Computer-Aided Model for Skin Diagnosis Using Deep Learning

Doaa A. Shoieb[1], Sherin M. Youssef[1,2], and Walid M. Aly[1,2]
[1]Department of Computer Engineering, Arab Academy for Science and Technology, Alexandria, Egypt
[2]Department of Computer Science, Arab Academy for Science and Technology, Alexandria, Egypt
Email: doaa.shoieb@gmail.com, {sherin, walid.ali}@aast.edu

***Abstract*—Melanoma is the deadliest kind of skin cancer. However, it's hard to identify melanoma during its early to mid stages by visual examination. So, there is a call for an automated model which assists in early diagnosis of skin cancer. This paper introduces an enhanced automated computer-aided model for skin diagnosis using deep learning. The model integrates an enhanced segmentation phase for locating the infected lesion of the skin and a Convolution Neural Network (CNN) is designed as a feature extractor. A classifier model has been designed based on multiclass linear Support Vector Machine (SVM) trained with CNN features extracted from the digital skin images dataset. The experimental results show an outstanding performance in the terms of sensitivity, specificity and accuracy compared with others in literature.**

***Index Terms*—computer-aided model, convolutional neural network feature, deep learning, digital skin image, and support vector machine**

I. Introduction

Melanoma is considered the deadliest form of skin cancer [1]. This is because 75% of deaths associated with melanoma skin cancer [2]. In 2013, it is estimated that 76,690 people will be diagnosed with melanoma and 9,480 people will die of melanoma in the United States [3]. According to the statistical data from the World Health Organization (WHO) [4], between 2 and 3 million non-melanoma skin cancers and 132,000 melanoma skin cancers occur globally each year. Due to recognizing the increasing in skin cancer, modern medical science is seeking to assist the dermatologists in their diagnosis without the need for special or expensive equipment. This model will help remote patient with fast and accurate diagnosis of their skin case.

The proposed model works by acquiring skin image, then apply some preprocessing techniques to improve the quality of the image and removing the skin parameters which may affect the skin images and will end with missing diagnosis. Improving the image quality by applying one of the morphological technique for thick hair removal. Then use a median filter to remove structure elements and very thin hair. Finally, the image is ready for segment the infected lesion from the skin image. Segmentation is one of the most important phases, as the resulting segmented lesion is used as an input to feature extraction and disease classification phases [5]. Existing well known feature including shape, color and texture sets combine many ad-hoc calculations and are unable to easily provide intuitive diagnostic reasoning. This paper presents the design and evaluation of a set of features for objectively detecting melanoma in an intuitive and accurate manner by applying deep feature extraction using pre-trained CNN. The last phase to classify skin infected lesion based on CNN features to train a multiclass SVM classifier.

The paper is organized as follows; Section 2 is a literature review about some of the related work done in computer-aided systems for skin diagnosis. Section 3 defines the deep learning, its functionality and how deep learning is suitable for biomedical applications. Section 4 details about the proposed model and the problem solving methodology. Section 5 presents some experimental results obtained from using the proposed approach and assesses its performance compared to the methods currently used. At last, Section 6 illustrates the conclusion of the paper by summarizing the study and by discussing possible directions for the future research.

II. Literature Review

The automated computer-aided skin diagnosis model is a three stage problem; the skin image capturing serves as input to segmentation phase to locate the infected skin lesion, then input the segmented lesion to feature extraction phase, and then the features extracted are fed into a classifier. The segmentation methods exploited by researchers include; region-based segmentation such as region growing method by automatic initialization of seed points [6]. Also texture segmentation proposed in [7] for locating skin lesion. The features extraction methods exploited by researchers include; set of High-Level Intuitive Features (HLIFs) introduced in [8].

III. Deep Learning

Deep learning is Machine learning technique that can learn useful features directly from images, text and sound [9]. Deep learning exploits many layers of non-linear information processing for supervised or unsupervised feature extraction and transformation, and for pattern analysis and classification. Deep Learning algorithms are

Manuscript received June 19, 2016; revised September 20, 2016.

doi: 10.18178/joig.4.2.122-129

largely motivated by artificial intelligence field, which attempt to emulate the human brain's ability to observe, analyze, learn, and make decisions [9]. Deep learning aims to emulate the hierarchical learning approach of the human brain to directly extract features from unsupervised data without human interference [9].

The essence of deep learning is to compute hierarchical features or representations of the data, where the higher-level features are defined from lower-level ones. Also standard machine learning techniques don't really work well when applied directly to images because they ignore the structure and composition nature of an image. Deep learning learns and extract features automatically from image [9]. An interesting characteristic of this approach is that feature extraction is also considered as a part of the learning process.

The key to the success of many biomedical *imaging* processing is the features used to characterize the input images. A major limitation of the features widely used for biomedical imaging processing (such as Haar wavelet and histogram of oriented gradient (HOG)) is their inability to extract and organize salient information from the data. So by using deep learning with its automatically extracted features can solve the limitation of features in biomedical applications.

IV. The Proposed Model Implementation

The proposed model seeks to design and develop a computer vision based system for segmentation and classification of skin lesions along with the extraction of discriminating set of features from skin lesions for efficient classification. The overview of the proposed model is shown in the Fig. 1.

There are many challenges in dealing with the skin digital images that taken from a normal camera as it contains noises such as hair and air bubbles. These noises may lead to inaccuracy of the classification and model will give the wrong predication result. In order to avoid that, images are exposed to various image processing techniques.

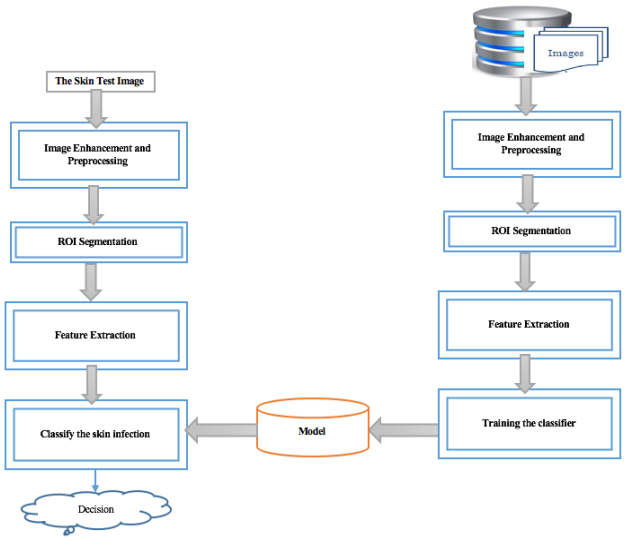

Figure 1. The architecture of the proposed model.

A. Image Enhancement and Preprocessing

Preprocessing is done to remove the background noises such as hair and air bubbles and other noises in the skin image. First use simple morphological closing operation with a disk-shaped structuring element. Based on the assumption that hair segments are thin structures, a simple morphological technique is applied; next, a hair mask is retained by using a global automatically threshold over the image intensity information. Each hair pixel from the resulted mask is replaced by an average mean of the neighbor's pixels as shown in Fig. 2.

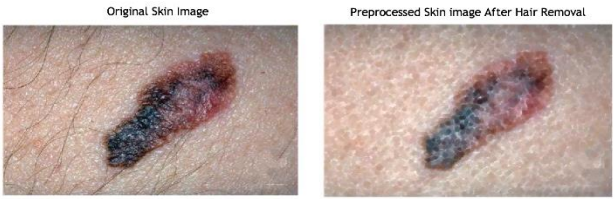

Figure 2. Original image and image after hair removal.

Noise elimination and image smoothing done by using median filtering. Median filtering is used for reducing the effect of small organizations like thin hairs and air bubbles and finally ends up with a smooth image as shown in Fig. 3.

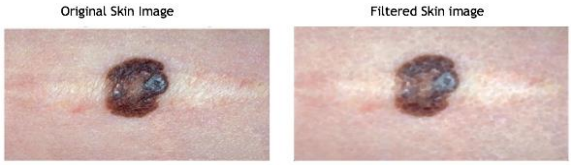

Figure 3. Original images and images after median filter.

B. Region of Interest (ROI) Segmentation

Region of Interest (ROI) segmentation is one of the most important tasks in image processing and machine vision. Image segmentation methods can be broadly classified into three categories: Edge-based methods, Region-based methods and Pixel-based direct classification methods. In the proposed system, the pixel-based direct classification method is used for segmentation. This involves three major steps: convert color space, feature extraction and clustering.

1) Convert to HSV color space

Red, Green, Blue or RGB color space is not preferred for color based detection and color analysis because of mixing of color (chrominance) and intensity (luminance) information and its non uniform characteristics. Hue, Saturation, Value or HSV is a color space that describes colors (hue or tint) in terms of their shade (saturation or amount of gray) and their brightness (value or luminance). HSV color space representing color in a way similar to how human perception. The HSV color space provides an intuitive representation and approximates the way in which human perceive. As HSV color space composed of Hue; represents color tone (e.g.: red, pink and blue), Saturation; represents the amount of color (e.g.: bright) and Value; represents the amount of light (e.g.: dark and light). Fig. 4 show the effect of converting RGB skin

images to HSV skin image and the corresponding H, S, and V channels.

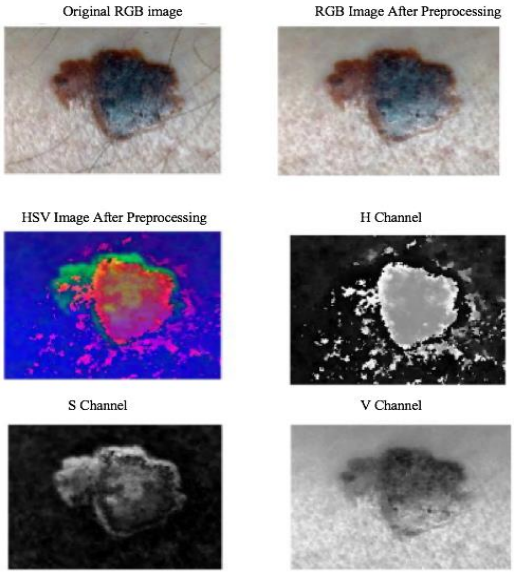

Figure 4. Original RGB image, RGB image after preprocessing, HSV image and its corresponding separate channels.

2) Texture feature of ROI region

There are two kinds of features that can be extracted from skin images color and texture. Some algorithms use color features only for segmentation. But lesion classification or clustering is difficult when relying solely on color features. Texture segmentation is a spatial arrangement of local intensity attributes which are correlated within areas of the visual scene corresponding to surface regions [10]. People are sensitive to three texture properties: repetition, directionality and complexity this indicated by Lohse [11]. The characteristic of texture repetition can be distinguished by its spatial frequency, and directionality by its orientation. So, texture segmentation requires simultaneous measurements in both the spatial and the spatial-frequency domains [11]. Gabor filters have the ability to perform the multi-resolution decomposition due to its localization both in spatial and spatial- frequency domain. A two dimensional Gabor function g(x,y) and its Fourier transform G(u,v) can be written as:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi jWx\right) \quad (1)$$

$$G(u,v) = \exp\left(\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u{}^2} + \frac{v^2}{\sigma_v{}^2}\right]\right) \quad (2)$$

where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$, and W is a constant of the highest frequency in the of the filter bank located at the center frequency of the filter bank [5]. In frequency domain this present a bandpass filter, where the bandwidth of the filter is controlled by the standard deviation of the Gaussian function and center frequency controlled by the frequency of complex sinusoid. There are number of bandpass filters with varying center frequencies, bandwidths, and orientations is bandpass filters, A Gabor filter bank having a number of controlled by the parameters of Gabor wavelets. An input image, $\xi(x,y)$ when filtered by the set of Gabor wavelets g(x,y) is given as:

$$R_{mn}(x,y) = \int \xi(x,y) g^*_{mn}(x\text{-}x_1, y\text{-}y_1) dx_1 dy_1 \quad (3)$$

where $R_{mn}(x,y)$ is the filter response at the spatial location (x,y), m=1, 2, 3, ……., M is the number of scales and n=1, 2, 3,…….., N is the number of orientations. Manjunath *et al.* [12] consider that the mean and standard deviation of the magnitude of the filter responses are used to represent the region for classification as he assumed that local image regions are spatially homogeneous [5].

$$\mu_{mn} = \iint |R_{mn}(x,y)| \, dxdy \quad (4)$$

$$\sigma_{mn} = \sqrt{\iint (|R_{mn}(x,y)| - \mu_{mn})^2 dxdy} \quad (5)$$

A feature vector is constructed using and components is known as the HT descriptor:

$$HT = [\mu_{11} \ \sigma_{11} \ \ \mu_{12} \ \sigma_{12} \ldots\ldots\ldots\ldots\ldots \ \mu_{MN} \ \sigma_{MN}] \quad (6)$$

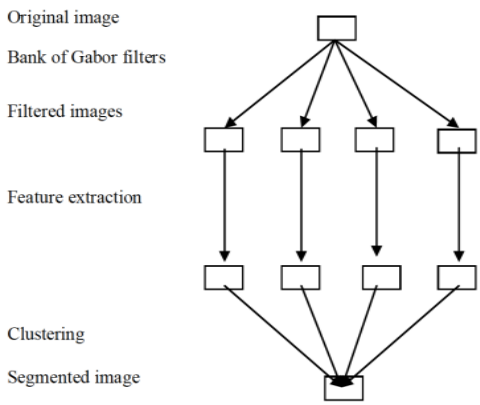

Figure 5. Texture segmentation [10].

Fig. 5 shows that the process applied to each channel in HSV color model to extract its texture features. In an image the low frequency components represent contrast and intensity, where as the high frequency components represents edges and sharp details present in the image. After extracting features from the image channels and then clustering of pixels in the feature space to produce the segmented image.

3) Region clustering using K-means

The final step in ROI segmentation is using k-means to identify similar groups of pixels based on the texture features. K-means is clustering technique that attempts to find a specific number of clusters (K), which are represented by their centroids. First, select K initial centroids theses represent the temporary means of the clusters, where K here is equal to two (normal skin lesion, infected skin lesion) based on the texture feature evaluated using a Gabor feature extractor. Then the Euclidean distance from each object to each cluster computed, and each object is assigned to the closest cluster. Followed by calculating the new centroid for each cluster. Then repeat the calculation of Euclidean distance and mapping object to appropriate cluster N times where N is number of iteration repeated to minimize error.

Pseudo code for the k-means clustering algorithm as following:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

Fig. 6 shows a simple for segmented image after applying K-means to the extracted texture feature from Gabor filter to correctly segment infected lesion from the original skin image. Which means partitioning an image into different regions of similar textures based on a specified.

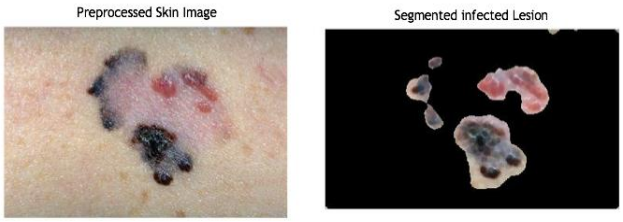

Figure 6. Original images and their corresponding segmented infected skin.

C. Feature Extraction

Feature extraction is one of the most important phases of image processing which requires extensive domain knowledge to help in classification phase. Deep Convolutional Neural Networks (CNNs), a specific type of deep learning algorithm, overcome the problem in traditional machine learning algorithms which required manual feature extraction before the classification process. CNNs not only perform classification, but they can also learn to extract features directly from raw images [10]. CNN provides the flexibility of extracting intrinsic and discriminating features from images which are the most suitable for classification [13].

CNNs consist of four types of layers are *convolution layers*, *pooling/subsampling layers*, *non-linear layers*, and *fully connected layers*. The convolution layer of CNN uses multiple learned filters to obtain multiple filter maps detecting low-level filters, and then the pooling layer combines them into higher-level features.

The CNNs used in the proposed model by using a pertained model and adapt it for the proposed system. There are several pre-trained networks that have gained popularity. Most of these have been trained on the ImageNet dataset, which has 1000 object categories and 1.2 million training images [14]. "AlexNet" is one such model and can be downloaded from MatConvNet [15]. AlexNet is CNNs that trained using two different GPUs it has 23 layers, five convolutional layers, three max-pooling layers, and three fully-connected layers. The architecture of this CNN is shown in Fig. 7.

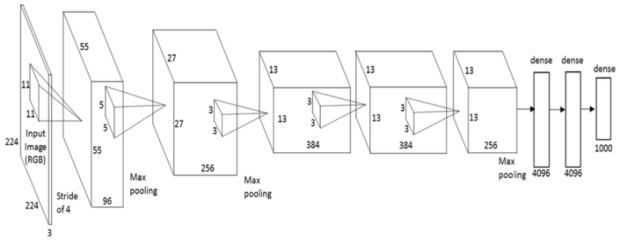

Figure 7. An illustration of the architecture of AlexNet CNN [14].

1) Convolution layers

Play the role of feature extractor but they are not hand designed. Through the training process the convolution filter kernel weights are decided on. The first convolution layer extracts low-level features like edges, lines, and corners. Higher-level convolution layers extract higher-level features [16]. First, acquire preprocessed segmented image and resize it to 227×227×3 which represent width, length and the three color channels in RGB color space. Then convolve 96 kernels with the input preprocessed image for for extraction low level features and output 96 sub images of size 55×55. Fig. 8 shows filters learned by Krizhevsky *et al.* [14]. Each of the 96 filters shown here is of size [11×11×3], and each one is shared by the 55*55 neurons in one depth slice. These kernels attempt to extract the edges and blobs. These 96 kernels divided into two sets each of them 48 kernel run on specific Graphical Processing Unit (GPU) [17]. The first set is color-agnostic and second set is color-specific.

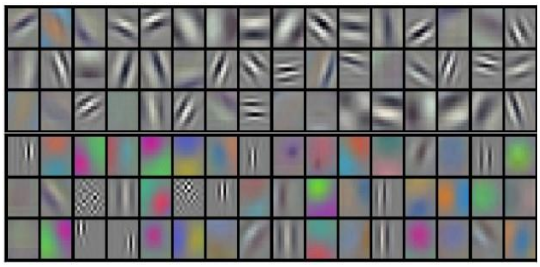

Figure 8. Convolutional kernels of size 11×11×3 learned by the first convolutional layer on the 227×227×3 input images.

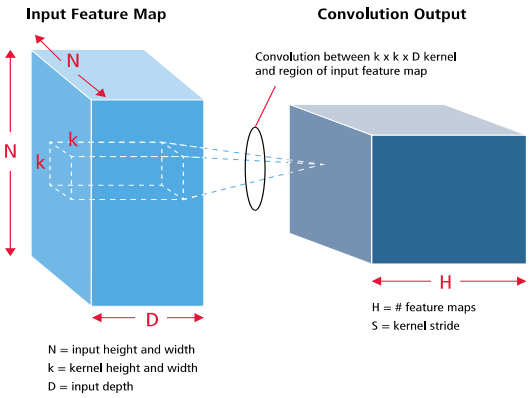

Figure 9. Pattern of 3D convolution [13].

Fig. 9 illustrates the process of 3D convolution, used in CNNs. The input image is of size 227×227× 3 is convolved with 96 kernels, each of size 11×11×3 separately. Convolution of an input segmented image

with one kernel produces one output feature, and with 96 kernels independently produces 96 features.

2) Pooling/subsampling layers

Makes the feature representations smaller and more manageable. It makes the features robust against noise and distortion. Fig. 10 shows how the pooling layer down-sampling the volume spatially by show the effect of pooling one of the 96 convolved images which generated after first convolution layer. The input volume of size [55×55×96] is pooled with filter size 2, stride 2 into an output volume of size [27×27×96], the depth is preserved. Fig. 11 is an example to illustrate the operation of max pooling as the input is of size 4x4 is divided into four non-overlapping matrices of size 2×2 by choosing the maximum value of the four values in the 2×2 matrix is the output.

3) Non-linear layers

CNNs in particular rely on a non-linear "trigger" function to signal distinct identification of likely features on each hidden layer. CNNs may use a variety of specific functions such as rectified linear units (ReLUs) and continuous trigger (non-linear) functions. A ReLU implements the function y = max(x,0), so the input and output sizes of this layer are the same. It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields of the convolution layer [18]. ReLU functionality is illustrated in Fig. 12 with its transfer function plotted above the arrow.

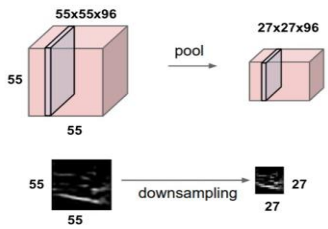

Figure 10. Effect of the pooling layer on one of the convolved images.

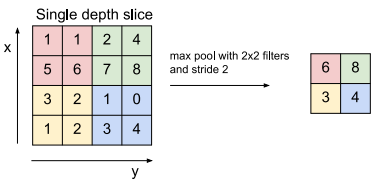

Figure 11. Example of max pooling.

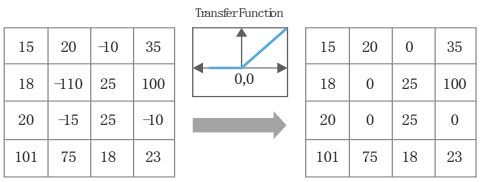

Figure 12. Representation of ReLU functionality [18].

4) Fully connected layers

In a fully connected layer, all the elements of all the features of the previous layer get used in the calculation of each element of each output feature. Fig. 13 explains how the features in the fully connected layer calculated. As shown in Fig. 13 the fully connected layer Layer L-1 has two features, each of which is 2x2, Layer L has two features, each having a single element. Each of this single element calculated by defined relation between neurons weight and values in layer L-1.

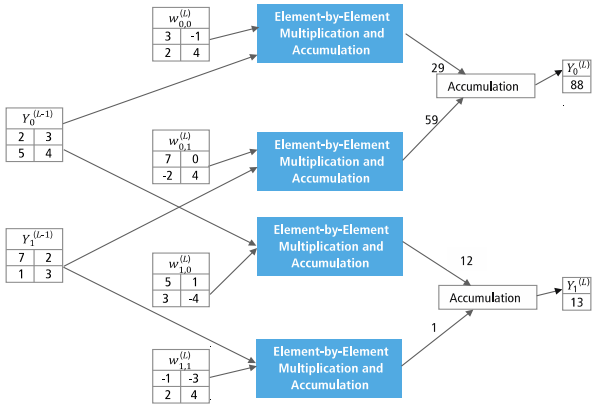

Figure 13. Processing of a fully connected layer [18].

Selecting which of deeper layers to choose for feature extraction is a design specific. There are three fully connected layers (fc6, fc7 and fc8) which combine features from all previous layers. Last layer (fc8) for classification purpose. Typically, fc7 is a good place for extracting the features. These features are 4096 calculating from combining the lower level features with connected weights.

D. The Classifier Model Using Support Vector Machine (SVM)

Once features are extracted, a classifier can be trained to classify a test sample as a member of one of the known classes. In this work the images have been classified using linear Support Vector Machine (SVM). SVM is a supervised learning technique that seeks an optimal hyper-plane to separate two classes of samples [19]. SVM performs classification by finding the hyper plane that maximizes the margin between the two classes as shown in Fig. 14 Support vectors are the data points that lie closest to the decision surface. They are the most difficult to classify. The basic idea of support vector machine is to find an optimal hyper-plane for linearly separable patterns.

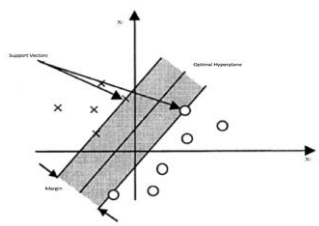

Figure 14. SVM classification including a hyper plane that maximizes the separating margin between two classes [19].

SVM have some remarkable characteristics such as, its ability to learn independently of the dimensionality of the feature [19]. This classifier has been chosen due to its robustness, simplicity and does not tend to over fit training data.

V. Experimental Analysis

To evaluate the performance of the proposed model, we conduct a set of experiments by comparing the proposed model to several state-of-the-art skin diagnosis models.

A. Dataset Description

We collect 337 images of skin lesion with ground truth, which were obtained using standard consumer- grade cameras in varying and unconstrained environmental conditions. These images were extracted from two different databases. First dataset from the publicly available online databases Dermatology Information System [20]. The dataset comprises 69 images 43 of them represent melanoma and 26 for not melanoma skin cancer. In Fig. 15 some sample images from the dermis dataset.

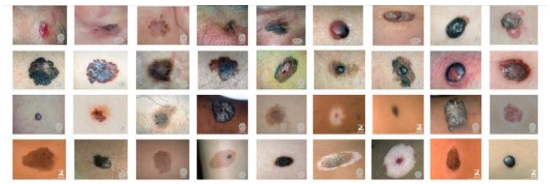

Figure 15. Samples of DermIS dataset 1st and 2nd rows images for melanoma skin cancer, 3rd and 4th rows images for Non Melanoma skin cancer.

Second dataset is DermQuest [21]. The dataset comprises 134 images 76 of them represent melanoma and 58 for not melanoma skin cancer. In Fig. 16 some sample images from the dermis dataset.

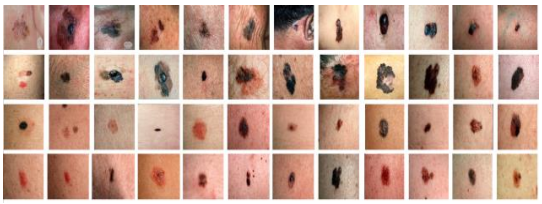

Figure 16. Samples of DermQuest dataset 1st and 2nd rows images for melanoma skin cancer, 3rd and 4th rows images for non melanoma skin cancer.

Third dataset for validation the proposed system for multi-class disease. This dataset obtained from DermNet Skin Disease Atlas website [22] which has 27 disease classes. Four classes of skin disease have been selected for system validation. These classes are Melanoma, Basal Cell Carcinoma, Eczema, and Impetigo. The dataset comprises 134 images 72 of them represent melanoma, 64 for Basal Cell Carcinoma, 74 Eczema and 31 for Impetigo skin disease. In Fig. 17 some sample images from the DermNet dataset.

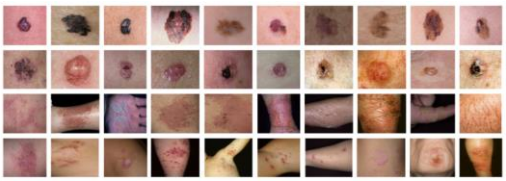

Figure 17. Samples of DermNet dataset, 1st row images for melanoma skin disease, 2nd rows images for basal cell carcinoma skin disease, 3rd row images for eczema skin disease and 4th rows images for impetigo skin disease.

B. Experimental Results

Experiments have been carried out to validate the efficiency of the proposed model. The experiments were carried out on a core i5, 2.3GHz processor with 8GB RAM using MATLAB 9.0. Comparisons with other models have been conducted. To measure the performance of classification, different sets of experiments have been conducted. The performance of the classification is evaluated in terms of classification sensitivity, specificity and accuracy from the confusion matrix of classification. The measures are computed by using the equations described below with the following conventions.TP (True Positive) is Positive samples classified as positive. TN (True Negative) is Negative samples classified as negative. FP (False Positive) is Negative samples classified as positive. FN (False Negative) is Positive samples classified as negative.

Sensitivity: Is the ability of the test to correctly identify those patients with the disease as given by the equation below.

$$Sensitivity = \frac{TP}{TP+FN} \tag{7}$$

Specificity: Is the ability of the test to correctly identify those patients without the disease classified as given by the equation below.

$$Specificity = \frac{TN}{TN+FP} \tag{8}$$

Accuracy: It is the total number of samples correctly classified to the total number of samples classified as given by the equation below.

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

The proposed system was tested on three different datasets and the performance measures evaluated on each of them. Table I and Table II show the average sensitivity, specificity and accuracy across different datasets of images or for just the melanoma or non-melanoma images. Table III shows the proposed system performance when used for diagnosis multiclass skin disease instead of only binary classification.

TABLE I. Classification Results of Melanoma and Non Melanoma Skin Lesion Obtained from DermIS Dataset

Skin disease	Accuracy	Specificity	Sensitivity
Melanoma	93.75%	100%	87.5%
Non Melanoma	93.75%	87.5%	100%

Table I Shows that the proposed system able to correctly identify all patients with non melanoma disease also can diagnosis 87.5 % of patients with melanoma correctly. Finally, the overall accuracy of the correctly diagnosis patients is 93.75%.

TABLE II. Classification Results of Melanoma and Non Melanoma Skin Lesion Obtained from DemQuest Dataset

Skin disease	Accuracy	Specificity	Sensitivity
Melanoma	94.12%	94.12%	94.12%
Non Melanoma	94.12%	94.12%	94.12%

Table II Shows that the proposed system able to correctly identify 94.12 % patients with non melanoma disease also can diagnosis 94.12 % of patients with melanoma correctly. Finally, the overall accuracy of the correctly diagnosis patients is 94.12 %.

TABLE III. CLASSIFICATION RESULTS OF MELANOMA BASAL CELL CARCINOMA, ECZEMA AND IMPETIGO SKIN LESION DISEASE OBTAINED FROM DERMNET SKIN DISEASE ATLAS WEBSITE

Skin disease	Accuracy	Specificity	Sensitivity
Melanoma	98.04%	100%	88.89%
Basal Cell Carcinoma	96.15%	96%	88.89%
Eczema	94.12%	100%	77.78%
Impetigo	91.42%	88.46%	100%

Table III Shows that the proposed system can produce high accuracy also when applied to multiclass skin disease. The results show that the proposed system correctly identifies all patients without Melanoma and Eczema diseases. Also, it can correctly identify all patients with the Impetigo disease non melanoma disease also can diagnose. Finally, the melanoma skin disease has the highest accuracy compered with other disease.

Table IV and Table V show the performance of the proposed system in comparison with other state of art skin diagnosis systems.

Table IV shows that the proposed system has higher performance in terms of accuracy, specificity and sensitivity than Jeffrey Glasister's system which proposed an enhanced system by extracting high level of features.

TABLE IV. COMPARISON BETWEEN PROPOSED MODEL AND JEFFREY GLAISTER [7] IN TERMS OF ACCURACY, SPECIFICITY AND SENSITIVITY. EXPERIMENTAL CONDUCTED ON DERMQUEST DATASET

Performance Measures	Proposed system		Jeffrey Glaister, 2014 [5]	
	Melanoma	Non Melanoma	Melanoma	Non Melanoma
Accuracy	94.12%	94.12%	83.3%	84.1%
Specificity	94.12%	94.12%	83.6%	82.9%
Sensitivity	94.12%	94.12%	83.1%	85.7%

TABLE V. COMPARISON BETWEEN PROPOSED MODEL AND ROBERT AMELARD [8] IN TERMS OF ACCURACY, SPECIFICITY AND SENSITIVITY. EXPERIMENTAL CONDUCTED ON BOTH DATASETS TOGETHER

Performance Measures	Proposed system		Robert Amelard, 2015[10]	
	Melanoma	Non Melanoma	Melanoma	Non Melanoma
Accuracy	94%	94%	83.59%	83.59%
Specificity	88%	100%	73.45%	73.45%
Sensitivity	100%	88%	92.52%	92.52%

Table V shows that the proposed system has higher performance in terms of accuracy, specificity and sensitivity than Robert Amelard's model which proposed an enhanced system by proposing an enhanced texture-based segmentation then extracting high level of features from the segmented lesion.

VI. CONCLUSION AND FUTURE WORK

In this paper, an enhanced robust model has been proposed for skin diagnosis using skin lesion image obtained from a standard camera. Using CNN as representative and discriminative feature extractor allows the model to represent its diagnosis in the effective solution for automated recognition of skin diseases. On one hand, this would be useful for dermatologists to reduce diagnostic errors and help remote patients to diagnosis and their skin lesion at a reduced cost while reducing over dependence on medical experts. Experimental results indicate that CNN features easily outperform hand-engineered features in terms of better sensitivity, specificity, and accuracy. The results of extensive experimental trials revealed that the proposed model produced a significant improvement of around 11% in diagnostic accuracy compared to the best of the other state of the art computer aided skin diagnosis.

As a further development to the model we are targeting to expand the multi-platform capability through mobile support. Experimenting with datasets of collection from hospitals. Enable doctors from given feedback about classification results.

REFERENCES

[1] Public Health Agency of Canada. [Online]. Available: www.publichealth.gc.ca

[2] A. F. Jerant, J. T. Johnson, C. D. Sheridan, and T. J. Caffrey, "Early detection and treatment of skin cancer," *American Family Physician*, vol. 62, no. 2, pp. 1-6, July 2000.

[3] SEER cancer statistics review, 1975-2012. [Online]. Available: http://seer.cancer.gov/statfacts/html/melan.html

[4] World Health Organization. [Online]. Available: http://www.who.int/en/

[5] F. Riaz, A. Hassan, S. Rehman, and U. Qamar, "Texture classification using rotation- and scale-invariant Gabor texture features," *IEEE Signal Processing Letters*, vol. 20, no. 6, pp. 22-26, June 2013.

[6] R. Sumithra, M. Suhilb, and D. S. Guruc, "Segmentation and classification of skin lesions for disease diagnosis," in *Proc. International Conference on Advanced Computing Technologies and Applications*, 2015, pp. 76-85.

[7] J. Glaister, A. Wong, and D. A. Clausi, "Segmentation of skin lesions from digital images using joint statistical texture distinctiveness," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 4, pp. 1220-1230, April 2014.

[8] R. Amelard, J. Glaister, A. Wong, and D. A. Clausi, "High-level intuitive features (hlifs) for intuitive skin lesion description," *IEEE Transactions on Biomedical Engineering*, vol. 62, pp. 820-831, February 2015.

[9] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *Journal of Big Data*, February 2015.

[10] M. Sivalingamaiah and B. D. V. Reddy, "Texture segmentation using multichannel Gabor filtering," *IOSR Journal of Electronics and Communication Engineering*, vol. 2, pp. 22-26, 2012.

[11] A. R. Rao and G. L. Lohse, "Identifying high level features of texture perception," *CVGIP: Graphical Models and Image Processing*, vol. 553, pp. 5218-5233, 1993.

[12] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 837-842,1996.

[13] M. D. Zeiler and R. Fergus, *Visualizing and Understanding Convolutional Networks*, Springer International Publishing, 2014, pp. 818-833.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, no. 2, 2012.
[15] A. Vedaldi and K. Lenc. "MatConvNet-Convolutional neural networks for MATLAB," *arXiv Preprint*, pp. 1412-4564, 2014.
[16] M. Srinivas, D. Roy, and C. K. Mohan, "Discriminative feature extraction from X-ray images using deep convolutional neural networks," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2016, pp. 917-921.
[17] K. Ovtcharov, O. Ruwase, J. Kim, J. Fowers, K. Strauss, and E. S. Chung, "Accelerating deep convolutional networks using specialized hardware," *Microsoft Research*, Feb. 22, 2015.
[18] S. Hijazi, R. Kumar, and C. Rowen, *Using Convolutional Neural Networks for Image Recognition*, Cadence Design Systems Inc., 2015.
[19] C. Cortes and V. Vapnik, "Support-Vector networks," *Machine Learning*, vol. 20, pp. 273-297, November 1995.
[20] Dermatology information system. (2012). [Online]. Available: http://www.dermis.net
[21] DermQuest. [Online]. Available: http://www.dermquest.com
[22] DermNet. [Online]. Available: http://www.dermnet.com

Doaa A. Shoieb was born in Alexandria, Egypt, in 1992. She received the B.S. degree from the Arab Academy for Science and Technology (AAST), Alexandria, in computer engineering. She is currently pursuing the M.S. degree with the Department of Computer Engineering, AAST. Her research interests include image processing, pattern recognition, Artificial intelligence, biomedical engineering, and mobile application development.

Sherin M. Youssef is currently a professor & head of computer Engineering Department, Arab Academy for Science & Technology (AAST). She received her PhD degree from University of Nottingham (UK, 2004) in Intelligent distributed swarm intelligence and optimization systems. She received her Master degree (MSc) from University of Alexandria (Egypt, 1995) in machine intelligence. She received her BSc from the department of computer Science and Automatic control, University of Alexandria, Egypt.
She was an IEEE senior member and a member in the IEEE/ACM technical committee. Her main research interests lie in the areas of Artificial intelligence, Intelligent agent based systems, digital signal processing, image processing, Video surveillance systems, mobile applications, biometrics and biomedical engineering.

Walid M. Aly has acquired his Ph.D. on 2004 from faculty of Engineering, Electrical Engineering, Alexandria University, Egypt. He is a Computer Science professor at the College of Computing and Information Technology (CCIT), Arab Academy for Science, Technology & Maritime Transport (AASTMT). His research interests include intelligent systems, soft computing, modeling, simulation and machine learning. He is mainly teaching courses related to Programming, Artificial Intelligence and Mobile Computing.