

# Enhanced K-mean Using Evolutionary Algorithms for Melanoma Detection and Segmentation in Skin Images

Asmaa Aljawawdeh  
Department of Computer Science  
The University of Jordan  
Amman, Jordan

Esraa Imraiziq  
Department of Business Information Technology  
The University of Jordan  
Amman, Jordan

Ayat Aljawawdeh  
Master of Science in Nursing

**Abstract**—Nowadays, Melanoma has become one of the most significant public health concerns. Malignant Melanoma (MM) is considered the most rapidly spreading type of skin cancer. In this paper, we have built models for detection, segmentation, and classification of Melanoma in skin images using evolutionary algorithms. The first step was to enhance the K-mean algorithm by using two kinds of Evolutionary Algorithms: a Genetic Algorithm and the Particle Swarm Algorithm. Then the Enhanced Algorithms and the default k-mean separately were used to do detection and segmentation of skin cancer images. Then a feature extraction step was applied on the segmented images. Finally, the classification step was done by using two predictive models. The first model was built using a Neural Network back-propagation and the other one using some threshold values for some selected features. The results showed a high accuracy using Neural Back-propagation for the Enhanced K-mean by using a Genetic Algorithm, which achieved 87.5%.

**Keywords**—Melanoma; genetic algorithm; K-mean; particle swarm optimization; classification; segmentation

## I. INTRODUCTION

Nowadays, Melanoma has become one of the most significant public health concerns. Malignant Melanoma (MM) is considered the most rapidly spreading type of skin cancer. Yet, melanoma is the most serious and fatal skin cancer in the world. Melanoma prognosis can be greatly improved and surgically cured with 100% if early detected and diagnosed [1], [2]. Dermoscopy or “Skin Surface Microscopy” is non-invasive technology has been traditionally used and contributed significantly to improved early detection and survival rates in melanoma patients. Nonetheless, many researchers had argued that clinical inspection of melanoma is the most accurate method used in diagnosing melanoma, yet new technologies such as artificial intelligence using dermoscopic images can be a doable substitute [1], [2]. Dermoscopy, which has a higher sensitivity than naked eye in detecting melanoma, consists of two main processes. These processes are: 1) optical magnification and liquid immersion with angle-of- incidence lighting; and 2) cross-polarized lighting. This technique makes skin contact area become translucent, consequently allowing visualization of subsurface structures of the skin. Interestingly, relying on dermoscopic diagnosis solely is very challenging and can lead to poor and irrelevant results.

Therefore, many researchers are recently becoming interested in developing and employing “automfacatic digital dermatoscopic image analysis” methods to enhance the diagnostic accuracy of melanoma all over the world and improve clinical outcomes. Using dermoscopic images alone, the differentiation of benign versus malignant lesions will not be an easy task. Accordingly, a further detailed analysis is desperately needed [2]. Dermoscopic Image analysis typically consists of four main steps: 1) image acquisition; 2) lesion segmentation (border detection); 3)feature extraction; and 4) classification. What is more important in image analysis is the segmentation step as the accuracy of other steps is highly dependent on it [1]. In dermoscopic image analysis, implementing the segmentation step is quite challenging for many reasons. Examples of factors that may impact the accuracy of this step are : a low contrast between the lesion and the surrounding skin; irregular lesion borders and skin texture; presence air bubbles and hair; and lastly presence of multiple colors in the lesion [1], [2]. The purpose of this paper is to detection and segmentation Melanoma in skin images using evolutionary algorithms.

This paper is organized as follows: Section 2 reviews some of the related works, Section 3 discusses the background of some common algorithms and techniques. In Sections 4 & 5, we present the methodology and the Segmentation system approach. The experiment Details and the Results obtained are discussed in Section 6 & 7. Finally, Findings of this work conclude and future work in Section 8.

## II. RELATED WORKS

In [1], the authors developed a model for the automatic segmentation of dermoscopic images by using Self-generating neural networks (SGNN) and Genetic algorithm. To optimize and stabilize the clustering result in their model, the GA is combined with SGNN. SGNN is generalized from self-generating neural tree to Self-Generating Neural Forest, then a group of optimal seed samples is selected by GA. These seeds are used by SGNN to generate an optimal clustering partitioning of the dermoscopy images. Their model delivered more accurate segmentations as compared with other automatic methods.

In [2], the authors proposed an algorithm for detection of skin lesion from digital images using Genetic Algorithm. The

lesion segmentation has been compared to the results of other algorithms. Their proposed segmentation algorithm had higher segmentation Sensitivity, Specificity and Accuracy compared to other segmentation algorithms.

In [3], the authors proposed an evolutionary strategy based segmentation algorithm and apply it to skin lesion. It could detect the lesion automatically without setting parameters manually. Their Segmentation method was flexible to adopt other fitness functions.

In [4], the authors divided the image into their respective RGB channels to obtain the spectral properties of each channel. They used the green channel according to fact that its contains more information. The authors identified skin cancer based on analysis of frequencies found in the green channel with k-law nonlinear filter. They analyzed the different types of skin cancer (basal cell carcinoma, squamous cell carcinoma and melanoma), introducing different range of classification for each type.

In [5], the Authors computerized machine learning of skin cancer using convolutional neural networks technique combined with portioning algorithm on dermoscopy Images dataset, which contains 2,032 with different disease label, the using convolutional neural networks technique achieve 72% for three-way classification and 55.4% for nine-way classification.

In [6], the authors provide an overview of automatic Dermoscope images by lesion segmentation, feature extraction then they applied a machine learning algorithm for cancer skin detection and classified to benign or malignant for early diagnosis.

In [7], the researchers implemented and designed an automated algorithm for the diagnosis of melanoma from dermoscope medical images. In order to evaluate their results, they used confusion matrix (accuracy, sensitivity, specificity, precision and AUC) also used another measurement techniques such as Jaccard and Dice.

In [8], the authors proposed a computerized model for prediction and classification the skin cancer into two types: benign and malignant melanoma using neural networks techniques. Their model depends on images preprocessing, extract features and then they applied neural networks algorithms. They achieved 84% accuracy.

In [9], authors used the PSO to search for the best centroid which have the minimum mean error and nearest distances.

In [10], The researchers in this paper used multi-approaches for MR Image segmentation for Brain melanoma detection such as: k-mean, particle swarm optimization and genetic segmentation. Their results showed that particle swarm optimization achieve better results. The authors evaluated their proposed approaches by using rand index (RI), Variation of informatics (VOI) and Global consistency error (GCE).

In [11], the authors proposed a hybrid algorithm known as dynamic particle swarm optimization and k-mean (DPSOK) to improve image segmentation quality and efficiency. DPSOK improved results compared with k-mean.

In [12], the authors reviewed particle swarm optimization (PSO) based on different approaches such as neural networks,

rough set, clustering, threshold, Genetic algorithm, wavelets and fuzzy system and applied it on several images segmentation domain. This approach approved that PSO based on different algorithm as hybrid could be more efficient.

In [13], the authors suggested an automatic method for segmentation, which contained many steps, start from reduced the color image into an intensity image, then segmented the image by using an intensity thresholding. After that the authors smoothing the segmentation using image edges. They used a double thresholding to focus on an image area. Finally, they used an elastic curve model to represent the final segmentation. Their proposed method depends on three parameters: Standard Deviation of the Gaussian Smoother, image gradients to determine threshold, and sharpness of color changes. The results showed an average error for 20 random selected images, which considered the same as four experts manually segmented the images.

### III. BACKGROUND

#### A. Genetic Algorithm (GA)

GA is efficient and robust adaptive search techniques based on the idea of natural selection. The relevant steps of GA are [1], [2]:

Step 1: Randomly generate an initial population  $G(0)$ .

Step 2: Evaluate the fitness  $f(m)$  of each individual  $m$  in the current population  $G(t)$ .

Step 3: Execute genetic operators including selection, crossover and mutation.

Step 4: Generate then next population  $G(t+1)$  using genetic operators.

Step 5: Return to Step2 until the maximum of the fitness function is obtained or reaches the last population.

#### B. Neural Network - Back-Propagation (NN)

The back-propagation neural network is the most famous architecture in the artificial neural network world. It is known as a strong function estimation for both prediction and classification problems. Back-propagation (BP) containing a training and learning technique for Multilayer perceptron neural network. Normally, the dataset is divided into two subsets: training and testing. BP works by submitting every input sample to the network where the estimated output is calculated by performing weighted sums and transfer functions [14].

#### C. Particle Swarm Optimization (PSO)

Particle swarm optimization is one of the most popular evolutionary computation technique forms inspired by nature social behavior and dynamic movements with communication of birds and fish. It was developed by Kendy and Ebrahart to achieve some objective such as searching best food source [10]. In PSO, there are a huge number of particles search for the optimum solution and communicates with other particles at the same time. Each particle updates its position according to best location, global best location, and velocity, by the following formula [10]:

$$V_i^{t+1} = V_i^t + c_1 U_1(p_i^t - p_i^t) + c_2 U_2(g_i^t - p_i^t) \quad (1)$$

p: particles position, v: path direction, c1: weight of local information, c2: weight of global information, pBest: best position of the particle, gBest: best position of the swarm, U1 and U2: random variable.

#### D. K-mean Clustering Algorithm

One of the unsupervised machine learning algorithms that used to solve clustering problems which doesn't have a label or clear class. It is a technique that is based on the gravity centroids of the segment elements represent the cluster. The K-Means algorithm calculates the distance between cluster points and the cluster centers of the objects. The main problem of this algorithm that initial centroid selected randomly [10].

#### E. Image Segmentation

Segmentation is an essential step in image analysis, it can be defined as the process of separating and dividing the region of interest from the image. The Image segmentation is the most studied and interested Area in Computer Vision. A segmentation method is usually built and designed taking into consideration the properties of a particular class of images [13]. The lesion Border can be estimated well if the segmentation was accurate and correct. Segmentation method used here depends on k-mean clustering. We have used the centers as Threshold.

### IV. METHODOLOGY

The K-mean Algorithm is a very sensitive algorithm in centers initialization step and, usually, reach the local minima, so in this paper, we have solved a k-mean problem by using two kinds of the evolutionary algorithm for centers initialization step: a Genetic algorithm and particle swarm optimization, to be applied later to the Image Segmentation step. The other Contribution in this paper is to use RGB model measurements in Feature Extraction Step. Measurements contain Statistical Features such as Entropy and Correlation. Also, Image processing Features such as smoothness and Outs Threshold. These features were used to produce and build a dataset which was used to train and test various classification Models.

Our proposed method is to build a classification system which distinguishes Malignant Melanoma from benign Melanoma. Our methodology depends on Digital Image Processing techniques and Artificial Intelligence for both segmentation and classification steps. The Pseudo code for Methodology that we have used in this paper as below:

- 1) Applying the Genetic Algorithm and PSO Algorithm on the images in order to get three suitable cluster points for the images.
- 2) The 3 cluster Points being used as initialization inputs for the K-mean Algorithm in Segmentation.
- 3) Run the k-mean algorithm to get a segmented image for the infected area in images.
- 4) Extract Features from the segmented images and uses it to continue building the dataset.

- 5) Using an artificial Techniques in order to do the classification step to determine and classify the benign or malignant infection depending on feature extractions and statistical measurements.
- 6) Do Evaluation for the used classification methods depending on three measurements: Recall, Precision, and accuracy.

#### A. Fitness Function (for GA & PSO)

The fitness function used for both GA and PSO is the same. It is shown in Table I, as below:

TABLE I. GA & PSO PARAMETERS

Parameter	Value
The length of every individual	6
Upper Bound	The size of image (X, Y)
Lower Bound	1

The length of every individual used into both PSO and GA as in Table II. It is used to represent the best solution, where v1 = x1 and v2 = y1 represent the first center, v3 = x2 and v4 = y2 represent the second center, and v5 = x3 and v6 = y3 represent the third center, those three centers used to evaluate the fitness function.

TABLE II. SOLUTION REPRESENTATION IN GA & PSO

V1	V2	V3	V4	V5	V6
----	----	----	----	----	----

The chosen index for the three centers from upper and lower boundaries was random. Every center contains three values, which is used as a reference after transfer the image into RGB model (Red-Green-Blue) as shown in Table III.

TABLE III. RGB MODEL REPRESENTATION

	R Values	G Values	B Values
Center 1	R(V1,V2)	G(V1,V2)	B(V1,V2)
Center 2	R(V3,V4)	G(V3,V4)	B(V3,V4)
Center 3	R(V5,V6)	G(V5,V6)	B(V5,V6)

Then start calculate the distance and measures the difference for every pixel in the image (R, G, B) within the three chosen centers. Let the Pixel = p (1,1), then:

$$DIS_{Center1} = \sqrt{(r_{p(1,1)} - R(V1, V2))^2 + (g_{p(1,1)} - G(V1, V2))^2 + (b_{p(1,1)} - B(V1, V2))^2} \quad (2)$$

$$DIS_{Center2} = \sqrt{(r_{p(1,1)} - R(V3, V4))^2 + (g_{p(1,1)} - G(V3, V4))^2 + (b_{p(1,1)} - B(V3, V4))^2} \quad (3)$$

$$DIS_{Center3} = \sqrt{(r_{p(1,1)} - R(V5, V6))^2 + (g_{p(1,1)} - G(V5, V6))^2 + (b_{p(1,1)} - B(V5, V6))^2} \quad (4)$$

$$Dis\_final = Minimumof(DIS_{Center1}, DIS_{Center2}, DIS_{Center3}). \quad (5)$$

After that, we create a matrix of distance for every center, which includes the pixel and its distance value that assigned to

this center cluster, this done for whole pixels starts from (1,1) and ends with the size of image (m, n), as shown in Table IV.

TABLE IV. MATRIX OF DISTANCES

	Center 1 value of Distance	Counter 1	Center 2 value of Distance	Counter 2	Center 3 value of Distance	Counter 3
Pixel (1,1)	DIS(center)	1				
Pixel (1,2)			DIS(center)	1		
Pixel (1,3)	DIS(center)	2				
Pixel (1,4)					DIS(center)	1

Then Calculate the Mean Square Error (MSE), separately for the three matrices as shown in (6) are represented in Table V.

$$MSE = \left(\frac{1}{MN}\right) \sum_{i=1}^M \sum_{j=1}^N (X_{ij} - X_{ij}^1)^2 \quad (6)$$

TABLE V. MSE CALCULATION

	Center 1 value of Distance	Counter 1	Center 2 value of Distance	Counter 2	Center 3 value of Distance	Counter 3
Pixel (1,1)	DIS(center)	1				
Pixel (1,2)			DIS(center)	1		
Pixel (1,3)	DIS(center)	2				
Pixel (1,4)					DIS(center)	1
Total	(DIS(center) <sup>2</sup> + DIS(center) <sup>2</sup> )	2	(DIS(center) <sup>2</sup> )		(DIS(center) <sup>2</sup> )	
MSE	(DIS(center) <sup>2</sup> + DIS(center) <sup>2</sup> )/2		(DIS(center) <sup>2</sup> )/1		(DIS(center) <sup>2</sup> )/1	

Then Calculate the Mean Square Error (MSE) total for the whole three matrices as shown in (7).

$$MSE(Total) = MSE1(center1) + MSE2(center2) + MSE3(center3) \quad (7)$$

So, the final fitness function or the objective function is to Minimize the Mean Square Error Total.

$$F(Y) = Minimize(MSETotal)...FitnessFunction \quad (8)$$

## V. SEGMENTATION

For this process, the k-means algorithm used to segmented the images. Three models of segmentation were done as below:

- 1) The output centers from GA was used as input centers for k means.
- 2) The output centers from PSO was used as input centers for k means.
- 3) The default k means algorithm (Initialization Randomly).

The number of clustering centers that used in k-means algorithm was 3. The number of iterations that getting stable when using k-means for both GA and PSO for the 32 images are shown in Fig. 1. It shows the difference between the two algorithms GA and PSO.

The output segmented images are shown in Fig. 2.

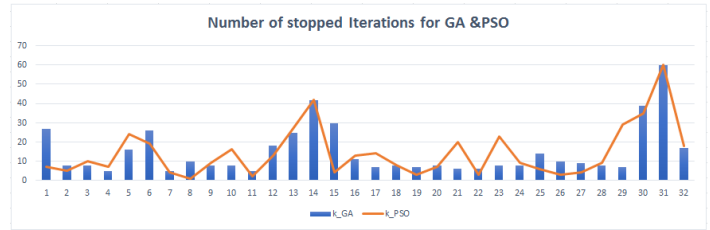


Fig. 1. K-mean-G &K-mean-PSO - The number of iterations to get stable in our experiment.

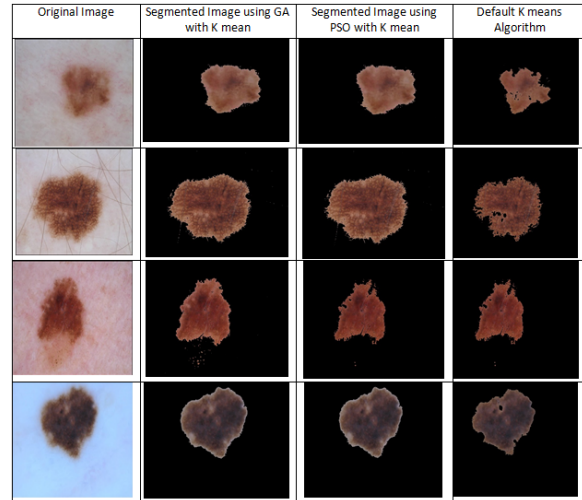


Fig. 2. Segmented images results.

## VI. EXPERIMENT

### A. Experiment Setup

Regarding the GA parameters, the population size is 20, the cross over probability Pc and the mutation probability Pm were taken to be 0.9 and 0.1 respectively, and the maximum number of generation is 50. Regarding the PSO parameters, the population size is 20, the Inertia weight and the correction factor were taken to be 1.0 and 2.0 respectively, c1 and c2 is 2.0 for both constants, and the maximum number of generation is 100.

### B. DataSet

The Data Set from the “ISIC 2017: Skin Lesion Analysis Towards Melanoma Detection” challenge dataset [15]. It was collected from the below website (International Skin Imaging Collaboration: Melanoma Project): <https://isic-archive.com/#images>.

The images were taken and saved in a folder within its data label into excel sheet (benign or malignant). The total number of images that were collecting is 32, it's the first images in their orders from image number 0 to image number 31. The resolution for images as below: (1022 \* 767) for totally 17 images and (1504 \* 1129)for the rest images (Totally 15 images). The dataset that being collected from the website manually contains only the following attributes: Location, Sex, approximation age, and the label attribute (Benign or Malignant). The remaining attributes, totally 32 attributes were

being calculated after the completed segmentation (in the feature extraction step) depending on the output images. The completed dataset which contains 36 attributes was the used dataset for classification step.

C. Feature Extraction

Feature extraction is the technique to extract the unique and useful features from a segmented image [14]. By extracting features, the classification becomes very easy and simple. Features obtained here are shown in Table VI:

TABLE VI. FEATURES

Type	Feature	Description
Manual Features (from Website)	Location	Determines the infected area of the Human Body (Abdomen, Back, Breast, Upper Limb, or Lower Limb)
	Sex	Male Or Female
	Approximation Age	The Age Range (25- 70) years old
	Benign-Malignant	Determines the infection classification (Benign or Malignant)
Statistical Features	Entropy-Original Image	Entropy of grayscale image for Original Image
	Entropy-Segmented Image	Entropy of grayscale image for Segmented Image
	Mean-Original Image	Computes the mean of the values in Original Image
	Mean-Segmented Image	Computes the mean of the values in Segmented Image
	Mean-RedMatrix-Original Image	Computes the mean of the values in Red Matrix for Original Image
	Mean-RedMatrix-Segmented Image	Computes the mean of the values in Red Matrix for Segmented Image
	Mean-GreenMatrix-Original Image	Computes the mean of the values in Green Matrix for Original Image
	Mean-GreenMatrix-Segmented Image	Computes the mean of the values in Green Matrix for Segmented Image
	Mean-BlueMatrix-Original Image	Computes the mean of the values in Blue Matrix for Original Image
	Mean-BlueMatrix-Segmented Image	Computes the mean of the values in Green Matrix for Segmented Image
	SD- Original Image	The standard deviation of the values for Original Image
	SD- Segmented Image	The standard deviation of the values for Segmented Image
	Correlation-Original Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Correlation for Original Image
	Correlation-Segmented Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Correlation for Segmented Image
	Contrast-Original Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Contrast for Original Image
	Contrast-Segmented Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Contrast for Segmented Image
	Energy-Original Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Energy for Original Image
	Energy-Segmented Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Energy for Segmented Image
	Homogeneity-Original Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Homogeneity for Original Image
	Homogeneity-Segmented Image ( Statistics from Gray-level Co-occurrence Matrix)	Calculates Homogeneity for Segmented Image
MSE	Calculates the mean-squared error (MSE) between the Original Image and Segmented Image	
2-D correlation	The correlation coefficient between the Original Image and Segmented Image	
Image Processing Features	Ots Threshold-Original Image	Computes the Global threshold of Original Image using Otsu's method
	Ots Threshold-Segmented Image	Computes the Global threshold of Segmented Image using Otsu's method
	Smoothness-Original Image	Computes the Smoothness of Original Image
	Smoothness-Segmented Image	Computes the Smoothness of Segmented Image
	Kurtosis-Original Image	Computes the Kurtosis of Original Image
	Kurtosis-Segmented Image	Computes the Kurtosis of Segmented Image
	Skewness-Original Image	Computes the Skewness of Original Image
	Skewness-Segmented Image	Computes the Skewness of Segmented Image
	Peak-SNR	Compute peak signal-to-noise ratio (PSNR) between the Original Image and Segmented Image
	SNR	Compute the signal-to-noise ratio (SNR) between the Original Image and Segmented Image

Entropy, Mean, Mean for Red matrix, Mean for Green matrix, Mean for Blue matrix, Standard Deviation, the Ots Threshold, Contrast, Correlation, Energy, Homogeneity, Smoothness, Kurtosis, and Skewness for both segmented images and original images (separately). Mean-Squared Error, Peak-SNR, SNR, and 2-D correlation between both segmented images with its original images. Adding to it the features that were collected manually from the website: Location, Sex, and approximation age. Totally with 36 features. The last feature contains a label feature for benign or malignant. Texture Analysis Using the Gray-Level Co-Occurrence Matrix (GLCM) includes the following measurements: Contrast, Correlation, Energy, and Homogeneity. Those four statistics measurements were used by authors in [14]. While Standard Deviation and Mean-Squared Error were used by authors in [1]. This step (Features Extraction) was done for the three datasets, which were created after segmentation process: GA with K-means, PSO with K-means, and default K-means.

D. Classification

Here we have built two models, the first one depends on Neural Network Back-Propagation algorithm and another one depend on some threshold values for selected features.

- 1) We have proposed a model that depends on just two selected features and its values:
  - a) The Segmented Entropy images.
  - b) The two-dimension (2D) correlation between the original image and its segmented output image.

The model is simple as below:

$$\text{The Malignant (INFECTED)} = (\text{The segmented Entropy} > 2.3) \ \& \ \& \ (2\text{-D} > -0.75)$$

Otherwise it will be a benign (NOT INFECTED) Where Entropy is a statistical measure of randomness, which can be used to characterize the texture of the image. Entropy is defined as [16]:  $-\sum(p_i \cdot \log_2(p_i))$  Where p is containing the histogram counts.

- 2) We have proposed another model that depends on Neural Network Back-Propagation algorithm.

Here we have used the Extracted Features from the step before to build this classification model. The used parameters are shown in Table VII.

TABLE VII. NEURAL NETWORKS PARAMETERS IN CLASSIFICATION MODEL

Parameter	Value
Learning Rate	0.3
The input layer neurons	36
Hidden layer neurons	(Input+output)/2 = 19
Output	2
Momentum	0.2

E. Confusion Matrix & Evaluation Metrics

		Predicted class	
		Positive	Negative
Actual class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Fig. 3. Confusion matrix.

The Confusion matrix is shown in Fig. 3. The Evaluation metrics that we have used here are the following:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$Precision = \frac{FP}{FP + TN} \tag{11}$$

- 1) Accuracy: Evaluation approach for compute the rate of correctly predicted examples of classes.
- 2) Recall: Known as true positive rate uses to compute class predicted for both classes.
- 3) Precision: Compute the rate of predicted class that was correctly classified.

## VII. RESULTS

The results for both models are shown in Fig. 4 & 5 for the three algorithms. In the predicted model using Neural Network-Back Propagation algorithm, the high accuracy (87.5 %) was for K-mean enhanced using Genetic Algorithm comparison with K-mean enhanced using PSO and with default k-mean. The same for other measurements (Precision and Recall). In the predicted model using threshold values, the high accuracy (84.375 %) was for K-mean enhanced using Genetic Algorithm comparison with K-mean enhanced using PSO and with default k mean. The same for other measurements (Precision and Recall). But it gives less measurements value, compared with the previous model in the two algorithms: K-mean enhanced using Genetic Algorithm and the default k-mean.

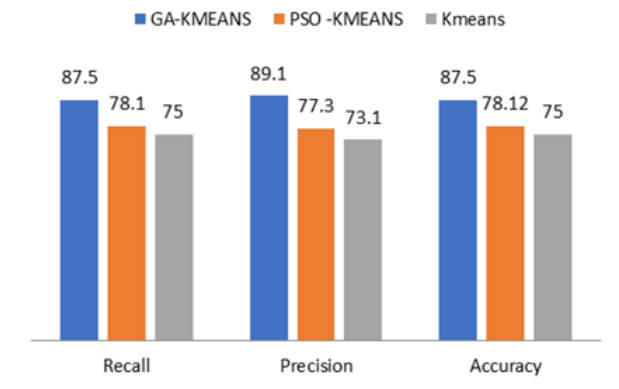


Fig. 4. Result of neural network (back propagation) model.

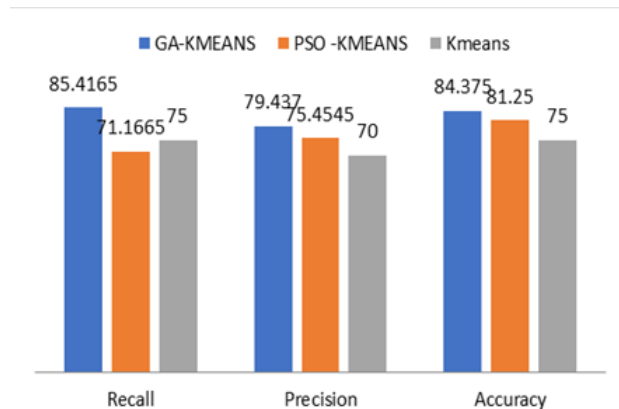


Fig. 5. Result of threshold values model.

## VIII. CONCLUSION AND FUTURE WORK

In this paper, we have built some models for detection, segmentation, and classification Melanoma in skin images using evolutionary algorithms. We have enhanced the K-mean algorithm by using two kinds of Evolutionary Algorithms; a Genetic Algorithm and Particle Swarm Algorithm. The Enhanced Algorithms and the default k-mean separately were used to do a segmentation of skin images. Then a feature extraction step was applied. After that, the classification step was done by using two predictive models. One of the predictive models was built using a Neural Network -back-propagation and the other one using some threshold values for selected features. The results showed a high accuracy using Neural Back-propagation for the Enhanced K-mean by using Genetic algorithm, which achieved 87.5%.

In future work, we can use other Evolutionary Algorithms to initialize k-mean. We can implement this work and apply it to other applications that contain and depends on images.

## ACKNOWLEDGMENT

The authors would like to thank Dr.Hossam Faris and Dr.Bashar Al-Shboul for their support with this project .

## REFERENCES

- [1] Xie, F., & Bovik, A. C. (2013). Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm. *Pattern Recognition*, 46(3), 1012-1019.
- [2] Sood, H., & Shukla, M. Segmentation of Skin Lesions from Digital Images using an Optimized Approach: Genetic Algorithm.
- [3] Situ, N., Yuan, X., Zouridakis, G., & Mullani, N. (2007, September). Automatic segmentation of skin lesion images using evolutionary strategy. In *Image Processing, 2007. IICIP 2007. IEEE International Conference on* (Vol. 6, pp. VI-277). IEEE.
- [4] Guerra-Rosas, E., lvarez-Borrego, J., & Coronel-Beltrn, . (2014, October). Diagnosis of skin cancer using image processing. In T. E. Simos, Z. Kalogiratos, & T. Monovasilis (Eds.), *AIP Conference Proceedings* (Vol. 1618, No. 1, pp. 155-158). AIP.
- [5] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
- [6] Mishra, N. K., & Celebi, M. E. (2016). An overview of melanoma detection in dermoscopy images using image processing and machine learning. *arXiv preprint arXiv:1601.07843*.
- [7] Gutman, D., Codella, N. C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., & Halpern, A. (2016). Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC). *arXiv preprint arXiv:1605.01397*.
- [8] Jaleel, J. A., & Salim, S. (2013). Implementation of ANN Classifier using MATLAB for Skin Cancer Detection.
- [9] Lam, Y. K., Tsang, P. W. M., & Leung, C. S. (2013). PSO-based K-Means clustering with enhanced cluster matching for gene expression data. *Neural Computing and Applications*, 22(7-8), 1349-1355.
- [10] Kkklahli, E., Erdomu, P., & Polat, K. (2016). Brain MRI Segmentation based on Different Clustering Algorithms. *Brain*, 155(3).
- [11] Li, H., He, H., & Wen, Y. (2015). Dynamic particle swarm optimization and K-means clustering algorithm for image segmentation. *Optik-International Journal for Light and Electron Optics*, 126(24), 4817-4822.
- [12] Kaur, A., & Singh, M. D. (2012). An overview of pso-based approaches in image segmentation. *International Journal of Engineering and Technology*, 2(8), 1349-1357.

- [13] Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., ...& Huntley, A. (1999). Segmentation of skin cancer images. *Image and Vision Computing*, 17(1), 65-74.
- [14] Aswin, R. B., Jaleel, J. A., & Salim, S. (2014, July). Hybrid genetic algorithmArtificial neural network classifier for skin cancer detection.In *Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 2014 International Conference on (pp. 1304-1309). IEEE.
- [15] Codella N, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza S, Kallou A, Liopyris K, Mishra N, Kittler H, Halpern A. "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)". arXiv: 1710.05006 [cs.CV]
- [16] Gonzalez, R.C., R.E. Woods, S.L. Eddins, *Digital Image Processing Using MATLAB*, New Jersey, Prentice Hall, 2003, Chapter 11.