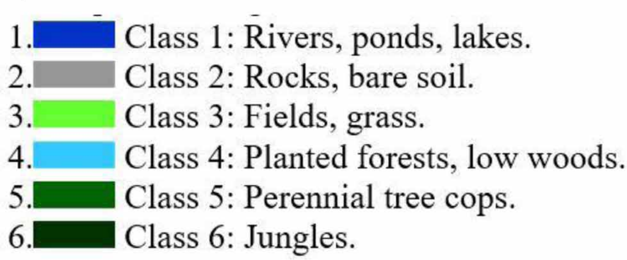
where C is the number of clusters, N_i is the number of element in i^{th} cluster and v_k is the centroid of i^{th} cluster.

EXPERIMENT

In the experiments, authors have selected the problem of classification on satellite imagery to test the proposed algorithm. In that, step 1 is the initial pre-processing step, select the processing area on the satellite image and Image geometry correction. Implement the DFCM algorithm to classifying into six clusters, based on the centroid of each cluster to assign corresponding to six types of land cover. The final step is to evaluate the quality of the classification results. The details of proposed algorithm for land cover classification from SPOT 5 and Landsat 7 ETM+ multispectral images consist of the following three main steps:

Step 1: Multispectral image preprocessing.
Step 2: Apply proposed method for classifying land cover in remotely sensed imagery. This multispectral image will be classified into six classes representing six types of land covers:

Key to be inserted under “EXPERIMENT” section



Step 3: Evaluate the accuracy of classification results.

The proposed method is programmed in Visual Studio C ++ 2010 with *m=2*.

Case #1

In this experiment, Landsat 7 Enhanced Thematic Mapper Plus (ETM+) image taken in 30 September 2009 covering a part of Hanoi city, Vietnam (see Figure 1) was used. Landsat 7 ETM+ image consists of eight spectral bands with a spatial resolution of 30 meters for band 1 to 5 and 7, 10 meters for panchromatic band (band 8). Spatial resolution for thermal infrared band (band 6) is 120 meters, but is resampling to 30 meters pixels.

The results of land cover classification were shown in Figure 2, in which 2a, 2b, 2c and 2d are classification results of K-means, DBSCAN, FCM and DFCM proposed algorithm, respectively. Table

Figure 1. Landsat 7 ETM+ image of the Hanoi region

1 showed the comparison of classification results obtained by using K-means, DBSCAN, FCM and DFCM algorithm. As can be seen, there was a significant difference on the area of regions classified by the aforementioned algorithms. The biggest difference was between the regions classified by K-means and DFCM algorithms.

In this study, to evaluate the quality of clusters, we considered the different validity indices, such as Mean Squared Error index (MSE) and Image Quality Index (IQI). It can be seen that the accuracy of land cover classification using K-means algorithm was very low. Many objects, such as bare soil and water, bare soil and sparse vegetation were misclassified. The accuracy of land cover classification was improved when using DBSCAN and FCM algorithms; however, it was not so high. The results of calculation of IQI and MSE indices by 4 algorithms K-means, DBSCAN, FCM and DFCM were shown in Table 2. It can be seen that the FCM and DFCM algorithms provided better classification result than the other algorithms, such as K-means and DBSCAN.

Case #2

The second experiment is selected in area of Chu Prong district, Gia Lai province (Central highlands of Vietnam). Remote sensing data used in the classification is SPOT 5 multispectral image in December 2009 (Figure 3). SPOT 5 multispectral images consist of five spectral bands with a spatial resolution of 10 meters for band 1 to 3, 20 m for short wave infrared (SWIR) band (band 4) and 5 meters (nadir) for panchromatic band (band 5).

In this example, we also classify the image into six classes as for Experiment 1. The classification results are shown in the Figures 4 (a-d). It can be seen, many objects, especially water, are misclassified when classification using K-means and DBSCAN algorithms (Figure 4a, 4b). These errors have significantly improved when using FCM, especially DFCM algorithms (Figure 4c, 4d).

Figure 2. Result of land cover classification from Landsat image: a) K-means; b) DBSCAN; c) FCM; d) DFCM proposed algorithm

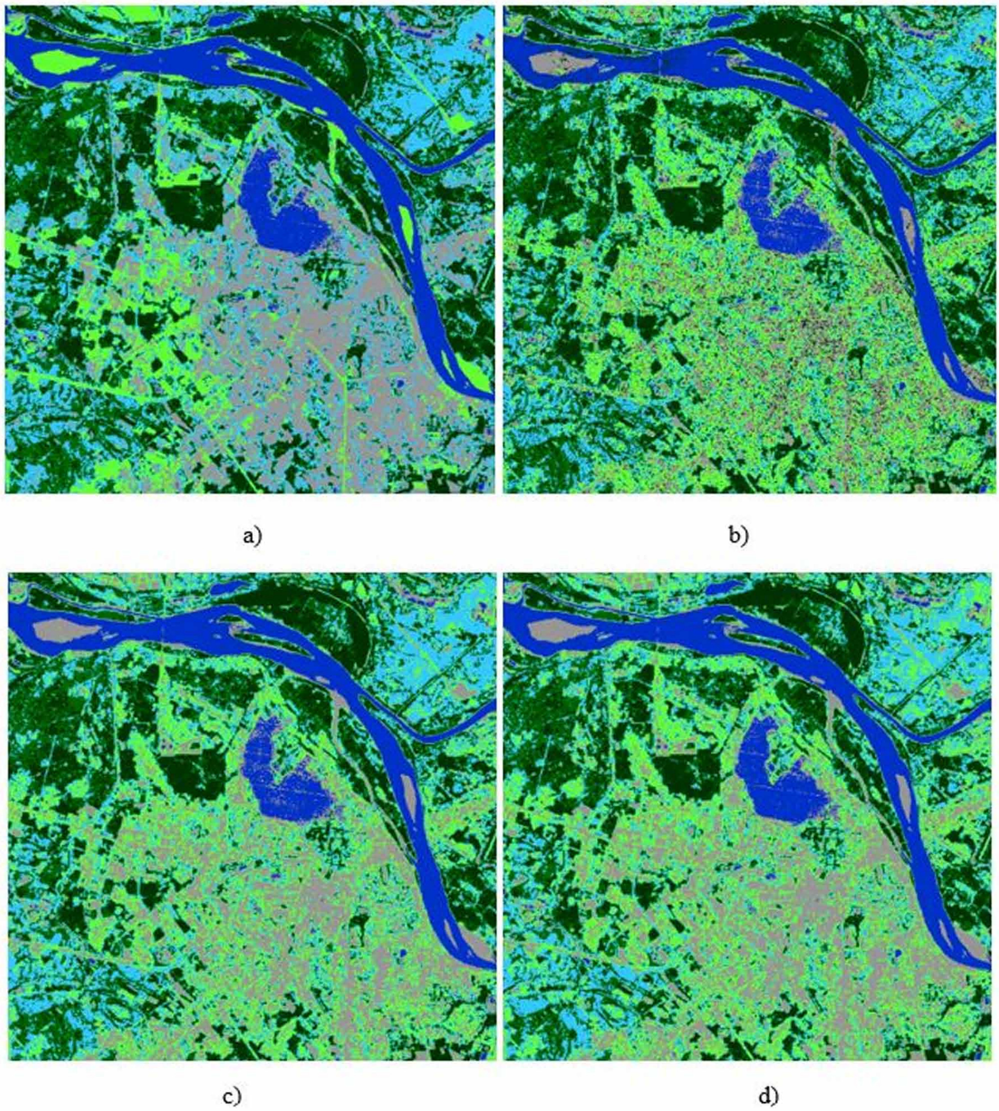


Table 1. Results of land cover classification in the experiment 1 region

Class	K-Means	DBSCAN	FCM	DFCM
1	8.753%	7.954%	7.055%	6.955%
2	14.447%	16.445%	18.443%	19.442%
3	13.218%	15.215%	20.210%	21.209%
4	25.659%	21.663%	15.670%	14.671%
5	23.651%	22.453%	19.656%	18.657%
6	14.272%	16.270%	18.967%	19.067%

Table 2. The various validity indices computed from Landsat image in the experimental case #1

Index	K-Means	DBSCAN	FCM	DFCM
MSE	0.2413	0.1721	0.0982	**0.0981**
IQI	0.2843	0.4183	**0.5643**	0.5631
SSE	147.9054	111.7842	86.4691	**69.4386**

Figure 3. SPOT 5 images in Chu Prong region, Gia Lai province

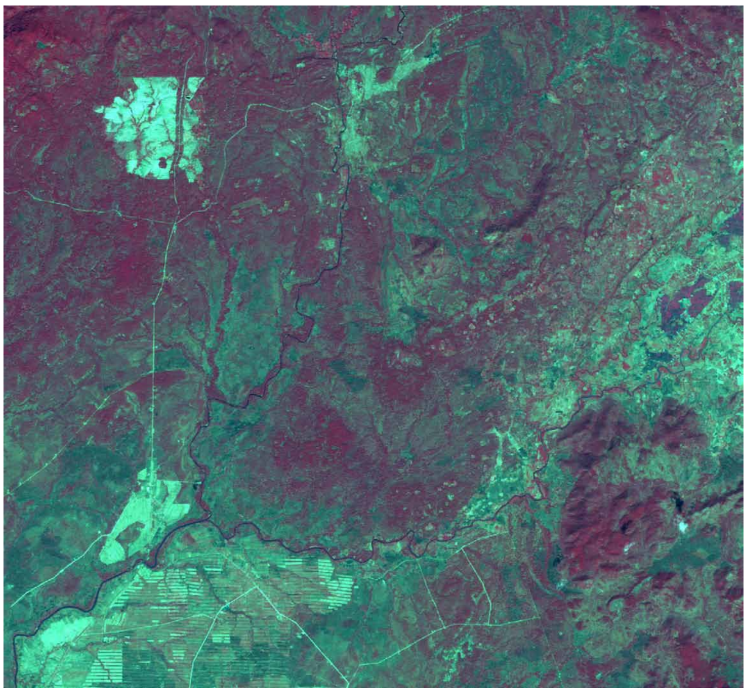


Table 3 shows the results of determining the area of six land cover classes. As compared with the results in the experiment 1, the results also showed that the area of land cover classes is different, in which the biggest difference between K-means and DFCM algorithms.

Table 4 presents the accuracy assessment of classification results using K-means, DBSCAN, FCM and proposed algorithms (DFCM). It can be seen that the K-means algorithm has the lowest accuracy in this data set. The accuracy of classification results is enhanced by using DBSCAN and FCM algorithms. The results also show that the accuracy of land cover classification is highest when using proposed algorithm (DFCM).

Case #3

In the experiment 3, we used SPOT 5 multispectral images (band 1 to 3) and short wave infrared (SWIR, band 4) bands in Tan An district, Long An province. The input image was taken in 2007 with spatial resolution 10 m for multispectral bands and 20 m for SWIR band (Figure 5).

As with the experiment 1 and experiment 2, six classes of land cover was classified by using K-means, DBSCAN, FCM and DFCM algorithms. The results of land cover classification from SPOT 5 image are shown in Figure 6 (a-d).

Figure 4. Result of land cover classification from SPOT image in Gia Lai province: a) K-means; b) DBSCAN; c) FCM; d) DFCM algorithm

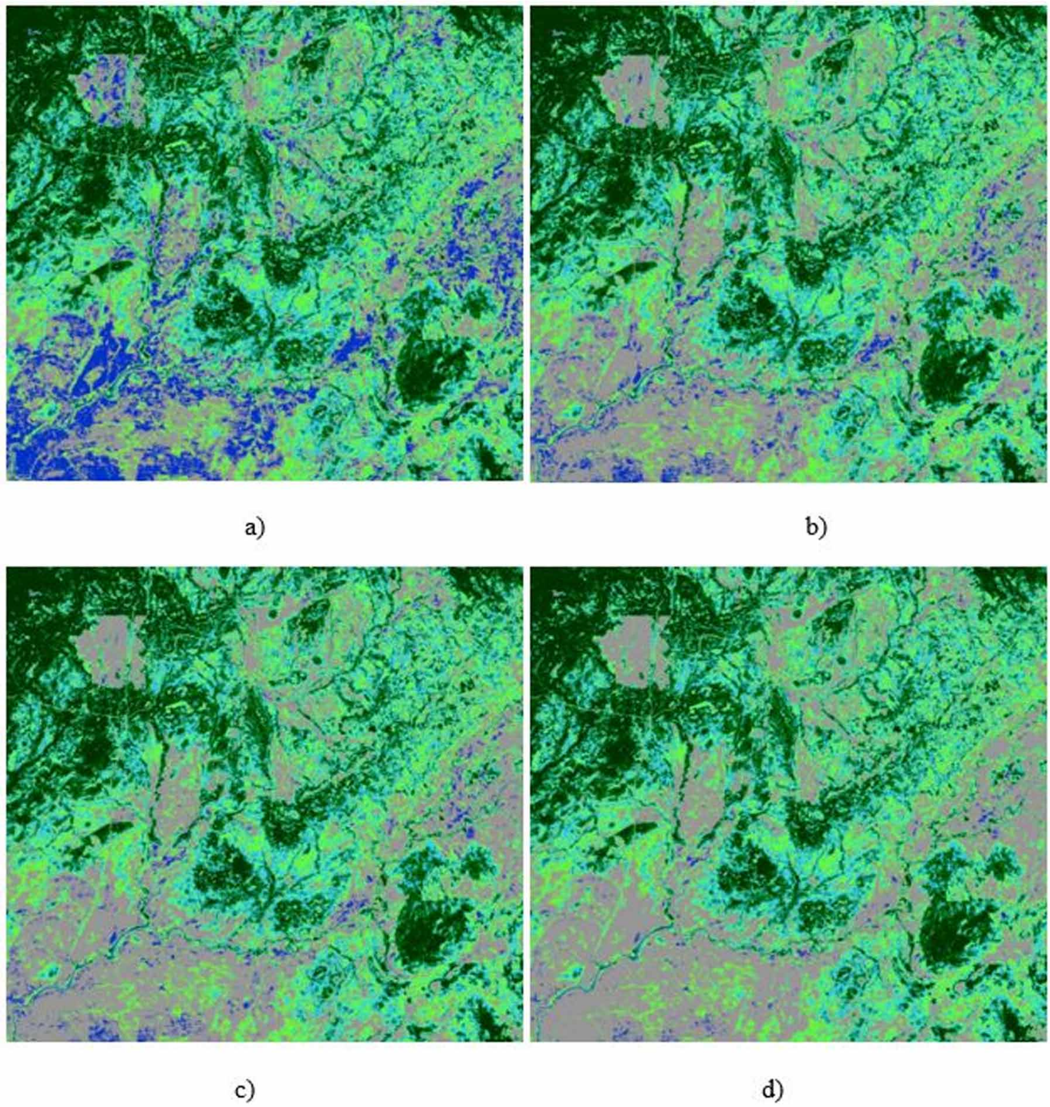


Table 3. Results of land cover classification in region of Dak Lak province

Class	K-Means	DBSCAN	FCM	DFCM
1	15.287%	10.530%	7.055%	6.016%
2	13.057%	16.294%	18.443%	19.638%
3	14.773%	14.631%	20.210%	21.423%
4	23.942%	20.831%	15.670%	14.819%
5	18.114%	21.590%	19.656%	18.845%
6	14.827%	16.125%	18.967%	19.259%

Table 4. The various validity indices from SPOT image in the experiment 2 region

Index	K-Means	DBSCAN	FCM	DFCM
MSE	0.3283	0.1982	0.1098	**0.0963**
IQI	0.1987	0.5762	0.6731	**0.6984**
SSE	132.9342	109.7648	79.7632	**65.9823**

Figure 5. SPOT 5 image in Tan an region, Long an province

The analytical results obtained show that class 1 (rivers, ponds, lakes) was misclassified into class 2 (rocks, bare soil). Table 5 indicates that there was a significant difference in the area of six land cover classes when classification using K-means, DBSCAN, FCM and DFCM algorithms. As well as the two experiments above, the biggest difference in area of six land cover classes is between K-means and DFCM algorithms.

Table 6 shows that the classification accuracy of K-means, DBSCAN, FCM and DFCM algorithms by using MSE and IQI indices. The obtained results show that the proposed algorithm has achieved the highest accuracy when compared to the other algorithms, such as K-means, DBSCAN and FCM.

In summary, from three test areas, the confusion in the classification is often between water and bare soil, especially wet soil, due to the difference in spectral characteristics is not great. The confusion is also found when vegetation classes classification, especially between grasses and trees. With satellite imagery average resolution, the differences of classification results can be acceptable in assessment of land cover on a large area, reducing costs compared to other methods.

Furthermore, the authors prove the robustness of DFCM algorithm by testing on images with different sizes (Table 7). The results showed that, the DFCM algorithm is rather stable and always faster

Figure 6. Result of land cover classification from SPOT image in Tan An province: a) K-means; b) DBSCAN; c) FCM; d) DFCM algorithm

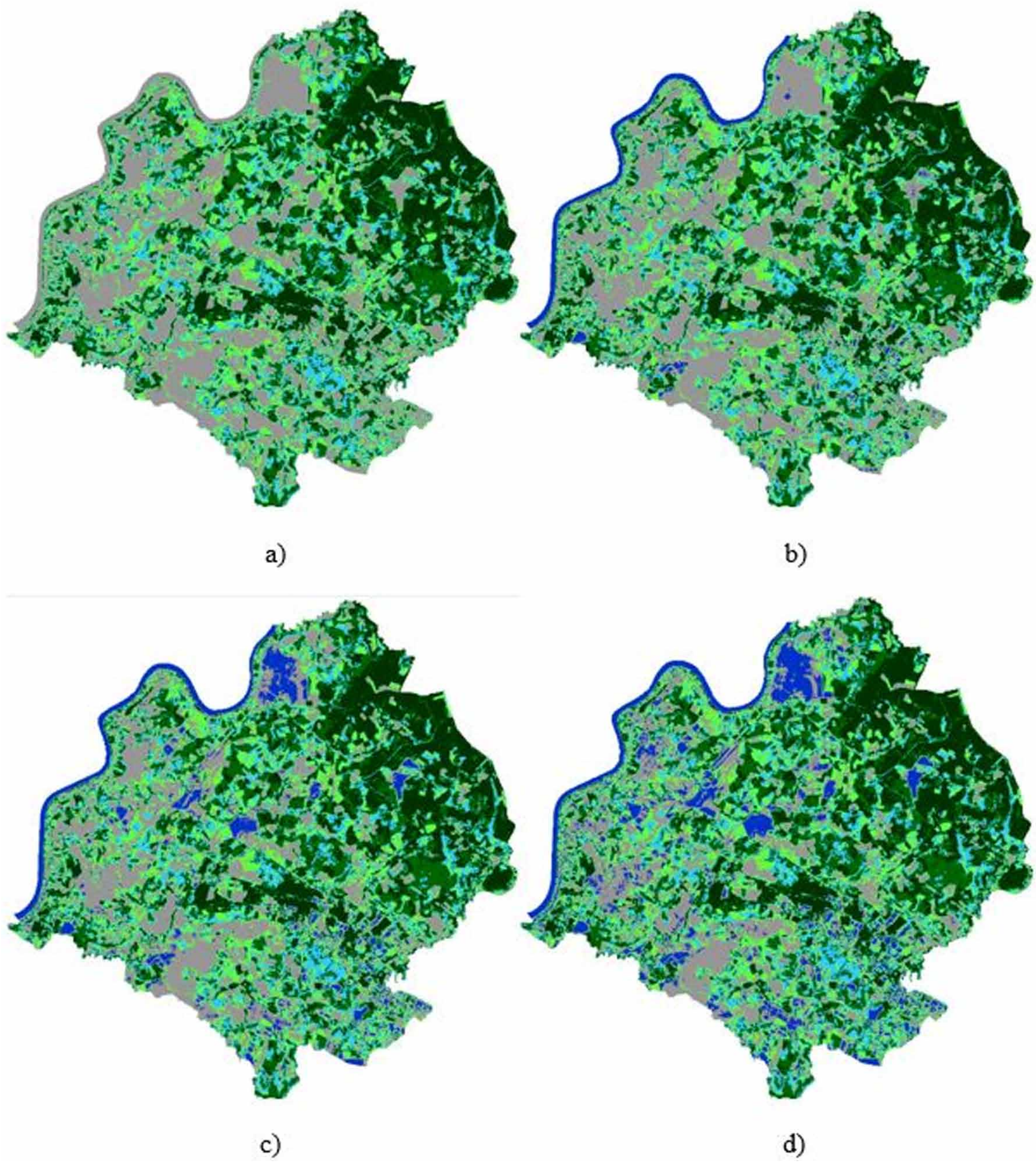


Table 5. Results of land cover classification in the experiment 3 region (Tan An area)

Class	K-Means	DBSCAN	FCM	DFCM
1	1.082%	2.978%	6.770%	12.457%
2	28.082%	26.186%	22.394%	16.706%
3	24.059%	22.164%	18.372%	16.476%
4	17.897%	17.518%	16.950%	16.570%
5	13.469%	15.365%	19.157%	21.053%
6	15.410%	15.789%	16.358%	16.737%

Table 6. The various validity indices computed from the SPOT image in the experiment 3 region (Tan An area)

Index	K-Means	DBSCAN	FCM	DFCM
MSE	0.5421	0.3761	0.1762	**0.1287**
IQI	0.3651	0.5198	0.6261	**0.7198**
SSE	126.8947	99.6823	102.7836	**75.8729**

Table 7. Comparison of computational time in seconds by various algorithms

STT	Size of Image (Pixel)	Computational Time (s)			
		K-Means	DBSCAN	FCM	DFCM
1	512×512	10.462	24.784	2.561	**1.873**
2	1024×1024	57.387	108.323	9.104	**3.562**
3	2048×2048	112.644	192.873	21.082	**5.213**

than the other three algorithms. Finally, the above experiments show that the DFCM algorithm the authors proposed is better than K-means, DBSCAN and FCM algorithms. Moreover, the computational time of the DFCM algorithm is also faster than the remaining 3 algorithms.

CONCLUSION

This paper presents a new method for land cover classification by combined pixel density and Fuzzy c-means algorithm. The results showed that the proposed algorithm has significantly improved the land cover classification accuracy. In all three experiments with other remotely sensed data, we carry out 30 times classifications with K-means, DBSCAN and FCM algorithms and then choose the best results. Meanwhile, we only perform once with proposed algorithm (DFCM).

In all three classification experiments from remotely sensed data, the MSE and SSE indices when using proposed algorithm are always smaller than using K-means, DBSCAN and FCM algorithms. Meanwhile, the IQI value when using DFCM algorithm is greater than using the three remaining algorithms (case #2 and case #3). For the case #1, the difference between IQI values when using FCM and DFCM algorithms is negligible, and much greater than using K-means and DBSCAN algorithms.

The proposed approach can be applied to other types of satellite images, which saves costs and time compared to other ways of land cover classification.

REFERENCES

Benmouiza, K., & Cheknane, A. (2016). Density-based spatial clustering of application with noise algorithm for the classification of solar radiation time series. In *2016 8th International Conference on Modelling, Identification and Control* (ICMIC) (pp. 279-283). doi:10.1109/ICMIC.2016.7804123

Bezdek, J., & Ehrlich, E. W. (1984). FCM: The Fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*(2), 1–11. doi:10.1016/0098-3004(84)90020-7

Chen, D., Yan, Y., & Wang, D. (2014). Density Clustering Based on Border Expanding. In *2014 10th International Conference on Natural Computation* (pp. 670-674). IEEE. doi:10.1109/ICNC.2014.6975916

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *1*(2). doi:10.1109/TPAMI.1979.4766909 PMID:21868852

Geerken, R., Zaitchik, B., & Evans, J. P. (2005). Classifying rangeland vegetation type and coverage from NDVI time series using Fourier Filtered Cycle Similarity. *International Journal of Remote Sensing*, *26*(24), 5535–5554. doi:10.1080/01431160500300297

Genitha, C. H., & Vani, K. (2013). Classification of satellite images using new Fuzzy cluster centroid for unsupervised classification algorithm. In *IEEE Conference on Information and Communication Technologies* (pp. 203-207). doi:10.1109/CICT.2013.6558090

Gordo, O., Martinez, E., Gonzalo, C., & Arquero, A. (2013). Classification of Satellite Images by means of Fuzzy Rules generated by a Genetic Algorithm. Latin America Transactions. *Revista IEEE América Latina*, *9*(1), 743–748. doi:10.1109/TLA.2011.5876414

Han, J. G., Chi, K. H., & Yeon, Y. K. (2005). Land Cover Classification of IKONOS Multispectral Satellite Data: Neuro-fuzzy, Neural Network and Maximum Likelihood Methods. In International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing, *LNCS* (Vol. *3642*, pp. 251-262). doi:10.1007/11548706_27

Jog, S., & Dixit, M. (2016). Supervised Classification of Satellite Images. In *2016 Conference on Advances in Signal Processing (CASP)* (pp. 93-98). doi:10.1109/CASP.2016.7746144

Khan, K., Rehman, S. U., Aziz, K., Fong, S., & Sarasvady, S. (2014). DBSCAN: Past, present and future. In *2014 Fifth International Conference on the Applications of Digital Information and Web Technologies (ICADIWT)* (pp. 232-238). IEEE. doi:10.1109/ICADIWT.2014.6814687

Li, D. F. (2004). Some measures of dissimilarity in intuitionistic fuzzy structures. *Journal of Computer and System Sciences*, *68*(1), 115–122. doi:10.1016/j.jcss.2003.07.006

Li, D. F. (2017). Theoretical and Practical Advancements for Fuzzy System Integration. Hershey, PA: IGI Global. doi:10.4018/978-1-5225-1848-8

Li, D. F., & Cheng, C. T. (2002). New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions. *Pattern Recognition Letters*, *23*(1-3), 221–225. doi:10.1016/S0167-8655(01)00110-6

Mai, D. S., & Ngo, L. T. (2015), Interval Type-2 Fuzzy C-Means Clustering with Spatial Information for Land-Cover Classification. In *The 7th Asian Conference on Intelligent Information and Database Systems (ACIIDS), LNAI* (Vol. 9011, pp. 387-397). Springer. doi:10.1007/978-3-319-15702-3_38

Mai, D. S., & Ngo, L. T. (2018). *Multiple kernel approach to semi-supervised fuzzy clustering algorithm for land-cover classification. Engineering Applications of Artificial Intelligence, 68*, 205–213. doi:10.1016/j.engappai.2017.11.007

Mai, D. S., Trinh, L. H., & Ngo, T. L. (2016). Combining Fuzzy probability and Fuzzy clustering for multispectral satellite imagery classification. *Vietnam Journal of Science and Technology*, *54*(3), 300–313. doi:10.15625/0866-708X/54/3/6463

Ngo, L. T., Mai, D. S., & Pedrycz, W. (2015). Semi-supervising Interval Type-2 Fuzzy C-Means clustering with spatial information for multi-spectral satellite image classification and change detection. *Computers & Geosciences, 83*, 1–16. doi:10.1016/j.cageo.2015.06.011

Ngo, T. L., Mai, D. S., & Nguyen, M. U. (2012). GPU-based Acceleration of Interval Type-2 Fuzzy C-Means Clustering for Satellite Imagery Land-cover Classification. In *12th International Conference on Intelligent Systems Design and Applications (ISDA)* (pp. 992-997). IEEE. doi:10.1109/ISDA.2012.6416674

Peherstorfer, B., Pflüger, D., & Bungartz, H. J. (2012). Clustering Based on Density Estimation with Sparse Grids. In *KI 2012: Advances in Artificial Intelligence*, *LNCS* (Vol. *7526*, pp. 131-142). doi:10.1007/978-3-642-33347-7_12

Rauf, K. S., Valentin, V. G., & Leonid, P. P. (2009). Fuzzy clustering methods in Multi-spectral Satellite Image Segmentation. *International Journal of Computing*, *8*, 87–94.

Torahi, A. A., & Rai, S. C. (2011). Land cover classification and forest change analysis using satellite imagery - A case study in Dehdez Area of Zagros Mountain in Iran. *Journal of Geographic Information System*, *3*(01), 1–11. doi:10.4236/jgis.2011.31001

Wang, Z., & Bovik, A. C. (2002). A universal image quality index. *IEEE Signal Processing Letters*, *9*(3), 81–84. doi:10.1109/97.995823

Wang, Z., & Bovik, A. C. (2009). Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, *26*(1), 98–117. doi:10.1109/MSP.2008.930649

Yang, J., He, Y., & Caspersen, J. (2016). A self-adapted threshold-based region merging method for remote sensing image segmentation. In *2016 IEEE International Geosciences and Remote Sensing Symposium* (pp. 6320-6323). doi:10.1109/IGARSS.2016.7730652

Le Hung Trinh, Dr., Associate Professor, is now the Head of Department of Geodesy and Cartography of Department of Geodesy and Cartography, Le Quy Don Technical University, Hanoi, Vietnam. Le Hung Trinh received his PhD degree in Remote Sensing and Photogrammetry from Moscow State University of Geodesy and Cartography in 2012. His research interests focus on application of remote sensing and GIS techniques in natural resources and environment management. He has authored two books and published more than 60 technical papers in the refereed journal and conference proceedings. He is corresponding author for this paper. Email: trinhlehung125@gmail.com Le Quy Don Technical University Address: 236 Hoang Quoc Viet Street, Bac Tu Liem district, Hanoi, Vietnam.

Dinh Sinh Mai is a lecturer at Department of Geodesy and Cartography, Le Quy Don Technical University (LQDTU), Hanoi, Vietnam. Now, he is a PhD student in Fundamentals of Mathematics for Informatics at Department of Information System, LQDTU. He received his B.S. (2009) and M.S. degrees (2013) in GeoInformatics and Computer Science from LQDTU. His research interests are fuzzy clustering, remote sensing image processing techniques, and pattern recognition and geographic information system (GIS) technologies. Email: maidinhsinh@gmail.com, Le Quy Don Technical University, Address: 236 Hoang Quoc Viet street, Bac Tu Liem district, Hanoi, Vietnam.