Computer Vision for Skin Cancer Diagnosis and Recognition using RBF and SOM

Abrham Debasu Mengistu *abrhamd@bdu.edu.et*
Bahir Dar University, Bahir Dar institute of technology
Faculty of computing
Bahir Dar, Ethiopia

Dagnachew Melesew Alemayehu *dagnachewm@bdu.edu.et*
Bahir Dar University, Bahir Dar institute of technology
Faculty of computing
Bahir Dar, Ethiopia

Abstract

Human skin is the largest organ in our body which provides protection against heat, light, infections and injury. It also stores water, fat, and vitamin. Cancer is the leading cause of death in economically developed countries and the second leading cause of death in developing countries. Skin cancer is the most commonly diagnosed type of cancer among men and women. Exposure to UV rays, modernize diets, smoking, alcohol and nicotine are the main cause. Cancer is increasingly recognized as a critical public health problem in Ethiopia. There are three type of skin cancer and they are recognized based on their own properties. In view of this, a digital image processing technique is proposed to recognize and predict the different types of skin cancers using digital image processing techniques. Sample skin cancer image were taken from American cancer society research center and DERMOFIT which are popular and widely focuses on skin cancer research. The classification system was supervised corresponding to the predefined classes of the type of skin cancer. Combining Self organizing map (SOM) and radial basis function (RBF) for recognition and diagnosis of skin cancer is by far better than KNN, Naïve Bayes and ANN classifier. It was also showed that the discrimination power of morphology and color features was better than texture features but when morphology, texture and color features were used together the classification accuracy was increased. The best classification accuracy (88%, 96.15% and 95.45% for Basal cell carcinoma, Melanoma and Squamous cell carcinoma respectively) were obtained using combining SOM and RBF. The overall classification accuracy was 93.15%.

Keywords: SOM, RBF, KNN, Digital Image Processing, Dermofit.

1. INTRODUCTION

Human Skin Cancer is a disease that appears on the outer layers of the skin which are caused when the skin cells are dead or damaged due to over exposure to Sun's ultraviolet radiation. But Skin cancer can also occur on areas of one's skin not ordinarily exposed to sunlight. The human skin is the largest organ in our body which provides protection against heat, light, infections and injury. It also stores water, fat, and vitamin D. [1] The Human skin is composed of two major layers called epidermis and dermis. The top or the outer layer of the skin which is called the epidermis composed of three types of cells flat and scaly cells on the surface called SQUAMOUS cells, round cells called BASAL cells and MELANOCYTES, cells that provide skin its color and protect against skin damage. The inner layer of the skin known as the dermis is the layer that contains the nerves, blood vessels, and sweat glands. There are three type of skin cancer Melanoma, Basal cell carcinoma and Squamous cell carcinoma. Skin cancer is diagnosed by physical examination and biopsy. Biopsy is a quick and simple procedure where part or all of the spot is removed and sent to a laboratory. It may be done by doctor or you can be referred to a

dermatologist or surgeon. Results may take about a week to be ready [1]. Dermatology imaging research believes that diagnosis of skin cancer can be automated based on a certain physical feature and color information which is the characteristics of different category of skin cancer.

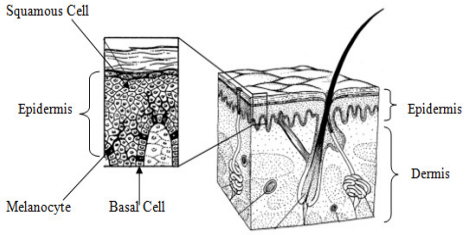

FIGURE 1: Layers of Skin.

2. LITERATURE REVIEW

Regarding with the benefits of early detection of human skin cancer, several dermatologists, medical professionals, medical industries, clinicians, computer scientists, academic researchers and technical experts have dedicated time and efforts to improve the early detection of human skin cancers. And some researchers have been done on the application on neural classifiers to skin injury classification purposes on this research paper presents a computer vision approach based on image processing algorithms and supervised learning techniques to help detecting and classifying wound tissue types which play an important role in wound diagnosis. The system proposed involves the use of the k-means clustering algorithm for image segmentation and a standard multilayer perceptron neural network to classify effectively each segmented region as the appropriate tissue type. [11]. Common classification method like statistical and rule based ones were applied in the researches to describe the diagnostic performance of Solar Scan, an automated instrument for the diagnosis of primary melanoma Images from a data set of 2430 lesions (382 were melanomas) were divided into a training set and an independent test set Solar Scan is a robust diagnostic instrument for pigmented or partially pigmented melanocytic lesions of the skin. [11].

More advanced techniques such as Neural Network were presented in the research [12] The aim of the study was to provide mathematical descriptors for the border of pigmented skin lesion images and to assess their efficacy for distinction among different lesion groups. New descriptors such as lesion slope and lesion slope regularity are introduced and mathematically defined descriptors was tested on a data set of 510 pigmented skin lesions, composed by 85 melanomas and 425 nevi, by employing statistical methods for discrimination between the two populations.

K-nearest neighborhood as another classification method was employed in the research of [12] on this research paper presents an algorithm for classification of non melanoma skin lesions based on a novel hierarchical K- Nearest Neighbors (KNN) classifier. The KNN classifier here, skin lesions are characterized by their color and texture. Finally, towards identification of human skin cancer uses the following common steps is image acquisitions, preprocessing, Segmentation, feature extraction, classification and finally the result will display to the user. [13]
In the recent years computational vision based diagnostic systems for dermatology have demonstrated significant progress. In this work, they review these systems by firstly presenting the visual features used for skin lesion classification and methods for defining them. Then they describe how to extract these features through digital image processing methods, i.e., segmentation, registration, border detection, color and texture processing) [14]. However, these imaging technologies are still expensive and may require specialized training to read the resulting images. Dermoscopy is the methodology for the exanimation of skin injuries based on imaging. This method provides a good and detailed view of the injuries. The imaging equipment used for taking the images is called Dermatoscope. It is handheld equipments which is compact and easy

to use. An oil film is placed between the lens of detematoscope and skin injuries. Main purpose of placing oil film is to obtain the magnified view of skin tissues [15].

The conventional diagnosing method for skin cancer is biopsy. It is a painful and time consuming technique. By incorporating Artificial intelligence and Digital Image Processing for skin cancer detection, it is possible to do the diagnosis without any physical contact with the skin. This can be implemented in a computer with the help of some software. Skin cancer detection system implemented using computer and software is known as Computer Aided Detection. The detection system is mainly based on Artificial intelligence and Digital Image Processing. Artificial intelligence has proven to be very efficient in decision making and pattern recognition applications. In this paper, the ANN Classifier is implemented in MATLAB software for skin cancer detection [16]. First stage in the skin cancer detection system is the input image. Dermoscopic image in digital format is given as input to the system. Dermoscopic image in digital format is given as input to the system. Next stage is the noise removal. The image contains hairs and other noises. These noises cause errors in classification. The noises are removed by filtering. Filtering method implemented here is the Median Filtering [16].

3. PROBLEM STATEMENT

On the previous researches there is a scope for the design of classifier to detect the type of cancer this provide a better and more reliable results for the patients, so that more patients can be diagnosed and cured. In line with this, human skin cancer identification is very useful in encouraging good quality in skin cancer diagnosis. There is a need for automated in recognition of human skin cancer systems so that the abuses during diagnosis and treatment can be minimized. Therefore, this thesis work will initiate a model for human skin cancer recognition which is consistent, efficient and cost effective by exploring the technology of image processing techniques. The ultimate goal is to ease the doctor's role in the recognition of skin cancer mentioned above by providing better and more reliable results, so that more patients can be diagnosed. The work on classifier design to detect the type of cancer will be taken in future [31]. Skin cancer is diagnosed by physical examination and biopsy. In case of physical examination the doctors will try to see the physical properties of skin cancers. When we see biopsy it is the procedure that the dermatologist takes some part or all of the spot and sent to a laboratory. It may be done by doctor or you can be referred to a dermatologist or surgeon. Results may take about a week to be ready [1]. To this end this study answers the following research question:

- ✓What appropriate image processing techniques used for human skin cancer recognition?
- ✓To what extent recognition effectiveness is registered for the human skin cancer?
- ✓What are the features that distinguish the three type of skin cancer?
- ✓What are the common features that the three type of skin cancer shares?
- ✓How to develop an automatic skin cancer recognition system based on image processing techniques?

4. DESIGN OF HUMAN SKIN CANCER RECOGNITION

The task of recognition occurs in wide range of human activity. The problem of recognition is concerned with the construction of a procedure that will be applied to differentiate items, in which each new item must be assigned to one of a set of predefined classes on the basis of observed attributes or features.

Accordingly, image analysis or computer vision is used in the recognition of human skin cancer to predefined classes. The predefined classes are the feature or attributes are computed from skin cancer images. These observed features of skin cancer were used to decide the class or the type of skin cancer. Hence, in this research the main interest is to differentiate the type of skin cancer varieties by using image analysis technique this is because of in order to maximize the curability of the disease if we identify the type of skin cancer where it belongs to it is very simple to cure the disease otherwise it is difficult.

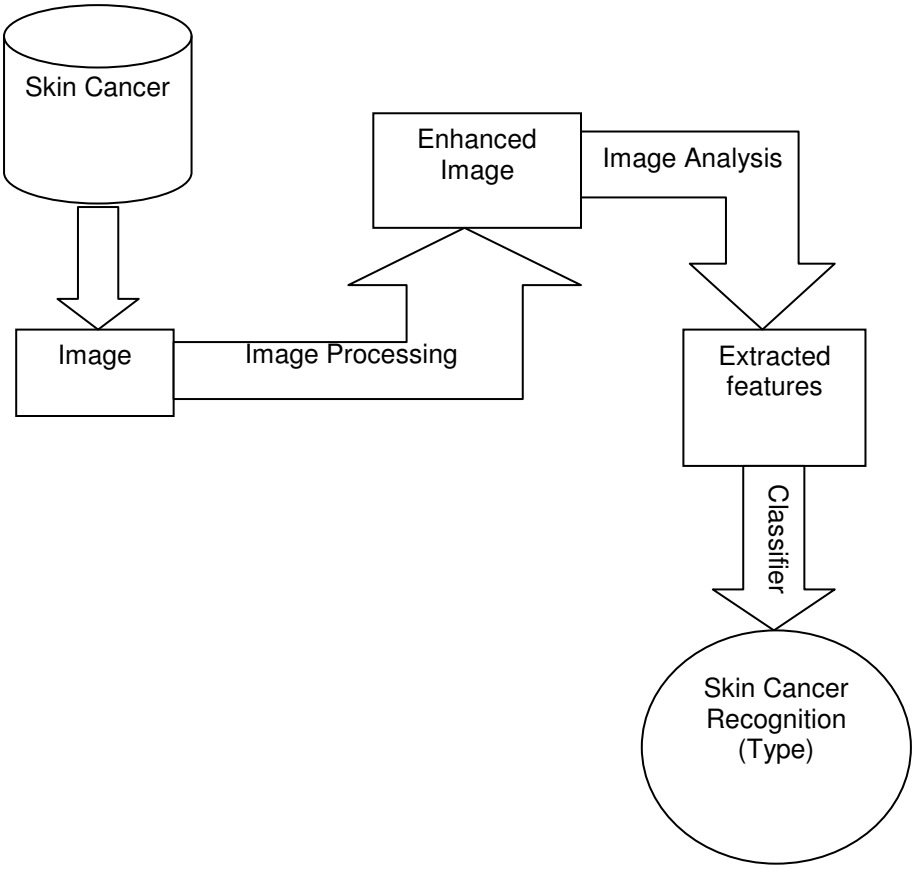

FIGURE 2: Skin Cancer Recognition Process Model.

As shown in the above figure 2, classification of human skin cancer involves the following activities:

Image acquisition of human skin cancer, an image processing techniques is applied on the acquired image to enhance the quality of image so as to reduce noises, appropriate feature are extracted from the enhanced image by using image analysis techniques that is used to classify dermoscopy images of skin cancer, the classification model of training and testing data of dermoscopy images of skin cancer will be developed, finally suitable pattern classifiers are selected to classify dermoscopy images of skin cancer to the predefined classes of skin cancer.

A. Image Acquisition

Image analysis starts with image acquisition this involves all aspects that have to be addressed in order to obtain dermoscopy image of human skin cancer the selection of radiation (light) sources and sensors (such as cameras) has to be considered very carefully. For this study, images have been taken from the following websites:

- ✓ https://www.dermquest.com/image-library/
- ✓ http://www.dermnet.com/images/

B. Image Processing

Image processing is a mechanism that focuses on the manipulation of images in different ways in order to enhance the image quality. Images are taken as the input and output for image processing techniques it is the analysis of image to image transformation which is used for the enhancement of image i.e. to increase the contrast for the input image and also restoration for geometrical distortion. [10]. Image segmentation is one of the most important tasks in image processing. It is the process of dividing an image into different homogeneous regions such that

the pixels in each partitioned region possess an identical set of properties or attributes [10]. The result of segmentation is a number of homogeneous regions, each having a unique label. Image segmentation is basically used to isolate region of interest from the background noise. For image processing techniques, we have used Matlab R2013a in which MATLAB is a high performance language for technical computing. MATLAB (MATRIX LABORATORY) is an interactive system for matrix based computation designed for scientific and engineering use. It is good for many forms of numeric computation and visualization. Hence, MATLAB was used for image processing tasks of Human skin cancer images to enhance the quality of image and to change images to binary for feature extraction purposes.From the original skin cancer images, the image is filtered in order to avoid other noises that are formed due to illumination effects as shown figure below.

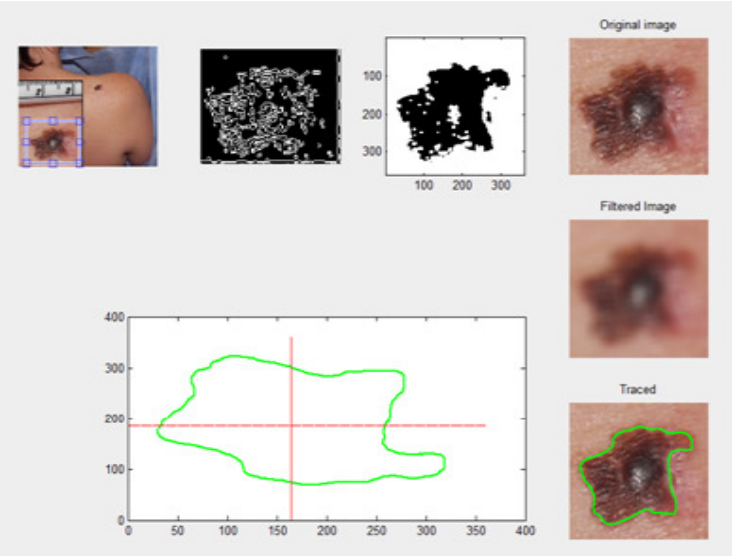

FIGURE 3: Image Processing.

C. Median Filtering

The Dermoscopic Image in Digital format is subjected to various Digital Image Processing Techniques. The standard image size is taken as 360x360 pixels [27]. Usually the image consists of noises in the form of hairs, bubbles etc. These noises cause inaccuracy in classification. In order to avoid that, images are subjected to various image processing techniques. One of the key element in image processing is filtering of image pre processing is done to removes the noise, fine hair and bubbles in the image. For smoothing image from noise, median filtering is used. Median filtering is a common step in image processing. Median filtering is used for minimizing the influence of small structures like thin hairs and isolated islands of pixels like small air bubbles [30].

D. Feature Extraction

Feature extraction is the method by which unique features of skin lesion images are extracted. This method reduces the complexity in classification problems. The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. We have the following three groups of features:

GLCM (Texture features of skin cancer): GLCM is a powerful tool for image feature extraction by mapping the grey level co occurrence probabilities based on spatial relations of pixels in different angular directions.

Morphological features: Morphology is the geometric property of a given image, in our case it is the size and shape characteristics of human skin cancer image.

Color features: Color is one of the features of skin cancer, they have different color variation of each cancer type and color analysis computed by taking the mean value of RGBs (Red, Green and Blue) components and the mean value of HSIs (Hue, Saturation and Intensity) components. Therefore, to compute the mean value of each component of these color spaces, we use matlab 2013 to split each component because matlab has built in function to convert to HIS or RGB color spaces. By using the built in function of matlab we can find RGB, the RGB color image stack is split to red, green and blue components. Hence, the color features are extracted by computing the mean values of RGBs and HSIs of Dermoscopy skin cancer images. That is, the mean value of red, mean value of green, mean value of blue, mean value of hue, mean value of saturation and mean value of intensity are computed from each component.

5. EXPERIMENTAL RESULTS

We extract 15 features (four morphology, five GLCM and six color features) were identified; hence, the total input features were fifteen. These features were used to classify different skin cancer image of human body. In line with this, we have designed experimental scenarios to test the classification performance by taking the extracted features of cancer images. The classifications were tested by four different algorithms namely ANN (Artificial Neural Network), KNN (Nearest Neighbor classification) , Naive Bayes and combining SOM and RBF classifiers in order to get a more accurate result. In order to train the classifiers, a set of training skin cancer image was required, and the class label where it belongs to, 235 skin cancer image were taken from American skin cancer society and DERMOFIT from the predefined three types of skin cancer i.e. Melanoma, Basal cell carcinoma and Squamous cell carcinoma.

There are two basic phases of pattern classification. They are training and testing phases. In the training phase, data is repeatedly presented to the classifier, in order to obtain a desired response. In testing phase, the trained system is applied to data that it has never seen to check the performance of the classification. Hence, we need to design the classifier by partitioning the total data set into training and testing data set. From the total data set of each skin cancer type 70% was used to build training and the remaining 30% of the total was used for testing data. From the total of 235 data sets, 162 were used for training and 73 were used for testing. In general, a classier has some input features based on the scenario of the designed experiment and some output features. In this study, there were three output classes, because the predefined human skin cancer images were three. The total numbers of exemplas or patterns were 235. This exemplas were normalized with mean 0 and variance 1.

KNN				ANN					Naïve				SOM		
	Basal cell	Melanoma	Squamous cell		Basal cell	Melanoma	Squamous cell		Basal cell	Melanoma	Squamous cell		Basal cell	Melanoma	Squamous cell
Basal cell	18	3	4	Basal cell	15	5	5	Basal cell	14	4	7	Basal cell	22	2	1
Melanoma	3	20	3	Melanoma	6	18	2	Melanoma	8	14	4	Melanoma	1	25	0
Squamous cell	6	2	14	Squamous cell	7	2	13	Squamous cell	7	2	13	Squamous cell	0	1	21
	Total	73			Total	73			Total	73			Total	73	
	correct	52			correct	46			correct	41			correct	68	
	not correct	21			not correct	27			not correct	32			not correct	5	
	%	71.232877			%	63.013699			%	56.164384			%	93.150685	

FIGURE 4: Summary result of KNN, ANN, Naïve and SOM classifier using extracted features.

As we have presented in detail in the previous section, the experiments were conducted under four scenarios by using feature sets of morphology, texture and color separately, and finally combining the three feature sets. Then, the experiment results were compared the performance of the Naïve Bayes classifier, KNN, neural network and SOM classification over the three

scenarios. The total number of data sets is 235. Out of these, 70% were used for training and the remaining 30% were used for testing. In general, the overall result showed that morphology and color features have more discriminating power than texture features and the classification performance of SOM is by far better than Naïve Bayes , artificial neural network classifier and KNN.

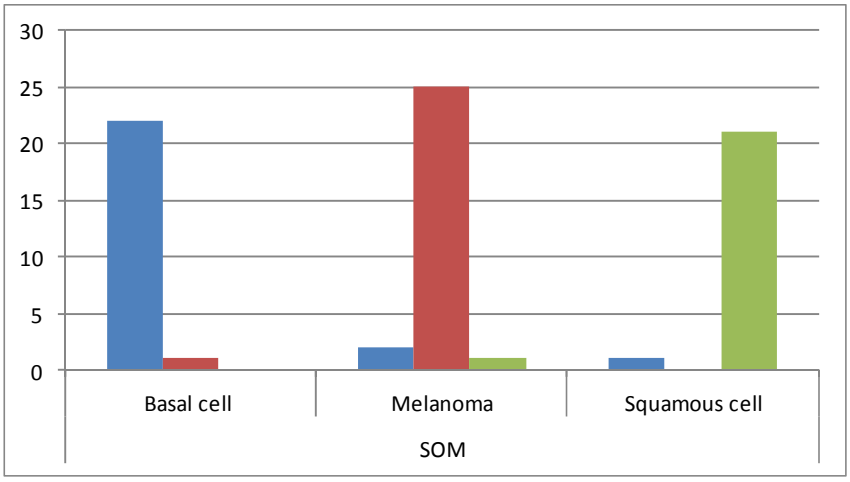

FIGURE 5: Overall Performance.

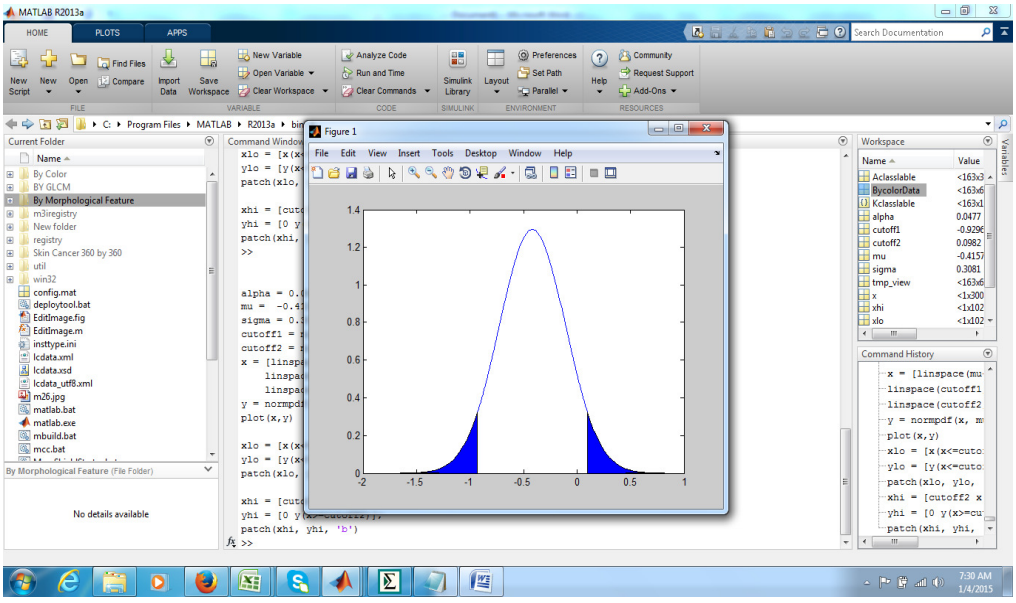

FIGURE 6: Confidence Interval of Detection.

6. CONCLUSION AND FUTURE WORK

In the classification problem of skin cancer recognition, morphological, GLCM and color features were extracted from a skin cancer images taken from three type of skin cancer Basal cell carcinoma, Melanoma and squamous cell carcinoma by using image analysis techniques. These selected features were used as input to the classification model. The experiment was conducted under four scenarios of the features data set such as GLCM, Morphology, Color and combining the three features. The result of the experimentation showed that the three varieties of Human skin cancer have been classified more accurately by SOM than using Naïve Bayes , ANN and KNN classifier. The image analysis for the recognition of the type of skin cancer can be further

investigated. The work can also be seen in depth and researched by the different characteristics of its physical and chemical in connection to image technology.

In light with this, the following recommendations are made for further research and improvements.

Identification of skin cancer type by exploring more features, Skin cancer recognition by levels of injuries using image analysis, implementing skin cancer recognition on mobile to make simplified for doctors, Computer vision for Recognition of leprosy and skin cancer.

7. REFERENCES

[1] http://www.cancer.org/cancer/skincancer

[2] Cancer in Africa, World Health Organization, International Agency for Research on cancer; 2009.

[3] http://www.skincancer.org/skin-cancer-information/skin-cancer-facts

[4] http://www.cdc.gov/cancer/skin/statistics/

[5] http://www.skincancer.org/skin-cancer-information/skin-cancer-facts

[6] World Health Organization, the global burden of disease: 2008 Update Geneva: World Health Organization,2008

[7] A Morrone, Skin Cancer in Ethiopia,21st world congress of dermatology, 2007

[8] http://www.skincancer.org/skin-cancer-information/skin-cancer-facts

[9] William K. Pratt: Digital image processing, PIKS Scientific inside, John Wiley, 4th Edition, 2007

[10] John Breneman towards early stages of malignant melanoma detection Using Consumer Mobile Devices

[11] Luís Filipe Caeiro M argalho Guerra Rosado, Automatic System for Diagnosis of Skin Lesions Based on Dermoscopic Images

[12] T Y Satheesha,, Dr. D Sathya Narayana, Dr. M N Giriprasad, review on early detection of melanoma:2012

[13] Ilias Maglogiannis, Dimitrios I. Kosmopoulos, Computational vision systems for the detection of malignant melanoma

[14] Ioana Dumitrache, Alina Elena Sultana, and Radu Dogaru Automatic Detection of Skin Melanoma from Images using Natural Computing Approaches

[15] Aswin.R.B, J. Abdul Jaleel, Sibi Salim, Implementation of ANN Classifier using MATLAB for Skin Cancer Detection,2013

[16] Ioana Dumitrache, Alina Elena Sultana, and Radu Dogaru Automatic Detection of Skin Melanoma from Images using Natural Computing Approaches

[17] Nilkamal S. Ramteke and Shweta V. Jain ABCD rule based automatic computer-aided skin cancer detection using MATLAB

[18] Santosh Achakanalli and G. Sadashivappa Skin Cancer Detection and Diagnosis Using Image Processing and Implementation Using Neural Networks and ABCD Parameters

[19] Mariam A.Sheha and Mai S.Mabrouk Automatic Detection of Melanoma Skin Cancer using Texture Analysis

[20] Catarina Barata, Margarida Ruela, Mariana Francisco, Teresa Mendonça, and Jorge S. Marques Detection of Melanomas in Dermoscopy Images Using Texture and Color Features

[21] Santosh Achakanalli & G. Sadashivappa, Skin Cancer Detection And Diagnosis Using Image Processing And Implementation Using Neural Networks And ABCD Parameters

[22] Luís Filipe Caeiro M argalho Guerra Rosado, Automatic System for Diagnosis of Skin Lesions Based on Dermoscopic Images

[23] Mariam A. Sheha , Mai S. Mabrouk , Amr Sharawy, Automatic Detection of Melanoma Skin Cancer using Texture Analysis

[24] Lin Li, Qizhi Zhang, Yihua Ding, Huabei Jiang, Bruce H Thiersand, Automatic diagnosis of melanoma using machine learning methods

[25] Sarika Choudhari, Seema Biday, Artificial Neural Network for SkinCancer Detection

[26] Dr. J. Abdul Jaleel, Sibi Salim, Aswin.R.B, Diagnosis and Detection of Skin Cancer Using Artificial Intelligence

[27] Peyman Sabouri, Hamid GholamHosseini, John Collins, Border Detection of Mela noma Skin Lesions, 2013

[28] Dr. J. Abdul Jaleel, Sibi Salim, Aswin.R.B, Artificial Neural Network Based Detection of Skin Cancer,2012

[29] Snehal Salunke, Survey on Skin lesion segmentation and classification,2014

[30] Nandini M.N., M.S.Mallikarjunaswamy, Detection of Melanoma Skin Disease using Dermoscopy Images

[31] Tinku Acharya and Ajoy K. Ray, Image Processing Principles and Applications, Jhon Wiley, 2005.

[32] William K. Pratt: Digital image processing, PIKS Scientific inside, John Wiley, 4th Edition, 2007.