An Efficient De noising Based Clustering Algorithm for Detecting Dead Centers and Removal of Noise in Digital Images

Lakshmana Phaneendra Maguluri, Naga Srinivasu Parvathanni, Ravikiran Karri

Abstract: As of now, several improvements have been carried out to increase the performance of previous conventional clustering algorithms for image segmentation. However, most of them tend to have met with unsatisfactory results. In order to overcome some of the drawback like dead centers and trapped centers, in this article presents a new clustering-based segmentation technique that may be able to overcome some of the drawbacks we are passing with conventional clustering algorithms. Clustering algorithms are used for segmenting Digital images however noise are introduced into images during image acquisition, due to switching, sensor temperature. They may also occur due to interference in the channel and due to atmospheric disturbances during image transmission and affecting the segmentation results Noise reduction is a pulmonary step prior to feature extraction attempts from digital images. In order to overcome this drawback, this paper presents a new clustering based segmentation technique that can be used in segmenting noise Digital images. We named this approach as De noising based Optimized K-means clustering algorithm (DOKM).where De noising is fully data driven approach. The qualitative and quantitative analyses have been performed to investigate the robustness of the OKM algorithm. And this new approach is effective to avoid dead centre and trapped centre in segmented Digital Images.

Keywords: limitations of conventional clustering algorithms; dead center problem; Salt-and-Pepper Noise; Image segmentation;

I. INTRODUCTION

Clustering is an unsupervised classification (grouping).attach label to each data points in a set, so that Object in each set can share some common trait. i.e. maintaining students (name, Roll-id, Branch, collage name).Segmenting images via clustering algorithm has been applied in various fields including in the medical field specifically in the biomedical image analysis. Clustering is iterative process.

Clustering based Image segmentation remains one of the major challenges in image analysis. Image segmentation is a process of partitioning an image into group similar regions of interest Segmenting the image into homogeneous is one of the important tasks in image analysis and is widely used in a variety of consumer electronics applications such as Object recognition geographical imaging robot vision and medical imaging.

Manuscript received on April, 2013.

LakshmanaPhaneendraMaguluri, M.Tech (Information Technology) In GITAM Institute Of Technology, GITAM University, Andhra Pradesh, India.

Nagasrinivasuparvathanni, M.Tech (Computer Science Technology) In GITAM Institute Of Technology, GITAM UNIVERSITY, and MBA in Human resource management in Andhra University Andhra Pradesh, India.

Ravikiran Karri, M.Tech (Information Technology) In GITAM Institute Of Technology, GITAM University, Vishakhapatnam,Andhra Pradesh, India.

The main feature of Image. Segmentation is to simplify the representation of an image into something that is easier to analyze. Segmentation algorithms can be broadly classified into some special categories based on segmentation techniques used such as the' thresholding, template matching, region based technique and clustering those techniques have their own limitations and advantages in terms of suitability, performance and computational cost. The thresholding is one of the simplest methods of image segmentation and is based on threshold value to turn a gray-scale image into a binary image. In industry several popular methods are used including the maximum entropy method, method, but all these methods are sensitive to noise. Template matching is a technique in digital image processing_for finding small parts of an image which match a template image .It is used to detect edges in images but however template matching become time consuming and region based technique is a approach that examines neighboring pixels of initial seed points and determines whether the pixel neighbors should be added to the region.

In generally, clustering algorithms are based upon the index of similarity or dissimilarity between pairs of data points. Many clustering algorithms are proposed with the objective to produce better segmentation. Despite the fact that the previous conventional K−means (KM) algorithm could significantly reduce or avoid the former problems. The OKM is a new approach is effective to avoid dead centre and trapped centre in Digital images. Several Clustering algorithms are used for segmenting Digital images however noise is introduced into images during image acquisition. To overcome this problem, we propose a new clustering algorithm for segmenting noise Digital images by incorporating the noise detection stage within the clustering algorithm to introduce an adaptive clustering based segmentation technique. The behavior enables to segment the noisy image properly even in the occurrence of noise without going through any filtering stage beforehand. The inherited noise detection behavior will improve the segmentation results of Digital image by only selecting noise-free pixels for the process of segmentation.

This paper is organized as follows. In Section II Limitations of Conventional Clustering Algorithm is presented. The Optimized K-means clustering Algorithm is presented in section III. The proposed adaptive clustering based segmentation technique is explained in section IV. Section V analyses the results obtained from the proposed algorithm with different noise densities by using qualitative and quantitative methods.

II. LIMITATIONS OF CONVENTIONAL CLUSTERING ALGORITHMS

In clustering area, one of the most important and widely used algorithm in computer vision as a form of image segmentation is K-means clustering algorithm. K−means (KM) algorithm is numerical, unsupervised, non-deterministic and iterative conventional method which is familiar for simple implementation. However, the k-mean clustering algorithm has many weaknesses which are as follows:

1. The number of clusters K must be determined before the algorithm is executed and it is time consuming process.
2. The algorithm is sensitive to initial conditions .Unfortunately, without proper initialization process, in some cases, the cluster centers are trapped at local minima, leading to them to lose the chance to be updated in the next iteration.
3. Poor pixel assignment could occur if the pixel with the same minimum Euclidean distance to two or more adjacent clusters. And it may be assigned to the higher variance cluster leading to dead center problems.

To overcome the aforementioned problems, the soft membership based called the Fuzzy C- Means (FCM) clustering algorithm is proposed. The FCM algorithm is an iterative unsupervised clustering algorithm. In fuzzy C-means each pixel has simultaneously belong to a degree of clusters rather than completely belongs to one cluster called member ship and distributes membership values in normalized fashion. The Fuzzy C-means clustering algorithm has some of the weaknesses which are as follows:

1. It becomes sensitive to outliers and could not homogenously segment the images.
2. However, it may also converge to local optimum location.
3. Fuzzy C-means algorithm is unsuitable for the images corrupted by impulse noises such as salt and pepper noise.

To overcome the aforementioned problems, the fitness concept has been introduced in the moving K−means (MKM) algorithm. The MKM algorithm has the capability to overcome the three basic problems algorithm which minimizes dead centers and center redundancy problems as well as indirectly reducing the effect of trapped center at local minima problems faced so far. The Moving K-means algorithm has the following drawbacks:

1. The Moving K-means algorithm is sensitive to noise.
2. For some cases of Moving k-means, the clusters or centers are not located in the middle or centroid of a group of data, leading to imprecise results.
3. To ensure all the clusters are active during the updating process the cluster with highest fitness value is forced to share its members with the cluster of lowest fitness value.

One of the obligations of MKM algorithm is overcome in AMKM algorithm, in ADMKM instead of moving the members of the center with the largest fitness Value to become a member of the center with the smallest fitness value, AMKM by simply assigning the members of the center with the largest fitness value if to the nearest cluster depending on the minimum Euclidean distance. The fuzzy concept was also introduced into the AMKM algorithm. The modified version of AMKM algorithm is adaptive fuzzy moving K−means (AFMKM) algorithm. However, they are also facing the same of the drawbacks; these algorithms failing to significantly update the lowest fitness cluster during the iteration and are also sensitive to initial parameters' values.

To overcome the aforementioned problems, enhanced moving K−means (EMKM) algorithm.In this EMKM algorithm, highest fitness cluster which is within the range will be assigned to the nearest neighboring cluster. Likewise, lowest fitness cluster obtains the members of the nearest neighboring cluster which lie outside of the range. However, the EMKM algorithm is less sensitive to the initial parameters. The fitness condition of the conventional MKM and its modified algorithms could not differentiate the dead centre and the cluster with zero variance during the process which results in a poor distribution of data. The pixel with equal distance to two or more adjacent clusters could be assigned to the higher variance cluster and the lower variance cluster will not be trained in the learning process which leading to hard membership problems. In such a situation the algorithms could fail to homogenously segment an image.

So in this article we proposed the optimized K−means (OKM) to overcome those weaknesses and homogenously segment the image. Optimized K-means algorithm is used to avoid dead centre and trapped centre at local minima which leads to producing better results and more homogenous segmented images.

III. PROPOSED OPTIMIZEDK-MEANS ALGORITHM:

In this paper, the OKM algorithm has been introduced as the modified version of the conventional K−means (KM) algorithm. The OKM algorithm fully concentrates on differentiating between the dead centers and cluster with similar intensity pixels. The pixel with the same Euclidian distance to two or more adjacent clusters is initially assigned to the dead centre and in upcoming next iteration it is assigned to the cluster with lower variance cluster, till up to no dead centre could be encountered. OKM Algorithm designed to ensure both types of centers or clusters are continuously updated in every iteration till up to the of pixels assigned to the fitting clusters. In the starting stage of the OKM algorithm, based on Euclidean distance we are going to assign all pixels to the adjacent cluster. The conflict pixels having the same Euclidean distance, the grey intensity of member are to be stored in ascending order according to their distance from the cluster with the highest fitness value and we assign the name of the array as E_r. Furthermore, the clusters are also stored in ascending order based on fitness value F_q.

As the literature study we observed one of the main problems in K-means algorithm cannot differentiate between clusters with similar intensity and the clusters without any members. In order to differentiate those clusters. In the proposed OKM algorithm, if we observe empty cluster and zero variance cluster, these clusters are sorted again in the ascending order according to the number of members in those clusters we named this array as *Hw*. Now we need to compare and map both data sets the *Er* and *Hw*. On the other hand, for the case where no dead centre is found,

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

the pixels with the same Euclidean distance to two or more nearby clusters are directly assigned to the lowest fitness value cluster.

From the mentioned description, the implementation of the OKM clustering algorithm could be outlined as follows:

1. Initialize the all the cluster centre value and α, and let iteration t = 0. (Note: α is the constant and its value within range 0 <α≤1).
2. Based on Euclidean distance, assigning all pixels to the nearby clusters. Except conflicting pixels.
3. For starting cluster we have measure Mean Square Error value from the following equation (Note: this step is only implemented for iteration t = 0).

   $MSE = \frac{1}{n}\sum_{j=1}^{k}\sum_{i\epsilon c_j}\left\|p_{i(x,y)-}c_j\right\|^2$---- (1)

   Where $pi(x, y)$ is the i–th pixel with the coordinate (x, y) to be segmented and cj is the j–th centre or cluster and n is the no of pixels in the image.
4. For all the clusters we have to calculate fitness value based on following equation and we have to find cluster with high fitness value.

   $f(c_j) = \sum_{i\in c_j}\left\|p_{j(x,y)-c_j}\right\|^2$--------- (2)
5. For the pixels having the same Euclidean distance to two or three adjacent clusters, sort the grey intensity of these pixels in the ascending order according to their distance from c_l and denote the sorting array as Er, where r = 1,2, 3… (K–1).
6. Find the empty cluster
   i) If cluster without members is found
      a. Sot all clusters in the ascending order according to their fitness values and name the sorting array asFq, where q = 1, 2, 3 …k.
      b. for all clusters obtained in step 6i)(a), sort these clusters in ascending order according to the number of pixels or members in the clusters and denote the sorting array as Hw, where w = 1, 2, 3….z (z is the total number of empty and zero variance clusters)
      c. Begin with assign the pixels with grey intensity of Eb to Hb where b = 1 and continue until the value of b in Hb equals to z or the value of b in Eb equals to (k–1).

   ii) If cluster with member is found
      a. We begin to assign the pixels with grey intensity Eb to the clusters with the lowest fitness value among their adjacent clusters, where $b = r$ = 1, 2..k–1
7. Increase iteration by $t = t + 1$, and update the centre positions, and measure the Mean Square Error value using following Eqs

   . $c_j = \frac{1}{n_j}\sum_{i\epsilon c_j} c_j(x,y)$ --------(3)

   Where, $n\ j$ is the number of pixels in the j–th cluster

   $MSE = \frac{1}{n}\sum_{j=1}^{k}\sum_{i\epsilon c_j}\left\|p_{i(x,y)-}c_j\right\|^2 - -(2)$.
8. Repeat steps 2 to 7 (except step 3) until the condition

   $\left\|MSE^{(t+1)} - MSE^{t}\right\|$<α---- (4)

   Is fulfilled. Where 0 <α≤1. The typical value of α to obtain a good segmentation performance should be close to 0.

IV.ADAPTIVE CLUSTERING BASED SEGMENTATION TECHNIQUE

We propose a new version of adaptive clustering-based segmentation technique for the digital images corrupted with salt and paper noise. For suppose assume that we are dealing with grayscale digital image where intensity is stored in an 8-bit integer, giving a possible256 gray levels. Around that interval the salt and paper noise takes on the maximum and minimum intervals.. to overcome from salt and paper noise we proposed advanced adaptive clustering based segmentation technique for digital images. We named it as adaptive clustering based segmentation techniques which consist of two stages. The first stage involves the noise detection in the image and in the second stage, we perform the clustering process. The noise free pixels will give full contribution on the clustering process, whereas for the noisy pixels, the fuzzy concept is applied to determine the degree of contribution on the clustering process. To improve the performance and robustness of the clustering algorithms we propose a technique, as De noising Based (DB) clustering is introduced to overcome the problem of noise.

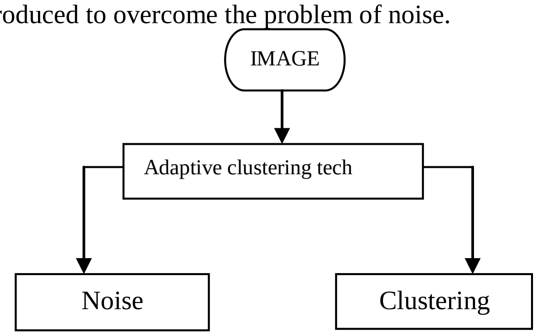

FIG: General flow of De noising Techniques:

First Stage: Noise Detection:

In salt and pepper type of noise, the noisy pixels takes either salt grey level value- 225(black) or pepper grey level value -0(white). The noise can be either positive or negative .Positive impulse appears as white (salt) points with intensity 255. Conversely, negative impulse appears as black (pepper) points with intensity 0. In this step, a binary noise mask is constructed for the noise Digital image Y. The dynamic range $[I_{max}, I_{min}]$ provide information about the noisy pixels in the image. The binary noise mask is computed from the noisy image as follows:

$$N(i,j) = \begin{Bmatrix} o, if y(i,j) = \mathrm{Im}\,ax \\ o, if y(i,j) = \mathrm{Im}\,in \\ 1, otherwise \end{Bmatrix}$$ -----(5)

Where $Y(i,j)$ is the pixel at location (i,j) with intensity Y, $N(i,j)=1$ represents the ‘noise-free’ pixel while $N(i,j)=0$, represents ‘noise ’pixels.

Second Stage: Clustering Process

In the second stage, a more powerful of clustering-based segmentation method as been proposed for removal of noise in Digital image is developed. Before assigning each pixel in the Digital image to their nearest center, check whether the pixel is ‘noise-free’ or ‘noise’ by the use of binary noise mask. If the pixel $Y(i,j)$ is noise, the absolute difference $G(i+k, j+l)$ between the neighboring pixels and the central pixel $Y(i,j)$ in 3X3 window is calculated using :

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$G(i+k, j+1) = \|y(i+k, j+1) - y(i,j)\| with\ k, 1 \in (-1,0,1) ---(6)$$

And

Y(i+k,j+1)≠y(i,j)

The Imax value of the absolute difference among the eight neighboring pixels of *Y(i,j)* in the 3x3 window will be used as the fuzzy gradient value. The fuzzy set processes the neighborhood information represented by the fuzzy gradient value to estimate a correction term which aims at cancelling the noise. Mathematically, the fuzzy set *F(i,j)* which is shown as follows:

$$f(i,j) = \begin{cases} o : o \leq \max(g(i,j)) < T_1 \\ \dfrac{\max(G(i,j)) - T_1}{T_2 - T_1} : T_1 \leq \max(G(i,j) < T_2 \\ 1 : \text{otherwise.} \end{cases} \quad \text{-------(7)}$$

Where T_1 $and T_2$ are the thresholds to perform partial correction, and set to 20 and 40 respectively as described The correction term *Y'(i,j)* for replacing the current pixel *Y(i,j)* is given by:

$y_1(i,j) = (1 - F(i,j) * Y(i,j) + F(i,j) * mij$---(8)

Where m_{ij} is the median of noise-free pixels in the 3x3 window. For each detected "noise pixel", the size of Filtering window is initialized to 3x3.

If the present filtering window does not have a min number of at least one "noise-free pixel", then the filtering window will be expanded by one pixel at each of its four sides. This procedure is repeated until minimum of one "noise–free pixel" criterion is met. The search for the noise free pixels is halted when the filtering window reached a size of 7x7.When the filtered window has reached the size of 7x7 although no "noise-free pixel" is detected, then the first four pixels in the 3x3 filtering window is used to compute m_{ij}.;

$$m_{ij} = median\{Y(i-1, j-1), Y(i, j-1), Y(i+1, j-1), Y(i-1, j)\} --- (9)$$

To increase the robustness of FOKM clustering towards noise, these corrected values (for noise pixels) are used to replace the original pixels values during the process of assigning the data to their nearest center.

$$v_i = \begin{Bmatrix} y(i,j), if N(i,j) = 1 \\ y1(i,j), if N(i,j) = 0 \end{Bmatrix} \quad \text{-----(10)}$$

By employing this concept into Fuzzy Optimized K-means Clustering Algorithm, the new proposed algorithm is called De noising Fuzzy Optimized k-means Clustering Algorithm

V. EXPERIMENTAL RESULTS

Qualitative Analysis:

In this section, we present the experimental results of the proposed algorithm on a Digital image contaminated by different levels of salt-and-pepper noise. The proposed method is performed on a Digital image that consists of a total of 38808 pixels. The segmentation results by the proposed algorithm are shown in figure 3.Figure 1 shows the Digital image corrupted with 10 %,20%, 30%, 40% and 50% density of salt-and-pepper noise. Figure 2 shows the segmentation results produced by K-means clustering algorithm for the noisy images in figure1. From the results, we can visualize that the results produced in segmenting the Digital image by the clustering algorithms (K-means, Fuzzy C-means) is are influenced by the noise, which indicates that the algorithms are less robust to noise mixture. In this paper, we use K-means clustering algorithm for comparison with the proposed algorithm. The noise contamination has highly affected the clustering process in the clustering algorithm which contributes to poor segmentation results. The proposed algorithm is seen to have successfully minimized the influence of noise from affecting the segmentation process and achieved satisfactory results.

Quantitative Analysis:

In this section, we have tabulated a quantitative Evaluation of the clustering result by using the function

$$F(I) = \sqrt{p}\frac{\sum_{i=1}^{p} e_i^2}{\sqrt{a_i}} --- (11)$$

Where I is the clustered image to be evaluated, p is the Number of regions found, a_i is the size of i-th region, $p(a_i)$is the number of regions with the area a_i and e_i is defined as the sum of Euclidean distances between the features of pixels of region I and the corresponding region in the clustered image. The values of the function for the clustered microarray image using k-means and proposed method are tabulated in table 1. The findings can therefore conclude that the proposed method serve as a better approach for segmenting noise Digital image.

Table 1: Evaluation of function F(I) for two different Algorithms:

Noise Density	K-means (2-clusters)	Proposed method (2-clusters)
10	75.8761	21.4956
20	79.2985	29.0326
30	83.5486	38.4573
40	87.3892	46.5894
50	94.3857	59.3216

VI. CONCLUSION

This paper presents new adaptive clustering based Segmentation technique for segmentation of noise corrupted Digital image. The robustness of the proposed algorithm can be observe red by using the qualitative and quantitative analysis favor, producing better results as compared to the K-means clustering algorithm through its inclusion of the noise detection stage in the clustering process. This stage can reduce the effect of noise during the clustering process. In addition the proposed algorithm has also successfully detected dead centers in Digital images. This method is a good technique for segmentation of noise Digital image without going through any filtering stage.

REFERENCES

1. S.N. Suleiman and N.A.M. Isa, "Adaptive fuzzy−K−means clustering algorithm for image segmentation", *IEEE T.Consum. Electr.* **56**, 2661–2668 (2010).
2. N.A.M. Isa, S.A. Salamah, and U.K. Ngah, "Adaptive fuzzy moving K−means clustering algorithm for image segmenta−tion", IEEE T. Consum. Electr. 55, 2145–2153 (2010).
3. F.U. Siddiqui and N.A.M. Isa, "Enhanced moving K−means(EMKM) algorithm for image segmentation", IEEE T.Consum. Electr. 57, 833–841 (2011).
4. J.−W Jeong, D.C. Shin, S.H. Do, and V.Z. Marmarelis,"Segmentation methodology for automated classification and differentiation of soft tissues in multiband images of high−resolution

ultrasonic transmission tomography", IEEE T. Med. Imaging 25, 1068–1078 (2006).

5. W.Lou, " Efficient Removal of Impulse Noise From Digital Images", IEEE Transaction on Consumer Electronics, Vol. 52, no.2, pp.523-527, 2006.
6. S.N. Sulaiman and N.A.M. Isa, Denoising−basedclusteringalgorithms for segmentation of low level salt−and−peppernoise−corrupted images", IEEE T. Consum. Electr. 56,2702–2710 (2010).
7. K.K.V.Toh, H.Ibrahim, and=M.N.Mahyuddin, " Salt andPepperNoise Detection And Reduction UsingFuzzy switching Median Filter", IEEETransactions on Consumerelectronics, vol.54, no.4, pp.1956-1961, 2008.
8. Spot Edge Detection of Microarray Images Using B1idimensional Empirical Mode Decomposition",C3IT -2012,ELSEVEIR, Procedia Technology, Vol 4: pp227-231.
9. V.P Dinesh Kumar, and T. Thomas, "Clustering of invariance improved Legendre moment descriptor for content based image retrieval," IEEE International Conference on Signal Processing, Communications and Networking, pp. 323-327, 2008.
10. F.U. Siddiqui and N.A.M. Isa, "Optimized K-means (OKM) clustering algorithm for image segmentation" in singer by OPTO−ELECTRONICS REVIEW **20**(3), 216–225 DOI: 10.2478/s11772−012−0028−8 in 2012

AUTHORS PROFILE

LakshmanaPhaneendraMaguluri has obtained his bachelor's degree from Gudlavalleru Engineering Collage Under JNTU-K in Computer Science and Engineering in 2011and Currently doing his M.Tech (Information Technology) In GITAM Institute Of Technology, GITAM UNIVERSITY, ANDHRA PRADESH, INDIA.

Nagasrinivasuparvathanni, has obtained his bachelor's degree from Sri Sarathi Institute of Engg. and Technology Under JNTU-K in Computer Science and Engineering and Currently doing his M.Tech (Computer Science Technology) In GITAM Institute Of Technology, GITAM UNIVERSITY, and MBA in Human resource management in Andhra University ANDHRA PRADESH, INDIA.

Ravikiran Karri, has obtained his bachelor's degree from Sri Vasavi Engineering College Under JNTU-K in Information Technology and Currently doing his M.Tech (Information Technology) In GITAM Institute Of Technology, GITAM UNIVERSITY, Vishakhapatnam,
ANDHRA PRADESH

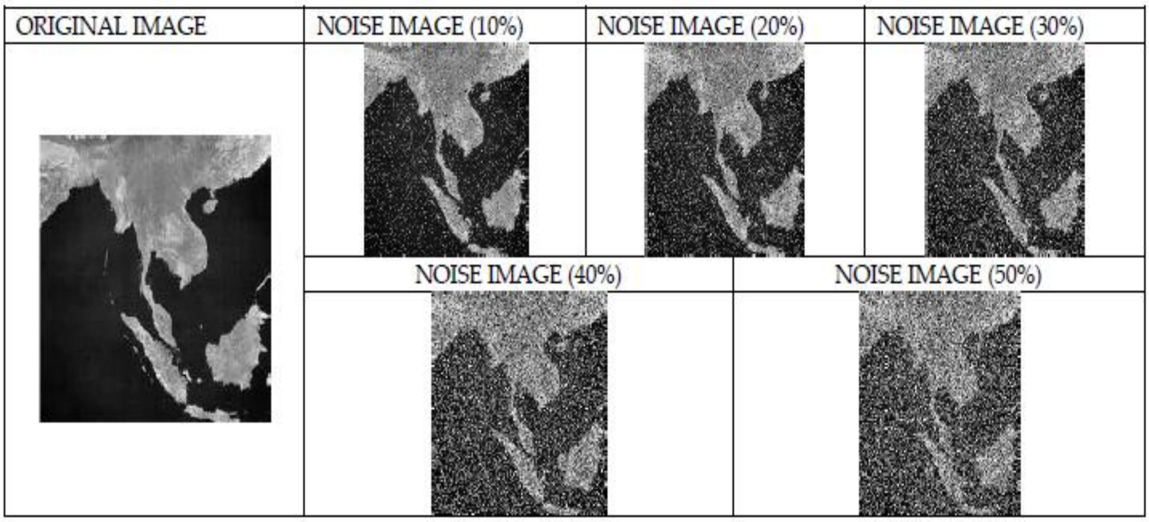

FIG: Digital Image corrupted by salt-and-pepper noise with different probabilities

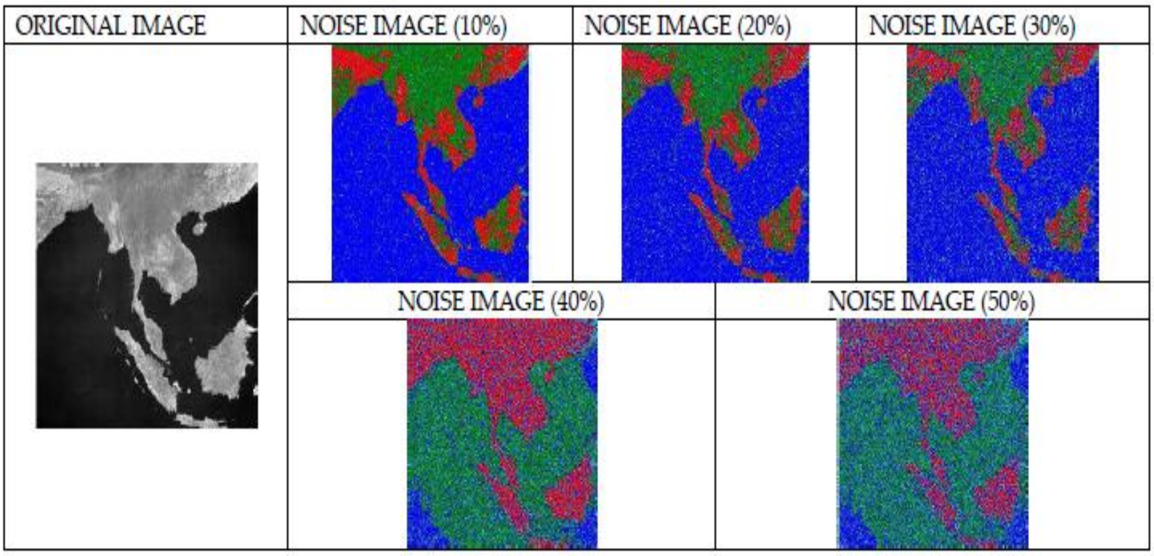

FIG: Results by K-means Clustering Algorithm on noise images (2 clusters).

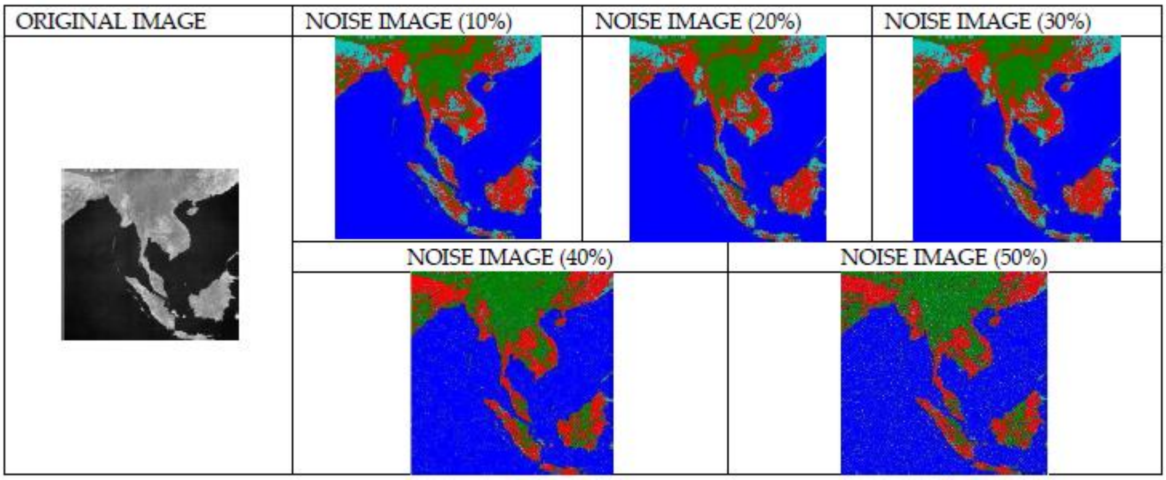

FIG: Results obtained by proposed method on noise images (2 clusters)

FIG: OKM ALGORITHM FOR DEACTING DEAD CENTERS