International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Skin Disease Detection System Using Image Processing Machine Learning

Dr.Dnyande Hire[1], Vedant Pahune[2]

[2]Associate Professor

D. Y. Patil Institute of Engineering Management and Research, (DYPIEMR) Akurdi, Pune

Abstract

Skin diseases are very common and affect people of all ages. These diseases happen more often as compared to other types of illnesses. There are a lot of reasons for this, such as Fungal infections, bacteria, allergies, and even viruses. Thanks to new lasers & special medical technology, we can now find skin diseases faster than before. But here's the thing: these tests can be super costly and are not always available. That's why dermatology needs an automated screening system. This system can use image processing techniques to help. Image processing is great because it helps us figure out what skin disease someone has. Many computer vision techniques look for skin problems, too. In places like Saudi Arabia, with its hot deserts, skin diseases are pretty common. So, our research aims to help with this issue. We've come up with a method that uses image processing to detect skin diseases! It starts by taking a digital picture of the affected skin area. Then we analyse that image to figure out what type of disease it is. Our method is super simple & quick! The best part? You don't need fancy, expensive tools—just a camera and a computer. This method uses colour images. First, we resize the image to pull out features using something called a pre-trained convolutional neural network that sounds complicated, but it's really just tech that helps us. After that, we use Multiclass SVM to classify those features

1. INTRODUCTION

It has been estimated that 10% to 12% of the population suffers from skin problems in India. This is the largest internal organ of the body, covering or protecting the body while also absorbing sensations from the environment. The structures consist of seven layers of ectodermal tissue, which act like a camouflage for bones, muscles, and other inner organs. Poor hygiene, rising pollution levels, changing climate patterns, and exposure to harmful rays of UV also cause several diseases related to the skin. For example, a drop of just 1% in the ozone layer can cause a rise of up to 2-3% in skin cancer patients.

Photo-sensitive skin diseases and infections are more common in India. These problems must be taken care of on time because untreated skin diseases not only affect the physical health of a person but also trigger mental health and even the quality of life. There is a need for easy and proper treatments that can deliver effective and timely care to everyone in a population that is catching up with growth.

2. PROBLEM STATEMENT

Like tuberculosis and AIDS, untreated severe skin conditions can lead to further complications. The costs of even the simplest skin care therapies can be high. It is necessary, therefore, to have diagnostic methodologies that are efficient yet affordable for skin diseases. It is essential to have dermatologists who recognize and treat disorders that affect the skin, hair, and nails. Technology is changing the face of health care by replacing traditional ways of doing things with automated systems in most fields nowadays.

Dermatology is the area of research scientists are looking to ease the hastening of doctors to diagnose patients of their skin ailments with high-speed technology. It uses primarily digital image processing so that an enhanced image can be obtained and pertinent information extracted through computer algorithms. This would encompass a few steps: getting the image, analyzing and enhancing it, and then drawing reports based on the analysis. Segmentation in image processing is the critical component that breaks down the image into distinct sections based on pixel characteristics. The technique of segmentation is useful in diagnosing ringworm, eczema, chickenpox, and psoriasis because the images become easier to analyze. These analyses are highly dependent on the accuracy of feature measurement. This project intends to identify the most appropriate image processing technique to find a specific skin condition through the following stages:

1. Data Collection: It collects images of various conditions

.2. Pre-processing: The enhancement in quality and elimination of the noise from the images.

3. Segmentation: Segmentation, i.e., segmentation of images into regions for further analysis.

4. Feature Extraction and Classification: The features of each condition have to be identified and categorized.

3. LITERATURE REVIEW

This paper proposes a recognition method that classifies three particular skin diseases as herpes, dermatitis, and psoriasis. After preliminary filtering and transformation based on the removal of noise and irrelevant backgrounds from the skin images, for segmentation purposes, it utilizes the GLCM that has the ability to capture texture and color features in skin disease images. After segmentation, the SVM classification methodology is applied, and thus, it is possible to identify the disease with very high accuracy. Experimental results confirm the feasibility and correctness of this approach.

Sumithra et al. - Automatic Segmentation and Classification of Dermatological Lesions for Disease Detection

The paper brings out an autonomous technique for the segmentation and classification of dermatological lesions. Then, it uses filtering to remove hair and noise from the images, then proceeds with segmentation using a region-growing technique. In this case, automatic seed point initialization has been done. Then color and texture feature extraction followed by classification with SVM and k-NN and their combinations. The proposed approach was tested on a database of 726 samples from 141 images representing five skin diseases. With SVM, an F-measure of 46.71% was achieved, whereas with k- NN, it remained at 34%, but when the classifiers are combined, it goes up to 61%. The results are very promising.

Kolkur et al., "Human Skin Detection using RGB, HSV and YCbCr Color Models

To this effect, this work approaches an identification mechanism for human skin using RGB, HSV, and YCbCr color models that turns out to be efficient for the detection of skin with invariance in orientation and size. This algorithm comes out as an accurate method for the detection of skin pixels by combining some definite ranges of a color parameter for accuracy.

Kalaiarasi et al., Dermatological Disease Detection Using Image Processing and Neural Networks

Kalaiarasi et al. suggest a two-stage approach that incorporates computer vision and machine learning for more accurate detection of dermatological diseases. Features are obtained from pre-processed skin images in the first stage, and the second stage applies algorithms from machine learning toward developing capabilities of disease identification based on histopathological characteristics. It was tested on six diseases with a very high accuracy rate of 95%.

Hameed et al. - Skin Disease Classification Using CNN and SVM Hybrid Approach

Hameed suggests a diagnostic system that combines deep convolutional neural network (CNN) with error-correcting code (ECOC) SVM for the classification of skin lesions. On a 9,144 images dataset, the hybrid used Alexnet for feature extraction and ECOC SVM for classification achieving 86.21% accuracy over five different skin diseases: healthy, acne, eczema, benign, and malignant melanoma.

Shanthi et al. CNN-based Detection of Skin Diseases This paper detected four types of skin diseases, such as acne, keratosis, eczema herpeticum, and urticaria, using an 11-layer convolutional neural network. It used images from the DermNet database to test the model, which achieved its accuracy within the interval of between 98.6% and 99.04%. The limitation of this study is that it was based on a relatively small sample size of four diseases.classes.

4 METHODOLOGIES

4.1 Image Dataset

A dataset for skin disease detection includes images showing the normal skin and abnormal skin conditions. The input data are the images that aid the system in seeking specific diseases. To counter the differences in image sizes for the dataset, images are resized by either zooming in or compressing. Resizing will standardize dimensions for images thus ensuring that features derived from all images are consistent for the analysis to be taken uniformly. Moreover, resizing reduces processing time hence a better performance of the system. For example, a source image of 260×325 pixels can be resized to 227×227 pixels for uniformity.

4.2 Preprocessing

In preprocessing, images have to be cleaned, smoothed, and filtered so that they can be analyzed. Preprocessing must be accurate and meaningful, especially where the context of the work is related to the detection of skin disease. Accuracy may solely depend on the quality of the images and their standardization. The data should be cleaned to correct an error or fill in missing information

4.3 Segmentation

Image segmentation is the process of splitting an image into meaningful regions or classes based on factors such as g ray level, brightness, colour, contrast, and texture. Since it distinguishes lesions from the rest of the healthy skin, segmentation is very critical in ensuring the accurate analysis of the images since this procedure directly impacts the following processes, which will be less precise unless the input is very accurate. Segmentation is indeed very difficult in microscopic images due to the size, shape, colour, and thin contrast with surrounding skin. Techniques for segmentation include threshold-based, region-based, cluster-based, and edge- based methods

4.4 Feature Extraction

Feature extraction is an important step in an analyse and establishing correlations between the objects within images, through which algorithms can understand image data by transforming it into a usable format. While dermo images abound with several traits, not all of them are necessarily relevant to the skin disease classification task. Considering too many nonsensical features complicates the classifier and increases computational requirements but eventually decreases the classification accuracy. For effective classification of skin disease, only the most relevant features should be extracted for accurate representation of specific region

4.5 Statistical Features

Apart from the features acquired from the GLCM of the image texture, statistical features from the RGB color space are also extracted. These help further in analyse the distribution of color in the skin images for better classification. Among these statistical features, the following are computed for every image:

• Mean: Average pixel intensity value in an image.

• Variance: It measures the spread or variation of pixel intensities.

• Standard Deviation: It indicates the level of variability or dispersion of pixel values.

• Root Mean Square (RMS): It returns a measure of average intensity, which turns out to be quite useful for general interpretation of brightness and contrast.

4.6 Classification

After the feature extraction of GLCM features and statisticalfeatures, classification remains to be done. Classification can be divided into several steps:

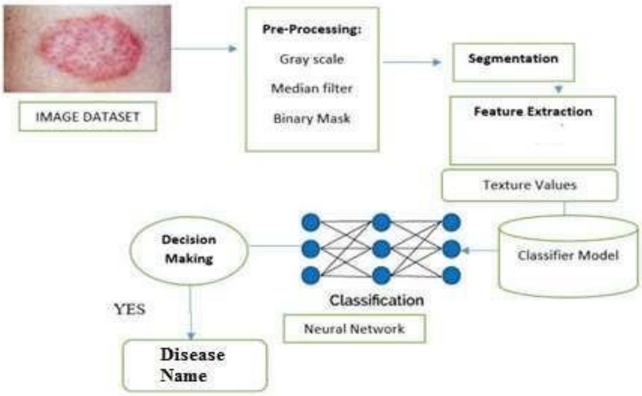
1. Select a Model: Select a suitable algorithm or machine learning model for that particular type of data that needs to be classified.

2. Training the Model: The selected model is trained by using features that were derived from the dataset.

3. Model Evaluation: Test data may be used to classify the performance of the respective model.

4. Parameter Tuning: Tighten the model parameters to boost accuracy.

5. Prediction: Use the learned model to make predictions on new images.



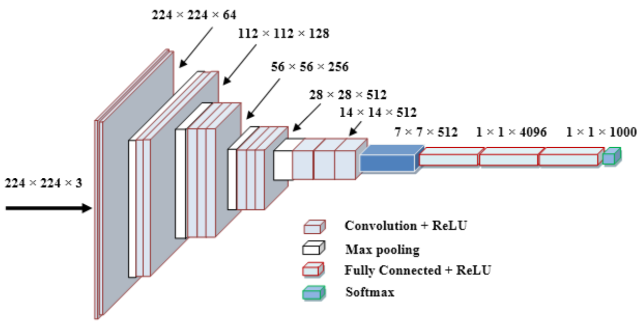
<u>Visual Geometry Group VGG</u>

The Visual Geometry Group at the University of Oxford and Google DeepMind created VGGNet CNN. The network design of the VGGNet, which is shown in below and is characterized by **3** × **3** convolutional kernels and **2** × **2** pooling layers, can be deepened by employing smaller convolutional layers to improve feature learning. The concept of a much deeper network with much smaller filters is known as a VGG network. In comparison to AlexNet's eight layers, VGGNet has more layers. VGGNet-16 and VGGNet-19 are now the two most popular VGGNet versions. In ImageNet, the VGG16 model achieves top-5 test accuracy of about 92.7 percent. A dataset called ImageNet has over 14 million

International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

images that fall into almost 1000 types. It was also among the very well models submitted at ILSVRC-2014. It significantly outperforms AlexNet by substituting a number of **3 × 3** kernel-sized filters for the huge kernel-sized filters. The VGG19 model (also known as VGGNet-19) has the same basic idea as the VGG16 model, with the exception that it supports 19 layers. The numbers "16" and "19" refer to the model's weight layers (convolutional layers). In comparisonVGG16, VGG19 contains three extra convolutional layers.



Architecture of VGG

Input: The VGGNet accepts 224×224 pixel images as input. To maintain a consistent input size for the ImageNet competition, the model's developers chopped out the central 224×224 patch in each image.

Convolutional Layers: VGG's convolutional layers use the smallest feasible receptive field, or 3×3, to record left-to-right and up-to-down movement. Additionally, 1×1 convolution filters are used to transform the input linearly. The next component is a ReLU unit, a significant advancement from AlexNet that shortens training time. Rectified linear unit activation function, or ReLU, is a piecewise linear function that, if the input is positive, outputs the input; otherwise, the output is zero. In order to maintain the spatial resolution after convolution, the convolution stride is kept at 1 pixel (stride is the number of pixel shifts over the input matrix).

Hidden Layers: The VGG network's hidden layers all make use of ReLU. Local Response Normalization (LRN) is typically not used with VGG as it increases memory usage and training time. Furthermore, it doesn't increase overall accuracy.

Fully Connected Layers: The VGGNet contains three layers with full connectivity. The first two levels each have 4096 channels, while the third layer has 1000 channels with one channel for each class.

RESULT

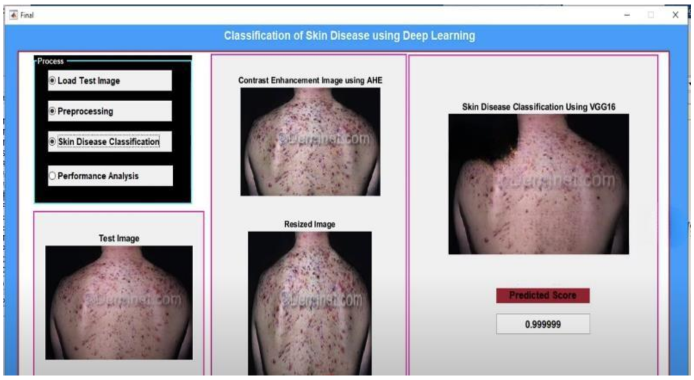
shows accuracy graphs comparing CNN training accuracy and validation accuracy over a series of epochs. The figure presents graphs comparing the loss values for different CNN classification algorithms. Training and validation accuracy graph: There usually is a graph showing two curves: The red line in the above figure shows how the training accuracy increases with more epochs. The green line

International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

represents the validation accuracy, that should ideally have the same trend. but will be different if overfitting exists

Final Output



4. CONCLUSIONS

We had a method that depended on image processing for the detection and classification of skin diseases. The method works on an analysis of the digital images of affected skin areas to indicate the type of disease. The approach may be simple, fast, and cheap because it uses only a camera and a computer without requiring expensive medical equipment. Since our model's learning path depends on whether it is capable of generalizing to unseen data, the precise selection of a proper loss function, such as categorical cross entropy and the Adam optimizer, is highly influential to our efficiency in training and performance improvements. This project phase, therefore, has not only deepened our understanding of deep learning and image processing but also enabled further development of blood group detection technology.

5. ACKNOWLEDGEMENT

We sincerely appreciate everyone who contributed to the successful completion of our Skin Disease Detection System First and foremost, we extend our heartfelt gratitude to our mentor, Dr. Dnyanda Hire, for her invaluable guidance, expertise, and unwavering support. Her insightful feedback and encouragement were instrumental in overcoming challenges and refining our work.

We are also thankful to Dr. D.Y. Patil Institute of Engineering, Management, and Research and its faculty members for providing the necessary resources and a supportive learning environment that enabled us to develop and implement this project.

A special acknowledgment goes to our dedicated team members, whose collaboration, problem-solving skills, and commitment to excellence played a crucial role in achieving our objectives. This project would not have been possible without the collective efforts and encouragement of everyone involved, and we deeply appreciate their contributions.

REFERENCES

1. Armaanzas R, Iglesias M, Morales DA, Alonso- nanclaresL.Application of voxel-based models for Alzheimer disease diagnosis using classifier combinations. IEEE Journal of Biomedical and Health Informatics. 2016 Mar 4;21(3):778-84.
2. Aruchamy S., Haridasan, A., Verma, A., Bhattacharjee, P., Nandy S. N., and Vadal S. R. K. February 2020.Detection of Alzheimer's disease using machine learning techniques in 3D MR images. Proceedings of the 2020 National Conference on Emerging Trends in Sustainable Technology and Engineering Applications, NCETSTEA, IEEE (pp. 1-4).
3. Bui DT, Hoang ND. A Bayesian model using Gaussian mixture and radial-basis-function Fisher discriminant analysis for spatial prediction of flood patterns. Geoscientific Model Development. 2017;10(9):3391.
4. Zhu X, Sobhani F, Xu C, Pan L, Ghasebeh MA, Kamel IR. Quantitative volumetric functional MR imaging as a biomarker for assessing early treatment response in patients with hypo-vascular liver metastases following yttrium-90 transarterial radioembolization. Abdominal Radiology. 2016 Aug 1;41(8):1495-504. [1]
5. Manerkar Mugdha S, Harsh Shashwata, Saxena Juhi, Sarma Simanta P, Dr. U. Snekhalatha, Dr. M. Anburajan. (2016). Skin disease classification through a multi SVM classifier. 3rd International Conference on Electrical, Electronics, Engineering Trends, Communication, Optimization, and Sciences (EEECOS), 363-368. [2]
6. Alam, Md, Munia, Tamanna Tabassum Khan, Tavakolian, Kouhyar, Vasefi, Fartash, MacKinnon, Nicholas, & Fazel- Rezai, Reza. (2016). Automatic measurement and assessment of eczema severity using image processing. Engineering in Medicine and Biology Society Conference August 2016, 1365-1368. [3]
7. Sumithra R, Mahamad Suhil, D. S. Guru. (2015). Segmentation and classification of skin lesions for diagnosing diseases. International Conference on Advanced Computing Technologies and Applications (ICACTA2015), 45, 76-85. [4]
8. Sheha Mariam A., Mabrouk Mai S, Sharawy Amr. (2012). Automatic detection of melanoma skin cancer. International Journal of Computer Applications (0975- 8887), 42(20), 22-26