WJERT. 2026. Vol. 12 Issue 5. 46-56. Review Article ISSN 2454-695X

World Journal of Engineering Research and Technology

www.wjert.org

Impact Factor: 7.029 Coden USA: WJERA4

AN IMPROVED REMOTE SENSING IMAGE CLASSIFICATION ALGORITHM USING A HYBRID APPROACH

Trung Nguyen Tu*, Huy Ngo Hoang, Chinh Nguyen Van

[1]Department of Information Technology, Thuyloi University, Hanoi, Vietnam.

[2]Intracom University, Hanoi, Vietnam.

[3]Halong University, Quangninh, Vietnam.

Article Received on 19/03/2026 Article Revised on 09/04/2026 Article Published on 01/05/2026

***Corresponding Author**

Trung Nguyen Tu

Department of Information Technology, Thuyloi University, Hanoi, Vietnam.

https://doi.org/10.5281/zenodo.19885800

How to cite this Article: Trung Nguyen Tu*, Huy Ngo Hoang, Chinh Nguyen Van. (2026). An Improved Remote Sensing Image Classification Algorithm Using A Hybrid Approach. World Journal of Engineering Research and Technology, 12(5), 46–56.

ABSTRACT

Remote sensing image classification is a problem of great interest to researchers in the field of remote sensing. There are two main approaches to this problem: pixel-based and object-based approaches. Remote sensing images may contain multiple spectral bands and very high spatial resolution. The pixel-based approach often achieves high accuracy but encounters difficulties when classifying large-scale images such as remote sensing data. Meanwhile, the object-based approach overcomes the problem of large image size but usually provides lower accuracy compared with the pixel-based approach. This paper proposes a hybrid approach that combines both methods to develop a more effective remote sensing image classification algorithm. The proposed algorithm is evaluated on remote sensing images acquired over Hoa Binh province.

KEYWORDS: Remote sensing, image classification, pixel-based approach, object-based approach, hybrid approach.

I. INTRODUCTION

The spatial resolution of multispectral images has been continuously increasing. Recently, remote sensing images have achieved resolutions of several meters. At present, with new satellites reaching spatial resolutions of up to 60 centimeters, the level of detail has increased by a factor of ten. With such high-resolution images, each pixel can be considered as part of a simple object. Consequently, image heterogeneity increases significantly.

Satellite images are mainly used in Geographic Information Systems (GIS). Their classification is highly beneficial for cartographic studies. For low-resolution multispectral images, pixel intensity values are sufficient for classifying individual pixels independently. In contrast, high-resolution image classification is much more challenging. Increased scene complexity introduces various levels of detail. For example, a tree within a field or shadows of objects may be visible, and contextual information becomes essential for accurate classification. Most GIS classification software and medical image analysis software generally apply similar methods to both low- and high-resolution images. While satisfactory results can be obtained for low-resolution images, their effectiveness for high-resolution images still requires further investigation. Therefore, manual classification is sometimes preferred over automated methods to ensure high accuracy.

In remote sensing image classification, two primary approaches exist: the pixel-based approach and the object-based approach. The pixel-based approach[9] often provides high accuracy but faces challenges when classifying large images such as remote sensing data. The object-based approach consists of two main stages.[2] First, objects are defined as regions (clusters) using unsupervised classification algorithms (segmentation or clustering). Clustering is a process used to extract the main characteristics of background objects by defining corresponding regions. Various segmentation methods exist, such as morphological methods, K-means-based methods, Finite Gaussian Mixture Models (FGMM), split-and-merge methods, and Markov models. Some algorithms incorporate contextual information to reduce segmentation heterogeneity.[2] In[4], Chen et al. presented a KMeans clustering algorithm using cluster center displacement. In[5], Balaji et al. proposed a new clustering algorithm based on transforming images from RGB color space to L*a*b space and performing clustering in that space. Second, object classification is performed using supervised classification algorithms. In[6], neural network approaches were applied to classify Landsat images. One of the most widely used classification methods in remote sensing is the Maximum Likelihood classification method[1], which belongs to the pixel-based approach. In[7], Nedeljkovic proposed an image classification algorithm based on fuzzy logic and Maximum Likelihood classification. Although the object-based approach addresses the large image size problem, it generally yields lower accuracy compared to the pixel-based approach.

In this study, we propose a new remote sensing image classification approach. The algorithm derived from this approach is an improved version of the Maximum Likelihood classification method, combining both pixel-based and object-based strategies.

II. PIXEL-BASED AND OBJECT-ORIENTED APPROACHES

A. *Pixel-based approach*

The classical remote sensing image classification approach is pixel-based.[9] In this approach, only spectral information is used for classification.[8] It includes traditional supervised and unsupervised classification methods.[8][9] The Maximum Likelihood classification method belongs to this category. Figure 1 illustrates the pixel-based multispectral image classification process.

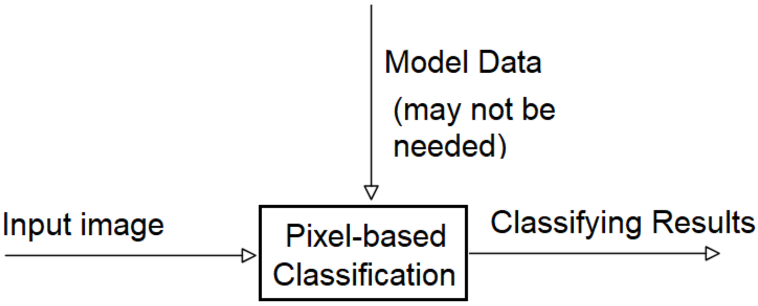

Figure 1: Pixel-oriented classification process for multispectral images.

B. *Object-oriented approach*

In the object-based approach, the processing unit is no longer individual pixels but image objects.[8] First, the image is segmented into meaningful groups of pixels. Second, a set of knowledge-based segmentation rules describing each class is defined. These rules include spectral, spatial, contextual, and textural information.[8] Finally, a classifier assigns each segment to the appropriate class according to these rules.[10] Figure 2 illustrates the object-based multispectral image classification process.

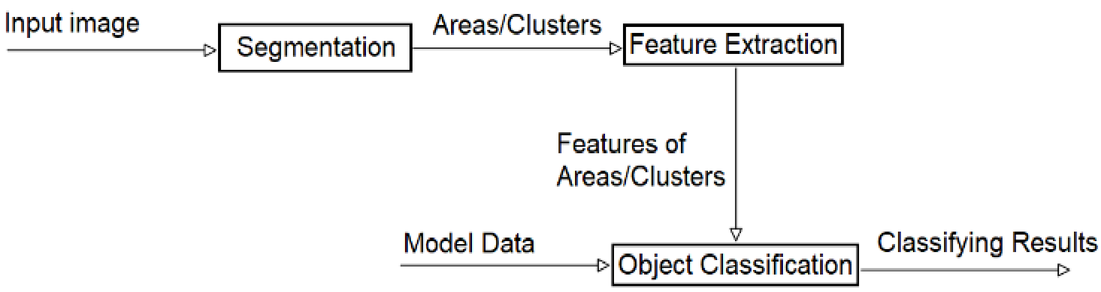

Figure 2. Object-oriented classification process for multispectral images.

III. MAXIMUM LIKELIHOOD CLASSIFICATION

In [1], the Maximum Likelihood classification method is presented in detail.

A. Bayesian Classification

Let the spectral classes of an image be denoted as ωi,i=1,…,M. where M is the total number of classes. To determine the class to which a pixel vector x belongs, the conditional probability $p(\omega ix/)$, i=1,…,M is considered. he measurement vector x is a column of brightness values for the pixel. It describes the pixel as a point in multispectral space with coordinates defined by brightness. The probability $p(\omega ix/)$ represents the likelihood that the correct class is ωi for a pixel at position x. Classification is performed according to:

$$x \in \omega_i, \qquad \text{nếu } p(\omega_i/x) > p\big(\omega_j/x\big) \quad \forall j \neq i \qquad (1)$$

Meaning, the pixel at x belongs to class ωi if $(\omega ix/)$ is the largest.

B. Maximum Likelihood Decision Rule

Although simple $p(\omega ix/)$ is unknown. However, if we assume sufficient training data is available for each land cover type, this can be used to estimate a probability distribution for a cover type describing the chance of finding a pixel from class ωi at position x.

The $p(\omega ix/)$ desired in part 1 and the available $p(\omega ix/)$ – estimated from training data – are related by Bayes' theorem (Freund, 1992):

$$p(\omega_i/x) = p(x/\omega_i)\, p(\omega_i)/p(x) \qquad (2)$$

Where $p(\omega i)$ is the probability that class ωi occurs in the image. $p(x)$ is the probability of finding a pixel in any class at position x, where:

Where (ωi) is the probability that class ωi occurs in the image. If, for example, 15% of the pixels in the image belong to class ωi, then $p(\omega i)$=0.15; $p(x)$ in (2) is the probability of finding a pixel in any class at position x. It is considered because:

$$p(x) = \sum_{i=1}^{M} p(x/\omega_i)\ p(\omega_i), \qquad (3)$$

Although $p(\omega i)$ itself is not important in what follows. $p(\omega i)$ is called the prior probability, since it represents the probability of class membership of a pixel before classification. By comparing posterior probabilities, and using (2), it can be shown that the classification rule in Part 1 becomes:

$$x \in \omega_i, \qquad nếu\ p(x/\omega_i)p(\omega_i) > p(x/\omega_j)p(\omega_j) \quad \forall j \neq i \qquad (4)$$

Where $p(x)$ has been removed as a common factor. Rule (4) is more practical than rule (1) when $p(x\omega i/)$ is known from the training data, and it can be assumed that $p(\omega i)$ is also known or can be estimated from image analysis knowledge. A mathematically equivalent result is obtained if in (4) we define:

$$g_i(x) = \ln\{p(x/\omega_i)p(\omega_i)\} = \ln p(x/\omega_i) + \ln p(\omega_i) \qquad (5)$$

is used, where ln denotes the natural logarithm, and (4) can be rewritten as:

$$x \in \omega_i, \qquad nếu\ g_i(x) > g_j(x) \quad \forall j \neq i \qquad (6)$$

That is, with the above transformation, the decision rule used in Maximum Likelihood classification refers to the discriminant function.

C. Multivariate Normal Class Model

At this stage, it is assumed that the probability distribution of the classes follows a multivariate normal model. This is an assumption rather than a provable property of natural spectral classes. However, it leads to mathematical simplifications. Moreover, it is a well-known multivariate distribution for such attributes.

Therefore, in (4), it is now assumed for N spectral bands that (see Appendix E):

$$p(x/\omega_i) = (2\pi)^{-N/2}|\Sigma_i|^{-1/2} exp\left\{-\frac{1}{2}(x - m_i)^t \Sigma_i^{-1}(x - m_i)\right\} \qquad (7)$$

$$\ln p(x/\omega_i) = -\frac{N}{2ln(2\pi)} - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(x - m_i)^t \Sigma_i^{-1}(x - m_i) \qquad (7.1)$$

Where μi and Σi are the mean vector and covariance matrix of the data in class ωi. The term common to all classes does not contribute to discrimination. Therefore, this factor is omitted and the final form of the discriminant function for Maximum Likelihood classification, based on the normal distribution assumption, is:

$$g_i(x) = \ln p(\omega_i) - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(x - m_i)^t \Sigma_i^{-1}(x - m_i) \qquad (8)$$

In practice, the analyst often has no useful information about prior probabilities. In such cases, equal prior probabilities are assumed; consequently, P(ωi) can be removed from (7) since it is identical for all i. In this case, the factor 1/2 can also be omitted, leading to the discriminant function:

$$g_i(x) = -\ln|\Sigma_i| - (x - m_i)^t \textstyle\sum_i^{-1}(x - m_i) \qquad (9)$$

The implementation of the Maximum Likelihood decision rule involves using either (8) or (9) in (6). However, an additional consideration concerns whether any available label or class is meaningful.

IV. PROPOSED CLASSIFICATION METHOD BASED ON A HYBRID APPROACH

In this study, we propose an image classification approach that combines both of the above approaches, which we refer to as a hybrid approach. First, the original image is segmented. Instead of extracting object features as in the object-oriented approach, the segmented regions are fed into the hybrid classifier to produce the final classification decision. Figure 3 illustrates the classification process based on the hybrid approach.

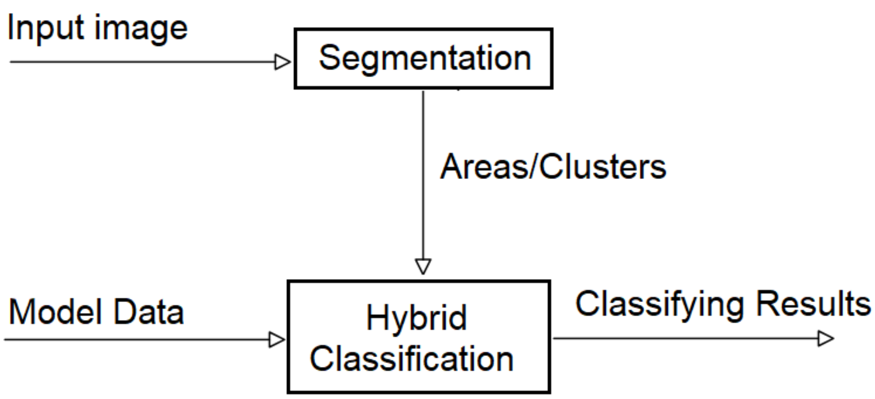

Figure 3: Multispectral image classification process based on the hybrid approach.

A. Unsupervised Segmentation

Segmentation is a process used to extract the main characteristics of background objects by defining corresponding regions. The task of image segmentation is to process the original multispectral image and divide it into different regions or clusters. Currently, there are various segmentation methods such as morphological methods, K-means-based methods, Finite Gaussian Mixture Models (FGMM), split-and-merge techniques, and Markov models. Most methods use only pixel intensity to define regions, which often results in highly heterogeneous segments, especially for high-resolution multispectral images. Some recent algorithms incorporate contextual information to reduce segmentation heterogeneity, where contextual features are extracted directly from the image.

A common issue in speech processing systems is that channel characteristics may vary from one session to another. One method used to minimize the impact of such variations on recognition performance is Cepstral Mean Normalisation (CMN).[12] This method has been widely and effectively applied in digital signal processing and speech recognition. However, when applied in real-time signal processing, the parameters and validity of CMN are often selected and verified experimentally on specific real signals without strict formal mathematical proofs. This paper presents an application of CMN in remote sensing image clustering.

In this study, we implemented the kMeans algorithm[11] and improved it to develop kMeans. The kMeans algorithm is described as follows:

Table 1: The Kmeans Algorithm.

Input: n objects and number of clusters k Output: Clusters C_i (i=1..k) such that the following objective function E is minimized: $E = \sum_{i=1}^{k} \sum_{x \in C_i} d^2(x, m_i)$ (9)
Step 1: Initialization Select k objects C_j (j=1..k) as the initial centers of the k input data clusters (randomly or empirically chosen). ***Step 2: Assign cluster centers based on distance*** For each object x_i ($1 \le i \le n$), compute its distance to each center Cj with j = $1..k$. The object belongs to cluster C_S whose corresponding center has the minimum distance to that object. $d(x, C_S) = \min d(x, C_j), 1 \le j \le k$ (10) ***Step 3: Update cluster centers*** For each j = $1..k$, update cluster center C_j by computing the arithmetic mean of the data object vectors assigned to that cluster. $C_j = \frac{\sum_{x \in cluster(j)} x}{count(cluster(j))}$ (11) ***Step 4: Iterate and check stopping condition*** Repeat Steps 2 and 3 until the cluster centers do not change between two consecutive iterations.

B. The hybrid classification

The discriminant function as presented in Section 3.3 of the Maximum Likelihood classification method applies only to individual pixels. In this section, we propose an improvement to this discriminant function to classify object-clusters obtained from the unsupervised segmentation stage in Section 4.1.

After clustering the original image I, we obtain a set O of object-clusters as follows:

$$O = \{oi: 0 \leq i \leq K\} \quad (12)$$

Where $Oi \cap Oj = \emptyset$ for $i \neq j$; $i,j = 1,2,..,K$; and $o1 \cup o2 \cup \ldots \cup oK = I$.

We construct the discriminant function for each cluster as follows:

$$fi(o) = \Sigma gi(x) x \in o / count(o) \quad (13)$$

From (6), we propose the class decision rule for each cluster as follows:

$$o \in \omega i, \text{ nếu } fi(o) > fj(o)\ \forall j \neq i \quad (14)$$

V. EXPERIMENTS

The dataset used for the experiments consists of two types. First, LANDSAT ETM+ images captured over Hoa Binh province, including district boundary images of Hoa Binh province. Second, SPOT images with high spatial resolution, consisting of four bands: Green, Red, Near-Infrared, and Infrared, captured over Hoa Binh and Son La. Due to the limited scope of the paper, the experiments are presented with two different input image samples.

The experimental results present a comparison between the classification algorithm based on the Maximum Likelihood classification method (the original algorithm), implemented in Grass software, and the algorithm based on the hybrid approach (the improved algorithm).

The first sample is a SPOT image with a size of 2201 × 2101. The input image is classified into three classes: soil (brown color in the result image), water (yellow color in the result image), and forest (blue color in the result image). The classification results are shown in Figure 7. From the results in Figure 7, observing the circled region in the image, the forest area is mixed with soil with very high heterogeneity. However, the original algorithm does not clearly reflect this intermixing and assigns it to a single class, while the improved algorithm clearly distinguishes it.

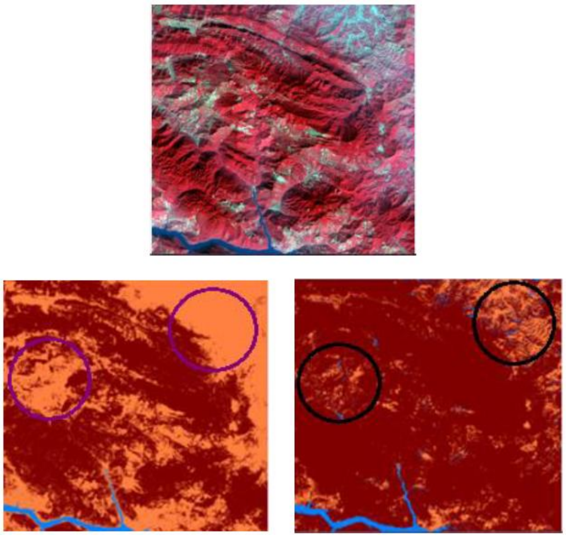

Figure 6: Input image, classification results of the original and improved algorithms.

Table 2: Compare Accuracy.

Class	MLK	Hybrid
Soil	95%	98%
Forest	100%	100%
Water	98%	98%

The second sample is a LANDSAT image with a size of 1596 × 1333. The input image is classified into three classes: Soil (black color in the result image), water (blue color in the result image), and forest (yellow color in the result image). The results are shown in Figure 7. The input image also has high heterogeneity, which is more clearly shown in the results of the proposed classification algorithm.

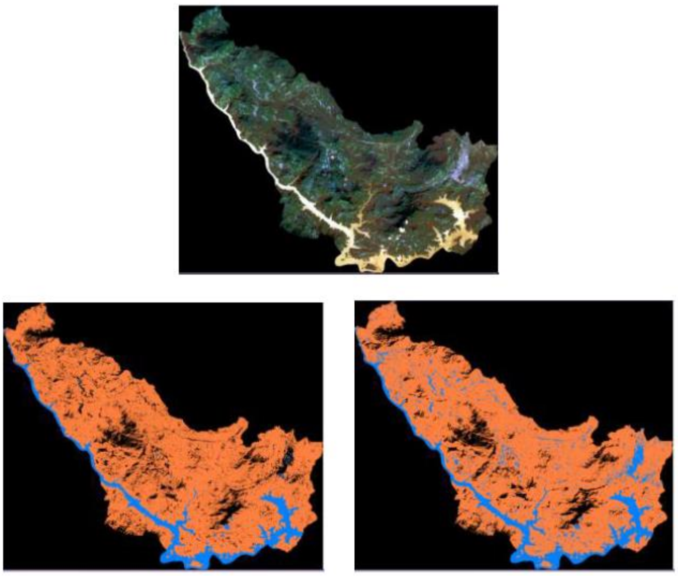

Figure 7: Input image, classification results of the original and improved algorithms.

Table 3: Compare Accuracy.

Class	MLK	Lai
Soil	85%	84%
Forest	90%	93%
Water	100%	100%

VI. CONCLUSION

In this study, we proposed a hybrid classification approach combining both pixel-based and object-oriented approaches. The classification process of this approach consists of two stages. First, image segmentation; in the improved algorithm of this approach, we use the KMeans algorithm. Second, hybrid classification; the regions are then classified using a hybrid classifier with a new class decision function that we propose to classify the object-clusters. This class decision function is improved based on the discriminant function of the Maximum Likelihood classification method. Experimental results show that the classification results of the proposed approach achieve higher accuracy than those of the Maximum Likelihood method.

REFERENCE

1. John A. Richards, Xiuping Jia, *Remote Sensing Digital Image Analysis*, Springer-Verlag Berlin Heidelberg, 2006.
2. Meritxell Bach Cuadra, Jean-Philippe Thiran, *Satellite Image Segmentation and Classification*, Fall., 2004.
3. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol.1. University of California Press, 1967; 281-297.
4. Chih-Tang Chang, JIMZ. C. LAI, MU-DERJENG, “A Fuzzy K-means Clustering Algorithm Using Cluster Center Displacement”, JOURNAL OF INFORMATION SCIENCE AND ENGINEERING, 2011; 27: 995-1009.
5. Balaji T., Sumathi M., “Relational Features of Remote Sensing Image classification using Effective K- Means Clustering”, International Journal of Advancements in Research & Technology, August-2013; 2(8): 103-107.
6. Smriti Sehgal, “Remotely sensed Landsat Image Classification using Neural network approaches”, International Journal of Engineering Research and Applications, 2012; 2(5): 043-046.
7. Nedeljkovic, “Image Classification based on fuzzy logic”, The International Archives of

the Photogrammetry, Remote sensing and Spatial Information Sciences, 2003; 34: XXX.

8. Sun Xiaoxia, Zhang Jixian, Liu Zhengjun, "A comparison of object-oriented and pixel-based classification approachs using quickbird imagery", Chinese Academy of Surveying and Mapping, Beijing, China.
9. H.R. Matinfar, F. Sarmadian, S.K. Alavi Panah, R.J. Heck, "Comparison of object-oriented and pixel-based classification on Lansadsat7, Etm^+ Spectral Bands (Case Study: Arid Region of Iran)", American-Eurasian J. Agric. & Environ, 2007.
10. Leukert k., "Transferability of knowledge-based classification rules", ISPRS2004, Istanbul, 2004.
11. http://www.onmyphd.com/?p=k-means.clustering.
12. http://recognize-speech.com/preprocessing/cepstral-mean-normalization
13. Intan aidha yusoff, Nor ashidi mat isa, *Two-Dimensional Clustering Algorithms for Image Segmentation*, WSEAS Transactions on Computers, October 2011; 10(10).
14. J. Liu, and Y. H. Yang, *Multiresolution color image segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Jul 1994; 16(7): 689-700.
15. R. H. Haralick, and L. G. Shapiro, *Image segmentations techniques*, Computer Vision Graphics Image Processing, 1985; 29: 100-132.