Journal of Electrical Engineering, Electronics, Control and Computer Science – JEEECCS, Volume 7, Issue 23, pages 1-4, 2021

Development of a Skin Cancer Detection Classifier using Artificial Neural Network (ANN)

Emmanuel Eragbe Sule
Department of Electrical and Electronic Engineering
The Federal University of Technology
Akure, Nigeria.
sule59002018db@futa.edu.ng

Abstract – **In recent times, dermatological diseases pose one of the biggest medical challenges in the 21st century resulting in high medical cost and painful diagnostic procedures (biopsy). Cancer affecting the skin is commonly referred to as Melanoma. This condition is curled from the cell that gives rise the disease known as melanocyte. The most dangerous form of skin cancer is malignant melanoma but early detection plays an essential part in its control and cure [1]. In 2018, it was estimated that about 9,320 deaths resulted from melanoma. In this work, a computer-based system to classify histopathological images of skin tissues using Artificial Neural Network (ANN) was implemented on MATLAB. Performance measures of the proposed system are encouraging and there is no evidence of over-fitting. Therefore, an extended version of this classifier system could be used to assist patients and hospital pathologists to increase efficiency in healthcare delivery.**

Keywords-Melanoma, Diagnostic, Biopsy, Histopathological, Artificial Neural Network, Classifier

I. INTRODUCTION

People of all ages are infected by various skin diseases and melanoma is adjourned the most pernicious form of cancer affecting the skin. Its emergence is from cancerous growth in pigmented lesion. The chances of patients recovering are tied to early diagnosis and detection of the ailment. This is dependent on attention paid to the patients and accuracy of the medical personnel. Early detection increases the chances for it being managed and cured.

The American Cancer Society (ACS) [2] in 2016 reported that close to 76,380 (61% men and 39% women) were diagnosed with melanoma. In the same period under review, ACS also reported approximately 10,130 fatalities (67% men and 33% women). Since then, the relative occurrence of melanoma continues to soar each year. Early detection fosters cure and lots of lives can be preserved hence many research using various methodologies are being carried out around the world to aid in the early diagnosis of melanoma.

Melanoma starts like a small mole with most victims paying little or no attention to it. When left unbridled, it begins to spread to various sections of the body system which becomes calamitous. The readily available nature of dermoscopic images and its economic value make it a practicable option for usages in image processing and artificial intelligence. Hence, using dermoscopic images for melanoma detection has high potency for replacing the present clinical epitome of holding off until the cancer is at its final stage and carrying out a sizable number of biopsies.

Therefore, digital image processing becomes our option for early diagnosis of the skin cancer as it is an economically viable technique. The recognition of the edges of an object in an image scene is a fundamental part of man's optical system as it provides information on the basic structure of the object from which an interpretive match can be achieved. Similarly, the partitioning of an image into a composite of edges is a utile requirement for object recognition. This research used monitored machine learning proficiencies; ANN is used as a classifier.

Other similar recognition and assortment techniques, such as K-Nearest Neighbors (KNN) also classify pixels based on their resemblance in each feature image and this aid to sort the normal/abnormal dataset of images [3].

In a bid to assist dermatologist in analysis, Filali, Sabri and Aarab [4] sort to carry out skin lesion detection by preprocessing input images into geometrical and texture components with the former used in segmentation stage and the latter for feature extraction. The classification using Gray Level Co- Occurrence Matrix (GLCM) had 80 percent accuracy while Local Binary Pattern yielded an accuracy of 81.23 percent.

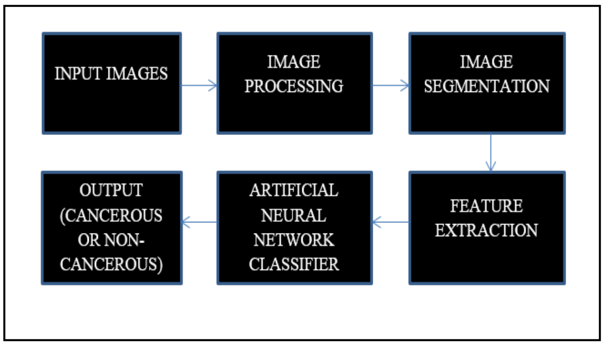
Vijayalakshmi [5] presented an automated classifier combining multiple algorithms of CNN and Support Vector Machine (SVM) with image processing tools leading to a system of overall accuracy of 85 percent.

This study extracts features that are closely associated with cancerous tissues and uses an ANN classifier to make predictions with an accuracy of 98.7 percent, which is higher than in previous studies using other classification methods.

II. MATERIALS AND METHODS

The methodology of this research work is divided into 6 stages. 1) Input Images 2) Image Processing 3) Image Segmentation 4) Feature Extraction 5) Artificial Neural Network Classifier 6) Output. Fig 1. illustrates the whole process in a block diagram.

Figure 1. Block diagram representation of methodology



A. *Input Images*

Histological images of skin tissue were obtained using an optical microscope with a 400x lens. Samples were hematoxylin-eosin stained during Mohs surgery, a surgical technique that involves removing thin layers of skin with cancer while simultaneously analyzing it in the microscope so that the tumor is removed progressively until only healthy tissue remains.

The foremost stage in a diagnostic system is the acquisition of input images. The dataset used for the image processing was provided by the Pathology – Unit Hospital del Mar – Parc de Salut Mar (Barcelona, Spain) which were to be used solely for academic and research purposes. The image dataset used for the research consist of 76 histopathological images (44 healthy samples, 32 samples with malignant melanoma) [6].

TABLE I. PROPERTIES OF IMAGES IN THE DATASET

Image Type	Height	Width	Bit Depth
Bitmap(.bmp)	480	640	24

B. *Image Processing*

The aim of digital image processing is to improve the quality of image and subsequently aid in features extraction thereby resulting in high classification accuracy. Image processing is carried out to remove unwanted objects or artifacts and perform color space transformation.

This stage comprises;

- Sizing of image.
- Noise Filtering.
- Red Green Blue (RGB) to Gray scale conversion.
- Contrast Enhancement.

Filtering technique adopted during this stage is Median Filtering. The median value of each image pixel replaces neighboring pixel value. This filtering technique is a statistical filter where f(x, y) relies on the pixel values order of g in the function S(x, y). The median filter produces an output that is a 50% rank of the ordered value [7].

$$f(x,y) = median\ \langle g(s,t)\rangle \in S(x,y) \qquad (1)$$

For instance, a 3 x 3 median filter with,

$$g(x,y) = \begin{matrix} 1 & 5 & 20 \\ 200 & 12 & 25 \\ 25 & 9 & 100 \end{matrix}$$

We must first order the value in the window as; 1, 5, 9, 12, 20, 25, 25, 100, 200. The number 20 corresponds to the 50% ranking (5^{th} value).

Then f (x, y) = 20.

Finally, to improve automatic segmentation and fill the holes in the segmented and yet keep the nuclei size unchanged a binary dilation followed by binary erosion was performed.

C. *Image Segmentation*

Segmentation have to do with the breakdown of image into distinct sections that are of the same kind with respect to selected features that include, color, luminance, texture amongst others [8]. The end result of segmentation is to modify and alter the make-up of an image into something with more meaning and then making the image a lot easier to analyze.

The Segmentation stage is a vital stride in the analysis of dermoscopic lesion images; thus, it has gained relevance in the areas of research with many algorithms and methodologies available for the diagnosis process. However, many factors should be considered before the selection of a suitable methodology. Priorities should be placed on automatic versus semiautomatic, vector vs scalar processing and other arguments whose values is to be evaluated [9]

Segmentation methods used for this work are:

- Clustering Means Clustering;
- Otsu Thresholding.

Clustering refers to sorting patterns or objects so that samples of similar groups are classified together. An enhanced Fuzzy Clustering Means (FCM) clustering method is posited. It integrates Otsu thresholding with traditional FCM and shrink its susceptibility to lower limits. It also reduces its propensity to obtain a threshold that is slanted to the portion with a greater chance of occurrence and obtain a threshold value with much improved veracity.

Otsu's thresholding involves repeating through all the potential threshold values and quantifying the amount of spread at each pixel level of threshold; this including foreground and background pixel. The goal is to obtain the threshold measures where the sum of background and foreground spreads are at their lowest points. Thresholding simply refers to representing grayscale with two units (binary; 0, 1) as deplicted in (2) [10]

$$S(x,y)\begin{cases}0, & g(x,y) < T(x,y)\\ 1, & g(x,y) \geq T(x,y)\end{cases} \qquad (2)$$

In (2), function S(x, y) is the measure of segmented datasets, g(x, y) is gray level of the pixel with values (x, y) and T(x, y) is the threshold measure of the coordinates (x, y). Lastly, function T(x, y) is the main coordinate which remains same for the whole image.

D. *Feature Extraction*

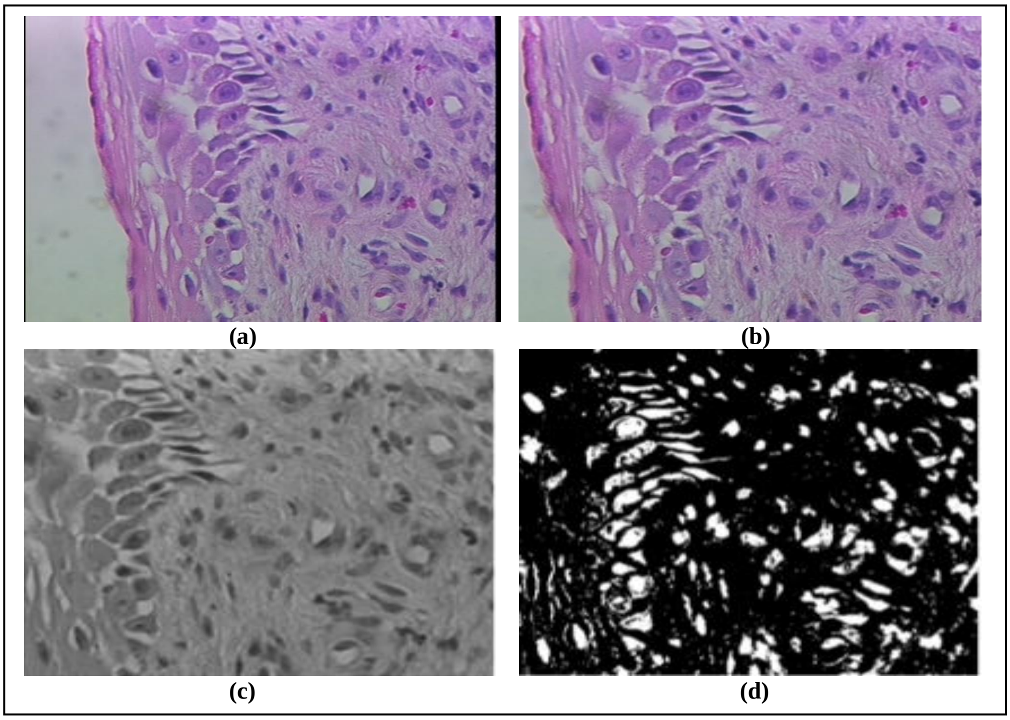
Feature extraction refers to the way singular features of the histological images are obtained. This stage reduces complication problems in sorting and classification.

Cancer is characterized by the abnormal and uncontrolled growth of cells and these invade adjacent tissues. The properties of cancerous tissues were observed by pathologists through histological analysis and used for diagnosis purposes [11]

In this research, tissue features were extracted for the subsequent classification task. The features extracted are (i) Nuclear-Cytoplasmic Ratio, (ii) Nuclei Number, (iii) Nuclei Size and Pleomorphism (size variance), (iv) Mean and (v) Standard Deviation.

The nucleus to cytoplasm ratio is a function defined by the size of a cell’s nucleus compared to its cytoplasm. Because of the uncontrolled growth of cancer cells, the NCR is increased. To compute it, simply count the number of 1s present in the binary image (which represent the nuclei) and divided them by the total image size. The nuclei usually increase in cancerous tissue because of the uncontrolled growth of its cells. To calculate the numbers of nuclei in the histology sample the algorithm counts the number of connected components in an 8-pixel neighborhood.

Figure 2. Histological image pre-processing showing (a) original (b) cropped (c) gray scaled (d) final image



E. *Artificial Neural Network Classifier*

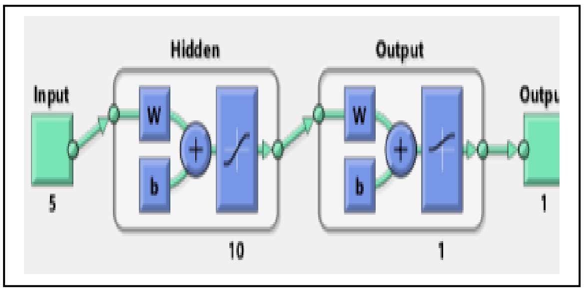
The classification process always comprises of the training, testing and final validation of the dataset in various cases. Each case in the training set comprises a target value with various features (in this project, five). The classifier was implemented using MATLAB. With the 5 features extracted, the classifier network takes in 5 input values. The hidden neurons in this classifier are 10 and a single output node.

The inputs are weighted at the entry of the neuron, every value obtained after weighing is multiplied by a weighing factor. At mid-section of the neuron network, all the weights are summed. The exit provides the weighted sum with bias by using a transfer function called activation function [12].

F. *Output*

The ANN model is assembled on MATLAB software. The model is trained, tested and validated using values obtained from data images alongside desired outputs. The seventy-six (76) images feature values were fed as input to the neural network to aid classification. The training is brought to an end only when the Mean Square Error (MSE) is at its lowest point. The ANN model classifies the dataset via output neuron into 1 and 0 with 1 representing cancerous and 0 non- cancerous.

Figure 3. Architecture of the ANN classifier



III. RESULTS

The results from the classifier were obtained by using 70 percent of the total sample images for training the classifier, 15 percent for testing and the remaining 15 percent for validation. Upon feeding the extracted features as inputs for the ANN classifier, the classifier classifies the dataset into 1 and 0 where **1** represents cancerous and **0** represents non-cancerous.

Figure 4. ANN training on MATLAB

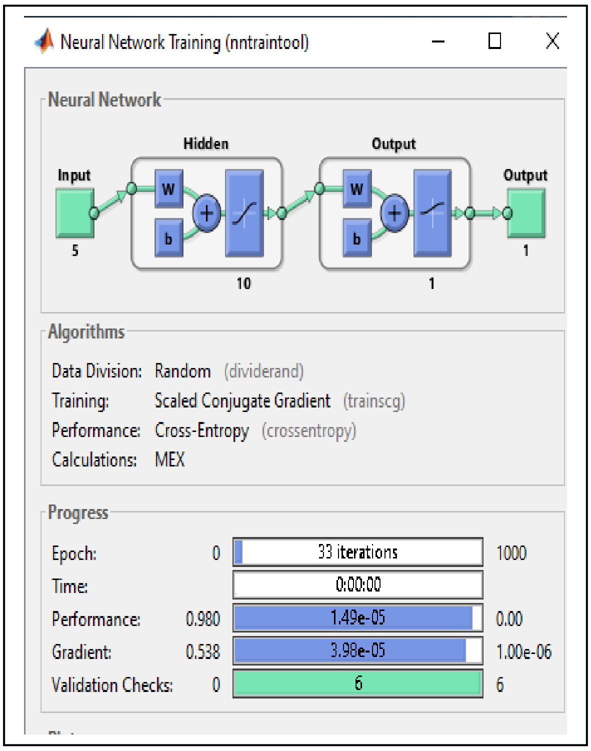
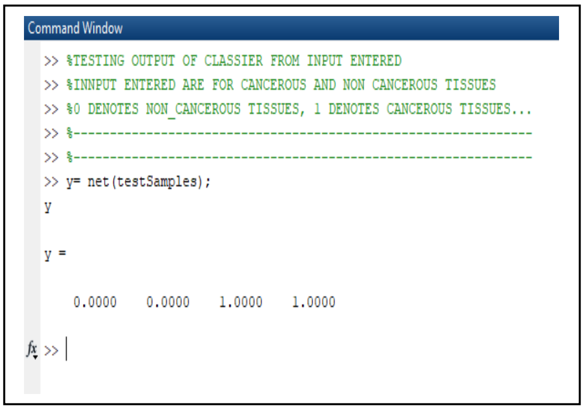


Figure 5. MATLAB window showing classification results for 2 non-cancerous tissues and 2 cancerous tissues



Rozeira and Jorge's technique [13] for detecting the pigment network in a dataset of 55 different dermoscopy images produced a sensibility SE=80.0% and specificity SP=67.5%.

Similarly, Arthur Tenenhaus' technique [14] for diagnosis of melanoma performed on another dataset of 227 dermoscopic images acquired under unrestrained conditions showed similar performances with respect to dermatologists having specificity 60% and sensitivity 95%.

The proposed method used in this research has an accuracy of 98.7%, specificity of 97.8% and sensitivity of 100%. These are higher than the conventional methods stated above.

TABLE II. PERFORMANCE INDICES OF THE CLASSIFIER

Performance Measure	Value
Sensitivity	1
Precision	0.978
Specificity	0.978
Accuracy	0.987

IV. CONCLUSION

It is indeed possible to design a computer-based system to classify histopathological images of skin tissue using ANN. This research will be of immense help to dermatologists as it reduces errors arising during diagnosis. Similarly, it serves as a first step to treatment for patients in rustic areas with sharp decline in the number of medical professionals [15]. An improved version of such a system could be used to assist pathologists in their day-to-day hospital tasks so that they can spend their time efficiently and better contribute to increase the effectiveness of our healthcare delivery system.

V. ACKNOWLEDGEMENT

My sincere gratitude goes to Dr. J. J. Popoola the head of the department, electrical/electronic engineering, The Federal University of Technology, Akure Nigeria, for his support and encouragements. I also want to thank the reviewers of this journal for their inputs and corrections.

VI. REFERENCES

[1] M. K. Chaithanya, S. Ranganayakulu and P. Venkatesan, "Skin cancer detection and feature extraction through clustering technique.," *International Journal of Innovative Research in Computer and Communication Engineering,* vol. 4, no. 3, 2016.

[2] American Cancer Society, "Cancer facts and figures," Atlanta, Georgia, 2018.

[3] B. Arunadevi and S. N. Deepa, "Extreme learning machine for classification of brain tumour in 3DMR images," *Informatologia,* vol. 46, no. 2, pp. 111-121, 2013.

[4] Y. Filali, A. Sabri and A. Aarab, "An improved segmentation approach for skin lesion classification," *Statistics, Optimization and Information Computing,* vol. 7, pp. 456-467, 2019.

[5] M. M. Vijayalakshmi, "Malanoma skin cancer detection using image processing and machine learning," *International Journal of Trend in Scientific Research and Development ,* vol. 3, no. 4, pp. 780-784, 2019.

[6] [Online], "Hospital del Mar. Hospital del Mar - Parc de Salut Mar. https://www.parcdesalutmar.cat/hospitals/hospital-del-mar/en_presentacio.html. Accessed June 7, 2017.," Barcelona.

[7] K. O. Boateng, B. W. Asubam and D. S. Laar, "Improving the Effectiveness of the Median Filter.," *International Journal of Electronics and Communication,* vol. 5, no. 1, pp. 85-97, 2010.

[8] O. S. Murumkar and P. P. Gumaste, "Feature extraction for skin cancer lesion detection," *International Journal of Science, Engineering and Technology Research (IJSETR),* vol. 4, no. 5, 2015.

[9] J. Burdick, O. Marques, J. Weinthal and B. Furht, "Rethinking Skin Lesion Segmentation in a Convolutional Classifier," *J Digit Imaging,* vol. 4, p. 34, 2018.

[10] T. Acharya and A. K. Ray, "Image Segmentation," in *Image Processing: Principles and Application,* New Jersey, John Wiley & Sons Inc, 2005, pp. 131-156.

[11] A. Masood, A. Al-Jumaily and K. Anam, "Texture Analysis Based Automated Decision Support System for Classification of Skin Cancer Using SA-SVM," in *International Conference on Neural Information Processing,* Switzerland, 2014.

[12] K. Andrej, J. Bester and A. Kos, "Introduction to Artificial Neural Network," in *Artificial Neural Network-Methodological Advances and Biomedical Applications,* China, InTech, 2011, pp. 1-18.

[13] C. Rozeira and M. Jorge, "Detecting the Pigment Network in Dermoscopy Images: a Directional Approach," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* August 2011.

[14] A. Tenenhaus, A. Nkengne, F. J. Horn, C. Serruys, A. Giron and B. Fertil, "Detection of melanoma from dermoscopic images of naevi acquired under uncontrolled conditions.," *Journal of Internation Society for Bioengineering and Skin(ISBS),* no. 16, pp. 85-97, 2010.

[15] E. U. Isilay and A. Cinar, "Classification of skin images with respect to melanoma and non-melanoma using deep neural network," *IOSR Journal of Engineering (IOSRJEN),* 2018.