DETECTION OF LESION USING SVM

Nivetha. S

Assistant professor, Department of CSE Dr. Sivanthi Aditanar College of Engineering, Tiruchendur, India

ABSTRACT

In recent days, skin cancer is seen as one of the most Hazardous form of the cancer found in Humans. Skin Cancer is a malignant tumor that grows in the skin cells. It can be affected mostly by the reason of skin burn caused by sunlight. Early detection and treatment of Skin cancer can significantly improve patient outcome. Automatic detection is one of the most challenging research areas that can be used for early detection of such vital cancer. A person's in which they have inadequate amount of melanoma will be exposed to the risk of sun burns and the ultra violet rays from the sun will be affected that body. Malignant melanomas is a type of melanoma that has irregular borders, color variations so analyze the shape, color and texture of the skin lesion is important for the early detection. It can have the components of an automated image analysis module, which contains image acquisition, hair detection and exclusion, lesion segmentation, feature extraction and finally classification. Finally the result show that the system is efficient achieving classification of the lesion as either melanoma or Non melanoma causes.

KEYWORDS

Image processing, classification, feature selection, support vector machine.

1. INTRODUCTION

Digital image processing deals with manipulation of digital images through a digital computer. Digital Image Processing is a subfield of signals and systems but focus particularly on images. DIP focuses on developing a computer system that is able to perform processing on an image. The input of that system is a digital image and the system process that image using efficient algorithms, and gives an image as an output. The most common example is Adobe Photoshop. Digital Image Processing is one of the widely used applications for processing digital images. Image processing basically includes the following three steps:

1. Importing the image with optical scanner or by digital photography.
2. Analyzing and manipulating the image which includes data compression and image enhancementand spotting patterns that are not to human eyes like satellite photographs.
3. Output is the last stage in which result can be altered image or report that is based on imageanalysis.

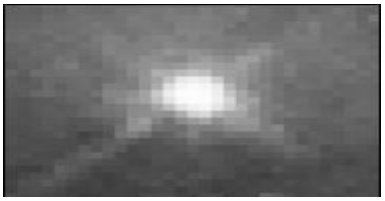

Figure 1: A Sample Image

An array or a matrix of pixels arranged in columns and rows. In the sample image figure1, there may be thousands of pixels that together make up this image. We will zoom that image to the extent that we are able to see some pixels division. It is shown in the image below.

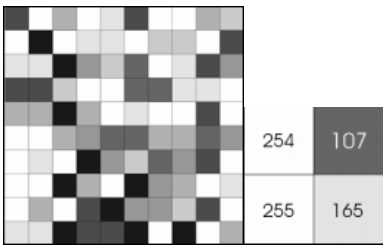

Figure 2: Pixelized Image

Each pixel figure 2 has a value from 0 (black) to 255 (white). The possible range of the pixel values depend on the color depth of the image, here 8 bit = 256 tones or gray scales.

The main objective of the proposed work is to automatically analysis the skin lesions with the help of image processing techniques. The manual segmentation as well as manual diagnosis of skin cancer lesions are very tedious process. The manual error will be unavoidable in those manual work. In order to reduce the manual error as well as manual diagnosis, the project is motivated to provide automatic solution for the skin lesion segmentation & classification.

2. RELATED WORK

Teck Yan Tan, Dermoscopic images are detected using an intelligent agent-based or robotic system toconduct long-term automatic health monitoring and robust efficient disease diagnosis as autonomous e- Careers in real-world applications [1]. In this research, we aim to deal with such challenges by presenting an intelligent decision support system for skin lesion recognition as the initial step, which could be embedded into an intelligent service robot for health monitoring in home environments to promote early diagnosis. The system is developed to identify benign and malignant skin lesions using multiple steps, including pre-processing such as noise removal, segmentation, and feature extraction from lesion regions, feature selection and classification. After extracting thousands of raw shape, colour and texture features from the lesion areas, a Genetic Algorithm (GA) is used to identify the most discriminating significant feature subsets for healthy and cancerous cases. A Support Vector Machine classifier has been employed to perform benign and malignant lesion recognition. Evaluated with 1300 images from the Dermo fit dermoscopy image database, the empirical results indicate that our approach achieves superior performance in comparison to other related research.

Eliezer Flore *set al.*, Pre-screening systems for the diagnosis of melanocytic skin lesions depend of the proper segmentation of the image region affected by the lesion. This paper proposes a feature learning scheme that finds relevant features for skin lesion image segmentation. This work introduces a new unsupervised dictionary learning method, namely Unsupervised Information-Theoretic Dictionary Learning (UITDL), and discusses how it can be applied in the segmentation of skin lesions in macroscopic images [2]. The UITDL approach is adaptive and tends to be robust to outliers in the trainingdata, and consists of two main stages. In the first stage, a textural variation image is used to construct an initial feature dictionary and an initial sparse representation via Non-Negative Matrix Factorization (NMF). In the second stage, the feature dictionary is optimized by selecting adaptively the number of dictionary atoms. The greedy approach used for dictionary optimization is quite efficient and flexible enough to be applied to

other dictionary learning problems. Furthermore, the proposed method can be easily extended for other image segmentation problems. The experimental results suggest that the proposed approach potentially can provide more accurate skin lesion segmentation results than comparable state-of-the-art methods [3]. The proposed segmentation method could help to improve the performance of pre-screening systems for melanocytic skin lesions, which can affect positively the qualityof the early diagnosis provided to skin lesion patients.

3. THE PROPOSED SYSTEM

Proposed System is an intelligent decision support system for benign and malignant skin lesions classification. The system includes the following key stages, i.e. pre-processing, skin lesion segmentation, feature extraction and classification. Figure 3 shows system architecture, which shows the principal processes of the proposed system. The test image is preprocessed for the removal of noise as well as contrast adjustment. The skin RGB image is initially converted into grayscale image for dimension reduction. The grayscale images undergoes for the hair removal process. The image is now subjected with the ROI segmentation based on active contour segmentation algorithm. The Segmentation result will be the separated skin lesion. The skin lesion area is subjected for the feature extraction technique. The texture, shape and color features are carried out from the skin lesion region. The Support vector machine classifier is already trained with the two set of features such as benign and malignant images. The test image features are given into the trained SVM model and the class will be recognized from the SVM. The following section describes the modules involved in the proposed work.

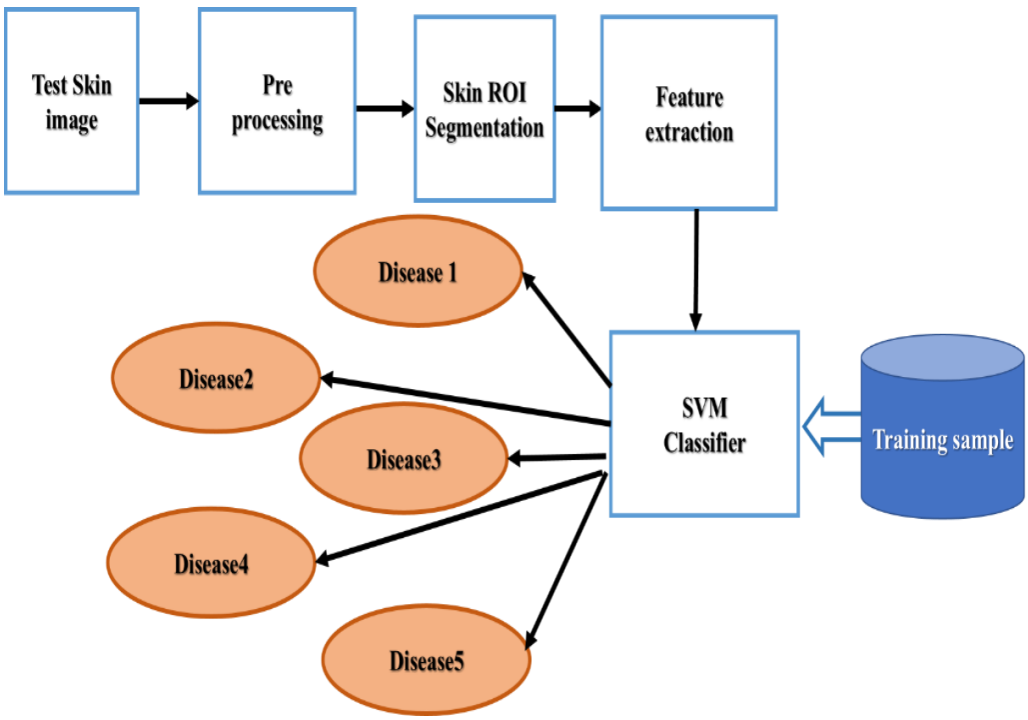

Figure 3: Block diagram for the proposed work

4. METHODS

PRE-PROCESSING

A pre-processing method is essential for benign and malignant skin lesions classification. This involves transforming raw data into an understandable format for further processing. In the real world, data are often inconsistent and incomplete and may contain many errors. Removing any noise and unnecessary features which cause confusion to classifiers is required. The following pre-processing is conducted in this research.

Hair Remover: The Enhanced Dull Razor algorithm figure 4 was used to remove hairs from images where morphological closing image processing was generalized to grey-level images, followed by identification of the narrow, elongated hair outline.

Bilinear interpolation was implemented to substitute the identified pixels of the hairs. This step resulted ina smooth fill inward from the borders of the region of interest.

Contrast Enhancement: Subsequently, the image clarity was enhanced by improving the shape andedges of the image. Image borders were sharpened using contrast enhancement. This process may also optimize subsequent segmentation accuracy.

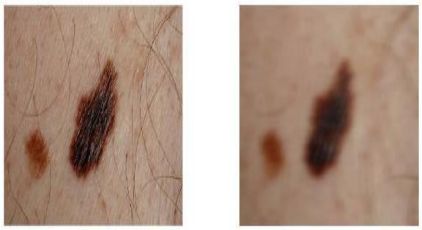

Figure 4: Hair removal result;

Left (original image) Right (After hair removal)

Grayscale Conversion: RGB images of lesions, with M× N pixels in size, were transformed to greyscale by removing hue and saturation using a process which computes the weighted sum of the colour components.

SEGMENTATION

Image segmentation figure 5 is a technique to determine the shape and size of the border, and to separate the object from its background based on different features extracted from the image. After removing the noise from the lesion area, the lesion needs to be separated from the skin, and therefore the analysis for diagnosis is conducted purely using the necessary area. Previous studies have proposed several different types of segmentation methods with high accuracies, such as clustering based, threshold- based and edge-based methods. In this research, the Adaptive Snake (AS) approach is chosen because of its efficiency indicated by previous research. AS inefficient for establishing a discriminating analysis that divides the image into two classes of pixels. In the first instance, the chosen color image is rendered in monochrome. Then corresponding threshold limits are set within the grey spectrum, and the pixels that occur within the range set by the limits are selected. Following this, non-lesion pixels are assigned with a value of zero. To extract multiple features [4], such as colour and area, the segmentation results of this threshold-based method are plotted into multiple images.

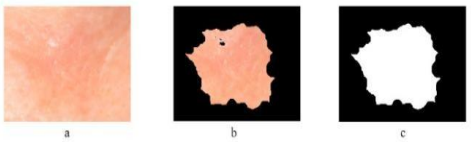

Figure 5: Segmentation result: a) original image b) Segmented lesion region

FEATURE EXTRACTION

After segmentation, image features are extracted for the subsequent classification. Several methods have been identified for feature extraction. Overall, the majority of related work employed the ABCDE rules of dermatology for feature extraction. In this research, measurements such as compactness index, fractal index, and edge abruptness are used in order to indicate border irregularity.

SHAPE

Asymmetry: A melanocyte lesion may be diagnosed by a number of identifiers, of which one of the most significant is a lack of symmetrical morphology. In dermatology terms, the ABCDE rule model rates this aspect as the most crucial factor. In consideration of the symmetry feature, a number of factors are concurrently relevant, including color, texture and morphology. A three-fold classification system can be derived from measuring symmetry, with three-class outputs representing total symmetry, a lack of symmetry along a single axis and a lack of symmetry along dual axe, respectively. The lesion asymmetry was evaluated by calculating the area with inner and outer of the lesion, using the formula shown as follows.

$$AI = \frac{\Delta AK}{AL} * 100$$

where AI represents asymmetry Index. ΔAK represents the area between the two halves of the lesion andAL denotes the lesion area.

Border Irregularity: Irregularities occurring in the edge of a malignant lesion offer useful information concerning that lesion's nature[5]. Typically, the edge of a malignant lesion usually exhibits four factors of interest, i.e. density, fractal dimension, radial variability and the extent to which its contourexhibits small irregularities. To identify the lesion border irregularity,

$$I = \frac{ab}{2\pi(a^2 + b^2)} \frac{P^2}{\Delta A}$$

where I represent irregularity with a and b representing the lengths of major and minor axes of the lesions.P represents the perimeter of the lesion and ΔA indicates the area of corresponding.

Compactness: Another relevant feature is the degree to which the lesion can be described as compact. In order to determine this aspect, a comparative analysis is performed between the lesion's boundary and a circle with a circumference of the same length. It is the former of these two numerical values that presents a challenge in its assessment. One solution to this issue is to use the proportions of the most easily measured values of maximum and equivalent lesion diameter as defined in Equation below

$$C = \frac{4\pi P^2}{\Delta A}$$

Color: The range of colour types utilized in diagnosing a melanocyte lesion can be broadly categorized into the following types: black, grey-blue, brown (dark), brown (light),red and white, which are indicators for a malignant skin lesion. The dermatological analysis allows for the determination of whether a colour category exists in a particular image and if so, where it exists [4]. This positional information is noted via a binary mask application, with image segmentation performed by the dermatological professional (see Figure 5 as an example with separate colour categories being present). Inthis paper, threetype of colour space including HSV, RGB and LAB are used.

Ratio of red, green and blue: In the case of red, the ratio represents the average of the red constituent present in a lesion divided by the mean colour of the surrounding non-lesion skin. The ratiofor red is expressed as follows:

Texture: The texture of a lesion can be estimated by a number of objective measures derived fromGeneralized Co-Occurrence Matrix (GCM). Through a body of existing research [7], Grey-Level Co- Occurrence Matrix (GLCM) has been intensively used as a widely-adopted and popular methodology. GLCM provides a number of numerical assessment measures, which are employed in this research with each being grey-level shift-invariant in nature. These enable sensitive linear shift recognition in terms of the intensity of illumination, such that texture can be categorized in these terms. Research [15] has demonstrated that a particular point exists beyond which an elevated G value leads to reduced ability to differentiate in disparity and contrast, despite maintaining an even level of the other measures[9]. So as to populate a matrix with a sufficient level of data, an equal quantization to 64 grey levels was carried out, with this number being above a lower bound of24 and selected based on the findings of existing studies [7].Such low values do, in addition, minimize the impact of image noise.

It is recommended that the GCM after normalization presents a strong level of density, so as to provide confident statistical estimation within the distribution of joint probability [10]. In this research, the measures consist of three color space (RGB, HSV, LAB) to reduce the impact of deference lighting before the color extraction, plus six color pair (RR, RG, RB, GG,GB, BB). Three grey level quantization, i.e. 64, 128, 256, are used for every color space. The 12 texture features represent autocorrelation, correlation, cluster prominence, dissimilarity, entropy, energy, maximum probability, contrast, homogeneity, cluster shade, inverse difference moment and variance as mentioned in Haralick [14] and six inter-pixel distances.

SUPPORT VECTOR MACHINE CLASSIFIER

Support vector machines (SVMs) are a set of supervised learning methods used for classification, regression and outliersdetection. More formally, a support vector machine constructs a hyperplane or set of hyper planes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks[6]. Intuitively, a good separation is achieved by the hyperplane[13] that has the largest distance to the nearest training-data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

SVM ALGORITHM

Classifying data is a common task in machine learning. Suppose some given data points each belong to one of two classes, and the goal is to decide which class a new data point will be in. In

the case of support vector machines, a data point is viewed as a p-dimensional vector, and we want to know whether we can separate such points with a (p-1)-dimensional hyperplane. This is called a linearclassifier [6].

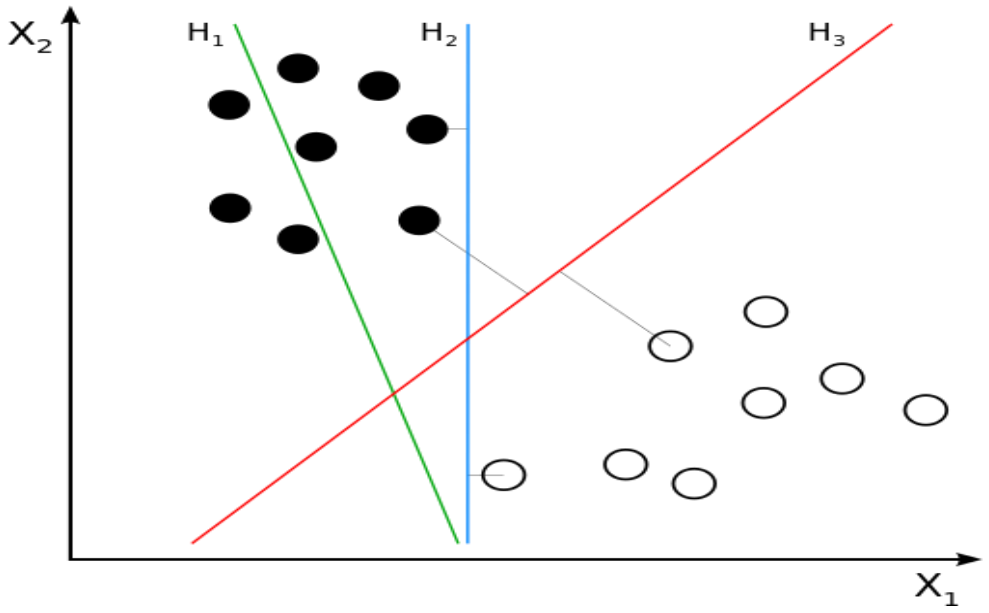

Figure 6: Linearly Separable patterns

There are many hyper planes that might classify the data. One reasonable choice as the best hyper plane is the one that represents the largest separation, or margin, between the two classes. So we choose the hyper plane so that the distance from it to the nearest data point on each side is maximized. If such a hyper plane exists, it is known as the maximum-margin hyper plane and the linear classifier it defines is known as a maximum margin classifier; [12]or equivalently, the perception of optimal stability (figure 6).

5. EXPERIMENTAL RESULTS

The test database images figure 7 are collected form the online database called university of Lowa health care. This work constitute the recognition of five different skin cancer disease automatically. The five disease names are listed over here.

1. Carcinoma
2. Lentiginous
3. Lupus
4. Melanoma
5. Nodular

The five disease images are differing from each other on the way of texture, color and shape. The following figure depicts the skin diseases[11] which are considered in this work.

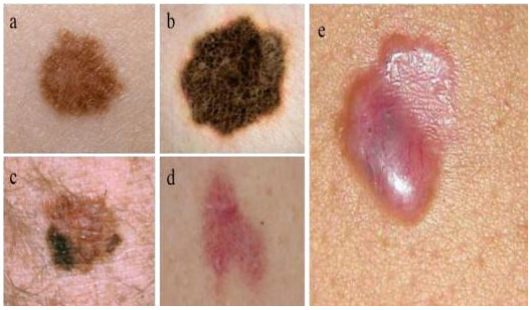

Figure 7: Sample database images:

(a) Carcinoma, (b) Lentiginous, (c) Lupus, (d) Melanoma, (e) Nodular

This work is implemented with total number of 45 skin cancer images. 16 Carcinoma images, 8 Lentiginous images, 10 Lupus images, 6 Melanoma images and 5 Nodular images are taken as the database images. Figure 8 shows the result of SVM classifier and the KNN classifier. Table 1 show automatic image classification system figure 8 is used to classify either as carcinoma, lentiginous, melanoma, lupus or nodular and the GUI displays what the type of skin cancer was.

Table 1: Overall feature matrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Label
1	60.2330 3	182.819 1	149.022 3	25	130.691 5	211	211	332 7.69 7	58512	2.34 E+09	671 46.7 8	28 .2 3	1.7 1E -05	225 74.6 2	- 1.0 6E -05	carci noma
2	120. 006 1	211. 078 2	166. 430 9	25	165. 838 4	139	139	211 5.10 8	24500	6.22 E+08	282 21.9	28 .2 3	4.0 8E -05	438 5.03	1.7 3E -05	carci noma
3	33.6 027 3	194. 022	154. 504 9	25	127. 376 6	265	265	392 4	86 65 5	4.08 E+09	993 43.2	28 .2 3	1.1 5E -05	624. 948 1	9.8 0E -07	carci noma
4	44.8 917 7	201. 854 9	170. 348	25	139. 031 5	105	105	237 6	20 79 0	5.93 E+08	239 21.1	28 .2 3	4.8 1E -05	341 4.17 1	- 1.9 5E -05	carci noma
5	189. 528 6	196. 350 4	141. 254	25	175. 711	191	191	252 2.13 6	40 14 4	1.21 E+09	461 65.2 6	28 .2 3	2.4 9E -05	536 7.50 3	- 9.1 1E -06	carci noma
6	115. 949	172. 820	119. 474	25	136. 081	26	26	295 6.74	66 03	2.34 E+0	758 41.5	28 .2	1.5 1E	487 6.62	- 4.1	carci nom

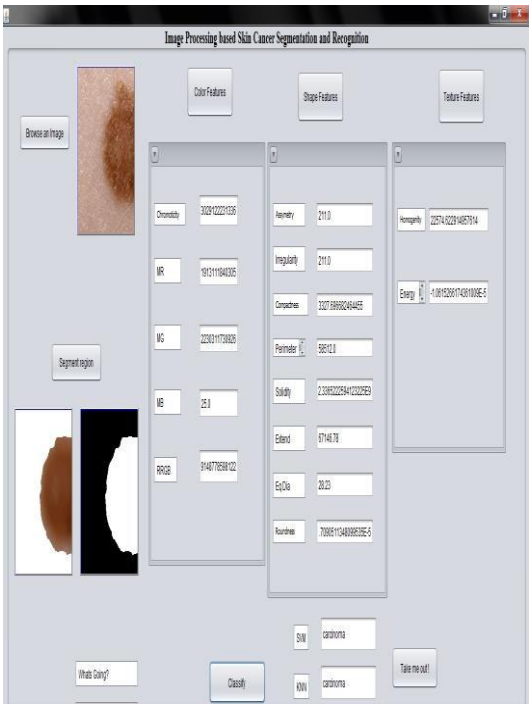

Figure 8: Classification results on prepared GUI

6. CONCLUSION

The image processing technique is used to detect and exclude the hair from the dermoscopy images, preparing it for further segmentation and analysis, resulting in satisfactory classification results. The proposed automated image analysis process includes image acquisition, hair detect ion and exclusion, lesion segmentation, feature extraction, and classification. The incidence of skin cancers has reached a large number of individuals within a given population , especially among whites, and the trend is still rising. Early detection is vital, especially concerning melanoma, because surgical excision currently is the only life saving method for skin cancer. The state of the art is used in the proposed system for the dermoscopy image acquisition, which ensures capturing sharp dermoscopy images with a fixed distance to the skin and consistent picture quality.

REFERENCES

[1] Rogers HW, Weinstock MA, Feldman SR, Coldiron BM.: "Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population", 2012. JAMA Dermatol 2015; 151(10):1081-1086.

[2] "Cancer Facts & Figures 2014". American Cancer Society, 2014.

[3] Siegel, R.L., Miller, K.D., and Jemal, A.: "Cancer statistics, 2016," CA: A Cancer Journal for Clinicians, vol. 66, no. 1, pp. 7-30, 2016.

[4] Brady, M.S., Oliveria, S.A., Christos, P.J., Berwick, M., Coit, D.G., Katz, J., Halpern, A.C.: "Patterns of detection in patients with cutaneous melanoma." Cancer. 2000 Jul 15;89(2):342-7.

[5] Kittler, H., Pehamberger, H., Wolff, K., Binder, M.: "Diagnostic accuracy of dermoscopy". In: The Lancet Oncology. vol. 3, no. 3, pp. 159.165. 2002.

[6] Vestergaard, M.E., Macaskill, P., Holt, P.E., et al.: "Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting." Br J Dermatol. Sep 2008;159:669-676.

[7] Bafounta, M.L., Beauchet, A., Aegerter, P., et al.: "Is dermoscopy (epiluminescence microscopy)

useful for the diagnosis of melanoma? Results of a meta-analysis using techniques adapted to the evaluation of diagnostic tests." Arch Dermatol. Oct 2001;137:1343-1350.

[8] Abder-Rahman, A.A., Deserno, T.M.,: "A systematic review of automated melanoma detection in dermatoscopic images and its ground truth data". In: Proc. SPIE 8318, Medical Imaging 2012: Image Perception, Observer Performance, and Technology Assessment.

[9] Carli, P., Quercioli, E., Sestini, S., Stante, M., Ricci, L., Brunasso, G., De Giorgi, V.: "Pattern analysis, not simplified algorithms, is the most reliable method for teaching dermoscopy for melanoma diagnosis to residents in dermatology". In: Br J Dermatol. vol. 148, no. 5, pp. 981-4. 2003.

[10] Ganster, H., Pinz, A., Rhrer, R., Wildling, E., Binder, M., Kittler, H.: "Automated Melanoma Recognition". In: IEEE Transactions on Medical Imaging, vol. 20, no. 3, 2001.

[11] Al-Masni MA, Al-Antari MA, Choi MT, Han SM, Kim TS: Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. Comput Methods Programs Biomed 162:221-231,2018

[12] Siegel RL, Miller KD, Jemal A: Cancer statistics, 2019. CA Cancer J Clin 69(1):7-34,2019

[13] Li Y, Shen L: Skin lesion analysis towards melanoma detection using deep learning network. Sensors 18:2,2018

[14] Unver HM, Ayan E: Skin lesion segmentation in dermoscopic images with combination of YOLO and GrabCut algorithm Diagnostics (Basel) 9(3),2019

[15] Bi L, Feng D, Fulham M, Kim J: Improving skin lesion segmentation via stacked adversarial learning. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp 1100–1103

[16] Jiang F, Zhou F, Qin J, Wang T, Lei B: Decision-augmented generative adversarial network for skin lesion segmentation. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp 447–450

[17] Khan MA, Javed MY, Sharif M, Saba T, Rehman A: Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification. In: 2019 International Conference on Computer and Information Sciences (ICCIS), 2019, pp 1–7

[18] Stefan Jianu SR, Ichim L, Popescu D: Automatic Diagnosis of Skin Cancer Using Neural Networks. In: 2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE), 2019, pp 1–4