https://dx.doi.org/10.22161/ijaers.6.5.53 ISSN: 2349-6495(P) | 2456-1908(O)

Skin Disease Detection And Classification

Amruta D. Joshi, Shreya S. Manerkar, Vinita U. Nagvekar, Kalpita P. Naik, Chinmay G. Palekar, Prof. V. Pugazhenthi, Prof. Sagar K. Naik

Department of Computer Engineering, Agnel Institute of Technology and Design, Assagao, Bardez-Goa

Abstract*—Skin diseases are most commonly occurring in people of all ages and are caused by bacteria, infection or radiation. These diseases have various dangerous effects on the skin and keep on spreading over time. A patient can recover from skin diseases if it is detected and treated in the early stages and this can achieve cure ratios of over 95%. Hence, it is important to identify these diseases at their initial stage to control them from spreading. Skin diseases are primarily diagnosed visually, beginning with an initial clinical screening and followed potentially by dermoscopic analysis. Such a system is often prone to errors. The main idea of this project is to improve the accuracy of diagnostic systems by using Image Processing and classification techniques. In the proposed system, an image captured on camera is taken as input. This image will be pre-processed in order to make it suitable for segmentation by using Contrast Enhancement and Grayscale Conversion. Global Thresholding technique is used to segment the pre-processed image through which the actual affected region is obtained. Texture features, such as Energy, Entropy, Contrast, IDM, are extracted from the segmented image using Grey Level Co-occurrence Matrix. Image Quality Assessment features such as MSE and PSNR are extracted. The extracted texture features will be used to detect the presence of skin disease and classify the disease as melanoma, leprosy or eczema, if present, using the Decision tree technique.*

Keywords*— *Global Thresholding, Grey Level Co-occurrence Matrix, Histogram Equalization, Image Processing, Texture features.

I. INTRODUCTION

1.1. Skin Diseases

Skin is the largest and most sensitive part of the human body which protects our inner vital parts and organs from the outside environment, hence avoiding contact with bacteria and viruses. Skin also helps in body temperature regulation. The skin consists of cells, pigmentation, blood vessels, and other components. It is comprised of 3 main layers, namely, the epidermis, the dermis, and the hypodermis.

Epidermis, being the outermost skin layer, forms a waterproof and protective sheath around the body's surface. The dermis, found beneath the epidermis, comprises of connective tissues and protects the body from stress and strain. A basement membrane tightly joins the dermis with the epidermis. The hypodermis, also called subcutaneous tissue, is not actually a part of theskin and lies below the dermis. It attaches the skin to the underlying bone and muscle and also supplies blood vessels and nerves to it.

Skin diseases occur commonly among humans. They are usually caused by factors like different organism's cells, a different diet, and internal and external factors, such as the hierarchical genetic group of cells, hormones, and immune system of conditions. These factors may act together or in a sequence of skin disease. There are chronic and incurable diseases, like eczema and psoriasis, and malignant diseases like malignant melanoma. Recent researchers have found the availability of cures for these diseases if they are detected in the early stages.

Atopic dermatitis, commonly called eczema, is a long-term skin disease whose common symptoms are dry and itchy skin, rashes on the face, inside the elbows, behind the knees, and on the hands and feet.

Hansen's disease, commonly called leprosy, is caused by slow-growing bacteria and can affect the body and facial parts like nerves, skin, eyes and nose lining.

Melanoma is severe and life-threatening skin cancer. The "ABCD's" of moles detected on the skin are Asymmetry, Border, Colour, and Diameter. Asymmetry implies that the shape of one half does not match the other half. Border means the edges of the mole are ragged, blurred, or irregular. Colour is uneven and may include shades of black, brown, and tan. The diameter of mole implies a change in size.

1.2. Skin Disease Detection System

Skin diseases are primarily diagnosed visually, beginning with an initial clinical screening and followed potentially by dermoscopic analysis. To ascertain what type of skin disease a person has, they must visit a dermatologist. The dermatologist then performs visual analysis using various tests, some of which include:

1. Patch test: Known allergens are applied to a patch of skin and left for some time. The skin is then tested for a reaction.

https://dx.doi.org/10.22161/ijaers.6.5.53 ISSN: 2349-6495(P) | 2456-1908(O)

2. Biopsy: Skin is removed using a scalpel, a blade or a biposy tool and taken to a laboratory for analysis.

3.Culture: Skin of affected area or hair or nails are cultured to determine which microorganism is causing the infection.

Such a system is often time consuming and requires a number of expert professionals. Since there are people involved in this process, it is prone to human errors. This system is also quite expensive as laboratories charge a lot of fees for the tests.

A skin disease detection and classification system is a system used for detecting whether a disease is present or not, and then classifying the type of disease, if present. The classification is based on decisions taken using the features extracted through the feature extraction methods.

In order to identify whether a disease is present or not, the system must be trained to recognize normal conditions of system activity. There are two main phases for this purpose: training phase (building a profile using data about a particular disease) and testing phase (comparing the current image data with the trained image data).

Md Nafiul Alam et al., "Automatic Detection and Severity Measurement of Eczema Using Image Processing", suggested an automatic eczema detection and severity measurement model using image processing and computer algorithm. The system basically allowed patients to take an image of the affected skin area and identify and determine the severity of eczema. This system included image segmentation, feature extraction, and statistical classification to identify and differentiate between mild and severe eczema. Once the eczema type was identified, a severity index was assigned to that image. Several common segmentation methods, like Otsu's, watershed, and region growing segmentation, were implemented, and since none of these methods gave correct outputs, color-based segmentation using k-means clustering was used. Feature extraction was based on color features, texture features, and border features and the classification was done using the SVM classifier method.

Benefits of this system were it provided a faster and easier way of detection of eczema, and its severity if detected. Its main limitation was that the system could be more accurate if the image database had a large number of calibrated images.[2]

Sumithra R. et al., "Segmentation and Classification of Skin Lesions for Disease Diagnosis", suggested a system for automatic segmentation and classification of skin lesions. Initially, skin images were pre-processed to remove visible skin hair and noise. These filtered images were then segmented using Region growing segmentation methodand feature extraction of color and texture features was performed. Based on the extracted features, the images were classified as a particular disease using Support Vector Machine(SVM) and K-Nearest Neighbour(KNN) classifier methods. The advantage that this system provides is better performance results by fusion of SVM and KNN methods instead of using either one of both.

Its disadvantage is that classification system performance reduces due to the effect of some classes, thereby affecting the overall performance of the system.[3]

Sheha et al.(March 2012) suggested an automated system for discrimination between melanocytic nevi and malignant melanoma avoiding segmentation process using texture analysis. Texture analysis refers to the characterization of regions in an image by their texture content. The study used 102 images from skin diseases Atlases and dermoscopy clinics. The process had 4 phases. First was pre-processing, which included resizing of images and RGB to grey conversion. The second phase was feature extraction done using grey level co-occurrence matrix (GLCM). Next phase was feature selection, which used The Fisher score ranking to select the most important texture features. The last phase was classification, using Multilayer Perceptron (MLPs). The advantage of this method was that it had a high accuracy of 93%.

In conclusion, this work showed that the combination between the co-occurrence matrix and ANN is a promising technique for discrimination between malignant melanoma and melanocytic nevi dermoscopy images.[4]

Table 1: Comparison of human skin diseases detection models

Title	Algorithms Used	Accuracy	Advantages	Disadvantages	Number of diseases classified
Classification of Skin Disease Using multi SVM Classifier	-C-means -Watershed -GLCM, IQA -Multi-SVM classifier using Matlab	High	-Gives high quality results even with a smaller training dataset	-Time complexity is high	2
Automatic and severity Measurement of Eczema Using Image Processing	-Region growing segmentation -Color based segmentation -Texture, color and	High	-A faster and easier way of detection with high accuracy	-Image database lowers the accuracy level	1

https://dx.doi.org/10.22161/ijaers.6.5.53 ISSN: 2349-6495(P) | 2456-1908(O)

	border based feature extraction -SVM classification				
Segmentation and Classification of Skin Lesions for Disease Diagnosis	-Region growing segmentation -Color and texture based feature extraction -SVM and KNN classification	Moderate	-Provides better performance results using a fusion of SVM and KNN methods	-Due to some classification classes, performance reduced -Either method alone couldn't perform better than the fusion of both	1
Automatic Detection of Melanoma Skin Cancer	-Fisher score ranking -Multi Layer Perceptron classifier method	High	-Gives better performance by skipping the segmentation method	-Can give errors if correct features are not selected	1

II. SYSTEM DESIGN

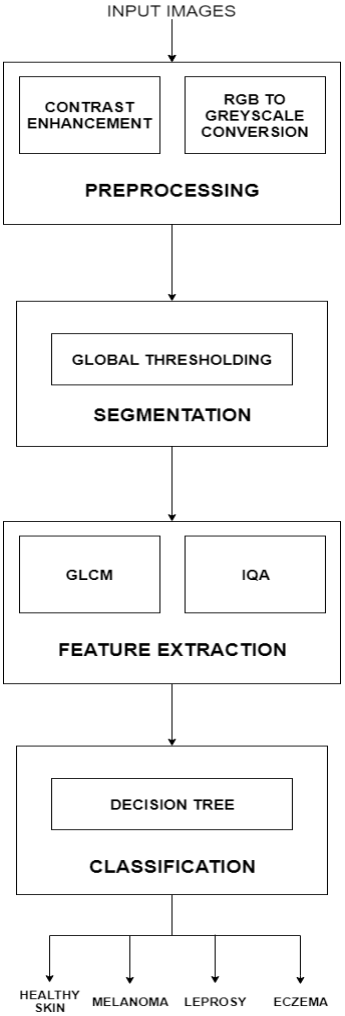


Fig. 1: flow diagram of the proposed system

2.1 Proposed System

The proposed idea aims at developing a system to provide early and easy detection of skin diseases, if any, using image processing and classification techniques. Here, the image data will be pre-processed and given to the segmentation algorithm. The segmented data will be passed to feature extraction methods, and the features thus extracted will be used by the classification algorithm to detect the presence of skin disease.

2.2 System Flow

The system takes as input an image of the affected skin region and feeds it to the pre-processing techniques, namely, Contrast enhancement in order to improve clarity and contrast of the image, followed by RGB to Grayscale conversion. The processed image is segmented in order to distinguish the actual affected part(foreground) from the remaining area(background). Next, the segmented image is sent for feature extraction. These features, in turn, will be used as input to classify which disease the image is referring to.

2.2.1 Image pre-processing

Image pre-processing consist of two sub-processes, Contrast Enhancement, and Grayscale Conversion. A raw binary image is converted into the RGB (red/green/blue) Matrix form. The RGB Matrix is first processed for Contrast Enhancement and converted to contrast enhanced RGB Matrix, this is done so as to distinguish each pixel with its neighbour. Contrast enhancement is performed using Histogram equalization[5]. The contrast enhanced RGB Matrix is converted into Grayscale Matrix[6].

2.2.2 Segmentation

Thresholding

The selection process is called Thresholding. Image Thresholding is a simple technique to partition an image into a foreground and background. It is one of the simplest approaches for image segmentation based on intensity levels. Thresholding can be implemented locally or globally.

Global Thresholding

Global thresholding[7] partitions an image into objects and background. It is the simplest of all thresholding. Segmentation is then achieved by scanning each pixel and

https://dx.doi.org/10.22161/ijaers.6.5.53 ISSN: 2349-6495(P) | 2456-1908(O)

labeling it as background or foreground depending on the grey level of that pixel.

2.2.3 Feature Extraction

This feature extraction is the process where we will be representing a raw image in a reduced form to help decision making pattern detection or classifications. The process involves reducing the amount of the resources which are being required to describe that in the larger sets of the data.

This process is widely used in machine learning. This is the general terms of the methods where we can construct combinations of the variable to get along with these problems while we can describe the data with high accuracy. Hence, we have divided this process into two methods for better feature extraction:

1. GLCM (Grey Level Co-occurrence Matrix)

In this method,the texture of the image is analyzed. The grey level co-occurrence matrix[8] is created by counting the number of times each pair of those specific values in a specified spatial relationship occur in the image. The texture features extracted from the grey level co-occurrence matrix are energy, entropy, contrast, IDM, correlation, and ASM.

2. Image Quality Assessment.

Image quality assessment[9] is one of the quality assessment methods.These are defined by the Full reference method. Image Quality Assessment Features MSE(Mean Square Error) and PSNR(Peak Signal to Noise Ratio) are extracted from the segmented image.

Full-reference method

Full reference method metric will always try to access the quality of the test images by calculating and comparing this with the referenced images, which can then be assumed to be having the perfect qualities.

This metric is the mean squared error (MSE), computed by averaging the squared intensity differences of distorted and reference image pixels, along with the related quantity of peak signal-to-noise ratio (PSNR).

2.2.4 Classification

Classification is the process of identifying to which category the input data belongs.

Decision Trees

Decision trees[10] use a tree like structure in which decisions and their possible outcomes are represented. A decision tree has a root node which is divided further into child nodes. A decision tree has three types of nodes: chance nodes, decision nodes, and end nodes.A chance node, denoted by a circle, represents the probabilities of certain results. A decision node, represented by a square, shows a decision to be made, and an end node shows the final outcome of a decision path.

ID3 Algorithm

In decision tree learning, ID3 (Iterative Dichotomiser 3) [11] is an algorithm used to construct a decision tree for a given dataset.In the beginning, set S is the root node. At each level of the tree, it iterates through every unused attribute of the set S and calculates the entropy(S) or information gain IG(S) of that attribute. It then selects the attribute which has the smallest entropy or largest information gain value. The set S is then split or partitioned by the selected attribute to produce subsets of the data.

III. RESULTS

The system proposed is a Skin Disease Detection System. This system uses images of skin captured with a camera to detect if it is healthy or not; if not, then classified as Melanoma, Eczema or Leprosy. The proposed system uses image processing and machine learning techniques. The process begins with pre-processing an input image using contrast enhancement and grayscale conversion. The contrast enhanced image is converted to a grayscale image. Global Value Thresholding technique is used to segment the grayscale image through which the actual affected region is obtained.

Output images of each of the above mentioned processes are mentioned displayed in the GUI in Fig 4. Grey Level Co-occurrence Matrix is created from the segmented image and Texture features such as Energy, Entropy, Contrast, IDM, correlation and homogeneity are extracted from it. Image Quality Assessment Features MSE(Mean Square Error) and PSNR(Peak Signal to Noise Ratio) are extracted from the segmented image. The values of all the features calculated are displayed in the GUI as well.

The texture features are used to classify the test images. The accuracy calculated using 45 test images is 87%, giving an error of 6 out of 45 images.

Input Image:

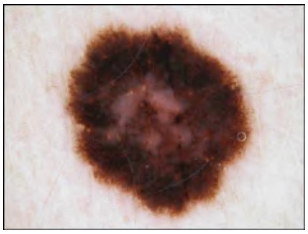
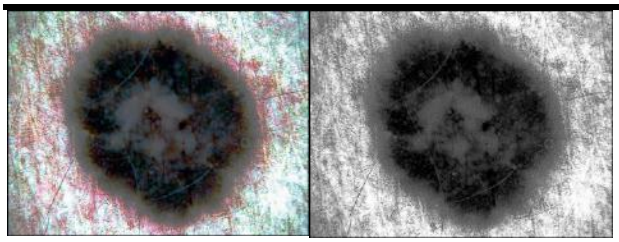


Fig 2: melanoma input image

Pre-processing:

https://dx.doi.org/10.22161/ijaers.6.5.53 ISSN: 2349-6495(P) | 2456-1908(O)



(a) (b)

Fig. 3: (a) contrast enhancement using histogram equalization
(b) grayscale conversion

Segmentation:

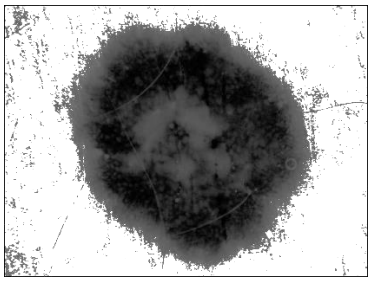


Fig 3: Segmentation using global thresholding

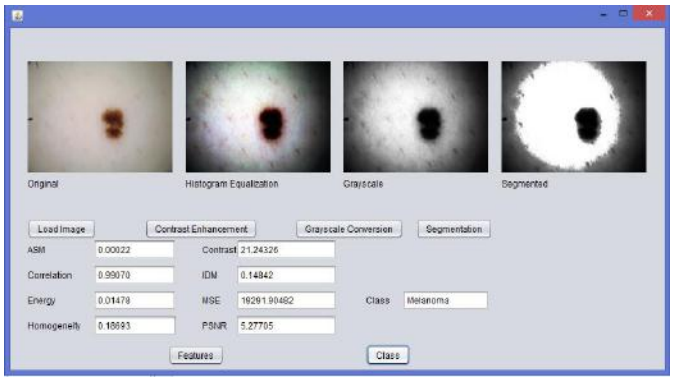


Fig 4: GUI for the system

IV. CONCLUSION

The system proposed is a Skin Disease Detection System. This system uses images of skin captured with a camera to detect if it is healthy or not; if not, then classified as Melanoma, Eczema or Leprosy.

The proposed system uses image processing and machine learning techniques. The process begins with pre-processing an input image using contrast enhancement and grayscale conversion. Global Value Thresholding technique is used to segment the pre-processed image through which the actual affected region is obtained. Features, such as Energy, Entropy, Contrast, IDM, Correlation, and ASM are extracted from the segmented image using Grey Level Co-occurrence Matrix. Image Quality Assessment features like MSE and PSNR are also extracted. These features will then be used to classify the image as healthy or into one of the 3 diseases: Melanoma, Eczema, and Leprosy.

This system can be used by dermatologists to give a better diagnosis and treatment to the patients. The system can be used to diagnose skin diseases at a lower cost. In future, this system can be improved to detect and classify more diseases as well as their severity.

REFERENCES

[1] Manerkar Mugdha S, HarshShashwata, SaxenaJuhi, SarmaSimanta P, Dr. U. Snekhalatha, Dr.M. Anburajan.(2016). Classification of Skin Disease Using multi SVM Classifier.*3rd International Conference on Electrical, Electronics, Engineering Trends, communication, Optimization and Sciences (EEECOS)*, 363-368.

[2] Alam, Md & Munia, Tamanna Tabassum Khan & Tavakolian, Kouhyar & Vasefi, Fartash & MacKinnon, Nicholas & Fazel-Rezai, Reza. (2016).Automatic and severity Measurement of Eczema Using Image Processing.*Engineering in Medicine and Biology Society Conference - August 2016*.1365-1368

[3] Sumithra R, Mahamad Suhil, D.S.Guru. (2015). Segmentation and Classification of Skin Lesions for Disease Diagnosis.*International Conference on Advanced Computing Technologies and Applications (ICACTA-2015)*, *45*, 76 – 85.

[4] Sheha Mariam A., MabroukMai S, Sharawy Amr. (2012). Automatic Detection of Melanoma Skin Cancer.*International Journal of Computer Applications (0975 – 8887)*, *42*(20), 22-26

[5] https://en.wikipedia.org/wiki/Histogram_equalization

[6] https://www.tutorialspoint.com/dip/grayscale_to_rgb_conversion.htm

[7] https://en.wikipedia.org/wiki/Thresholding_(image_processing)

[8] https://www.ucalgary.ca/mhallbey/glcm1

[9] https://en.wikipedia.org/wiki/Image_quality#Image_quality_assessment_methods

[10] https://en.wikipedia.org/wiki/Decision_tree

[11] https://www.cise.ufl.edu/~ddd/cap6635/Fall-97/Short-papers/2.htm