ISSN 2278-3091
Volume 11, No.6, November - December 2022
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse041162022.pdf
https://doi.org/10.30534/ijatcse/2022/041162022

Automatic Change Detection on Satellite Images using Principal Component Analysis, ISODATA and Fuzzy C-Means Methods

BEKKOUCHE Ibtissem[1], FIZAZI Hadria[2]
[1] Department of Computer Sciences, University of Science and Technology of Oran, Algeria,
ibtissem.bekkouche@univ-usto.dz
[2] Department of Computer Sciences, University of Science and Technology of Oran, Algeria,
hadria.fizazi@univ-usto.dz

Received Date: October 17, 2022 Accepted Date: November 23, 2022 Published Date: December 06, 2022

ABSTRACT

Change detection is the process of comparing two or more images and identifying the parts where a change has occurred. Difference detection processing between simple digital images, such as photographic images, is easy to implement. Whereas for satellite images, which compose of several images' grayscale and bands, this requires a methodological approach to image processing appropriate to the exploitation of these data because this will allow to follow the evolution over time of a region of interest through change detection techniques, so these images are a tool of choice in the management of natural resources. So, in this paper, we propose a hybrid automatic change detection approach for multi-temporal satellite images. It is based on several algorithms: ISODATA for automatic thresholding, Principal Component Analysis as transformation technique, Fuzzy C-Means as classification technique. Experiments were performed and assessed by their overall accuracy and results validated the effectiveness and efficiency of the proposed approach, named ISOFAP.

Key words : Change detection, Fuzzy c-means clustering, ISODATA, Principal component analysis.

1. INTRODUCTION

Change detection is a digital process that can be performed by traditional methods and using remote sensing technologies. The basic this process is to measure the change on the Earth's surface by jointly analyzing two or more two temporally separated images, in order to locate and quantify (automatically) the changes existing between these images [1]. This is a very active subject due to preoccupation about the consequences of global and local changes in the earth.

There are many change detection techniques in the literature: Algebraic methods such as Image differencing and Change vector analysis,…; Transformation methods such as Tasseled Cap Transformation, Principal Component Analysis…; Classification methods as Artificial Neural Networks, Comparison after classification…; Geographic Information System (GIS) as Integrated Method of GIS and Remote Sensing, GIS Approach…; Visual analysis using Visual interpretation; Hybrid approach in Combination of methods; Advanced models like Spectral Mixing Model, Li-Strahler Reflectance Model… [2][3]

In recent years, this technique has become one of the most interesting subjects in the extraction of information from satellite images and several researchers have opted, These last years, for the hybridization of methods such as: in 2020, Neelam Ruhil et al, have suggested an unsupervised change detection method based on wavelet fusion and the Kohonen Hybrid FCM-σ [4], in the same year, Mohan Singh et al, have proposed an image fusion using image fusion using image normalization and radiometric calibration and Particle Swamp Optimization Fuzzy C-Means (PSOFCM). In this article, an unsupervised change observation technique based on the PSOFCM [5], and in 2022, Abdelkrim Maarir et al, have proposed an unsupervised method of detecting change in satellite images by following two main steps: The first step focuses on data reduction using the Independent Component Analysis (ICA) algorithm to improve the efficiency of the classifier. The second stage for processing uses the Fuzzy C-Means classification method to find specified clusters [3].

So, for our paper, since there are several change detection techniques, the most interesting attitude would be to try to combine these techniques and develop a hybrid method, so for that we used: ISODATA for automatic thresholding, Principal Component Analysis (PCA) as transformation technique, Fuzzy C-Means (FCM) as classification technique.

After having tested several thresholding algorithms such as: Binary thresholding on the mean, OTSU thresholding and EM algorithm [6], we chose ISODATA. This is the algorithm that is used for automatic thresholding, we have chosen it because it is easy to implement more than it gives good results.
Principal Component Analysis (PCA) consists of transforming variables, interconnected, into new variables unsquared from each other for dimension reduction. So, we chose because it has been widely used for change detection

[7]. It has the ability to project the multi-dimensional original and it is less susceptible to image overlapping.

We found that the most used methods for change detection used a lot the K-means, so we tested it on our data but we chose to use the improved Fuzzy C-Means (FCM) version. FCM has reduced complexity and gives best result for overlapping data sets and comparatively better than k-means algorithm.

2. METHODS USED

2.1. Isodata

It is an image segmentation technique by clustering and it is an improved version of the k-means algorithm, which was first introduced by Velasco in 1980 [8] as a classic algorithm which makes it possible to carry out a very good categorization and which gives significant results. [9]

The thresholding of an image can be done by manual thresholding or by automatic thresholding.

Manual image thresholding involves 4 steps [9]:

1. Observing the histogram of the image
2. Choice of thresholds in the valleys
3. Definition of the classes of the regions by color range
4. Pixel classification

ISODATA thresholding is global thresholding, where a single threshold is used across the entire image to divide it into two clusters. It allows to find the value of the sought threshold in an automatic way following its steps [10]:

1. Choose an initial threshold T, for example T = the average intensity.
2. Divide the image into two groups G1 and G2 using T.
3. Calculate the average values of each region R1 and R2
4. Calculate the value $T = (R1+R2)/2$
5. Repeat steps 2 to 4 until the T value does not change.

2.2. Principal Component Analysis

Principal component analysis (PCA) is a mathematical technique used for data redundancy reduction by Jackson and Bund, 1983 [12]. It is useful when you have obtained data on a number of variables (perhaps a large number of variables), and there is some redundancy in these variables. [12]

PCA is one of the most popular multivariate analysis algorithms for change detection studies and can be performed on original or normalized data [7]. With this technique the digital images acquired by remote sensing, we can reduce its dimensionality such that the multispectral bands are the variables to be introduced.

There are certain steps to follow to implement PCA [7]:

- Take an original data set and calculate the mean of the data set.
- Subtract the mean for each dimension.
- Calculate the covariance matrix.
- Calculate the eigenvector and the eigenvalue of the covariance matrix.
- Extract the diagonal of the matrix as a vector.
- Variance sorting in descending order.
- Choose components and form a feature vector.
- Derivation of the new data set.

At the end, the number of PC is less than the number of variances in the original image. In CD studies, the consequence of this linearization is that the unchanged pixels or common information shared by a pair of images are assumed to be in a narrow and elongated space. Cluster along a principal axis equivalent to the first component (PC1). On the contrary, pixels containing a change would be more unique in their spectral appearance and should lie far from this axis (PC2). [13][14]

2.3. Fuzzy C-Means

Fuzzy C-Means (FCM), is an unsupervised fuzzy classification algorithm. Issued from the C-means algorithm, developed by Dunn in 1973 [15] and improved by Bezdek in 1981 [16], it introduced the notion of fuzzy set in the definition of classes: each point in the data set belongs to each cluster with a certain degree, and all clusters are characterized by their center of gravity [17].
The goal of Fuzzy C-Means clustering is to find the minimum of the following function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}\left\|x_i - c_j\right\|^2, \quad 1 \leq m \leq \infty \quad (1)$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j. As a rule, for each pixel, sum of all membership value belonging to all classes must be 1. [3][14]

3. METHODOLOGY

The Figure 1 shows and summarizes the organization of the processing steps adopted for the detection of changes on satellite images by our ISOFAP approach and this according to the basic procedure of an image change detection processing system [18].

This scheme is composed of several implementation phases:

1st step: This is a data preparation step and we must:

- ***Data acquisition and preparation:*** This is the stage of collecting and assembling data, which can be satellite images and field investigations. Then we can go to pre-processing, for example image cutting if necessary.
- ***Geometric correction:*** the verification of the geometric accuracy is essential for the detection of changes besides a bad georeferenced of more than one pixel would cause abnormal results for analyzes pixel by pixel.

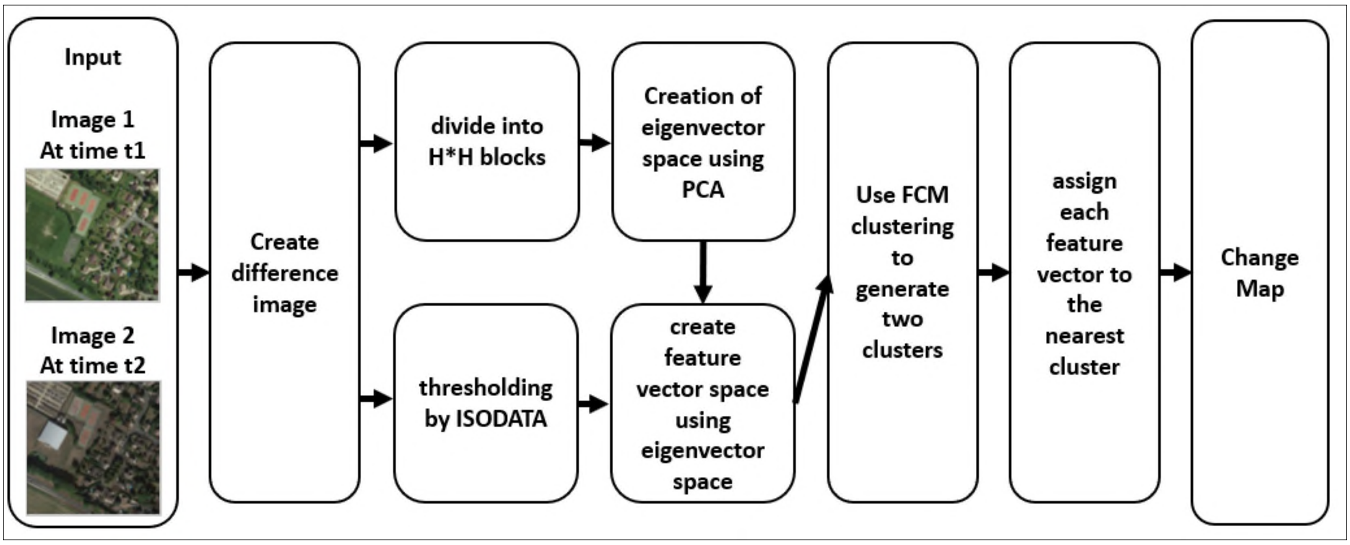


Figure 1: General scheme of the proposed ISOFAP method.

- ***Data normalization:*** Data must be normalized, especially for satellite images, in order to reduce the variability between multi-date images over the same geographical area.

2nd step: This is a digital processing step for the detection of changes and for this we will:

- Apply the difference image method
- After that, the obtained image is partitioned into H*H blocks, then the creation of eigenspace space using PCA and principal components are achieved.
- Apply the ISODATA algorithm for thresholding on X.
- After applying PCA and ISODATA, create the feature vector space using the eigenspace space.
- To determine the areas that have changed and not changed, we used the FCM algorithm to generate the two classes (k=2) and assign each feature vector to the class closest to either modified pixels or unchanged pixels.

3th step: This is a generalization of the final product, which is a map of land cover changes at a scale equivalent to that of the input data.

4th step: Once a model has been determined and implemented, the last step is to analyze and interpret the results to establish the quality of this model. There are various evaluation measures that can be used and chosen carefully, since the choice of measure can influence how performance is assessed and interpreted. For this we have chosen to use:

- ***Visual interpretation:*** the use of the human visual system as a quality judgment tool is not to be neglected but necessary to verify the quality of the images obtained by the classification. To evaluate this approach, we also used visual analysis according to the ground truth available in the area.
- ***Confusion Matrix:*** One of the most popular ways to measure the performance of a classification model. Each line corresponds to an actual class and each column corresponds to an estimated class and it includes the following values [19]:
  - True Positive, TP, when the actual class and the estimated class are both positive
  - True Negative, TN, when the actual class and the estimated class are both negative
  - False Positive, FP, when the actual class is negative but the estimated class is positive. This is called a Type 1 error.
  - False Negative, FN, when the real class is positive but the estimated class is negative. This is called a Type 2 error.

It can be used for more in-depth measurements to get a better assessment of the quality of the model. Among the classification measures used are accuracy, precision, error and specificity. [19]

Accuracy is the number of correct predictions made by the model.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2)$$

This measure is used when the number of True Positives and True Negatives are the most important.

$$\text{Error} = 1 - Accuracy \quad (3)$$

Precision is the number of correct elements rendered by the model.

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

This metric is used when the number of False Positives is highest.

Specificity is the number of negative classes predicted by the model.

$$Specificity = \frac{TN}{TN+FP} \quad (5)$$

4. IMAGES USED

Images database: A set of RGB images of 650 x 650 found in a database used for the detection of changes in satellite imagery using deep-learning, produced by Héloïse BAUDHUIN and Antoine LAMBOT. [20]

We applied our approach to two examples, presented in Figure 2 and Figure 3, of this database:

(a)

(b)

Figure 2: Images 1 (a) before change and (b) after change

(a)

(b)

Figure 3: Images 2 (a) before change and (b) after change.

Images of Boumerdes: Figure 4 shows a high-resolution satellite image of a Landsat 5 TM earthquake and both acquired in 2003, provided by the Center National des Techniques Spatiales d'Arzew and acquired by QuickBird.

The characteristics of the two images are: natural composition image with three channels: TM1, TM2 and TM3 bands and their size in pixels is 1002 x 1002. They contain different classes which are: asphalt, soil, vegetation and shade, and area of damage to the post-disaster image.

(a)

(b)

Figure 4: Images 3 (a) before disaster and (b) after disaster.

5. EXPERIMENT RESULTS ANS DISCUSSION

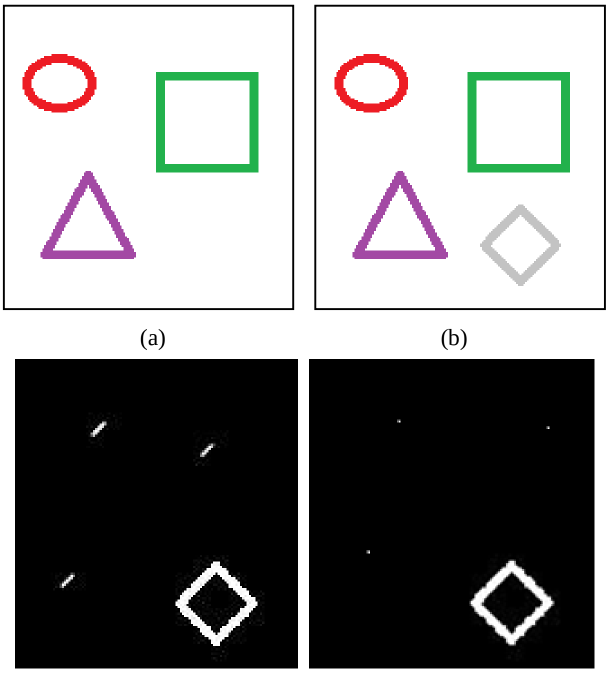
The first thing for our work, we did a study on a known change detection method to understand its concept and we chose PCA-FCM. After the study our proposal was to change the classification algorithm and replace it with another, so after several tries, we chose to work with the Fuzzy C-means (FCM) which is an improved version of the K-means algorithm. Then we did other tests with PCA-FCM but each time we had to initialize the threshold manually with the variation of the threshold parameters for each image. So, for that we have integrated an ISODATA thresholding algorithm to have an automatic threshold. In the rest of the article, some results of our proposed method "ISOFAP" in comparison with PCA-K-means.

To analyze and validate the proposed approach we used three different data images described in section 4 below.

For the PCA-K-means parameters, we varied the parameters as follows: number of classes is 2 classes, such as class 1 for changed pixels and class 2 for unchanged ones, number of iterations between 70 and 100, number of blocks: between 2 and 5, and Threshold: between 10 and 80.

And for ISOFAP we varied the parameters like the PCA-K-means except the threshold by ISODATA.

Figure 5 shows the first performance test, we created two examples of simple artificial images to do our tests, but we reduced the number of iterations to 20 and the threshold to 10 because they are simple images.



(a) (b)

(c) (d)

Figure 5: (a) artificial image, (b) artificial image changed, change map results (c) with PCA-K-means and (d) with ISOFAP.

By the visual interpretation, we notice that PCA-K-means has some confusions and that the detection is not so correct. However, ISOFAP gave us a better result than PCA-K-means although there are also confusions but it is minimal.

After confirming the correct operation of the two algorithms on a simple artificial image, we used the images 1 of the database (presented in Figure 2) and we applied for the two methods PCA-K-means and ISODATA. For initialization of the PCA-kmeans method, after several tests, the threshold was manually initialized to 60, the number of iterations to 80 and we varied just the number of blocks (h*h) between h equal to 2 and 5. We notice that he detected the changes for h equal 2.

The same image applied to our method ISOFAP, the number of iterations at 75, the variation also in the number of blocks between 2 and 5, and knowing that ISODATA has initialized the threshold to 43. from the results we also notice that at h equal to 2 the result is better than the others.

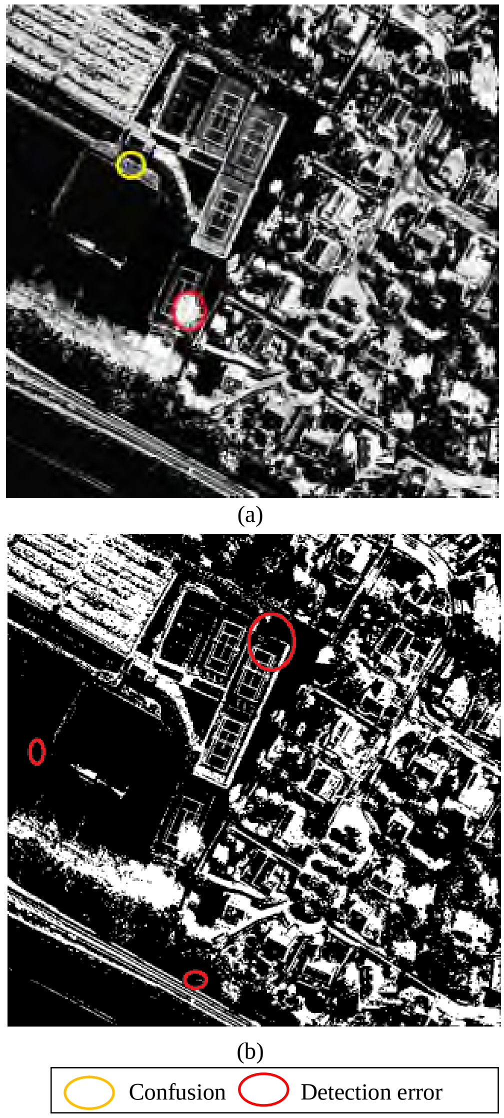


Figure 6: Visual interpretation of the best results by (a) ISOFAP and (b) PCA-K-means on images 1.

The previous Figures 6 present the best results given by ISOFAP and PCA-K-means applied to images 1. From our visual interpretation for comparing the best results of the two methods to the original change image, we notice that even if ISOFAP had confusions and detection errors caused by the conflict between the two changed and unchanged classes, it there are some details were better detected compared to PCA-K-means.

For images 2, we did the same tests as images 1. For PCA-Kmeans, after several tests we initialized the threshold at 75, the number of iterations at 85 and varied the number of blocks. we note that these results reinforce the change detection results on images 1 because the best result is given on the number of blocks equal to 2.

For ISODATA the best result for images 2 is the number of blocks at 2 because it gives more detailed detection and fewer conflicts, such as for this image the threshold at 20 and the number of iterations at 93.

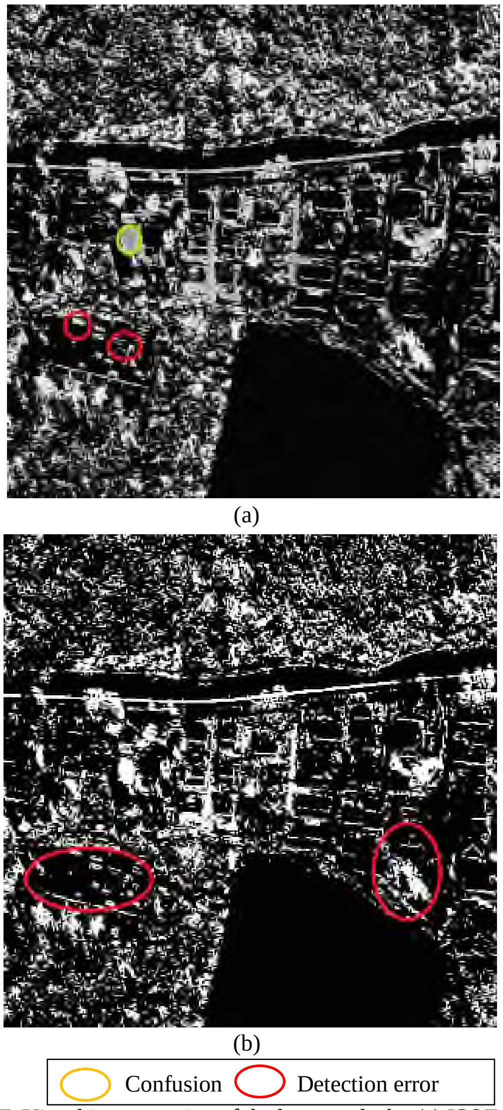


Figure 7: Visual interpretation of the best results by (a) ISOFAP and (b) PCA-K-means on images 2.

The comparison between the best result given by the two methods approves the previous result because we can say that the visual interpretation, presented in Figure 7, is the same and that ISOFAP has better detected the change even if there are conflicts and detection errors but they are less than PCA-K-means.

For the last tests, we chose the best parameters: number of blocks (h*h) h at 2, number of iterations at 90 and we applied them to the same part of the images 3 (In Figure 8). We notice that the PCA-K-means only gave us two classes: changed and unchanged, but ISOFAP gave a third class of pixels with the gray color, for the program it is conflicts and at the same time we notice that it is not totally unchanged but the change is not great.

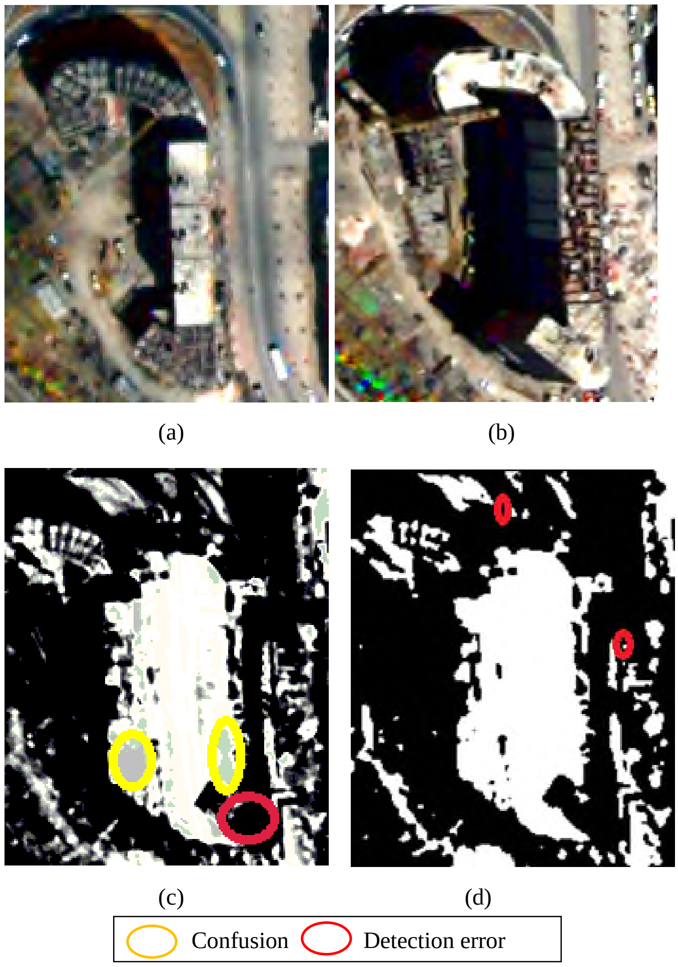


Figure 8: (a) image before, (b) artificial after, change map results (c) with ISOFAP and (d) with PCA-K-means.

From the visual interpretation of the results of our ISOFAP approach on the three groups of images, we notice that there are detection errors in the images, which can be caused by several reasons among them the resolution of the image because sometimes when the image is of lower quality poses conflicts between the pixels therefore gives errors.

Table 1: Evaluation metric results for images 1 by ISOFAP and PCA-K-means

Method	*Accuracy*	*Error*	*Precision*	*Specificity*	*Execution time*
ISOFAP	0.8556	0.1444	0.8918	0.1082	65.283269 s
PCA-K-means	0.7624	0.2376	0.6881	0.3119	57.882101 s

Table 2: Evaluation metric results for images 2 by ISOFAP and PCA-K-means

Method	*Accuracy*	*Error*	*Precision*	*Specificity*	*Execution time*
ISOFAP	0.6256	0.3744	0.7488	0.2512	75.243269 s
PCA-K-means	0.5824	0.4176	0.6181	0.3819	71.874101 s

Table 3: Evaluation metric results for images 3 by ISOFAP and PCA-K-means

Method	*Accuracy*	*Error*	*Precision*	*Specificity*	*Execution time*
ISOFAP	0.7356	0.2644	0.7918	0.2082	55.267169 s
PCA-K-means	0.6424	0.3576	0.5841	0.4159	47.817301 s

We finish our study by applying the confusion matrix, to the original images and the images of the best results, for extract the information that interests us and here are the results:

After the comparison by the evaluation metric (in Table 1, Table 2 and Table3) and the visual interpretation, we notice that our proposed ISOFAP approach gives good results to detect changes and the values of accuracy and precision are high for three different data images, so we conclude that ISOFAP is better than PCA-K-means with the exception of the execution time, it takes longer than PCA-K-means.

6. CONCLUSION

In this work, we have addressed one of the image processing operators which is the detection of changes in satellite images. We have proposed a hybrid method based on two techniques to do this treatment.

Our approach is based on a method already used for the detection of changes (PCA-K-means) and we have tried to improve it. We used ISODATA to make the thresholding automatic and kept the PCA because it is the most used for change detection and the least sensitive to image overlap, while we chose FCM because it has reduced complexity and it is an improved version of the k-means algorithm.

After the tests and the additions, we arrived at the implementation of our ISOFAP method, which is compared with the PCA-K-means, the results allowed us to conclude that our method can detect the change and that it gives better results.

REFERENCES

1. S. Achour. **Détection de changement en imagerie satellitaire multi-temporelle pour le suivi et la surveillance de l'environnement**, Ph.D. dissertation, Dept. Elect. Eng., Djillali Liabès Univ, Sidi Bel Abbes, Algeria, 2022.
2. D. Lu and P. Mausel, E. Brondizio, and E. Moran. **Change detection techniques**, Change detection techniques, International Journal of Remote Sensing, vol. 25, pp. 2365-2401, 2004.
3. A. Maarir and E. S. Azougaghe, and B. Bouikhalene. **Automatic Change Detection Based on the Independent Component Analysis and Fuzzy C-Means Methods**, In: International Conference on Business Intelligence, Springer, Cham, pp. 178–187, 2022.
4. N. Ruhil and M. Singh, D. Mitra, A. Singh, and K. K. Singh. **Detection of changes from Satellite Images Using Fused Difference Images and Hybrid Kohonen Fuzzy C-Means Sigm**, Elsevier, Procedia Computer Science, vol. 167, pp. 431-439, 2020.
5. M. Singh, K. D. Tyagi, A. Singh, and K. K. Singh. **Detection of changes in Landsat Images using Hybrid PSO-FCM**, Elsevier, Procedia Computer Science, vol.167, pp. 423-430, 2020.
6. B. Cayla. **Traitement d'images** (partie 3 : Seuillage d'image), available at https://www.datacorner.fr/image-processing-3/
7. M. Saravanan and M. A. Santhosh Sivan. **Efficient & analysis of PCA based image change detection algorithms**, International Journal of Advanced

Technology in Engineering and Science, vol.5, pp. 235-243, 2017.

8. F. R. D. Velasco. **Thresholding using the ISODATA Clustering algorithm**, IEEE Transaction on system, Man and Cybernetics, vol.10, pp. 771-774, 1980.
9. T. Baakek. **Segmentation et classification intélligente des images cérébrales**, Ph.D. dissertation, Dept. Electrical and Electronic Eng., Abou Bekr Belkaid Univ, Tlemcen, Algeria, 2015.
10. J. Neira, Course Program, **Lesson 1: Thresholding**, University of Zaragoza.
11. B. B. Jackson and B. Bund. **Multivariate Data Analysis: An Introduction**, McGraw-Hill, 1983.
12. R. D. Jeanson. **Détection de changement sur l'occupation du sol à l'aide de la méthode ACP et la logique floue**, Memory, Antananarivo Univ., R. D. Madagascar, 2015.
13. S. Panthan, and P. Thakre. **Fuzzy Clustering Technique and PCA Based Unsupervised Change Detection Method in Multitemporal SAR Images**, Vandana Publications, International Journal of Engineering and Management Research, vol.7, pp. 345-348, 2016.
14. M. H. Kesikoglu, U. H. Atasever, and C. Ozkana. **Unsupervised change detection in satellite images using fuzzy C-means clustering and principal component analysis**, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 7, P; W2, 2013.
15. J. C. Dunn. **A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters**, Taylor & Francis, Journal of cybernetics, vol 3, pp. 32-57, 1973.
16. J. C. Bezdek, R. Ehrlich, and W. Full, **FCM: The fuzzy *c*-means clustering algorithm**, Elsevier, Computers & Geosciences, vol. 10, pp. 191-203, 1984.
17. Z. Wang, N. Zhao, W. Wang, R. Tang, and S. Li. **A Fault Diagnosis Approach for Gas Turbine Exhaust Gas Temperature Based on Fuzzy C-Means Clustering and Support Vector Machine**, Hindawi, Mathematical problems in Engieering, vol. 2015, 2015.
18. T. T. H. Pham. **Détection des changements de l'occupation des terres de la zone littorale : cas du district de Tiên Hai (Viêt-nam)**, M. S. thesis, Sherbrooke Univ., Montréal, Canada, 2005.
19. A. A. Al Analyst. **Apprentissage Supervisé et classification**, available at https://fr.linedata.com/apprentissage-supervise-et-classification
20. H. Baudhuin and A. Lambot; **LamboiseNet**, *Master Thesis about Change Detection in Satellite Imagery using Deep Learning*, available at https://www.datacorner.fr/image-processing-3/