International Journal of Current Trends in Engineering & Technology
ISSN: 2395-3152
Volume: 02, Issue: 05 (SEPTEMBER-OCTOBER, 2016)

Intelligent Methodologies for Melanoma Diagnosis

Munya A. Arasi[1], El-Sayed A. El-Dahshan[2], El-Sayed M. El-Horbaty[3], Abdel-Badeeh M. Salem[4]
[1,3,4]Dept. of Computer Science, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
[2]Egyptian E-Learning University (EELU), 33 Elmeshah Street, Eldoki, , El-Geiza, Egypt.

***Abstract* — the most dangerous and the deadliest kind of skin cancer is malignant melanoma. Therefore the early stage of diagnosis is an important issue. This paper presents Computer Aided Diagnosis (CAD) system for melanoma diagnosis, the dermoscopy images are taken from the Dermatology Information System database and various pre-processing and image enhancement are applied on it. The extracted features are based on 2DWavelet Transform (DWT) and use Principle Component Analysis (PCA) in order to reduce the complexity of data. These features are given as the input to the various classifiers namely: Support Vector Machine (SVM), K-nearest Neighbor (K-NN), and Artificial Neural Network (ANN) to classify the given data set among benign and malignant melanoma. The results indicated that the proposed feature extraction method is a useful method for diagnosis of malignant melanoma with high accuracy, moreover; we compared the obtained results with other approaches. The comparative results indicates that the DWT and PCA with ANN give high diagnostic performance and robust about 96. 7%.**

***Keywords* — Malignant melanoma, Computer Aided Diagnosis, Feature Extraction, Classification.**

I. INTRODUCTION

In the past few decades, malignant melanoma was reported to be the deadliest of skin cancers. This disease is curable if detected at the right time. Therefore, early diagnosis is crucial for reducing of melanoma-related deaths [1]. The main issue of skin cancer diagnosis is recognition among benign and malignant melanoma in the early stages [2]. We cannot see the structures of a skin lesion by the naked eye, therefore the dermoscopy device allows for visualizing the structures of the skin clearly. The clinical algorithms are used for diagnosis of melanoma, such as image analysis, the ABCD rule, 7-point checklist, Menzies method [3]. However, accuracy of clinical algorithms for diagnosing melanoma is an issue of concern, especially with less experienced dermatologists. In the recent years, computerized methods are used in medical area, such as cancer research, heart diseases, brain tumors, etc. Many researchers are interested in Computer Aided Diagnosis (CAD) of melanoma, which removes uncertainty from the diagnostic process. CAD of melanoma generally contains four steps namely image acquisition, pre-processing, feature extraction, and classification. Image processing techniques such as image cropping, gradient operation, morphological operation, scaling color space transformation, contrast enhancement, filters, and color quantization are used for enhancing skin images in the pre-processing step [4]. The unique features of melanoma lesion images are extracted in feature extraction step like texture features, color features, Discrete Wavelet Transform (DWT) features, Gray level Co-occurrence Matrix (GLCM) features, high-level intuitive features (HLIF) [5], and Local Binary Pattern (LBP) [6]. Various classification methods such as k-Nearest Neighbors (K-NN), Support Vector Machines (SVMs), Naïve Bayes (NB), Artificial Neural Networks (ANNs), Convolutional Neural Network (CNN) [6], Decision Tree, Logistic Model Tree (LMT) and Hidden Naive Bayes (HNB) are used in the classification step for differentiating melanoma from benign. Classification can be made accurate for the most important features. Previous studies have shown that CAD systems for melanoma diagnosis are still a complex issue, Noises like presence of hair, illumination variation, and color that affect digital images lead to less accuracy. So it is important for the development of CAD systems to overcome these drawbacks. Amelard et al. [7] applied multistage illumination modeling (MSIM), which uses Markov chain Monte Carlo approach to correct the illumination variation. Median filtering is used for removing thin hairs [8]. Celebi et al. [9] presented an iterative algorithm based on the luminance component of the color space Hue-Saturation-Value (HSV) to produce accurate features with illumination variation. Lines of blood vessels are removed by a simple median filtering operation kernel size [3x3] or smoothing procedures. Schmid [10] and Lee et al. [11] used mathematical morphology methods for hair removal. A new methodology is porposed in this work. Attention was focused on improving quality of the images and removal of illumination variations. It employs bilinear interpolation and morphological closing operations, an adaptive median filter and contrast enhcnacement to preprocess the images, DWT to extract the features from the images and Principle Component Analysis (PCA) to reduce complexity of data. Different classifiers like SVM, K-NN, and ANN are used to classify the extracted features into benign and malignant melanoma lesions. Remaining parts of the paper are organized as outlined below: An overview of the existing CAD systems is provided in Section 2. The proposed methodology is given in Section 3. Section 4 explains the results and discussion. Conclusion and future work are given in Section 5.

II. RELATED WORK

The features extraction and the classification techniques are still on development. In this section a review of related works in the area of melanoma diagnosis using CAD systems is presented.

R.Garnavi et al. (2010) [3], applied wavelet transform on various color channels of skin images. Different statistical measures have been employed on wavelet coefficients. There are four classifiers namely SVM, Random Forest (RF), Logistic Model Tree (LMT) and Hidden Naive Bayes (HNB) are used. The results show that this approach gives an accuracy of 88.24%.

N.Fassihi et al. (2011) [12], applied the variance and mean of images by using wavelet coefficients into neural network as inputs. The accuracy of this proposed method is 90%.

Y. K. Jain et al. (2012) [1], developed a skin cancer system for non-experts to diagnosis among normal and abnormal cases. Features are extracted using wavelet transform, images are decomposed into different sub-bands. The classification system is based on the application of Probabilistic Neural Network (PNN) and Clustering Classifier. Average accuracy of the system using PNN is 97.5%, whereas it is 93.5% for Clustering Classifier.

M. Elgamal et al. (2013) [13], developed an automated medical system for skin cancer. First the discrete wavelet transformation was applied on the images to get the feature vectors, PCA reduces the dimensionality of the vectors. Afterwards, those vectors were used either with feed-forward NN or K-NN. The results show the best classifier is K-NN with 100% for sensitivity, 95% for specificity, and 97.5% for accuracy. ANN based classifier has an accuracy of 84% with 2DWavelet transform. A combination of both ABCD rules and wavelet coefficients has been shown to improve the image feature classification accuracy by 60% in diagnostic system [14].

S.Choudhari et al. (2014) [15], developed a CAD system for skin cancer using GLCM with ANN, the results showed 86.66% accuracy. Amelard (2015) **[7]**, presented HLIFs feature extraction technique for charachterizing the skin lesion, and used a linear SVM for diagnosing melanoma. The results showed improved classification accuracy.

Almaraz et al. (2016) [16], developed a CAD system to extract features from images, based on ABCD rule and textural features with SVM. The results show that the performance is 75.1 %.

III. Proposed Methodology

The proposed methodology of diagnosis among benign and malignant melanoma lesions is shown in Fig 1. It uses the steps of Pre-processing, feature extraction, classification and then evaluation. First, the input images are preprocessed by applying the contrast and filtering techniques. Then, the output becomes the input to the DWT, which decomposes the image and produces approximation coefficients. PCA is, then, applied to produce low dimension data. Finally, the system classifies the image using different classification methods like k-NN, SVM, and ANN and compare the performance based on certain parameters.

3.1 Data Description

Dermoscopy images for this study were obtained from the Dermatology Information System [17] (43 malignant melanomas, 26 nevi). They were taken in varying environments. Fig 2, presents sample images of skin lesions for this dataset. These images were divided into two datasets, 60% as a learning set and 40% as the test set. All images were resized within a 256 × 256 square prior to feature extraction. The proposed system is implemented in MATLAB software version 2013 for malignant melanoma diagnosis.

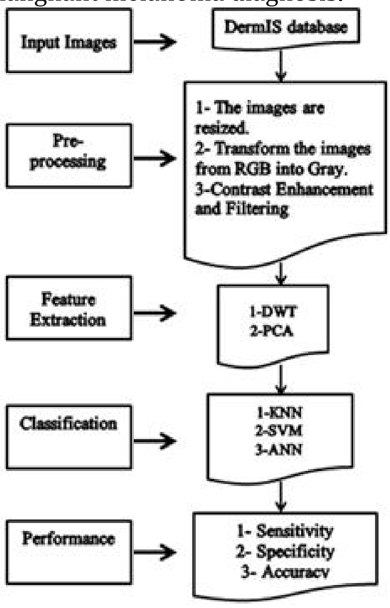

Fig 1. Proposed system

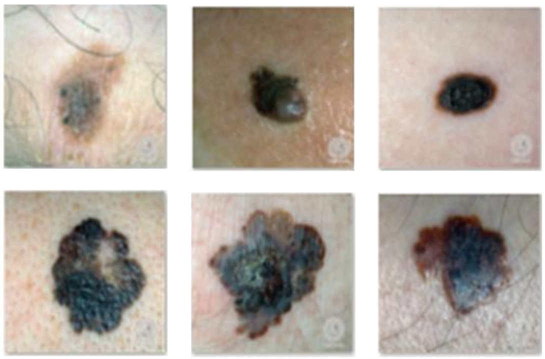

Fig 2. Images 1–3, and 4–6 show benign and malignant lesions, respectively

3.2 Pre-Processing

The pre-processing step is very important for noise removal and quality improvement of the images; Certain problems present in the image like the hair, variable lighting effects and air bubbles cause errors in classification [14]. The pre-processing performed here is shown in Fig 3, it involves four processes. First, the images are converted from RGB color space to gray level. Second, hair removal is done using bilinear interpolation and morphological closing operations.

After that, the images are smoothed using an adaptive median filter, it is a standard nonlinear image processing technique which is developed for removing the noise produced by the image capturing; also, it can be used to remove the fine hairs , it replaces each pixel with the mean value of the neighboring pixels [18]. Finally, the method enhances the image clarity by contrast enhancement; the contrast is increased by mapping the pixel values of the image to new values using gamma correction with value of 0.2.

3.3 Feature Extraction

The unique features of skin lesion images are extracted in feature extraction step, this reduces the original data and improves the accuracy of classification methods [19]. Feature Extraction technique performed here includes two steps. First, the images are transformed using 2D wavelet transform. It divides the image into approximate and three detailed images which show the basic information, vertical, horizontal and diagonal details, respectively. Second, the features are extracted from the transformed data by PCA. It reduces the dimensionality of the wavelet coefficients. This leads to more efficient and accurate classifier. DWT is a method which the wavelets transform are discretely sampled. It captures both frequency and time domain. This is an advantage over Fourier transforms [1]. Wavelets decompose an image into 4 sub-bands with low–low (LL), low–high (LH), high–low (HL), and high–high (HH) components at each scale, which correspond to approximation, horizontal, vertical and diagonal respectively. The sub-band LL is used for the next 2D DWT. The LL subband is the approximation component of the image, and the LH, HL, and HH subbands can be regarded as the detailed components of the image [4]. Fig 4, shows the decomposed wavelet. PCA is a method based feature reduction that has been used in the image analysis and compression. It is to reduce the high dimensionality of the data space to the smaller dimensionality of feature space. It is used when there is a strong correlation between observed data. It can be used as prediction, redundancy removal, feature extraction, data compression, etc. It finds the linear lower-dimensional representation of the data such that the variance [14]. So, the main goal of using PCA in our methodology is to reduce the dimensionality of the wavelet coefficients. The resulting feature vectors have a few components; means, less time and memory requirements.

3.4 Classification

Classification is the final step of the diagnostic process, where the extracted features are classified into malignant melanoma and benign [20]. We have used and compared various classifiers such as SVM, K-NN, and ANN. The SVM is a statistical learning technique, that can be worked by constructing a hyper-plane, which optimally separates the given learning vectors x_i into classes either 1 (one class) or −1 (the opposite class), with the maximal separation margin. It can be used for a kernel function $K(x,x_i)$ based on polynomial functions, radial basis functions that map the N-dimensional input vector into an L dimensional feature space. In this study, we implement sequential minimal optimization (SMO) algorithm for training a SVM classifier. The kernel Gaussian radial basis function (RBF) is used [3]. K-NN is one of the most successful techniques of object recognition, it is based on the nearest neighbor algorithm in the learning. In the training phase, it consists of simply storing all known instances and their class labels. In the testing phase, it finds the k nearest instances to the test instance and returns its class by identifying the single most frequent class label [14]. The Back propagation algorithm is used in ANN classifier; it consists of an input layer, the hidden layer and the output layer [16]. The beginning of training the weights are initialized randomly. There will be known inputs and desired output, during each epoch, according to the initial weights an activation function used, the network gives actual output of the network which is compared with the desired output. An error is generated when both are not similar, this error is propagated backwards and weights are updated so the training is stopped when the difference between the desired output and the actual output are minimized [4]. The classification results could have misclassification, the system may identify a normal lesion as malignant melanoma, or identify the malignant melanoma as normal.

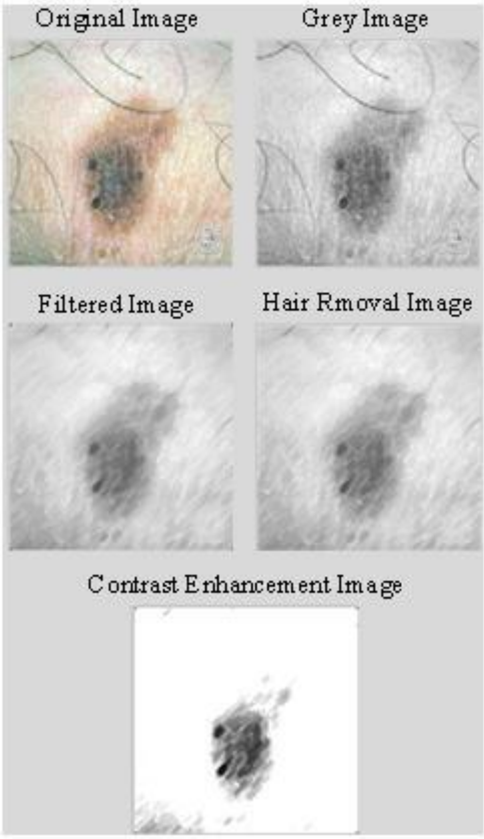

Fig 3. Pre-processing methods on dermoscopy images

$$\text{Error} = \text{Desired Output} - \text{Actual Output} \quad (1)$$

Performance of the classification is done according to sensitivity, specificity, and accuracy. Diagnosis result is given in terms of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) [9].

$$\text{Sensitivity (SE)} = \frac{TP}{TP+FN} * 100\ \% \quad (2)$$

$$\text{Specificity (SP)} = \frac{TN}{TN+FP} * 100\ \% \quad (3)$$

$$\text{Acuuracy (AC)} = \frac{(Tp+TN)}{(TP+TN+FP+FN)} * 100\ \% \quad (4)$$

They are calculated through the following equations (1, 2, 3). Where, TP refer to the cancerous image that is classified by the system as cancerous. TN refer to the non-cancerous image that is classified by the system as non-cancerous. FP refers to the image that is classified by the system as cancerous image, but in fact it is not cancerous. FN refers to the image that is classified by the system as noncancerous image, but in fact it is cancerous.

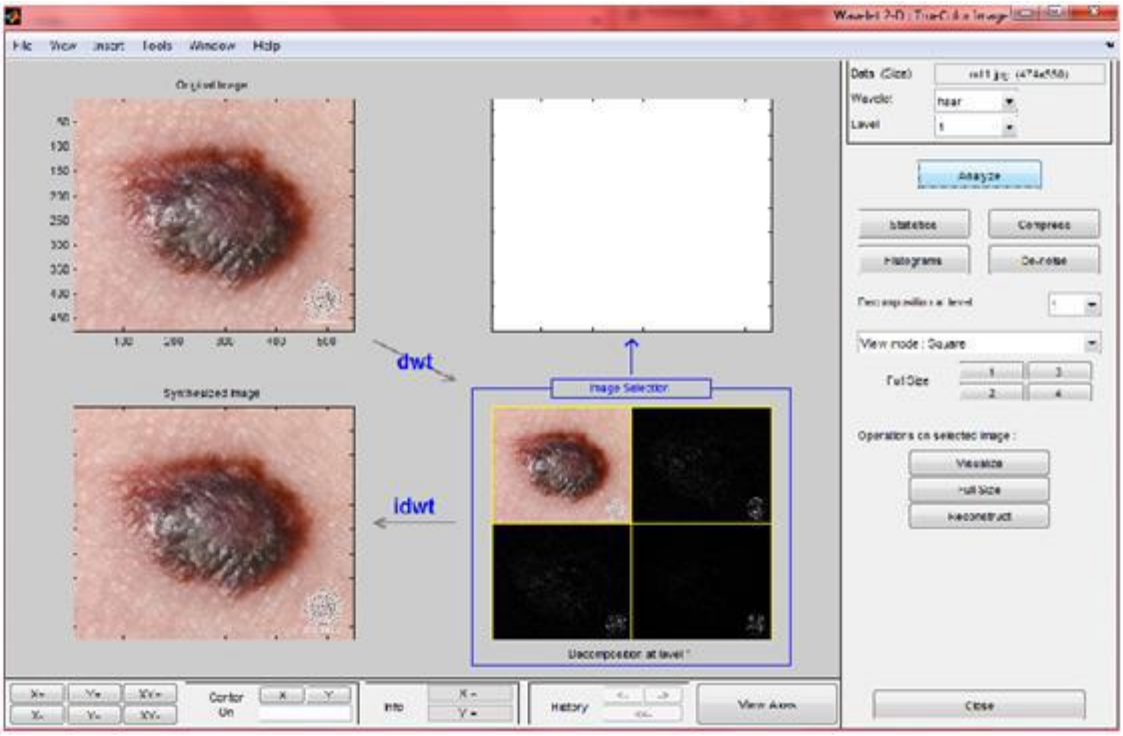

Fig 4. Feature extraction using 2D DWT in MATLAB

IV. RESULTS & DISCUSSION

The proposed work compares various classification methods for the diagnosis of malignant melanoma. The data set contains 43 images belonging to malignant melanomas and 26 images belonging to nevi. DWT is used for feature extraction and PCA is used for dimension reduction. Feature vector is calculated for each input image and these inputs fed to classify. There are different classification methods, i.e. SVM, K-NN and ANN. The ANN classifier is implemented in MATLAB software. The training datasets and desired outputs are trained. The training is stopped when the mean square error reaches a minimum value. The confusion matrix is shown in Fig. 5 which gives the errors in the classification. It displayed 1 misclassification, one non-cancerous image is classified as cancerous. Therefore, the accuracy of the system is 96.7%. The results from Table 1 and Table 2, illustrate the results of using different classifiers with DWT analysis on melanoma images. The performance is done by calculating from output results. Table 1, shows the computed True Positive, True Negative, False Positive, and False Negative values respectively on the testing data. The dataset consists of 30 images, SVM classifier erroneously classified 4 non-cancerous images as cancerous, K_NN classifier erroneously classified 3 cancerous images as non-cancerous, whereas ANN erroneously classified 1 cancerous image as non-cancerous.

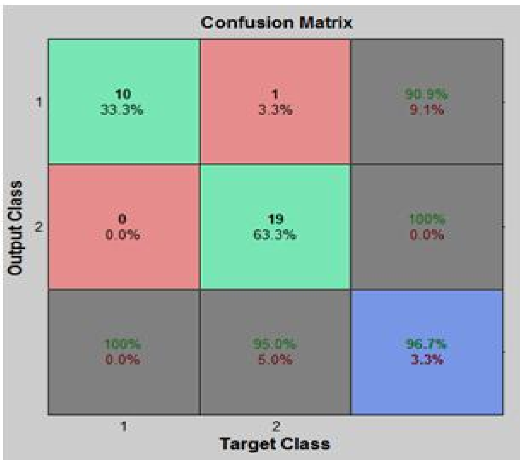

Fig. 5. Confusion Matrix

Table 2, shows the sensitivity, specificity and accuracy of the classification methods on the testing data. According to the results, ANN is more effective for detection of melanoma from benign. It is displayed the best performance among the three classification methods with an accuracy of 96.7%. It is followed by K-NN with the highest accuracy of 90% and SVM with the accuracy 86.67%. Fig. 6, shows the diagnostic results of different classifiers SVM, K_NN, and ANN in terms of sensitivity, specificity, and accuracy.

Table (1) Performance of the used classifiers

Classifiers	Test Set	True (+ve)	True (- ve)	False (+ ve)	False (- ve)
SVM	30	20	6	4	0
K-NN	30	17	10	0	3
ANN	30	19	10	0	1

Table (2) Comparison of classification results

Classifiers	Sensitivity	Specificity	Accuracy
SVM	60%	100%	86.67%
K-NN	100%	85%	90%
ANN	95 %	100%	96.7%

Table3. Comparative results of different studies

Ref.	Datasets	Methods	Accuracy
Proposed methods	DIS	DWT+PCA+SVM	86.67%
		DWT+PCA+K_NN	90%
		DWT+PCA+ANN	96.7%
[5]	DIS and DermQuest	HLIF+SVM	87.38%
[6]	DIS and DermQuest	LBP+ SVM	71.4 %
		LBP+ CNN	71.4%
[16]	DIS and DermQuest	(ABCD Rule And Textural Features) + SVM	75.1 %
[3]	Interactive Atlas of Dermoscopy	DWT+ LMT	88.24%
[4]	NA	DWT+ANN	84%
[12]	NA	DWT+ANN	90%
[19]	NA	GLCM+MLP	92%
Not Available (NA).			

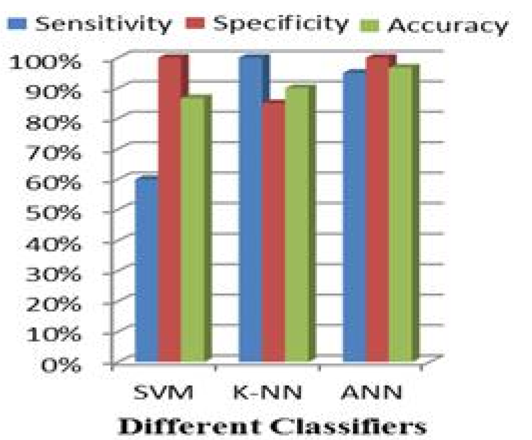

Fig. 6. Performance results of Different Classifiers

Table 3, shows comparison of the different classification methods based on accuracy for different studies. From Table 3, it can be seen that skin lesion images are extracted from Dermatology Information System (DermIS) and DermQuest in [5], [6], [16], they merge between two datasets. The results show that HLIF combined with SVM method gives better feature extraction results compared to LBP combined with SVM method. Apart from that, the best results are given by the MLP combined with GLCM [19]. DWT with ANN is better in diagnosing malignant melanoma giving the highest accuracy [12]. In conclusion our proposed method DWT+PCA+ANN give the highest accuracy, which means that the proposed approach is better in discerning among benign and malignant melanoma lesions.

V. CONCLUSION & FUTURE WORK

This paper presents a DWT and PCA hyper method for recognition among benign and malignant melanoma. The PCA has been employed on wavelet coefficients. The proposed feature extraction method has been applied on (43 malignant melanomas, 26 nevi) images. Then the three different classifiers SVM, K-NN, and ANN are used for classification. The best performance is obtained by applying ANN classifier, leading to an accuracy of 96. 7%. A comparative study has also been performed, which shows that the DWT and PCA methods is a useful method for diagnosis malignant melanoma with high accuracy. The results show that our proposed methods can discern malignant melanoma from benign successfully. We compared these results with other methods of different studies. It indicates that the proposed method gives high diagnostic performance and is robust. In future, we are looking towards developing a CAD system for diagnosis of melanoma from benign. The hybrid features the GLCM, color and 2D wavelet transform are extracted from dermoscopy images, we hope that improvement and get more accurate diagnosis. Integrating different machine learining techniques, our interst is to improve diagnosis accuracy with large dataset.

REFERENCES

[1]. Jain YK, Jain M, "Comparison between different classification methods with application to skin cancer", International Journal of Computer Applications, Vol. 53, Issue 11, pp. 18-24, 2012.
[2]. Iyatomi H, Oka H, Celebi ME, Hashimoto M, Hagiwara M, Tanaka M, Ogawa K, " An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm", Computerized Medical Imaging and Graphics, Vol. 32, Issue 7, pp. 566-579, 2008.
[3]. Garnavi R, Aldeen M, Bailey J, "Classification of melanoma lesions using wavelet-based texture analysis", In Digital Image Computing: Techniques and Applications (DICTA), pp.75-81, 2010.
[4]. Jaleel JA, Salim S, "Implementation of ANN Classifier using MATLAB for Skin Cancer Detection",

International Journal of Computer Science and Mobile Computing, pp.87-94, 2013.

[5]. Amelard R, Wong A, Clausi DA, "Extracting morphological high-level intuitive features (HLIF) for enhancing skin lesion classification", In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.4458-4461, 2012.

[6]. Karabulut, Esra M, Turgay I, "Texture analysis of Melanoma Images for Computer-aided Diagnosis", Annual Int'l Conference on Intelligent Computing, Computer Science & Information Systems (ICCSIS-16) pp. 26-29, 2016.

[7]. Amelard R, Glaister J, Wong A, Clausi DA, "High-level intuitive features (HLIFs) for intuitive skin lesion description", IEEE Transactions on Biomedical Engineering, Vol. 62, Issue 3, pp. 820-831, 2015.

[8]. Mengistu AD, Alemayehu DM, "Computer Vision for Skin Cancer Diagnosis and Recognition using RBF and SOM", International Journal of Image Processing (IJIP), Vol. 9, Issue 6, pp. 311- 319, 2015.

[9]. Celebi ME, Aslandogan YA, Stoecker WV, Iyatomi H, Hiroshi OH, Chen X, "Unsupervised border detection in dermoscopy images", Skin Research and Technology, Vol. 13, Issue 4, pp. 454–462, 2007.

[10]. Schmid P, "Segmentation of digitized dermatoscopic images by two-dimensional color clustering", IEEE Trans Med Imaging, Vol.18, Issue 2, pp. 164–171, 1999.

[11]. Mirzaalian H, Lee T K, and Hamarneh G , "Learning features for streak detection in dermoscopic color images using localized radial flux of principal intensity curvature", In: IEEE Workshop on Mathematical Methods for Biomedical Image Analysis, pp. 97–101, 2012.

[12]. Fassihi N, Shanbehzadeh J, Sarrafzadeh H, Ghasemi E, "Melanoma diagnosis by the use of wavelet analysis based on morphological operators", International MultiConference of Engineers and Computer Scientists (IMECS) I, 2011.

[13]. Elgamal M, "Automatic skin cancer images classification", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 4, Issue 3, pp. 287-294, 2013.

[14]. Ramteke NS, Jain SV, "Analysis of Skin Cancer Using Fuzzy and Wavelet Technique–Review & Proposed New Algorithm", International Journal of Engineering Trends and Technology (IJETT), Vol.4, Issue 6, pp. 2555-2566, 2013.

[15]. Choudhari S, Biday S, "Artificial Neural Network for SkinCancer Detection", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Vol. 3, Issue 5, pp. 147-153, 2014.

[16]. Almaraz-Damian JA, Ponomaryov V, Rendon-Gonzalez E, "Melanoma CADe based on ABCD Rule and Haralick Texture Features" In 2016 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW), pp. 1-4 , 2016.

[17]. Dermatology Information System. http://www. dermis. net. Accessed 26 September 2016

[18]. Abu Mahmoud M, Al-Jumaily A, Takruri M S, " Wavelet and Curvelet Analysis for Automatic Identification of Melanoma Based on Neural Network Classification", International Journal of Computer Information Systems and Industrial Management (IJCISIM) ISSN 2150-7988, Vol. 5, pp. 606-614, 2013.

[19]. Sheha MA, Mabrouk MS, Sharawy A, " Automatic detection of melanoma skin cancer using texture analysis", International Journal of Computer Applications, Vol. 42, Issue 20, pp. 22-26, 2012.

[20]. Mishra NK, Celebi ME, "An overview of melanoma detection in dermoscopy images using image processing and machine learning", arXiv preprint arXiv:1601.07843, pp. 1-15, 2016.

AUTHOR'S PROFILE

(1) Munya Abdulmajud Arasi is graduated from Aden University, Faculty of Engineering in 2003 with a B.S. in Computer Science and Engineering Department. In 2013 she received her M.S. degree in Information Technology Engineering from Aden University. Currently, she is doing her PhD at the Dept. of Computer Science, Faculty of Computer and Information Sciences Ain Shams University, Egypt. She worked on machine learning techniques for computer based decision support systems. In Yemen, Aden University, Faculty of Engineering, she holds the post of an instructor at Information Technology Engineering Department.

2) El-Sayed A. El-Dahshan He received his B.Sc. in Physics &Computer Science from Ain Shams University Cairo, Egypt. He received a post graduate diploma in electronics from Ain Shams University, Cairo Egypt. In 1990 he received his MSc. in the microwaves area from Ain Shams University Cairo, Egypt. He received his Ph.D. degree in thin films technology 1998 (cooperation system-Scientific Channel) between Claustahl-Zeller Field Teschneche Universtate -Germany and Ain Shams University –Egypt). He is currently an assistant professor of industry electronics at Faculty of Science-Ain Shams University.

(3) El-Sayed M. El-Horbaty received his **Ph.D. (1985)** in ***Computer science*** from ***London University, U.K.***, his **M.Sc. (1978)** and **B.Sc. (1974)** in ***Mathematics*** from ***Ain Shams University, Egypt***. His work experience includes 42 years as an academic in Egypt (Ain Shams University), Qatar (Qatar University), and Emirates (Emirates University, Ajman University, and ADU University). He Worked as Deputy Dean of the faculty of IT, Ajman University (2002-2008). He worked as a Vice Dean of the faculty of Computer & Information Sciences, Ain Shams University (2009-2011). He is working as Head of Computer Science Department, in faculty of Computer & Information Sciences, Ain Shams University (2012-2015)., EGYPT.

(4) Dr. Abdel-Badeh Mohamed Salem Professor of Computer Science, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt http://www.asu.edu.eg/staff/profile.php?action=show&pid=8256 Head of Artificial Intelligence and Knowledge Engineering Research Labs, http://aiasulab.ga2h.com/About/ Member of Euro Mediterranean Academy of Arts and Sciences http://www.euromediterraneanacademy.org/ He is a professor emeritus of Computer Science since September 2007 till now. He was a former Vice Dean of the Faculty of Computer and Information Sciences at Ain Shams University, Cairo-Egypt (1996-2007). He was a full professor Dr.of Computer Science at Faculty of Science, Ain Shams University from 1989 to 1996. He was a Director of Scientific Computing Center at Ain Shams University (1984-1990).In 1996 he was moved to the Faculty of Computer and Information Sciences at the same university.