A Study on Different Techniques for Skin Cancer Detection

Nikita Raut, Aayush Shah, Shail Vira, Harmit Sampat

[1,2,3,4] *Student, Dept. of Computer Engineering, K.J.Somaiya College Of Engineering, Maharashtra, India*

Abstract:- *The number of cases of melanoma skin cancer has been increasing year by year. Skin cancer is one of the most dangerous types of cancers, because it's much more likely to spread to other parts of the body if not diagnosed and treated early. About three million people are diagnosed with the disease every year in the United States alone. Early detection of Melanoma skin cancer is very much necessary for the patient for it to be curable. Today's technological advancements can make possible the early detection of skin cancer. As per the literature the lesion characteristics such as shape, color, structure etc. are the important parameters for detection of skin cancer. In this paper we review the various soft computing and artificial intelligence techniques for early stage melanoma skin cancer detection.*

***Key Words*: Melanoma, skin cancer, soft computing, artificial intelligence, neural networks, accuracy**

1. INTRODUCTION

Melanoma is one of the most deadlytypes of all skin cancers in humans thatappear on the skin as pigmented moles or marks. Melanoma occurs when something goes awry in the melanin-producing cells (melanocytes) that give color to your skin. Some of the risk factors of Melanoma are fair skin, a history of sunburn, genetic factors, weakened immune system, tanning beds and excessive ultraviolet light exposure.

According to the National Cancer Institute, about 87,110 new melanomas were expected to be diagnosed in 2017, and about 9,730 people were expected to die of melanoma. Melanoma is more than 20 times more common in whites than in African Americans. Overall, the lifetime risk of getting melanoma is about 2.6% (1 in 38) for whites, 0.1% (1 in 1,000) for blacks, and 0.58% (1 in 172) for Hispanics. The risk of melanoma increases as people age. The average age of people when it is diagnosed is 63. But melanoma is not uncommon even among those younger than 30. In fact, it's one of the most common cancers in young adults (especially young women)..

If melanoma is not detected, it grows and spreads along the first layer of skin before penetrating into the deeper layers and finally, comes into contact with the lymph vessels and the blood. The early detection of skin cancer increases the chances of a cure unlike when it is discovered in advanced stages. However, early detection of skin cancer is an expensive affair. As skin lesions look quite similar to each other, it is difficult to determine whether a lesion is benign or malignant. Normal moles are generally a uniform color — such as tan, brown or black — with a distinct border separating the mole from your surrounding skin. They're oval or round and usually smaller than 1/4 inch (about 6 millimetres) in diameter — the size of a pencil eraser. To help you identify characteristics of unusual moles that may indicate melanomas or other skin cancers, look for moles with irregular shapes, irregular border, changes in color and diameters greater than 6mm.

In order to analyse the skin sore and classify it as benign or melanoma, various non-invasive techniques are being proposed. Every parameter is undergoing evaluation with an end goal that the feature values will be utilized to predict which sort of skin cancer it is: Image acquisition, Pre-processing, Segmentation, Feature extraction, and Classification. These techniques have proved to be more efficient, less painful and less expensive than the medical detection techniques. In upcoming sections, we discuss some recent approaches, which perform accurate skin cancer detection with the help of various algorithms.

2. LITERATURE SURVEY

A. ABCD rule based detection

In this paper [9], Hardian et.al use ABCD rule of dermoscopy for extracting the skin lesion. A smartphone camera is used to capture the lesion image and this image is processed by using the ABCD rule. Feature extraction is performed on the pre-processed image where the four features - Asymmetry (A), Border (B), Color(C), and D (Diameter) are extracted in the following way. 1) Asymmetry – Melanoma lesions are asymmetric in nature. Asymmetry index is used for determining the level of symmetry of the object. This is done by dividing the image horizontally or vertically. 2) Border - In case of melanoma the border is irregular, ragged, and blurred. The compactness index is used to determine the border irregularity. 3)Color – Melanoma are not uniform in color unlike the benign mole. Normalized Euclidean distance between each pixel is used to determine the color uniformity. 4) Diameter - The melanoma lesion has larger than 6 mm.The diameter in the image is found out and compared to 6mm measurement.

The major drawbacks of the proposed method are:

1) According to [4] , in order to analyse the ABCD score, the criteria are assigned semi-quantitatively. Each of the criteria is then multiplied by a given weight factor to calculate a total dermoscopy score. The ABCD rule works appropriately for thin melanocytic wounds. The ABCD rule has about 59% to 88% accuracy in diagnosing melanoma, but biopsy is needed for more precise diagnosis.

2) According to [10], the ABCD method, in some cases, isn't helpful to classify malignant and benign skin moles. In addition, the method may not recognize some malignant moles at early stages, for example malignant melanoma with a small diameter than 6mm. This poses the challenge of differentiating malignant melanoma and birthmarks

B. Backpropogation Neural Networks based diagnosis

In this paper, Pratik et al. [3] propose the use of the ABCD rule for extracting features and a backpropogation neural network to classify the lesions as melonama or benign. The criteria that combine to create the ABCD rule of dermoscopy are asymmetry, border, color, and diameter. In this paper, they have extracted the features based on the ABCD rule after successful segmentation. As the number of classes increase, it becomes difficult to classify lesions into their appropriate classes accurately. Neural Networks inherently possess better capability of handling complex relationships between different parameters. Thus, the proposed model makes use of Backpropogation neural networks. However, the proposed method as certain drawbacks. The major drawbacks of this method are slow convergence rates and trapping in local minima. According to [4] the back-propagation algorithm is known as a local search algorithm which uses gradient descent to iteratively develop the weights and biases in the neural network

C. Hybrid artificial neural network and world cup optimization

NavidRazmjooy et al. [4] have proposed a new hybrid algorithm between the artificial neural network and world cup optimization for enhancing the back-propagation algorithm efficiency and for escaping from trapping in the local minima. World cup optimization (WCO) algorithm has been used to help ANN to find the initial optimal weights in the back-propagation algorithm, to speed up the convergence speed and to minimize the root mean square error between the actual output and the target output. In WCO algorithm, the best teams from each group arise to the next level and the rest are eliminated. They have used the same concept in Backpropagation algorithm to search around the best cost for some epochs and if the search result is better than the best cost, the output will be the achieved; otherwise, previous output will be selected. The biggest advantage of the proposed method is the use of world cup optimization to resolve the drawbacks of backpropogation algorithm.

D. Hybrid genetic algorithm- Artificial Neural Network

In this paper [7], Aswin et al. present a computer aided approach for skin cancer detection. The steps involved are 1) image processing 2) segmentation 3) feature extraction and 4) classification. The Image processing step includes image resizing and image hair removal using dull razor software, which is free medical imaging software. Even though the proposed model makes use of an efficient software tool, the paper neglects the negative effects that the tool may have. The software makes use of a mean filter which effectively smoothens the hair in the images; however, in this process it may smoothen the edges and thus compromise the quality of the image. The next step involves segmentation using the open source ImageJ software. However, the author of the paper ignores the fact that the ImageJ software has certain drawbacks. The major drawback is that if you're going to deal with large stacks of images, you may bump up against problems in the default memory configuration for the program. The Otsu colorthresholding which is a simple yet powerful thresholding technique is employed in this paper. Threshold level is set such that the background skin pixels are removed and only foreground lesion remains.

The major drawback of thresholding methods, according to [8] is that they can achieve good results only if there is a high contrast between the lesion area and the surrounding skin region, which may not always be the case.

They have then used GLCM for feature extraction and to obtain the features, namely Contrast, Correlation, Homogeneity and Energy. The selected features are given as input to classifier which classifies the images as cancerous or non-cancerous .The proposed system for feature extraction makes use of the fact that Malignant melanoma has a mix of red, green, and dark coloration whereas benign lesion has a uniform color pattern.

Finally, the Classification is done using a hybrid Genetic algorithm – Artificial neural network classifier. A conventional ANN classifier uses backpropogation algorithm for training. However, the major drawback of this conventional technique is that the solution may get trapped in the local minima instead of global minima. In order to eliminate this problem, a hybrid approach has been used where the weights of ANN classifier are optimized by the genetic algorithm to improve accuracy. This Paper shows that the results are better using the hybrid approach than the conventional approach.

E. Convolution Neural Network Based Diagnosis

In their paper [3], Pratik et al. encourage the use of Convolution Neural Networks (CNN) as future scope, since CNN models can be used for classification of the affected skin images without the need for performing segmentation and feature extraction independently. Mobeen et al. [5] have made use of custom made automated segmentation and have used a novel approach for implementation of the CNN methodology, where CNN has been used for feature extraction and ANN was used to classify those extracted features. The advantage of this proposed system is that CNN does not require any additional classifier like SVM, KNN since 3 fully-connected layers were used for training the classification model. This type of classification brings its own unique benefits, like it is possible to apply back-propagation algorithm, which adjusts the parameters of neurons in all layers to obtain better classification model.

In this paper [8], Aya Abu Ali and Hasan Al-Marzouqi have employed transfer learning by using and modifying the convolution neural network (CNN) architecture in the

LightNetpretrained model to achieve higher accuracy. They classified the images without applying lesion segmentation or without using complex image pre-processing techniques. The images were pre-processed and resized to create uniformity in the sizes of all the images. Each block of the 5 blocks had a convolution, relu, pooling and a dropout layer except for the last block where the Last block has fully connected layer and softmax layer. They calculated the error by comparing the predicted output with the actual output.

The advantage of the proposed model presented in this paper is that it is simpler as it does not use data augmentation and uses lesser no of parameters which is crucial in mobile applications where constraints on the size of the network and energy consumption affect the utility of developed tools. This model attains comparable results with fewer parameters. However, improved results can be obtained by using better feature extraction techniques like the ABCD rule and better image pre-processing.

The proposed system in [6] addresses 2 parts: lesion segmentation and lesion classification. The proposed solution makes use of a fully convolution-deconvolution network to segment the skin tumour from the surrounding skin. In order to address lesion classification, 2 solutions have been proposed. The first trains the dataset from scratch and second makes use of pre-trained VGG-16. They found that the results obtained from VGG-16 model were better than those obtained from the simple convNet architecture

F. Neuro Fuzzy System

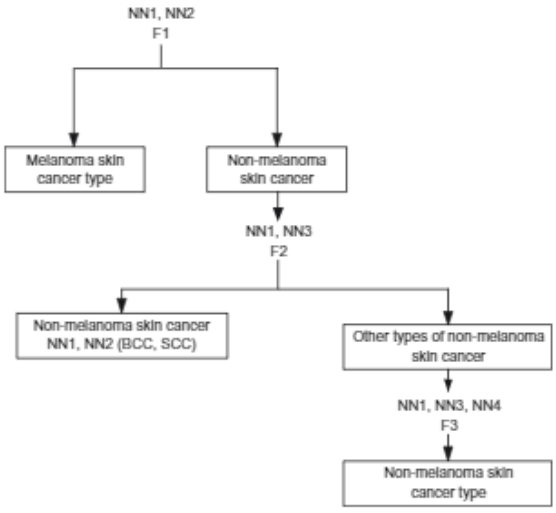
Bareqa Salah et al. [11] propose the use of neural networks with fuzzy logic to calculate the relevant features and determine the type of skin cancer.

The proposed system is composed of 4 neural networks (NN). 1) NN1 determines the main type of cancer: melanoma or non-melanoma. 2) NN2 determines the type of melanoma cancer. 2) NN3 distinguishes non-melanoma skin cancer and other types of non-melanoma. 4) NN4 classifies the type of non-melanoma skin cancer. The outputs were generated between negative infinity and positive infinity instead of 0 and 1. The proposed fuzzy logic system uses the IF-THEN rules over the outputs of the neural network.

They provided the output of NN1 and NN2 as input of the first Fuzzy Logic Inference System1 (FIZ1),FIZ2 used NN1 and NN3 outputs as input data, while F1Z3 used output of NN1, NN3, and NN4 as input.

The accuracy obtained by the neuro-fuzzy system was slightly greater than that obtained by the hierarchical Neural Network. However, the biggest disadvantage was that the increase was only found out to be 0.5%.

This might be due to fewer studied images. Additionally, the proposed system classifies only melanoma and non-melanoma cancers and does not take into consideration the benign cancers



G. Detection Using Delaunay Triangulation

In this paper [12], A. Pennisi et al. have proposed a method for Melanoma Detection Using Delaunay Triangulation. First, they have pre-processed the image to remove the hair on skin by morphological closing and then the image is equalized. The next step is detection of skin by defining boundaries in the chosen color space. The main advantage of this method is that it does not require a training phase. The next step involves extracting the contours of the lesion by edge detection using canny edge detector. The detected edges are used as input for the Delaunay Triangulation procedure which computes a triangular tessellation of the image.

The proposed approach is highly accurate when dealing with benign lesions, while the detection accuracy significantly decreases when melanoma images are segmented. Additionally, this algorithm is highly sensitive with respect to images containing irregular borders, multiple shades of pigmentation, and varying structure and thus, it presents a lesion area that is smaller than the actual area. This behaviour is the main reason as to why they had to consider geometrical and color features extracted from the output of the proposed algorithm as input for four different binary classifiers. However, the disadvantage of the proposed system is that it is not suitable for diagnostic applications.

H. Classification using Support Vector Machine (SVM)

Hiam et al. [1] classify images of skin lesion as Normal or Melanoma using SVM. The steps involved in this research are image pre-processing, segmentation using thresholding, statistical feature extraction using Gray Level Co-occurrence Matrix (GLCM), feature selection using Principal component analysis (PCA), and classification using Support Vector Machine (SVM). They have pre-processed the image by improving the contrast of the image, while noise filtering was done to remove the hair cover on the skin. The next step was image segmentation where the ROI (region of interest) was selected. In this step, they performed thresholding, image filling to remove the background pixels from inside

the object, image opening to remove the extra background pixels, smoothening the contour of the object's boundary and then finally cropping the image to appropriate size. The next step involved extracting the features of Asymmetry, border irregularity, Energy, correlation, homogeneity, entropy, skewness and mean. PCA was then used to reduce the number of features to the best 5 features: TDS, mean, standard deviation, energy, and contrast respectively. Finally, they used kernel SVM with radial basis function to classify the images as benign and malignant.

Kavitha et al. [2] perform texture and color feature extraction for classification of melanoma using SVM. Texture is an important feature that identifies the object present in any image. The texture feature extraction is done using Gray Level Co-occurrence Matrix (GLCM), while color histograms in different color spaces are constructed to identify the melanoma. For these histograms color bins are generated and it is considered as a single feature vector. They have then provided the feature vector as an input to the SVM classifier. The advantage of SVM is that in high dimensional spaces it works effectively and since it uses a subset of training points in the decision function it is considered as memory efficient

I. Detection and classification based on supervised and unsupervised learning

In this paper, Ms. H. R. Mhaske et.al have used supervised algorithms like neural network and support vector machines and the K-mean clustering unsupervised algorithm to detect and classify melanoma skin cancer. The results from the three methods: k mean clustering, backpropogation neural network, support vector machines have been compared in this paper. They have used the following steps. First, image pre-processing, segmentation and feature extraction have been performed on all the images in the dataset. Then classification is done which makes use of various classifiers. For Unsupervised learning using k-means algorithm the classification result obtained is 52.63%. In k-means algorithm the n data points are divided into k clusters. In case of melanoma skin cancer detection two clusters are formed: one cluster is for cancer and another one is for non-cancer. It has been observed that the accuracy for Back Propagation Neural Network is 60% to 75% and by using Support Vector Machine it is 80% to 90% . Thus, the paper shows that for the proposed system, Support Vector Machine gives better results than the K-means clustering technique and the Backpropogation Neural Network.

Technique Used	Advantages	Disadvantages
Backpropogation Neural Network	Weight Adjustment	Slow convergence rate and trapping in local minima
Convolution Neural Network	Eliminates the need for performing segmentation and feature extraction independently	Time consuming on a huge dataset
ABCD rule	Simplicity	Not useful in some cases Eg. When the diameter of lesion is less than 6mm
Neuro Fuzzy System	Accurate results by providing fuzzy values instead of crisp values	-
Hybrid Artificial Neural Network and world cup optimization	Solves the shortcomings of local minima that arise in backpropogation neural network	Complexity
Hybrid genetic algorithm – Artificial Neural Network	Uses genetic algorithm to escape local minima	-
Unsupervised algorithm – K means algorithm	Use of clusters to identify patterns. Thus, no need for labelled data	Less accuracy as compared to SVM and Neural networks
SVM	Accuracy	Not better than RF(Random Forest) in some cases since RF can handle categorical data, unbalanced data as well as data with missing values

3. PROPOSED METHODOLOGY AND FUTURE WORK

Skin Cancer Detection is a challenging task primarily due to two reasons: (i) the repercussions of incorrect detection and (ii) the need for excellent accuracy in detection. The problemof accuracy could be dealt with by using efficient soft computing and artificial intelligence techniques. We have seen that various techniques give fairly accurate results but they still left much to be desired. The SVM model that followed show a substantial improvement in the accuracy of skin cancer detection. By analysing the empirical results, we can see that unsupervised learning algorithms like k-mean clustering and certain neural networks like backpropogation

neural networks have certain disadvantages, SVMs are better at detection with a high degree of accuracy. A result having close to perfect accuracy is yet to be seen due to factors like insufficient database and failure of the proposed algorithms to achieve the ideal results.

Our study shows the use of SVMs and CNNs will be able to provide promising results if the results can be supplemented with additional intelligence. Hence, for future work in skin cancer detection, we recommend further research on using various soft computing and artificial intelligence algorithms with different classifiers like Random Forest (RF) classifier and analyzing the results for the same.

4. CONCLUSION

In this study, we have examined various non-invasive techniques for skin cancer classification and detection. The melanoma detection requires various stages like pre-processing, segmentation, feature extraction and classification. This survey focuses on different strategies like Genetic Algorithm, SVM, CNN, ABCD rule etc. As per the review, each algorithm is found to have its advantages and disadvantages. However, amongst the analysed algorithm, the SVM algorithm has the least amount of disadvantages and thus, outweighs other algorithms like K-mean clustering and Backpropogation neural networks. From all these existing approaches we can conclude that classification based on neural network technique is better than others.

REFERENCES

[1] Alquran, Hiam, et al. "The melanoma skin cancer detection and classification using support vector machine." Applied Electrical Engineering and Computing Technologies (AEECT), 2017 IEEE Jordan Conference on. IEEE, 2017.

[2] Kavitha, J. C., and A. Suruliandi. "Texture and color feature extraction for classification of melanoma using SVM." Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on.IEEE, 2016.

[3] Dubai, Pratik, et al. "Skin cancer detection and classification." Electrical Engineering and Informatics (ICEEI), 2017 6th International Conference on.IEEE, 2017.

[4] Razmjooy, Navid, Fatima Rashid Sheykhahmad, and NoradinGhadimi. "A hybrid neural network–world cup optimization algorithm for melanoma detection." Open Medicine 13.1 (2018): 9-16.

[5] urRehman, Mobeen, et al. "Classification of Skin Lesion by Interference of Segmentation and Convolotion Neural Network." 2018 2nd International Conference on Engineering Innovation (ICEI).IEEE, 2018.

[6] Quang, Nguyen Hong. "Automatic skin lesion analysis towards melanoma detection."Intelligent and Evolutionary Systems (IES), 2017 21st Asia Pacific Symposium on.IEEE, 2017.

[7] Aswin, R. B., J. Abdul Jaleel, and SibiSalim. "Hybrid genetic algorithm—artificial neural network classifier for skin cancer detection."Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014 International Conference on.IEEE, 2014.

[8] Ali, Aya Abu, and Hasan Al-Marzouqi. "Melanoma detection using regular convolutional neural networks."Electrical and Computing Technologies and Applications (ICECTA), 2017 International Conference on IEEE, 2017.

[9] Firmansyah, Hardian Robby, EntinMartianaKusumaningtyas, and FadilahFahrulHardiansyah. "Detection melanoma cancer using ABCD rule based on mobile device." Knowledge Creation and Intelligent Computing (IES-KCIC), 2017 International Electronics Symposium on IEEE, 2017.

[10] Marghoob, Ashfaq A., and Alon Scope. "The complexity of diagnosing melanoma." Journal of Investigative Dermatology 129.1 (2009): 11-13.

[11] Salah, Bareqa, et al. "Skin cancer recognition by using a neuro-fuzzy system." Cancer informatics10 (2011): CIN-S5950.

[12] Pennisi, Andrea, et al. "Melanoma detection using delaunay triangulation." Tools with Artificial Intelligence (ICTAI), 2015 IEEE 27th International Conference on.IEEE, 2015.

[13] Mhaske, H. R., and D. A. Phalke. "Melanoma skin cancer detection and classification based on supervised and unsupervised learning." Circuits, Controls and Communications (CCUBE), 2013 International conference on.IEEE, 2013.

[14] Lorena, Ana C., et al. "Comparing machine learning classifiers in potential distribution modelling." Expert Systems with Applications 38.5 (2011): 5268-5275.

[15] Brown, Iain, and Christophe Mues. "An experimental comparison of classification algorithms for imbalanced credit scoring data sets." Expert Systems with Applications 39.3 (2012): 3446-3453.