International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. #, No.# (Mon-20..)

Synergistic Exploration Combining Traditional And Evolutionary Methods To Improve Supervised Satellite Images Classification

Ismahane Kariche[1] and Hadria Fizazi[1]

[1] *Department of Computer Science, University of Sciences and Technology of Oran-Mohamed Boudiaf (USTO-MB), Oran, Algeria*

ismahane.kariche@univ-usto.dz, hadria.fizazi@univ-usto.dz

Received ## Mon. 20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##

Abstract: This paper aims to enhance the performance of supervised classification of satellite images by adopting a spectral classification approach, which often encounters the issue of class confusions due to its reliance solely on spectral information. The proposed approach, EAMD (Evolutionary Algorithm and Minimum Distance), integrates a Genetic Algorithm-based evolutionary method with the Minimum Distance method. During the training phase, the Genetic Algorithm generates an optimal set of subcategories to represent different object classes present in the image and identifies an optimal representative set of pixels for class assignment. Experimental tests conducted on various satellite images yield promising results, demonstrating the capability of Genetic Algorithms to enhance classification accuracy and effectively identify and exclude misleading pixels responsible for class confusions. This aspect is crucial, as the effectiveness of supervised classification heavily depends on the quality of the training samples. Validation of the approach was further reinforced by intentionally injecting erroneous data into the training data. Compared to the Minimum Distance method, the proposed approach successfully detects and avoids the erroneous pixels, a task unaccomplished by the Minimum Distance method. The obtained results demonstrate that the hybrid proposed approach offers significant potential for improving the accuracy and reliability of satellite image classification techniques.

Keywords: Genetic Algorithm, Minimum Distance, Satellite Images, Supervised Classification

1. INTRODUCTION

Remote sensing is a field of science wherein an information about an object, area, or phenomenon is acquired using remote sensors like those from aircrafts or satellites [1]. Every day a large number of Earth observation spaceborne and airborne sensors from many different countries provide a massive amount of remotely sensed data [2]. Remote sensing data play an indispensable role in crop growth monitoring, the change detection of land use/cover, and disaster monitoring [3]. Satellite images are among the most commonly used data in remote sensing for the study of earth surface, they can be optical or radar, panchromatic, multispectral or hyperspectral, with varying levels of spatial resolution, ranging from low to very high. In order to extract information from these data, various processing techniques are utilized. Pre-processing is used to improve the image data by removing unwanted distortions [4]. Pre-processing functions involves the operations required prior to the main data analysis and consists of processes aimed at geometric correction, radiometric correction and atmospheric corrections to improve the ability to interpret the image components qualitatively and quantitatively [5]. To make image easier for visual interpretation Enhancements are used [5]. Transformation operations and fusion methods are utilized to effectively exploit information derived from multiple images.

Classification is among the processing methods used to extract useful information from satellite images. It plays a major role in extract and interpretation of valuable information from massive satellite images [6]. Classification is applied to various areas like urban planning, disaster management, vegetation monitoring, and forest cover monitoring [7]. The classification is usually defined as a procedure of integrating the pixels with the significant classes [8]. When applying a classification algorithm to a satellite image, the data obtained by the satellite sensors as digital levels are changed into a categorical scale that is easily interpreted

ismahane.kariche@univ-usto.dz

by analyst experts. The resulting classified image is a thematic map of the original satellite image, and pixels belonging to the same class share similar spectral characteristics [9]. There are many different approaches to classification [6], [10], [11], however, in common there are two broad categories: unsupervised and supervised classification techniques [12]. Unsupervised classification, which is also known as clustering, is used to partition the satellite image into homogeneous clusters (classes) [13]. The user identifies the number of clusters and bands to be generated [5], later analyst assigns meaningful labels to the clusters and produces well classified satellite image [6]. Among the most common unsupervised methods, let's mention the ISODATA and the K-means clustering methods. Supervised classification methods require input from an analyst, known as training set [6]. In contrast to the unsupervised approach, these methods require prior knowledge (ground truth) about the study area, which is used to train the system. Once trained, the classifier is then used to attach labels to all the image pixels according to the trained parameters. The quality of a supervised classification depends on the quality of the training set [5]. Parallelepiped, Maximum Likelihood and Minimum Distance [13], [14], [15] and Mahalanobis Distance are the common supervised classifiers [13]. Supervised or unsupervised classification methods may be either purely spectral, based solely on pixel radiometry, or contextual, integrating additional information such as pixel neighborhood (texture information). Classifying satellite images presents challenges due to the large amount of data and spectral similarities, particularly in spectral classification, which is susceptible to confusion errors since it based solely on spectral information. The aim of this study is to enhance spectral classification results by combining a conventional method (Minimum Distance: MD) with an evolutionary approach based on a Genetic Algorithm (GA).

GAs are one of the most well-known natural-inspired algorithms [16]. GAs have been successfully used in various fields, including optimization, pattern recognition, and image processing [16], [17], they have demonstrated their effectiveness in exploring large and complex solution spaces, enabling the discovery of solutions that may be challenging to attain using traditional methods. GA are stochastic search techniques based on the mechanism of natural selection and natural genetics [18]. They abstract the problem space as a population of individuals, and try to explore the fittest individual by producing generations iteratively. GA evolves a population of initial individuals to a population of high-quality individuals, where each individual represents a solution of the problem to be solved [19]. GA is one of the most well-regarded evolutionary algorithms in the history [20]. In their review of metaheuristic algorithms, the authors of [21] noted that GAs are among the most frequently mentioned metaheuristics in recent years. This observation was made in a summary of top 10 metaheuristics algorithms with the highest number of citations according to Google Scholar (as of December 31, 2022). The most widely used algorithm is Particle Swarm Optimization (PSO), which has more than 75000 citations on its own. GA is ranked as the second most popular algorithm with more than 70000 citations [21].

In the field of satellite images classification authors referenced in [22] and [23] utilized GAs for unsupervised classification of satellite images. Additionally, the authors of [24] proposed an approach that combines a GA with SVM algorithm for classifying multisource data, which includes optical and radar images. In their study, the authors of [25] implemented a GA withing a supervised approach based on classifier systems. Their method searches for the best classifier for each class, independently [25], it is from this approach that the idea arises in our methodology to optimally create the set of subclasses for different categories, while considering the interdependence between the classes. This interdependence is of paramount importance for accurately modeling the relationships within our classification system and, consequently, improving machine learning results.

This paper is organized as follows: it starts by outlining the classical MD classification method, followed by a brief introduction to GAs. It then details the adopted methodology, including practical implementation on different satellite images. The in-depth discussion and conclusion synthesize the results, providing a comprehensive perspective on the relevance of the proposed hybrid approach.

2. SUPERVISED SPECTRAL CLASSIFICATION

Figure 1 illustrates the supervised classification process based solely on spectral data.

3. MINIMUM DISTANCE CLASSIFICATION METHOD (MD)

Based on the concept of radiometric distance and class mean vectors. During the training phase, mean vectors for each class are calculated from the training data.

Considering a class C*i*. The mean of C*i* in image *j* is calculated as follows:

$$M_{ij} = \frac{1}{nbp_i} \sum_{k=1}^{k=nbp_i} GL_{kj} \quad (1)$$

nbp_i : the pixel count of *Ci*, GL_{kj} : The grayscale level of the k^{th} pixel from *Ci* in the j^{th} image.

Classifying pixels involves calculating the Euclidean distance of each pixel from class centroids and assigning it to the closest class. A pixel $P:(GL_1,GL_2,...,GL_l)$ is assigned to *Ci* if and only if:

$$D(P, M_i) < D(P, M_k), \forall\, k \neq i, k = 1, \dots, nbc \quad (2)$$

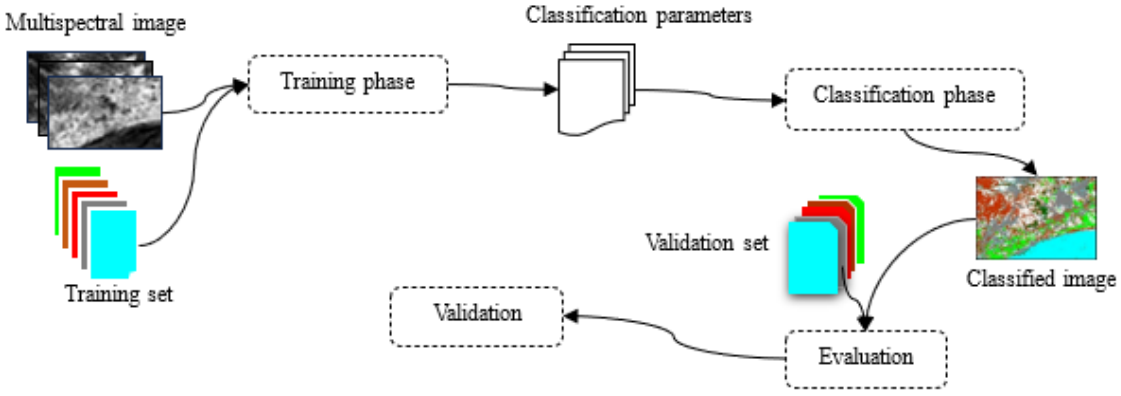

Figure 1. Supervised classification process.

$$D(P, M_i) = \sum_{j=1}^{j=l} (GL_j - M_{ij})^2 \quad (3)$$

GL_j: The grayscale level of P in the image j.

nbc : the number of classes.

4. GENETIC ALGORITHM (GA)

GA is a part of Evolutionary Computing, which is a rapidly growing area of Artificial Intelligence [26].

GAs are widely used in engineering, scientific as well as business applications. They are successfully applied to the problems which are difficult to solve using conventional techniques such as machine learning and optimization [27]. GA starts with an initial set of solutions (initial population) known as chromosomes. The new populations are produced by iterative use of genetic operators on individuals present in the population [28]. These operators enable both the exploration and exploitation of the solution space. Algorithm 1 illustrates a pseudocode of a classical GA.

A. *Genetic operators*

1) *Selection*

Before moving on to the selection process, individuals from the current generation must be evaluated by the fitness function. Fitness function is the criterion for judging the performance of chromosomes and the only criterion for optimizing algorithms [29]. The selection operation involves choosing elite individuals from the current population to serve as parents, resulting in the generation of offspring. Fitness values are used as criteria to judge whether individuals are elitist [19]. Different technical approaches can be employed for the selection process such as Elitism Selection [30], Rank Selection [19], [31] and Roulette Wheel Selection (RWS).

RWS assigns to each individual i of the current generation a probability $p(i)$, calculated as follows:

$$p(i) = \frac{f(i)}{\sum_{j=1}^{TPOP} f(j)} \quad (4)$$

A circle is then drawn and divided into n sectors (n is the number of individuals in the population), where each individual i occupies a portion proportional to its probability $p(i)$, this guarantees that the sectors of greater size are occupied by the higher-quality chromosomes. Then, spin the wheel n number of times. When the roulette stops, the sector on which the pointer point corresponds to the individual being selected [31]. It can be seen that individuals with high selection probability are more likely to be selected [32].

2) *Crossover*

Crossover is the main genetic operator. It operates on two chromosomes at a time and generate offspring by combining both chromosomes' feature [18].

Figure 2 illustrates various crossing techniques which are employed to perform genetic recombination between two individuals.

3) *Mutation*

Mutation is often performed after crossover [19].

Algorithm 1 Pseudo code of a classical GA

```
BEGIN
Iteration←1
Generate initial population
Evaluation (initial population)
While not (stopping criterion) do
   Select the best individuals for reproduction
   New population=crossover (population)
   New population=mutation (New population)
   Evaluation (New population)
   population=New population
   Iteration=Iteration+1
End while
END
```

This genetic operator randomly alters a portion of the population to explore the search space, the size of this portion is determined by a significantly lower mutation probability compared to the crossover probability. Figure 3 shows a simple mutation example applied to a randomly chosen individual from the new population.

5. METHODOLOGY

The implemented classification process: EAMD, utilizes a GA combined with the MD classification method to determine, during the learning phase from the training samples, the most effective pixel sets representing the classes. The hybridization with the MD method is also applied during the classification step, where the assignment of pixels to classes is based on the obtained sets. Each class is divided into several subclasses, the GA will determine, for each class, the best subset of representative subclasses. Subclasses are defined by grayscale level intervals in the different images. The GA is used to filter the training sample for each class, preserving only the most representative pixels (the elite), which serves as the basis for classifying pixels. The primary objective of the elite is to minimize class confusions and thereby maximize classification accuracy.

A. *Individual encoding*

As illustrated in Table 1, Individuals are defined by k disjunctions of grayscale level intervals (subclasses) in each image j and for each class C_i. The number k is among the parameters to initialize in the implemented GA.

B. *Implemented crossover*

The GA operates in a single population for all classes, in that case, the initial step involves selecting the crossover class number, this can be done in two ways:

- Once the two parents are chosen, the GA selects a crossover class number, then the spectral band and the crossover location.

  In this scenario the genetic recombination is restricted in a single class and the offspring preserve the values from the other classes.

- Once the two parents are selected, the crossover operation is applied to all classes. The band number and crossover location are randomly chosen, so they may differ for each class within the same individual.

In the first method, the values of most classes remain relatively stable across generations, which eventually leads to the selection of the second approach due to its faster convergence. The parent selection is carried out by a RWS method.

C. *Implemented mutation*

The mutation occurs as follows:

- select an individual
- randomly generate class, spectral band and subclass number.
- replace the selected interval with a newly generated one, ensuring that its boundaries remain within the extreme values, calculated during the creation of the initial population step.

D. *Evaluation process and proposed fitness*

Two pieces of information are computed for each individual in the evaluation step (Algorithm 2):

- Individual's fitness: Provides an assessment of the potential classification accuracy when based on the currently evaluated individual.
- The elite: Optimal subset of pixels from the training dataset, capable of attaining the fitness of the individual under evaluation (finest class representatives).

A pixel P is represented as a vector (GL_1, GL_2, GL_3), where GL_j denotes the gray level intensity of P in the image j.

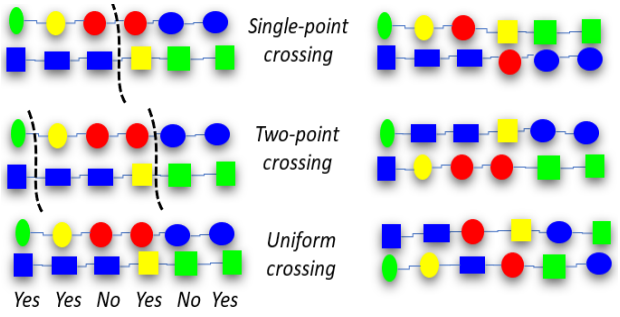

Figure 2. Crossover

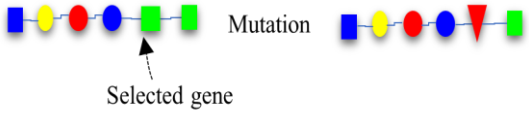

Figure 3. Simple mutation

TABLE I. INDIVIDUAL ENCODING

	Image *j*
C_1	$Int_{j1}^{1} \cup Int_{j2}^{1} \cup \ldots \cup Int_{jk}^{1}$
C_2	$Int_{j1}^{2} \cup Int_{j2}^{2} \cup \ldots \cup Int_{jk}^{2}$
⋮	⋮
C_n	$Int_{j1}^{n} \cup Int_{j2}^{n} \cup \ldots Int_{jk}^{n}$

Algorithm 2 Evaluation Process

```
BEGIN
for count class=1 to number of classes do
    for pixel=1 to number of pixels in the count class do
        if pixel can be classified by the GA in the count class then
            well classified (count class) = well classified (count class) +1
            add pixel to the vector elite class (count class)
        end if
    end for
    calculate centers (count class) the vector of class centroids in the three images
end for
for count class=1 to number of classes do
    for pixel=1 to number of pixels in the count class do
        if pixel is not member in elite (count class) then
            for c=1 to number of classes do
                calculate distances (c) as the Euclidian distance between pixel and centers (c)
            end for
            if index of min(distances(c)) ==count class then
                add pixel to the vector elite class (count class)
                well classified (count class) = well classified (count class) +1
            else
                misclassified (index, count class) =misclassified (index, count class) +1
            end if
        end if
    end for
    fitness (count class) = well classified (count class) /size (training sample (count class))
end for
for count class=1 to number of classes do
    calculate S as the sum of the elements in the count class row of the misclassified matrix
    fitness (count class) = fitness (count class)- (S / (S+ well classified (count class)))
end for
fitness (individual)=sum of classes' fitness / number of classes
elite (individual) = elite class
END
```

can be classified is a function that checks whether a pixel *P*, member of the training sample pixels for a class *C*, will be classified by the program within its own class or assigned to different one.

A pixel *P* will be assigned to *C* if and only if its gray level value in each spectral band matches at least one subclass of membership (GL_j belongs to at least one of the intervals ($Int_{j1}^{C}, Int_{j2}^{1}, \ldots, Int_{jk}^{C}$).

If the predicted class for *P* matches its true class, *P* is categorized as *well classified* and added to the class-specific elite group. For the remaining non-elite pixels, being part of the class's training dataset, an alternative test is employed, offering them a second chance to join the elite. To become an elite member of *C* and be counted among the *well classified* pixels, a non-elite pixel *P* must exhibit the minimum Euclidean distance to the centroid of the *C*'s elite (referred to as '*centers*' in Algorithm 2), in comparison to the centroids of elites from the other classes. Otherwise, the pixel *P* will retain its *non-elite status* and be counted as *misclassified*, which is essentially an error of confusion between *C* and the nearest elite class to *P*.

At the output, the evaluation process provides for each individual its *fitness* and an *elites set*. The fitness *F* of an individual *i* is calculated as follows:

$$F(i) = \frac{\sum_{j=1}^{j=nbc} f(C_j)}{nbc} \quad (5)$$

Where:

nbc: the total number of classes and $f(C_j)$: the fitness value of C_j, calculated as follow:

$$f(C_j) = T_1 - T_2 \quad (6)$$

The fitness of an individual is determined by its classification accuracy when considering its set of elites. The individual's accuracy represents the proportion of correctly classified instances it achieves, but this score is

negatively affected by the error rate resulting from misclassified pixels.

Therefore, in order to compute $f(C_j)$, two rates are calculated (1): the rate of correctly classified instances T_1, and the error rate T_2, determined as:

$$T_1 = \frac{well\_classified(C_j)}{nbp(C_j)} \quad (7)$$

$$T_2 = \frac{S_j}{S_j + well\_classified\ (C_j)} \quad (8)$$

Where:

$$S_j = \sum_{i=1}^{i=nbc} misclassified\ (C_i, C_j), i \neq j \quad (9)$$

$nbp(C_j)$: The pixel count in class C_j

E. Elitism strategy

The persistence of certain population elements in the next generation enables the GA to keep a stable fitness value when the best chromosome in the current generation is not of better quality than the one in the previous generation. It also promotes fitness improvement when the characteristics of individuals in the current generation are enhanced. This mechanism helps retain the best chromosomes from each generation. In this study, 10% of the best individuals are reproduced in each generation.

F. Pixels assignment

From the learning phase, we obtain the best individual from the last generation, representing the optimal subclass sets for each class in every spectral band, as well as the elite collections for each class. Based on these elements, the program classifies pixels using two methods:

- The first method utilizes the *can be classified* function from Algorithm 2 to determine the class of each pixel *P* in the image. This test considers the three grayscale intensities of the pixel, if each intensity identifies a subclass within every spectral band, *P* is assigned to that class (assign it the color of that class). In case where no class is found for *P* or when multiple classes are identified for *P*, the program calculate the Euclidean distance between *P* and the centroids of elite sets (generated by the best individual at the end of the learning phase) for each class. *P* is then assigned to the nearest class.
- In the second method, the algorithm uses only the elite set (without undergoing the test *can be classified).* Each pixel *P* in the image is assigned to the nearest class, by calculating its distance from the different elite set centers.

G. Classification quality assessment

The confusion matrix illustrated in Table 2 is used in this study to evaluate classification accuracy.

Where:

C_i: Class *i* and *nbc* the number of classes.

C_{ii} : Number of correctly classified pixels in *Ci*.

CO_i: Commission errors (total of pixels migrated to *Ci*):

$$CO_i = 100 \times \frac{1}{TL_i} \sum_{j=1}^{nbc} C_{ij}\ , j \neq i \quad (10)$$

OM_j: Omission errors (total of pixels migrated from *Ci* to other classes):

$$OM_j = 100 \times \frac{1}{TC_j} \sum_{i=1}^{nbc} C_{ij}\ , i \neq j \quad (11)$$

GA_j: Ground Accuracy of *Cj*:

$$GA_j = 100 \times C_{jj}/TC_j \quad (12)$$

IA_i : Image Accuracy of *Ci*:

$$IA_i = 100 \times C_{ii}/TL_i \quad (13)$$

$$The\ Precision\ P = 100 \times \frac{1}{T} \sum_{i=1}^{nbc} C_{ii} \quad (14)$$

Where: TL_i and TC_j respectively denote the total sum of elements in row *i* and column *j*, whith *i* and *j* range from *1* to *nbc*. *T* is the pixel count in the sample set.

6. DATA SET AND AREA OF STUDY

To validate our approach, experimental tests were conducted on two real satellite images:

- An image of 428 × 484 pixels, extracted from a 20m resolution SPOT satellite image, incorporating the spectral bands XS1, XS2 and XS3, captured in the Western region of Oran, Algeria. The color composition XS3, XS2, XS1 is illustrated in Figure 4 (a). The image belongs to the Center of Space Techniques (CTS), Algeria.

TABLE II. CONFUSION MATRIX

	C1	**…**	**Cnbc**	**CO**	**IA**
C1	C_{11}	…	C_{1nbc}	CO_1	IA_1
⋮	⋮	⋮	⋮	⋮	⋮
Cnbc	C_{nbc1}	…	$C_{nbc\ nbc}$	CO_{nbc}	IA_{nbc}
TC	TC_1	…	TC_{nbc}	**Precision P**	
OM	OM_1	…	OM_{nbc}		
GA	GA_1	…	GA_{nbc}		

- An image of 800 × 600 pixels, extracted from an ASTER scene with a 15m resolution, covering a region in Mostaganem, located in the West of Algeria. The color composition of the multispectral VNIR (Visible and Near Infrared) image is illustrated in Figure 4 (b). The image was downloaded from MADAS (METI AIST Data Archive System).

7. TESTING AND RESULTS

A. *MD classification results (SPOT image)*

Figure 5, Table 3 and Table 4 present the classification results obtained through the MD method.

B. *EAMD results (SPOT image)*

The size of an individual corresponds to the number of subclasses (intervals) within each class in each image. The graph in Figure 6 illustrates how classification quality evolves with individual size. After multiple trials, we opted for a 1000-individual population. The termination criteria depend on reaching a maximum fitness value of 1 (100%), completing a predefined number of generations or achieving process stability over several generations. It's worth noting that the graph was generated after multiple runs, as the GA's convergence strongly depends on the initial population.

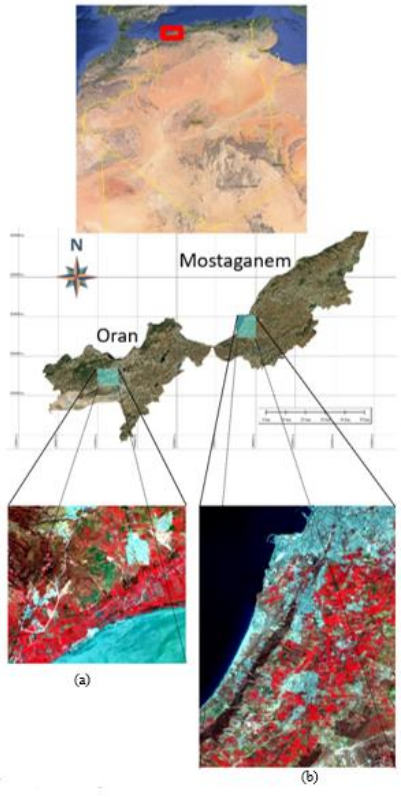

Figure 4. Area of study

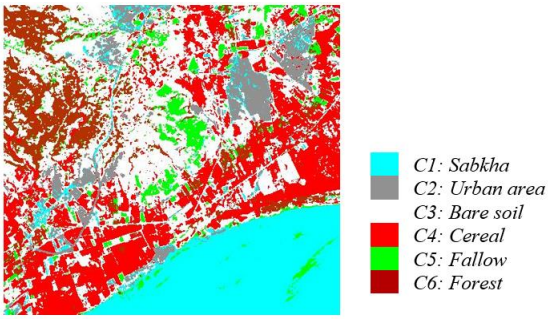

Figure 5. Classified image using MD method, Spot Image

TABLE III. CONFUSION MATRIX OF THE TRAINING SAMPLES USING MD METHOD, SPOT IMAGE. C1: SEBKHA, C2: URBAN AREA, C3: BARE SOIL, C4: CEREAL, C5: FALLOW, C6: FOREST

	C1	C2	C3	C4	C5	C6	CO	IA
C1	162	15	0	0	0	0	8.47	91.53
C2	0	134	0	0	0	0	0	100
C3	0	1	80	0	38	0	32.77	67.23
C4	0	0	0	177	0	0	0	100
C5	4	0	20	0	95	0	20.17	79.83
C6	0	0	0	0	16	139	10.32	89.68
TC	166	150	100	177	149	139	**Precision 89,33%**	
OM	2.4	10.67	20	0	36.24	0		
GA	97.59	89.33	80	100	63.76	100		

TABLE IV. CONFUSION MATRIX OF THE VALIDATION SAMPLES USING MD METHOD, SPOT IMAGE. C1: SEBKHA, C2: URBAN AREA, C3: BARE SOIL, C4: CEREAL, C5: FALLOW, C6: FOREST

	C1	C2	C3	C4	C5	C6	CO	IA
C1	826	31	1	0	2	0	3.95	96.05
C2	0	444	0	0	0	0	0	100
C3	0	0	208	0	141	2	40.74	59.26
C4	0	0	35	357	33	0	16	84
C5	175	0	20	0	255	0	43.33	56.67
C6	0	0	0	0	1	266	0.37	99.63
TC	1001	475	264	357	432	268	**Precision 84.23%**	
OM	17.48	6.53	21.21	0	40.97	0.75		
GA	82.52	93.47	78.79	100	59.03	99.25		

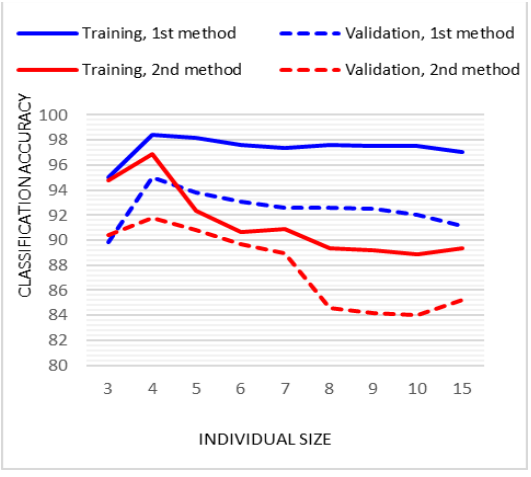

Figure 6. Evolution of classification accuracy according to the individual's size, Spot image

The graph also demonstrates the promising performance of the proposed approach, notably improving overall classification rates for different sizes. Results of the first pixel assignment method outperform the second. The best performance is achieved with 4 intervals for individuals. Table 5, Table 6 and Figure 7 show the best results obtained using the first pixel assignment method, with an individual size equivalent to 4 sub-classes.

The analysis of the results shows a significant improvement in overall classification rates, with a 9.08% increase for the training sample and a 10.8% increase for the validation set. Notably, the fallow class saw the most substantial enhancement, with the rate increasing from 63.76% to 99.33% for the training samples and from 59.03% to 84.26% for the validation set. This represents a 37.57% improvement for the training samples and a 25.23% improvement for the validation set. These results indicate that the hybrid approach has effectively reduced confusion errors compared to the traditional MD classification method. This suggests that the GA did not select the pixels responsible for these errors to be part of the elite members within each class, demonstrating the GA's ability to filter training samples by selecting the most representative pixels while generating optimal class subsets within the three spectral bands.

To reinforce this conclusion, the approach was tested on another image with a higher thematic complexity, the Aster image, as illustrated in Figure 4 (b), which encompasses 10 different object classes.

C. Aster Image classification

Table 7 and Table 8 depict the classification results obtained solely through the MD method. The classified Aster image is illustrated in Figure 8 (a).

Classification errors are particularly significant for the *Sand* class, which exhibited the lowest classification rate (32.35% in the training set and 44.19% for validation pixels). Omissions are distributed between the Urban and Bare Soil classes. Additionally, the Bare Soil class displayed low precision rates, falling below 50% for the entire validation set.

TABLE V. CONFUSION MATRIX OF THE TRAINING SAMPLES USING EAMD METHOD, SPOT IMAGE. C1: SEBKHA, C2: URBAN AREA, C3: BARE SOIL, C4: CEREAL, C5: FALLOW, C6: FOREST

	C1	C2	C3	C4	C5	C6	CO	IA
C1	166	0	0	0	0	0	0	100
C2	0	150	0	0	0	0	0	100
C3	0	0	87	0	1	0	1.14	98.86
C4	0	0	0	177	0	0	0	100
C5	0	0	13	0	148	0	8.07	91.92
C6	0	0	0	0	0	139	0	100
TC	166	150	100	177	149	139	**Precision 98,41%**	
OM	0	0	13	0	0.67	0		
GA	100	100	87	100	99.33	100		

TABLE VI. CONFUSION MATRIX OF THE VALIDATION SAMPLES USING EAMD METHOD, SPOT IMAGE. C1: SEBKHA, C2: URBAN AREA, C3: BARE SOIL, C4: CEREAL, C5: FALLOW, C6: FOREST

	C1	C2	C3	C4	C5	C6	CO	IA
C1	983	10	0	0	14	0	2.38	97.62
C2	16	465	6	0	0	0	4.52	95.48
C3	0	0	221	0	21	0	8.68	91.32
C4	0	0	37	357	33	0	16.39	83.61
C5	2	0	0	0	364	0	0.55	99.45
C6	0	0	0	0	0	268	0	100
TC	1001	475	264	357	432	268	**Precision 95,03%**	
OM	1.80	2.11	16.29	0	15.74	0		
GA	98.20	97.89	83.71	100	84.26	100		

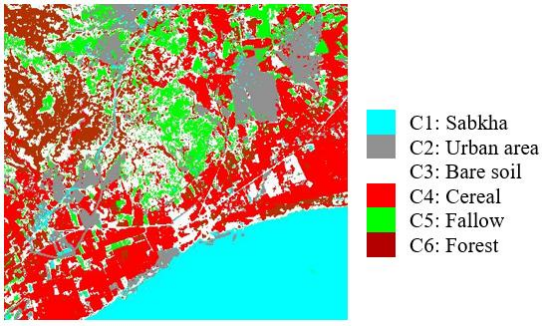

Figure 7. Classified image using EAMD, Spot image

TABLE VII. CONFUSION MATRIX OF THE TRAINING SAMPLES USING MD METHOD, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	335	0	0	26	0	0	0	4	0	0	8.22	91.78
C2	0	286	0	0	0	0	0	0	0	0	0.00	100
C3	16	0	221	0	0	0	0	0	0	0	6.75	93.25
C4	12	0	0	184	0	0	0	12	1	5	14.02	85.98
C5	0	0	0	0	318	3	0	0	0	0	0.93	99.07
C6	0	0	0	0	0	154	0	0	0	0	0	100
C7	0	0	0	5	0	0	33	1	34	87	79.38	20.63
C8	0	0	0	2	0	0	0	318	6	0	2.45	97.55
C9	0	0	0	0	0	0	32	0	131	4	21.56	78.44
C10	0	0	0	32	0	0	37	0	27	184	34.29	65.71
TC	363	286	221	249	318	157	102	335	199	280	**Precision 86.22%**	
OM	7.71	0	0	26.1	0	1.91	67.65	5.07	34.17	34.29		
GA	92.29	100	100	73.90	100	98.09	32.35	94.93	65.83	65.71		

The results obtained with EAM method, while adopting the first assignment method, are presented in the Table 9, Table 10 and Figure 8 (b). These results demonstrate a significant enhancement in overall classification rates through the hybrid approach, with improvements of 7.25% for the training set and 6.47% for the validation set. At the category level, notable improvements are observed, particularly for the categories Bare Soil, Sand, and Urban, showing respective increases of 17.59%, 35.3%, and 21.08% in the training sample, and 26.41%, 21.7%, and 12.27% in the validation set.

This, consequently, reinforces the conclusion that the proposed hybrid approach helps clean the training dataset of pixels that could disrupt the classification.

Next, we will put this approach to the test by applying it to incorrect samples. We intentionally introduced errors into the training dataset, this will allow us to assess the extent to which the approach can identify and avoid them. The validation set will be kept unchanged to ensure a stable basis for comparing the different results.

TABLE VIII. CONFUSION MATRIX OF THE VALIDATION SAMPLES USING MD METHOD, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	429	10	4	49	0	0	0	27	0	0	17.34	82.66
C2	9	537	0	0	0	0	0	0	0	0	1.65	98.35
C3	39	0	381	2	0	0	0	0	0	0	9.72	90.28
C4	29	0	0	360	0	0	0	82	0	67	33.09	66.91
C5	0	0	0	0	513	25	0	0	0	0	4.65	95.35
C6	0	0	0	0	0	338	0	0	0	0	0	100
C7	0	0	0	1	0	0	57	3	92	59	73.11	26.89
C8	0	0	0	3	0	0	0	551	31	3	6.29	93.71
C9	0	0	0	0	0	0	36	6	194	13	22.09	77.91
C10	0	0	0	19	0	0	36	4	92	412	26.82	73.18
TC	506	547	385	434	513	363	129	673	409	554	**Precision 83.58%**	
OM	15.22	1.83	1.04	17.05	0	6.89	55.81	18.13	52.57	25.63		
GA	84.78	98.17	98.96	82.95	100	93.11	44.19	81.87	47.43	74.37		

TABLE IX. CONFUSION MATRIX OF THE TRAINING SAMPLES USING EAMD METHOD, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	353	0	0	18	0	0	0	2	0	0	5.36	94.63
C2	0	286	0	0	0	0	0	0	0	0	0	100
C3	4	0	221	0	0	0	0	0	0	0	1.78	98.22
C4	6	0	0	207	0	0	0	4	0	1	5.05	94.95
C5	0	0	0	0	318	3	0	0	0	0	0.93	99.07
C6	0	0	0	0	0	154	0	0	0	0	0	100
C7	0	0	0	0	0	0	69	0	4	34	35.51	64.49
C8	0	0	0	2	0	0	0	329	4	1	2.08	97.92
C9	0	0	0	3	0	0	5	0	166	1	5.14	94.86
C10	0	0	0	19	0	0	28	0	25	243	22.86	77.14
TC	363	286	221	249	318	157	102	335	199	280	**Precision 93.47%**	
OM	2.75	0	0	16.87	0	1.91	32.35	1.79	16.58	13.21		
GA	97.25	100	100	83.13	100	98.09	67.65	98.21	83.42	86.79		

TABLE X. CONFUSION MATRIX OF THE VALIDATION SAMPLES USING EAMD METHOD, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	444	4	4	32	0	0	0	20	0	0	11.90	88.10
C2	9	543	0	0	0	0	0	0	0	0	1.63	98.37
C3	28	0	381	2	0	0	0	0	0	0	7.30	92.70
C4	25	0	0	380	0	0	0	48	0	35	22.13	77.87
C5	0	0	0	0	513	23	0	0	0	0	4.29	95.71
C6	0	0	0	0	0	340	0	0	0	0	0	100
C7	0	0	0	0	0	0	85	0	24	28	37.96	62.04
C8	0	0	0	2	0	0	0	596	23	4	4.64	95.36
C9	0	0	0	0	0	0	12	7	302	7	7.93	92.07
C10	0	0	0	18	0	0	32	2	60	480	18.92	81.08
TC	506	547	385	434	513	363	129	673	409	554	**Precision 90.05%**	
OM	12.25	0.73	1.04	12.44	0	6.34	34.11	11.44	26.16	13.36		
GA	87.75	99.27	98.96	87.56	100	93.66	65.89	88.56	73.84	86.64		

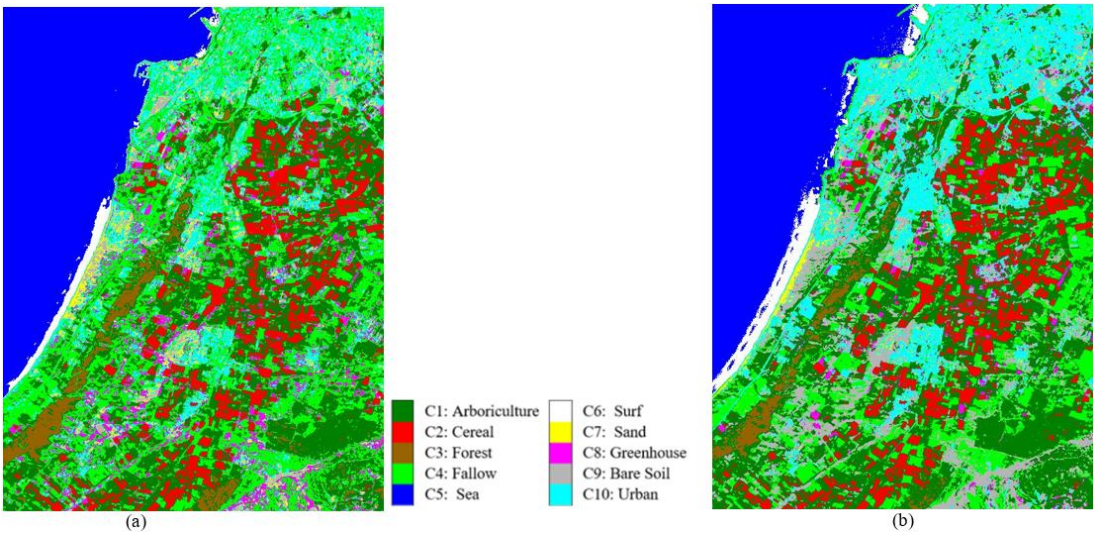

Figure 8. Classification Aster image: (a) with MD method, (b) with EAMD method

Figure 9 showcase examples of errors introduced in the training dataset of the Sea class. Some training samples (illustrated in blue) were taken from agricultural plots (Cereal), this consequently leads to detrimental noise between these two classes exhibiting opposing spectral characteristics.

D. MD classification results with the erroneous training sample

The detailed results represented in Figure 10, Table 11 and Table 12 highlight a poor classification accuracy. The Sea class was entirely omitted by the MD method in the classified image, all pixels of this category are assigned to the *Surf* class. This observation is confirmed by the confusion matrix of the validation set (Table 12), where a confusion error of 100% was recorded between these two classes, emphasizing that no pixel from the *Sea* class was correctly categorized. Furthermore, the confusion matrix (Table 11) indicates that 143 training set pixels (31,02%) from the Sea class were directed to the Cereal class and 318 (68,98%) pixels to the Surf class.

Graph in Figure 11 (b) illustrates the inaccurate training Sea set, comprising two sub-clouds: one with appropriately chosen pixels (bottom) and one with inaccurately selected pixels (top). The correctly chosen pixels are in closer proximity to the center of gravity of the Surf class than to that of its center of gravity: Sea 2, justifying their assignment to the Surf class. The upper sub-cloud is assigned to the Cereal class due to the proximity of its center of gravity to these pixels.

It is crucial to highlight that the lower sub-cloud from Figure 11 (b) represents the initially accurate training sample for the Sea class, represented in Figure 11 (a),

Figure 9. Example of Errors introduced in the training Sea dataset

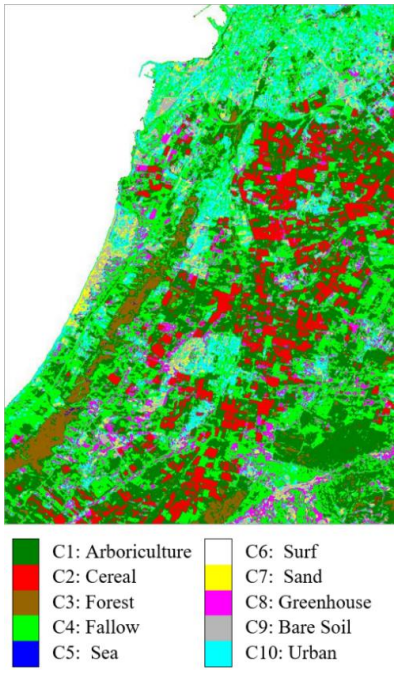

Figure 10. Classified Aster image using MD with erroneous samples

where we can observe that the cloud is well centered on its center of gravity: Sea 1.

Figure 11 (c) and Figure 11 (d) depict the validation set of the Sea class in relation to the centers of other classes. Figure 11 (c) illustrates that the cloud is centered around the Sea 1 center, computed from the accurate training set. However, in Figure 11 (d), the validation pixels are closer to the Surf center than the new center Sea 2, calculated from the inaccurate Sea training set, this justifies their complete migration towards the Surf class. The close proximity of the gravity centers Sea 2 and Forest led to the migration of some pixels from the Forest class to the Sea class, thus justifying the classification errors (Table 11 and 12) recorded between these two spectrally extremely different classes.

E. EAMD results with erroneous training sample

The results of the hybrid approach demonstrate a clear improvement in accuracy compared to the MD method, especially for the *Sea* class that disappeared during the MD classification. The confusion matrix shown in Table 13 reveals that 111 training set pixels from the *Sea* class were misclassified as *Cereal*. It is crucial to note that these pixels are incorrect and should belong to the *Cereal* class. The hybrid approach effectively detected and corrected these errors.

However, despite an overall improvement in results, some confusions remain inevitable. For example, in the classified image shown in Figure 12 (a), agricultural plots (*Cereal)* have been incorrectly assigned to the *Sea* class. Table 13 and 14 indicate that 37 pixels from the training set and 26 pixels from the validation set have migrated from the *Cereal* class to the *Sea* class.

The similar scenario repeats itself with the *Forest* class as well, where 13 pixels from the training set and 10 pixels from the validation set are classified as belonging to the *Sea* class.

This situation arises from the fact that pixels that were not assigned to a specific class in the test *can be classified* (Algorithm 2, line 4) are compelled by the GA to join the elite group, thereby offering them a second opportunity to be assigned to the nearest class (lines 13 to 19 in Algorithm 2). Thus, the approach compels pixels, even incorrect ones, to find a class affiliation, thereby providing certain incorrect pixels with the opportunity to join the *Elite* set. Although the error rate in this scenario is not very high, the accuracy is nonetheless reduced.

TABLE XI. CONFUSION MATRIX OF THE ERRONEOUS TRAINING SAMPLES USING MD METHOD, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	335	0	0	26	0	0	0	4	0	0	8.22	91.78
C2	0	286	0	0	**143**	0	0	0	0	0	33.33	66.67
C3	16	0	211	0	0	0	0	0	0	0	7.05	92.95
C4	12	0	0	184	0	0	0	12	1	5	14.02	85.98
C5	0	0	**10**	0	0	0	0	0	0	0	100	0
C6	0	0	0	0	**318**	157	0	0	0	0	66.95	33.05
C7	0	0	0	5	0	0	33	1	34	87	79.38	20.63
C8	0	0	0	2	0	0	0	318	6	0	2.45	97.55
C9	0	0	0	0	0	0	32	0	131	4	21.56	78.44
C10	0	0	0	32	0	0	37	0	27	184	34.29	65.71
TC	363	286	221	249	461	157	102	335	199	280	**Precision 69.32%**	
OM	7.71	0.00	4.52	26.10	**100**	0	67.65	5.07	34.17	34.29		
GA	92.29	100.00	95.48	73.90	0	100	32.35	94.93	65.83	65.71		

TABLE XII. CONFUSION MATRIX OF THE VALIDATION SAMPLES USING MD METHOD WITH ERRONEOUS TRAINING SAMPLES, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	429	10	4	49	0	0	0	27	0	0	17.34	82.66
C2	9	537	0	0	0	0	0	0	0	0	1.65	98.35
C3	39	0	378	2	0	0	0	0	0	0	9.79	90.21
C4	29	0	0	360	0	0	0	82	0	67	33.09	66.91
C5	0	0	3	0	**0**	0	0	0	0	0	100	0
C6	0	0	0	0	**513**	363	0	0	0	0	58.56	41.44
C7	0	0	0	1	0	0	57	3	92	59	73.11	26.89
C8	0	0	0	3	0	0	0	551	31	3	6.29	93.71
C9	0	0	0	0	0	0	36	6	194	13	22.09	77.91
C10	0	0	0	19	0	0	36	4	92	412	26.82	73.18
TC	506	547	385	434	513	363	129	673	409	554	**Precision 72.70%**	
OM	15.22	1.83	1.82	17.05	**100**	0	55.81	18.13	52.57	25.63		
GA	84.78	98.17	98.18	82.95	**0**	100	44.19	81.87	47.43	74.37		

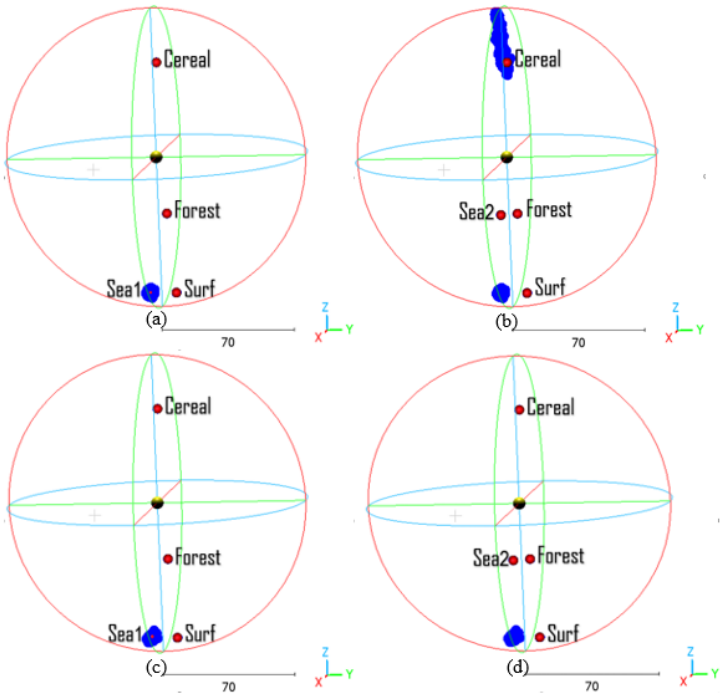

Figure 11. 3D visualization of class centroids and the pixel clouds of the Sea class: (a) initially accurate training sample for the Sea class, (b) the inaccurate training Sea set. (c) the validation set and Sea class center computed from the accurate training set. (d) the validation set and Sea class center computed from the inaccurate training Sea set. (These illustrations were created using the Cloud Compare software, v 2.13)

TABLE XIII. CONFUSION MATRIX OF THE ERRONEOUS TRAINING SAMPLES USING EAMD METHOD, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	353	0	0	18	0	0	0	2	0	0	5.36	94.63
C2	0	249	0	0	**111**	0	0	0	0	0	30.83	69.17
C3	4	0	208	0	0	0	0	0	0	0	1.89	98.11
C4	6	0	0	207	0	0	0	4	0	1	5.05	94.95
C5	0	**37**	**13**	0	350	3	0	0	0	0	13.15	86.85
C6	0	0	0	0	0	154	0	0	0	0	0	100
C7	0	0	0	0	0	0	69	0	4	34	35.51	64.49
C8	0	0	0	2	0	0	0	329	4	1	2.08	97.92
C9	0	0	0	3	0	0	5	0	166	1	5.14	94.86
C10	0	0	0	19	0	0	28	0	25	243	22.86	77.14
TC	363	286	221	249	461	157	102	335	199	280	**Precision 87.75%**	
OM	2.75	12.94	5.88	16.87	24.08	1.91	32.35	1.79	16.58	13.21		
GA	97.25	87.06	94.12	83.13	75.92	98.09	67.65	98.21	83.42	86.79		

TABLE XIV. CONFUSION MATRIX OF THE VALIDATION SAMPLES USING EAMD METHOD WITH THE ERRONEOUS TRAINING SAMPLE, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	441	4	4	32	0	0	0	20	0	0	11.98	88.02
C2	9	517	0	0	0	0	0	0	0	0	1.71	98.29
C3	28	0	371	2	0	0	0	0	0	0	7.48	92.52
C4	25	0	0	380	0	0	0	48	0	35	22.13	77.87
C5	3	**26**	**10**	0	497	**30**	0	0	0	0	12.19	87.81
C6	0	0	0	0	**16**	333	0	0	0	0	4.58	95.42
C7	0	0	0	0	0	0	85	0	24	28	37.96	62.04
C8	0	0	0	2	0	0	0	596	23	4	4.64	95.36
C9	0	0	0	0	0	0	12	7	302	7	7.93	92.07
C10	0	0	0	18	0	0	32	2	60	480	18.92	81.08
TC	506	547	385	434	513	363	129	673	409	554	**Precision 88.68%**	
OM	12.85	5.48	3.64	12.44	3.12	8.26	34.11	11.44	26.16	13.36		
GA	87.15	94.52	96.36	87.56	96.88	91.74	65.89	88.56	73.84	86.64		

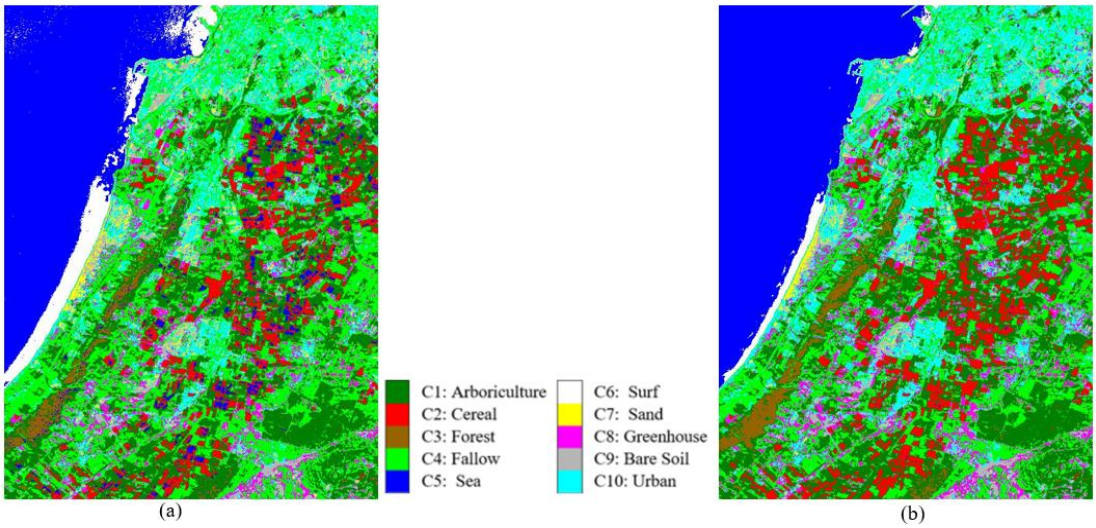

Figure 12. Classified Aster image using erroneous samples: (a) with EAMD 1st approche , (b) with EAMD 2nd approach

TABLE XV. CONFUSION MATRIX OF THE OF ERRONEOUS TRAINING SAMPLES USING EAMD METHOD 2ND APPROACH, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	348	0	0	13	0	0	0	2	0	0	4.13	95.87
C2	0	286	0	0	**142**	0	0	0	0	0	33.18	66.82
C3	4	0	221	0	0	0	0	0	0	0	1.78	98.22
C4	11	0	0	195	0	0	0	4	0	1	7.58	92.42
C5	0	0	0	0	319	0	0	0	0	0	0	100
C6	0	0	0	0	0	157	0	0	0	0	0	100
C7	0	0	0	0	0	0	59	0	4	35	39.80	60.20
C8	0	0	0	2	0	0	0	329	4	1	2.08	97.92
C9	0	0	0	3	0	0	5	0	166	1	5.14	94.86
C10	0	0	0	36	0	0	38	0	25	242	29.03	70.97
TC	363	286	221	249	461	157	102	335	199	280	**Precision 87.52%**	
OM	4.13	0	0	21.69	30.80	0	42.16	1.79	16.58	13.57		
GA	95.87	100	100	78.31	69.20	100	57.84	98.21	83.42	86.43		

TABLE XVI. CONFUSION MATRIX OF THE VALIDATION SET WITH RRONEOUS TRAINING SAMPLES, USING EAMD METHOD 2ND APPROACH, ASTER IMAGE. C1: ARBORICULTURE, C2: CERAL, C3: FOREST, C4: FALOW, C5: SEA, C6: SURF, C7: SAND, C8: GREENOUSE, C9: BARE SOIL, C10: URBAN

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	CO	IA
C1	434	5	60	34	0	0	0	21	0	0	21.66	78.34
C2	9	526	0	0	0	0	0	0	0	0	1.68	98.32
C3	28	0	325	2	0	0	0	0	0	0	8.45	91.55
C4	35	0	0	354	0	0	0	53	0	33	25.47	74.53
C5	0	16	0	0	513	21	0	0	0	0	6.73	93.27
C6	0	0	0	0	0	342	0	0	0	0	0	100
C7	0	0	0	0	0	0	77	0	26	31	42.54	57.46
C8	0	0	0	14	0	0	0	587	24	3	6.53	93.47
C9	0	0	0	0	0	0	12	10	262	8	10.27	89.73
C10	0	0	0	30	0	0	40	2	97	479	26.08	73.92
TC	506	547	385	434	513	363	129	673	409	554	**Precision 86.32%**	
OM	14.23	3.84	15.58	18.43	0	5.79	40.31	12.78	35.94	13.54		
GA	85.77	96.16	84.42	81.57	100	94.21	59.69	87.22	64.06	86.46		

What if we were to remove this opportunity for erroneous samples?

F. EAMD a Second Approach

In order to ensure that the GA avoids including incorrect pixels in the elite set, various tests were performed on these samples, following the same approach with the exception of the evaluation phase, where lines of code from 12 to 25 were excluded. Consequently, the GA provides only one chance for each pixel to become a part of the elite set. The results, illustrated in Figure 12 (b), Table 15 and 16, demonstrate the successful noise

filtration achieved by the GA in this approach. Notably, out of the 143 erroneously introduced pixels in the Sea class, only one was retained in the elite set. This effectively resolved the problem of confusions between the *Sea* and *Cereal* classes in the training data set, as evident in the confusion matrix (Table 15).

It is noteworthy that, notwithstanding the potential interpretation of these pixels' presence in the confusion matrix as indicative of confusion errors, it indeed underscores the approach's high level of precision. These erroneous incorporated pixels, intentionally introduced into the evolutionary process, have been accurately identified by the GA as incorrect data, while concurrently determining their true class membership.

The validation confusion matrix shows that all pixels of the *Sea* class were correctly classified, and the noise between it and the Surf class was eliminated.

This idea was also tested on the initial samples (without errors). The results match those shown in Figure 12 (b), Table 15 and 16 as the erroneous sample filtered by the GA is simply the initial sample (before the addition of noise).

The results show an improvement in classification performance compared to the classical approach, which relies solely on the MD method as shown in Figure 8 (a), Table 7 and 8. However, it did not achieve the success rates of the first approach (Figure 8 (b), Table 9 and 10). This result can be explained by the fact that it offers only one opportunity for pixels to join the elite set, and based solely on the optimal set of sub-classes. On one hand, this approach has the advantage of not including erroneous pixels, but on the other hand, it eliminates some deserving pixels that were selected by the first approach.

By using the elite set and the sub-classes generated by the GA of the second approach (Erroneous filtered samples) as the starting base for the classification according to the first approach, we achieve the same result as illustrated in the Figure 8 (b), Table 9 and 10, with a reduced number of generations, thereby less execution time. This is because in our experiments, once filtered by the second approach, the erroneous file remains identical to the initial sample set of the first approach before noise injection.

8. CONCLUSION

The evolutionary approach presented in this article demonstrates the power of GA in improving the accuracy of supervised classification, with a particular focus on the quality of the training samples, which significantly impacts the classification quality. The strength of this approach lies not only in the classification rates obtained but also in its ability to detect and eliminate potential errors within the training set, especially in cases where the samples have not been carefully chosen.

The proposed evaluation process generates, in output, two crucial sets for each individual: the set of sub-classes that best represents each class and the set of pixels called "Elites." These two sets play an essential role in the step of assigning pixels to classes.

The proposed fitness measure is linked to the precision of the classification, in case it is carried out by considering the individual under evaluation. However, even if this individual manages to achieve 100% accuracy for correctly classified pixels (0% omissions), he is penalized by the percentage of commission errors he may cause. This penalty is possible because classes are considered interdependent, with all of them being simultaneously included in the learning process, thereby enhancing individual performance and, consequently, the overall precision of the classification.

Two approaches have been proposed to evaluate individuals. The first aims to include all pixels from the training dataset in the elite set. Pixels that may cause confusion within a class are selected to be part of the elite set of the class closest to them.

This method has significantly reduced confusions related to spectral similarities between certain classes compared to the classical approach.

However, it has the disadvantage of encompassing potentially misleading pixels in the elite set, as each pixel in the dataset is obliged to find a class membership.

In the second approach, the genetic algorithm retains only relevant pixels. Consequently, some pixels from the training dataset may be excluded from the process without being assigned to a class, thereby increasing the risk of rejecting important pixels and depriving them of joining the elite. Experiments demonstrate the effectiveness of this approach in detecting and avoiding intentionally introduced erroneous pixels in the training dataset. However, since it does not require pixels to find a class membership, this approach has the drawback of rejecting certain representative pixels, potentially reducing confusions, especially in cases where classes exhibit identical spectral behavior in the utilized spectral bands.

The hybrid method proposed in this paper significantly enhances results compared to classical spectral classification. For optimal performance, adopting firstly the second approach is recommended for meticulous sample filtering, especially if the selection of training samples has not been rigorous. The obtained results will serve as the initial foundation for the first approach.

REFERENCES

[1] AT. Magpantay and P. L. Fernandez, “Improving the Classification of Landsat-8 OLI Images using Neighborhood Median Pixel Values”. In: 2020 International Conference on Communication and Signal Processing (ICCSP), IEEE, Chennai, India, pp. 1054–1058, July 2020. https://doi.org/10.1109/ICCSP48568.2020.9182359

[2] M. Chi, Plaza A, JA. Benediktsson, Z. Sun, J. Shen and Y.Zhu, "Big Data for Remote Sensing: Challenges and Opportunities", Proc IEEE, vol. 104, Num 11, pp. 2207–2219, November 2016. https://doi.org/10.1109/JPROC.2016.2598228

[3] H. Song, Q. Liu, G. Wang, R. Hang and B. Huang, "Spatiotemporal Satellite Image Fusion Using Deep Convolutional Neural Networks", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, pp. 1–9, february 2018. https://doi.org/10.1109/JSTARS.2018.2797894

[4] K. Sanghvi, A. Aralkar, S. Sanghvi and I. Saha , "A Survey on Image Classification Techniques", SSRN Electronic Journal, January 2021. https://doi.org/10.2139/ssrn.3754116

[5] S. Revanna, P. Deepa and K. R. Venugopal, "Remote Sensing Satellite Image Processing Techniques for Image Classification: A Comprehensive Survey", International Journal of Computer Applications, Vol. 161, pp. 24–37. March 2017. https://doi.org/10.5120/ijca2017913306

[6] S . Abburu and S. Babu Golla, "Satellite Image Classification Methods and Techniques: A Review". IJCA- International Journal of Computer Applications, Vol. 119, Num. 8, pp. 20–25, June 2015. https://doi.org/10.5120/21088-3779

[7] A. Singh and KK. Singh, "Satellite image classification using Genetic Algorithm trained radial basis function neural network, application to the detection of flooded areas", Journal of Visual Communication and Image Representation, Vol 42, pp. 173–182. January 2017. https://doi.org/10.1016/j.jvcir.2016.11.017

[8] S. Dhingra and D. Kumar, "A review of remotely sensed satellite image classification". International Journal of Electrical and Computer Engineering, Volum 9, pp. 1720–1731. June 2019.

[9] M. Espinola, J.A. Piedra-Fernandez, R. Ayala, L. Iribarne and J. Wang, "Contextual and Hierarchical Classification of Satellite Images Based on Cellular Automata", IEEE Transactions on Geoscience and Remote Sensing, Vol. 53, Num. 2, pp. 795–809. May 2015. https://doi.org/10.1109/TGRS.2014.2328634

[10] O. El Kharki, J. Mechbouh, D. Ducrot and M. Rouchdi and J.M. Ngono, "Panorama sur les méthodes de classification des images satellites et techniques d'amélioration de la précision de la classification ".Revue Française de Photogrammétrie et de Télédétection (RFPT), Num 210, pp. 23–38, April 2015. https://doi.org/10.52638/rfpt.2015.259

[11] A.B. Gavade and V.S. Rajpurohit, "Systematic analysis of satellite image-based land cover classification techniques: literature review and challenges". International Journal of Computers and Applications, Vol. 43, pp. 1–10. Febraury 2019. https://doi.org/10.1080/1206212X.2019.1573946

[12] A.K. Goswami, S. Sharma and P. Kumar, "Nearest Clustering Algorithm for Satellite Image Classification in Remote Sensing Applications". IJCSIT- International Journal of Computer Science and Information Technologies, Vol. 5 (3), pp. 3768–3772, 2014.

[13] S. Ahmed, "Comparison of Satellite Images Classification Techniques using Landsat-8 Data for Land Cover Extraction". International Journal of Intelligent Computing and Information Sciences-IJICIS Vol 21, Num 3, pp. 29–43. September 202. https://doi.org/10.21608/ijicis.2021.78853.1098

[14] S. Borra, R. Thanki, and N. Dey, "Satellite Image Analysis: Clustering and Classification". SpringerBriefs in Applied Sciences and Technology, Springer, Singapore, Febraury 2019, pp. 31 https://doi.org/10.1007/978-981-13-6424-2_3

[15] M. C. Girard and M. Girard, "Traitement des données de la télédétection , Environnement et ressources naturelles", 2nd edition, DUNOD., June 2010.

[16] Y.K. Hamad, A.Y. Yousuf and T.S. Atia, "An Evolutionary Quantum Scrambling Scheme for Medical Image with NEQR Representation", International Journal of Computing and Digital Systems, Vol. 13, Num. 1, pp. 1383-1396, May 2023. http://dx.doi.org/10.12785/ijcds/1301112

[17] J.H. Holland, "Adaptation in Natural and Artificial Systems: an Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence", MIT press, 1992.

[18] M. Gen and R. Cheng, "Genetic Algorithms & Engineering Design". A Wiley-Interscience Publication, John Wiley and Sons, INC, 1997.

[19] L. Haldurai, T. Madhubala and R. Rajalakshmi, "A Study on Genetic Algorithm and its Applications". International Journal of Computer Sciences and Engineering, Vol. 4, Num. 10, pp.139–143, 2016

[20] S. Mirjalili, J. Dong , and A. Lewis, " Nature-Inspired Optimizers Theories, Literature Reviews and Applications", Studies in Computational Intelligence, Vol. 811, Springer, Poland, 2020. https://doi.org/10.1007/978-3-030-12127-3

[21] K. Rajwar, K. Deep K and S. Das, "An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges", Artificial Intelligence Review, Vol. 56, pp. 13187–13257, April 2023. https://doi.org/10.1007/s10462-023-10470-y

[22] M.D. Yang, Y.F. Yang, T.C. Su and K.S. Huang, "An Efficient Fitness Function in Genetic Algorithm Classifier for Land use Recognition on Satellite Images". Hindawi Publishing Corporation, The Scientific World Journal, Vol. 2014, PP. 1–12. http://dx.doi.org/10.1155/2014/264512

[23] S. Kapil, M. Chawla and M.D. Ansari, "On K-means data clustering algorithm with genetic algorithm". In: 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC). In: IEEE, Waknaghat, India, pp. 202–206. , 2016. https://doi.org/10.1109/PDGC.2016.7913145

[24] C. Sukawattanavijit, J. Chen and H. Zhang, " GA-SVM Algorithm for Improving Land-Cover Classification Using SAR and Optical Remote Sensing Data". In: IEEE Geoscience and Remote Sensing Letters, Vol. 14, pp. 284–288. March 2017. https://doi.org/10.1109/LGRS.2016.2628406.

[25] A. Quirin and J. Korczak, "Discovering of Classification Rules from Hyperspectral Images". In: Eurasip Book Series on Signal Processing and Communications, January 2006, pp. 327–348.

[26] A. Haldar, S. Biswas and S. De, "A Review on Genetic Algorithm", International Journal of Engineering Sciences & Research Technology IJESRT, Vol. 3, pp. 369–372. January 2014.

[27] K.B. Mankad, "The Significance of Genetic Algorithms in Search, Evolution, Optimization and Hybridization: A Short Review", International Journal of Computer Science and Business Informatics, Vol. 9, pp. 103–115, January 2014.

[28] S. Katoch, S.S. Chauhan and V. Kumar, "A review on genetic algorithm: past, present, and future". Multimedia Tools and Applications, Vol. 80, pp. 8091–8126. October 2020. https://doi.org/10.1007/s11042-020-10139-6.

[29] L. Deng, P. Yang and W. Liu, "An Improved Genetic Algorithm". In: IEEE 5th International Conference on Computer and Communications, pp. 47–51. Downloaded from IEEE Xplore, 2019.

[30] P. Ghodmare, M. Jibhakate, N. Meshram , V. Bele and P. Ghir, "A Review Paper on Brief Introduction of Genetic Algorithm", IJSTE- International Journal of Science Technologie & Engineering, Vol. 4, Num. 8, pp. 42– 44, February 2018.

[31] A. Shukla, H.M. Pandey and D. Mehrotra, "Comparative Review of Selection Techniques in Genetic Algorithm". International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), Greater Noida, India, pp. 515-519, March 2015.

[32] Z. Lin and G. Zhang, "Genetic algorithm-based parameter optimization for EO-1 Hyperion remote sensing image classification", European Journal of Remote Sensing, Vol. 53, pp. 124–131, April 2020.

https://doi.org/10.1080/22797254.2020.1747949.

Ismahane Kariche was born in Blida, Algeria. She received the Engineer degree in Computer Science from the University of Sciences and Technology of Oran Mohamed Boudiaf (USTO-MB) and a Magister degree in Space Techniques and Applications from the Center of Space Techniques, Oran, Algeria. Currently, she serves as a lecturer researcher at the University of Oran 2 Mohamed Ben Ahmed, and is pursuing a doctoral degree in Computer Science at the USTO-MB. Her research activities focus on Artificial Intelligence, Metaheuristics and Bioinspired Algorithms, Satellite Image Processing, Remote Sensing and Geomatics.

Hadria Fizazi was born in Oran Algeria. Obtained Engineer degree in Electrical Engineering from the University of -Mohamed Boudiaf- in 1981, a doctor Engineer degree in Industrial Computer Science and Automation from the University of Lille 1 Franch in 1987 and the Doctorate degree in Computer Science from the University of -Mohamed Boudiaf - in 2005. Her research interests cover bioinspired algorithms, metaheuristics, image processing and the application of pattern recognition to remotely sensed data. She is currently a Professor in the Faculty of mathematics and Computer Science at the University of Science and Technology of Oran -Mohamed Boudiaf-.