Detection and Classification of Skin Diseases

Prof. Latha A[1], Akash Kumar Jaiswal[2], Bindushree G L[3]

[1]Dept. of Computer Science, Sapthagiri College of Engineering, India
[2]Dept. of Computer Science, Sapthagiri College of Engineering, India

---

Abstract-Skin diseases are among the foremost common health problems worldwide and are related to substantial burden. Chronic and incurable skin diseases, such as Dermatitis and Eczema, are associated with significant morbidity in the form of physical discomfort and impairment of patients' quality of life. This system proposes a skin disease detection method based on image processing technique. This system is Deep Learning and Machine Learning based and hence very accurate results are obtained. The patient provides a picture of the infected area of the skin as an input to the prototype. Deep Learning techniques are performed on this image and therefore detected skin disease is displayed as the output.

Keywords---Dermatology, **Computer Vision, Image Processing, Deep Learning, Machine Learning, Convolutional Neural Network.**

1. INTRODUCTION

In today's real time scenario of daily life, computer vision methodology has attracted researchers due to its nature of providing efficient information for better visual and experimental analysis. In computer vision approach, image classification is also a promising technique which is used for various applications such as pattern recognition, remote sensing application, medical image processing etc. It is a process of pixel sorting from image and accumulating into individual classes. For classification, various methods have been developed to classify and recognize the image class efficiently. These techniques are categorized as follows:

- Supervised image classification
- Unsupervised image classification
- Object based image classification

The Dermatology remains the foremost uncertain and sophisticated branch of science due to its complicacy in the procedures involved in diagnosis of diseases associated with hair, skin, nails. The variation in these diseases are often seen due to many environmental, geographical factor variations. Human skin is taken into account the most uncertain and troublesome terrains due to the existence of hair, its deviations in tone and other mitigating factors. The skin disease diagnosis includes series of pathological laboratory tests for the identification of the right disease. For the past ten years these diseases are the matter of concern as their sudden arrival and their complexities have increased the life risks. These Skin abnormalities are very infectious and need to be treated at earlier stages to avoid it from spreading. Total wellbeing including physical and mental health is also affected adversely. Many of those skin abnormalities are very fatal particularly if not treated at an initial stage. Human mindset tends to presume that the majority skin abnormalities aren't as fatal as described thereby applying their own curing methods. However if these remedies are not apt for that selective skin disease then it makes it even worse. The existing diagnosis procedure consists of long laboratory procedures but this paper proposes a system which will enable users to predict the disease of the skin using computer vision.

Computer vision, a scientific field dealing with the way computers will extract the required information from digital photos. Computer vision tasks embody different techniques for acquiring, processing, analyzing and understanding digital images. Image processing is a method to perform some operations on a image which includes many methods, in order to get an enhanced image or to extract some useful information from it. Image processing methods are becoming increasingly sophisticated and the tendency is to develop as much automation as possible.

2. MOTIVATION

In assigning health priorities, skin diseases are sometimes thought, as small-time players within the worldwide league of illness compared with diseases that cause significant mortality, such as HIV/AIDS. However, skin problems are the most common diseases generally seen in medical aid settings in tropical areas, and in some regions where transmissible diseases like tinea imbricata are endemic, they become the predominant exposition.

Although mortality rates are generally lower than for other conditions, people's needs for effective remedies for skin conditions should be met for a number of important reasons.

•First, skin diseases are so common and patients present in such large numbers in medical care settings that ignoring them isn't a viable option. Children, especially, tend to be affected, adding to the burden of disease among an already vulnerable group.

•Second, morbidity is significant through disfigurement, disability, or symptoms such as intractable itch, as is the reduction in quality of life.

•Third, the relative economic cost to the families of treating even trivial skin complaints limits the conduction of therapies. Generally, families must meet such costs from an overstretched household budget, and such expenses in turn reduce the capacity to purchase such items as essential food.

• Lastly, checking the skin for signs of disease is an important concept for a wide range of illnesses, such as leprosy, yet a basic knowledge of the simple features of disease whose presenting signs occur within the skin is usually lacking at the first care level.

3. LITERATURE SURVEY

Many cases of skin diseases within the world have triggered a requirement to develop an efficient automated screening method for detection and diagnosis of the world of disease [2]. In this system, a study of the role of color information in detecting the edges of images was conducted. Therefore another color space (HIS) is implemented. Several edge detection techniques are used like Laplace and Perwitt, the outputs shows that the Laplace operator is more efficient than Perwitt operator in the edge detection. Wavelet Transform plays a crucial role within the image processing analysis, especially in texture recognition of knowledge. For its fine result when using Multi resolution modeling. The texture images are going to be entered to Wavelet Mother Function; this may segment the feel into sub bands. These sub bands contain information about the feel , then this information are going to be entered to feature extraction, the output from them represent the input to the synthetic Neural Network (ANN) which represents powerful tool for handling problems of huge dimension.

Skin Diseases are becoming very common now days. Number of individuals affected by skin diseases is increasing rapidly. Human judgment on diagnosis of skin diseases is usually subjective and not reproducible. To achieve more reliable and objective accuracy computer aided diagnosis may be used [3]. With advancement in medical imaging, image based classification is highly used for the detection of diseases in medical field. Feature engineering is extremely important for any classifier to realize maximum results. Convolution Neural Networks (CNN) can learn features on its own reducing total time required for development of such systems and at an equivalent time increasing level of accuracy. They have acquired more than 850 original images for two skin diseases from department of Skin and VD, KEM Hospital, Mumbai. The images are processed manually such as cropping and resizing. The deep learning algorithm CNN, is used for feature extraction from input images of two skin diseases. The features obtained are then fed to Support Vector Machine (SVM) for classification. The results indicate CNN are often feasibly used of feature extraction in disease of the skin detection.

Figure 1: Image before and after cropping [3]

Median filter [1] is the nonlinear filter more used to remove the impulsive noise from an image. Furthermore, it's a more robust method than the normal linear filtering, because it preserves the sharp edges. Median filter may be a spatial filtering operation, so it uses a 2-D mask that's applied to every pixel within the input image. To apply the mask means to centre all the features during a pixel conversion, evaluating the covered pixel brightnesses and determining which brightness value is the median value. Figure.2 [1] presents the concept of spatial filtering supported a 3x3 mask, where I is that the input image and O is that the output image.

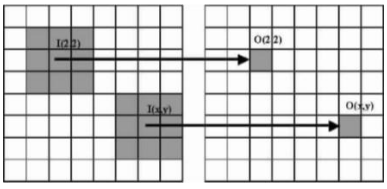


Figure 2: Median Filter Operation [1]

The system [7] provides an Expert System for diagnosis the skin disease for the existing problems of skin diseases. The system is divided into below mentioned major parts.

- Image preprocessing, segmentation and feature extraction.
- Classification model and skin disease predication.
- Medical treatment suggestions or advice.

The proposed expert system mainly consists of following components;

A. Image processing unit.
B. Data mining unit

Image processing is that the main a part of the planning process during this expert system. Various algorithms like median filter, Gaussian filter are used to removes the background noises such as hair and air bubbles and other noises in the skin disease image. Then image segmentation is done using threshold segmentation which helps to separate infected area of the image from healthy skin. Then Morphological transformations is used for eliminating similar type objects. watershed is used to identify the disease area of the skin and finally we extracted features (Color and shape features) of skin disease image.

In Data Mining Unit, the features obtained from the image processing unit will be examined and the disease will be identified and displayed by this unit. Then some questions like age, gender, affected areas within the body, how long it had been infected etc. will be asked from the user to provide medical treatment suggestions. Here because the classification model, they have used five different processing classification algorithms (AdaBoost, BayesNet, J48, MLP and NaiveBayes) and the one which yields the good results, the best classifier is choosen .

Another Architecture [4] provides an automated image based system for recognition of skin diseases using machine learning classification. This system will utilize computational technique to research , process, and relegate the image data predicated on various features of the pictures . Skin images are filtered to get rid of unwanted noise and also process it for enhancement of the image. Feature extraction using complex techniques like Convolutional Neural Network (CNN), classify the image supported the algorithm of softmax classifier and acquire the diagnosis report as an output. This system will give more accuracy and can generate results faster than the normal method, making this application an efficient and dependable system for dermatological disease detection. Furthermore, this will even be used as a reliable real time teaching tool for medical students within the dermatology stream.

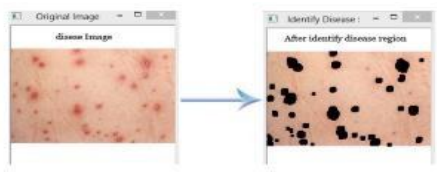


Figure 3: Disease region identification [4]

Another system [8] uses Computer Vision as the first stage in identification of the type of skin disease based on the numerous features extracted from the image using various image processing techniques. The computer vision stage itself consists of two phases. In the first phase, we pre- process the image taken through the camera of the smart phone in order to extract the necessary features. The second phase involves using the features extracted in order to identify the disease using various algorithms like Maximum Entropy Model and Artificial Neural Networks. The features that were extracted includes the color code of the inflicted area, size of infliction, contrast of the infliction with respect to the surrounding the healthy skin, shape through edge detection of the infected areas. The system uses various machine learning algorithms like ANN, Decision Tree and KNN in the second stage in order to refine the classification of the image. The second stage of prediction is made available to the medical professionals who have access to various histopathological attributes like exocytosis, Hyperkeratosis, acanthosis, parakeratosis and other attributes. The system takes these attributes as input from the user and gives a better classification of the diseases.

4. METHODOLOGY

Figure. 4 shows the overall architecture of the proposed system. The system can be broadly classified into Data Collection. Image Processing, Feature Extraction and Image Classification modules.

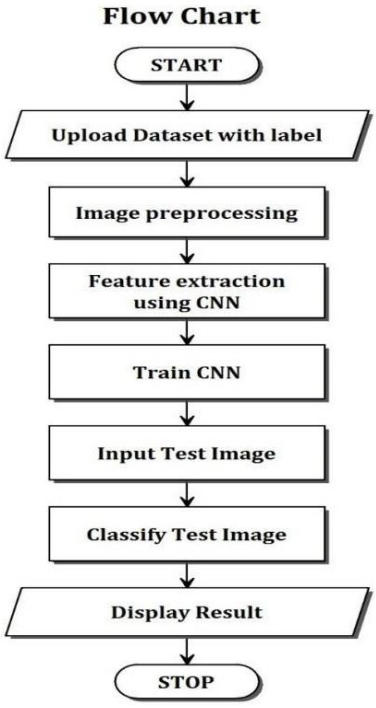


Figure 4: Flow diagram of the system

A.DATA COLLECTION

The diseases that are detected are listed below:

Acne, Benign, Dermatitis, Eczema, Melanoma, Nevus, Pigmented, Seborrhae Keratotis, Squamous Cell coranoma and Vascular Lesion.

Datasets are collected from the images available online and also the images from the dermnet are used (dermnet.com).We have used 3099 datasets as a training samples and 595 datasets as a testing samples.

B.IMAGE PREPROCESSING

There may be presence of some distortions while taking a picture, at times, you want to get rid of distortions caused by lights and shadows in an image. Normalizing the RGB values of an image can at times be a simple and effective way of achieving this. When normalizing the RGB values of an image, each pixel's value is divided by the sum of the pixel's value over all channels. So if we have a pixel with intensified R, G, and B in the respective channels, its normalized values will be R/S, G/S, and B/S (where, S=R+G+B).

C.FEATURE EXTRACTION USING CNN

Deep learning is a class of machine learning algorithms that uses multiple layers for extracting higher level features from the raw input. For example, in the technique image processing , lower layers are used to identify edges, while higher layers are used to identify the concepts relevant to a person's like digits or letters or faces.

The image is skilled a series of convolutional, nonlinear, pooling layers and fully connected layers, then generates the output. Figure.5 (google.com) depicts the layers in the CNN.

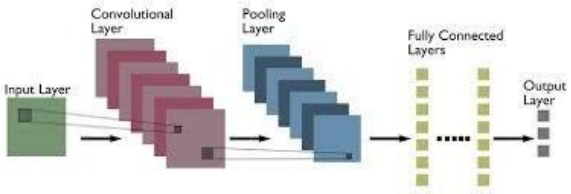


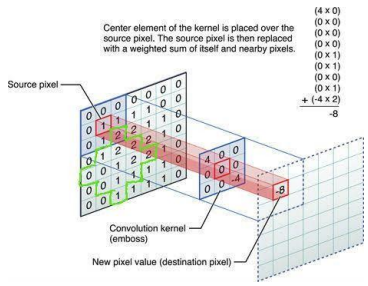
Figure 5: Layers in CNN

i. Covoutional layer:

The image (matrix with pixel values) is entered into this layer. Imagine that the reading of the input matrix begins at the highest left of image. Next the software selects a smaller matrix there, which is named a filter (or neuron, or core). Then the filter produces convolution, i.e. moves along the input image. The filter's task is to multiply its values by the first pixel values. All these multiplications are summed up. One number is obtained in the end. Since the filter has read the image only within the upper left corner, it moves further and further right by 1 unit performing an identical operation. After passing the filter across all positions, a matrix is obtained, but smaller than an input matrix.

The covoutional layer shown in the below figure (google.com) works on the following formula [4]:

$$I\_new(x,y) = \sum_{j=-1}^{1}\sum_{i=-1}^{1} a_{ij} I\_old(x-i, y-j)$$



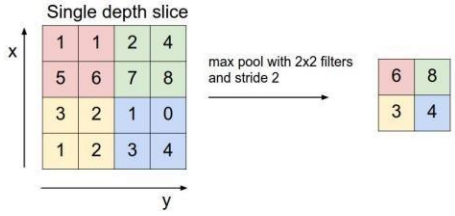
ii. ReLu Layer:

ReLU or Rectified Linear Unit, it is a non-linear operation. ReLU acts on an elementary level which means an activation function. In other words, it is an operation which is applied per pixel and supersedes all the non-positive values of each pixel in the feature map by zero.

Equation : $f(x) = ln(1+e^{x})$ [4]

iii. Pooling Layer:

Spatial Pooling which is also called sub sampling or down sampling is used to reduce the dimensions of each feature map but even while doing so, retains the most consequential information of the map. We have used max pooling, shown in the below figure (google.com). After pooling is done, eventually 3D feature map is converted to one dimensional feature vector.



iv. Fully Connected Layer:

After completion of series of convolutional, nonlinear and pooling layers, it's necessary to connect a totally connected layer. The FC layer (fully connected) takes the output information from convolutional networks. Attaching a totally connected layer to the top of the network leads to an N dimensional vector, where N is that the amount of classes from which the model selects the desired class.

D.CLASSIFICATION

A support vector machine (SVM) is a machine learning algorithm that recognizes data for classification and regression analysis. SVM may be a supervised learning method that appears at data and sorts it into one among two categories. A support vector machine is additionally referred to as a support vector network (SVN).

5. REFERENCE COMPARISION

Name and Authors	Accuracy and number of diseases	Conclusion	Future Work
Diagnosis of sin diseases using Convoutional Neural Networks, Jainesh Rathod, Visha Waghmode, Aniruddh Sodha.	5, 70%	This diagnosis work includes both detection and classification using convolutional neural network.	To achieve more than 78.2% accuracy and to increase the number of diseases.
Convolution Neural Network for Feature Extraction in Skin Disease Detection, Seema Kolkur, Dhananjay Kalbande, Vidya karkar	2, 91%	CNN is used for feature extraction and SVM as a classifier. Major advantage of this system is the saving of time.	To increase the number of diseases.
Dermatological Disease Detection Using Image Processing and Machine Learning, Sujay S Kumar Varun Saboo.	6, Up to 95%	The mobile application is developed using ANN, KNN and Decision Tree algorithm.	The system fails to detect the diseases of varying skin colors.
Expert System For Diagnosis Of Skin Diseases, A.A.L.C. Amarathunga, E.P.W.C. Ellawala, G.N. Abeysekara, C. R. J. Amalraj	3, 85%	This windows application includes 5 different algorithms and contains questionnaire section which helps in the initial treatment.	Adding more number of diseases and to develop application for Android, IOS and etc.

6. SOME MAJOR LIBRARIES USED FOR IMPLEMENTATION

(a) OpenCV : OpenCV is an Open Source Computer Vision library with Python interface along with Java, C and C++ interfaces too. It is supported on multiple platforms and is used for real-time image processing and computer vision.

(b) Scikit learn : A free machine learning library with various algorithms for tasks such as classification and regression.

(c) Keras : It is a deep learning library, capable on running on top of Tensorflow.

(d) Tensorflow : Developed by Google, it is an open- source library. Here it is used as a backend for Keras as it is useful for numerical computations and calculations. Also libraries such as numpy, pandas, etc are used.

7. RESULT

The system is able to classify 10 most common skin diseases efficiently. An initial training gives the output accuracy of 82% approximately. This can be definitely increased by increasing the training data set in the deep learning model. A large data set can increase the accuracy to more than 90 percent.

8. CONCLUSION

The above discussed image processing and deep learning algorithms are used to efficiently classify the diseases. Major advantage of the system is the saving of time and effort involved in feature engineering. CNNs learn features on their own. Hence skin diseases can be diagnosed using CNN and also be classified using the same. Using advanced computational techniques and large dataset, the system can match the results of a dermatologist thus improving the quality standards in the area of medicine and research.

9. REFERENCES

[1] M. Rodri&nhy; guez, "The median filter problem", Tracer.lcc.uma.es, 2016

[2] W. K. Saad, "Method for detection and diagnosis of the area of skin disease based on colour by wavelet transform and artificial neural network", Al-Qadisiya Journal for Engineering Sciences, 2009

[3] Seema Kolkur, Dhananjay Kalbande, Vidya karkar, "Convolution Neural Network for Feature Extraction in Skin Disease Detection", Journal of Advanced Research in Applied Artificial Intelligence and Neural Network Volume 5, Issue 1&2 - 2018, Pg. No. 1-5.

[4] Jainesh Rathod, Vishal Waghmode, Aniruddh Sodha, Dr. Prasenjit Bhavathankar, "Diagnosis of skin disease using Convolutional Neural Networks", ICECA IEEE Conference Paper,2018

[5] Sujay S Kumar Varun Saboo, "Dermatological Disease Detection Using Image Processing and Machine Learning",
IEEE Conference Paper, 2016

[6] Jyostna More, Prachi Bhadale, Girija Lonkar, "Method for Skin Disease Detection Using Color Images", International Journal of Engineering Science and Computing, 2019

[7] A.A.L.C.Amarathunga, E.P.W.C. Ellawala, G.N. Abeysekara, C. R. J. Amalraj, "Expert System For Diagnosis Of Skin Diseases", International journal of scientific & technology research volume 4, 2015

[8] Sermanet P, Chintala S, LeCun Y. "Convolutional neural networks applied to house numbers digits classification. InPattern Recognition (ICPR)", 2012 21st International Conference on 2012 Nov 11 (pp. 3288-3291). IEEE.

[9] Li TL, Chan AB, Chun A. "Automatic musicle pattern feature extraction using convolutional neural network", InProc. Int. Conf. Data Mining and Applications 2010 Mar.

[10] Athiwaratkun B, and K. "Feature representation in convolution neural netwoks", arXiv preprint arXiv:1507.02313, Jul 8 2015.