Skin Disease Detection Using Image Processing and Machine Learning

R. Vijaya Prakash[1], Vaishnavi Gantala[2], Rushikesh Ananthula[3], Hemantha Banda[4]

[1]SR University, Warangal, India
[2,3,4]SR Engineering College, Warangal, India
[1]vijprak.r@gmail.com,[2]vaishnavigantala@gmail.com, [3]rishie.ananthula@gmail.com,
[4]hemanthareddy11@gmail.com

Abstract.

Dermatological diseases are the most common type of disease on a global scale. Despite its prevalence, it is extremely difficult to diagnose and requires extensive domain experience. We present a method for detecting various types of these diseases in this research paper. To accurately identify the disease, we employ a two-stage approach that effectively combines Image Processing and Machine Learning on clinically evaluated histopathological attributes. Various pre-processing techniques are applied to the image of the skin disease before feature extraction is performed in the first stage. Using machine learning algorithms to identify diseases based on histopathological attributes found through skin analysis is the second stage of this process. The system was able to accurately identify up to 95% of the six diseases after training and testing.

Keywords: Skin Disease Detection, Machine Learning, Image Processing, Convolutional Neural Networks (CNN), Histopathology

1 INTRODUCTION

The human body's largest and most important organ is the skin. Skin covers the entire 20 square feet of a human's body. Protection from microbes and environmental elements, as well as the ability to feel heat and cold through the skin, are some of the functions of the human immune system. The skin, however, can be affected by a variety of external and genetic factors. One of the most common forms of skin disease is the viral type, followed by a fungus or an allergic reaction. Fungal and allergic-type diseases can be cured if they are diagnosed properly and recognized at the earliest possible opportunity. In the case of viral infections, early detection is critical. Machine learning and artificial intelligence (AI) have made rapid progress in medicine over the last few years. Detecting different skin diseases requires both physical symptoms and image analysis. Skin diseases can be more accurately detected using image analysis.

In a developing country, the widespread use of smart phones has opened up new avenues for inexpensive disease diagnosis. We can take advantage of the built-in camera and image processing capabilities of smartphones to make diagnoses. Using a two-step approach, we've developed an application to solve the problem. Image processing is used in the first stage for identification, and machine learning is used in the second stage to provide a nearly foolproof solution. The differential diagnosis is complicated by the fact that a disease may initially present with the symptoms of one condition but later develop the signs and symptoms of another.

2
In most cases, a biopsy is required to make the diagnosis, but there are also many similarities in the histopathology of these diseases. Using machine learning models on clinically evaluated features derived from microscopic examination of skin samples, this problem can be resolved.

The subjective nature of diagnosis makes it difficult for medical students to verify their findings. Using this system, students can verify their findings because they have access to clinical data. The University of California, Irvine's machine learning data repository provided the training data set. A combination of Image Processing and Machine Learning algorithms has helped us improve our precision. Psoriasis, Seborrheic Dermatitis, Lichen Planus, Pityriasis Rosea, Chronic Dermatitis, and Pityriasis Rubra Pilaris are six of the most common diseases that the system can detect.

There were three different machine learning processes and image processing techniques used in this research paper for the prediction of different skin diseases. A comparison analysis of the machine learning is developed in order to comprehend the entire experiment.

2 Literature Review

The dissection of skin diseases using colour images is proposed as a method that does not necessitate the involvement of a doctor. Color image processing techniques, k-means clustering, and colour gradient techniques are used to identify the diseased skin, and artificial neural networks are used to classify the disease using the system. Six types of skin diseases were used to test the system, and the first stage accuracy was 95.99 percent, and the second stage accuracy was 94.016 percent.[1]. The first step in the detection of skin diseases is the extraction of image features [2]. The more features that can be extracted from an image, the more accurate the system will be. As much as nine types of skin diseases were successfully treated [2]. Melanoma is a deadly form of skin cancer if it isn't diagnosed and treated early enough. Images can be segmented using image processing to detect melanoma [3]. To extract more features, a segmentation process that falls on the infected spot boundaries is described [3]. It was proposed by [4] to create an algorithm database that included images from a wide range of Melanoma resources to create a tool for diagnosing the disease in people with dark skin. As an example, [5] used the support vector machine technique to classify skin diseases like melanoma, basal cell carcinoma (BCC), nevus, and seborrheic keratosis (SK) (SVM). It's the most accurate method out of a variety of others. If chronic skin diseases become more common in different parts of the world, they could have dire consequences for people's health [5]. A computer system that automatically detects and measures the severity of eczema was proposed by [6]. Segmentation is accomplished by detecting skin, extracting a set of features (colour, texture, borders), and then determining the severity of eczema by using Support Vector Machines (SVM) [6].

A new skin disease detection method based on image processing was proposed in a research paper [7]. It was an RGB image of the skin disease area that they used as an input. Then, using a pre-trained convolution neural network model, resize the images and extract the features. For classification, they make use of a multi support vector machine [7]. They

demonstrate the simplicity and ease with which this method can be used, as well as its accuracy. In their review paper [8], researchers present a number of studies on skin diseases. Some skin diseases can be classified using machine learning processes in all studies. Those processes are described in great detail, including the mechanisms, algorithms, and accuracies used [8]. Five different machine learning algorithms were used by Bhadula et al. to identify skin diseases [9]. In addition to the random forest and SVM algorithms, they also used Nave Bayes, Logistic Regression, and Convolution Neural Networks. A neural network model based on convolutional recurrent connections was found to produce the best results from the confusion matrix [9]. A neural network-based artificial intelligence system was proposed by Jayachandra et al. [10]. The image acquisition process performed feature extraction in the first part of this system. The feed-forward neural network was used for classification in the second stage [10]. Six different data mining algorithms were proposed by Verma et al. [11] to classify the various types of skin diseases. The 15 most important features are chosen to predict disease classifications. In addition to these six algorithms, the authors also create an ensemble method using Bagging, AdaBoost, and Gradient Boosting classifier techniques for prediction. It is concluded that the ensemble method provides a more accurate and effective prediction for detecting skin disease [11].

Skin cancer detection could benefit from the use of an artificial intelligence system proposed by Hasan et al [12]. For skin cancer detection, they used image processing and a deep neural network that was trained over time. They first used an image processing method to segment and extract features from the affected area, and then used a convolution neural network for prediction. 93.7% of the training was accurate, and 89.55% of the testing was accurate [12].

3 Methodology

3.1 Dataset

. For the experimental purpose here various skin disease datasets are used. All the datasets are collected from Kaggle. The final dataset contains more than forty thousand images of skin disease data. All the images are divided into 12 skin disease classes. The sample images of the dataset are showing in figure 1. The info is split into, training set, validating/testing set. A training set is adopted for learning to suit the parameters and is specifically applied to change the varied weights and errors of the system in every coaching run. Validation/testing set tune the parameters and area unit used solely to assess the effectiveness and potency of the system. During this technique, the divide mode is ready to ninetieth for the coaching of the information, 100 percent for the validation/testing of the information.

A comparison of different machine learning approaches was given in this study.

3.2 Machine Learning Approaches

Machine learning [9][11][13][14] is a part of Artificial Intelligence (AI). It can learn automatically from input data and make an output decision. This process involves 4 steps 1) Data Collection, 2) Data Pre-processing, 3) Feature Extraction, and 4) Classification.

4

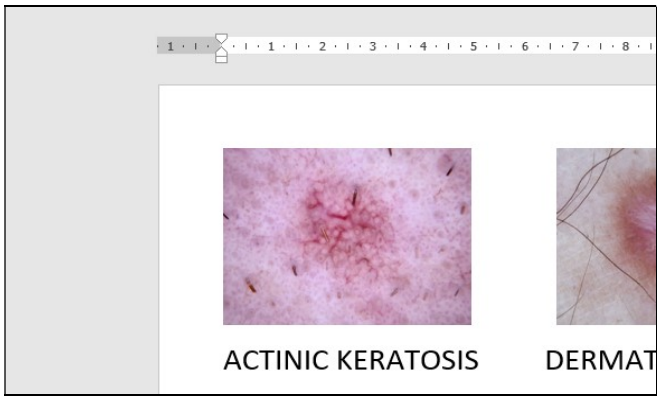

Fig.1 Sample Dataset

a. Data Collection

Section 3 describes the data collection procedure. The RGB colour space is used for all the input photos.

b. Data Pre-processing

The data processing procedures in this experiment are as follows:

Step 1: Accept and change the brightness and contrast of the input RGB images

Brightness = 7 < im < 35, contrast = 15 < im < 25

Step 2: Convert the RGB images into Grayscale images.

Step 3: Remove the noises from each R, G, B color channel using a median filter.

c. Feature Extraction.

The Gray Level Co-occurrence Matrix (GLCM) [15] is used to pick features for experimental purposes. It is a statistical texture analysis method. It is a statistical texture analysis method of the second order. This procedure computes the statistical information from each surrounding pixel in a grayscale image.

d. Classification

The next step is to classify the images using the GLCM features that were extracted from the input images. The classification process involves the following steps, 1) Choose a model, 2) Train the model, 3) Evaluate the model, 4) Tune the parameters, 5) Predict the outcome

4 Results and Discussions

Figure 2 describes the total classification process of the machine learning approach. And figure 3 describes the overall deep learning process.

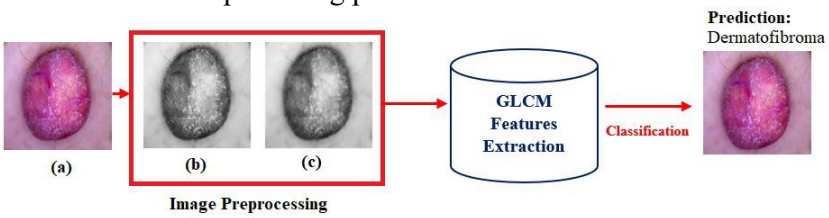

Figure 2: Overall Classification Process for Machine Learning Approach [a) Original Image, b) Grayscale image, c) After Noise Removal

The performance analysis of three machine learning processes is shown in table 1.

5

Table 1: Performance Analysis of Different Machine Learning Approaches

Methods	**Accuracy**	**Specificity**	**Sensitivity**	**Kappa Value**
CNN	0.9299	0.99927	0.98684	0.913
Ensemble Bagged Tree Algorithm	0.9252	0.9709	0.969	0.9071
Support Vector Machine	0.8353	0.75	0.8689	0.6036

Both the bagged tree ensemble method and the CNN classifier have the best accuracy in disease prediction, as shown in table 1. These algorithms have specificity, sensitivity, and Cohen's kappa values greater than 90%. This suggests that there are fewer false negatives and more accurate positives. Figures 3 and 4 show the accuracy and training loss for various epochs. Figure 3 shows that the accuracy improves as the number of epochs increases. Increasing the number of epochs reduces training loss in the same way. In general, the methods of machine learning and image processing show satisfactory results in the detection of skin diseases.

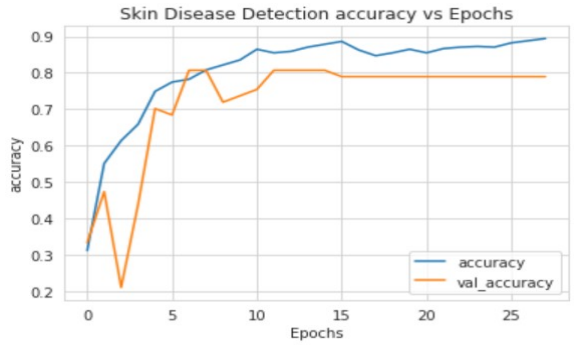

Fig.3 Accuracy trends for various Epoch values

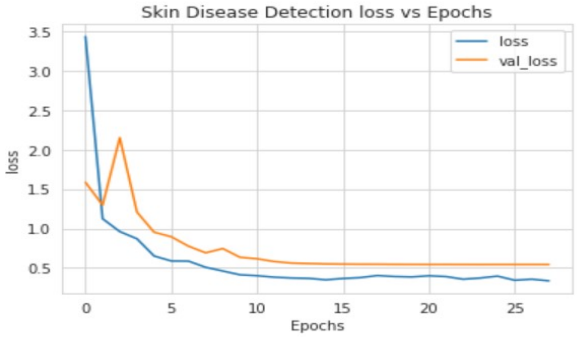

Fig. 4 Training Loss for different Epochs values

6

5 CONCLUSION

The largest organ of the body is the skin. Skin diseases can occur for a variety of reasons, including both external and internal. As a result, the study of skin diseases is a critical component of medical research. Because it has the potential to reduce the number of people who die from skin diseases or the spread of infectious diseases. The clinical process is extremely time-consuming, and it may not always be able to correctly identify the disease at hand. These situations call for the use of machine learning and image processing techniques.

Machine learning and image processing techniques for the detection of skin diseases were briefly discussed in this paper. The performance comparison was done using three different machine learning algorithms and image processing. Both processes were found to be excellent at detecting skin disease.

References

1. Patnaik, Sourav & Sidhu, Mansher & Gehlot, Yaagyanika & Sharma, Bhairvi & Muthu, P.. (2018). Automated Skin Disease Identification using Deep Learning Algorithm. Biomedical and Pharmacology Journal. 11. 1429-1436. 10.13005/bpj/1507.
2. Mayowa Oyedere, Ken Gray, ROP and TOB optimization using machine learning classification algorithms, Journal of Natural Gas Science and Engineering, Volume 77, 2020, 103230, ISSN 1875-5100, https://doi.org/10.1016/j.jngse.2020.103230.
3. Li-sheng Wei, Quan Gan, Tao Ji, "Skin Disease Recognition Method Based on Image Color and Texture Features", *Computational and Mathematical Methods in Medicine*, vol. 2018, Article ID 8145713, 10 pages, 2018. https://doi.org/10.1155/2018/8145713
4. Kwiatkowska D, Kluska P, Reich A. Convolutional neural networks for the detection of malignant melanoma in dermoscopy images. *Postepy Dermatol Alergol*. 2021;38(3):412-420. doi:10.5114/ada.2021.107927
5. A. K. Verma, S. Pal, and S. Kumar, "Comparison of skin disease prediction by feature selection using ensemble data mining techniques," Informatics Med. Unlocked, vol. 16, no. April, p. 100202, 2019, doi: 10.1016/j.imu.2019.100202.
6. M. Z. K. Patricia, V. Z. E. Santiago, and M. Y. F. Javier, "Expert System for the Prediagnosis of Skin Diseases," Int. J. Mach. Learn. Comput., vol. 10, no. 1, pp. 81–86, 2020, doi: 10.18178/ijmlc.2020.10.1.902.
7. Alkolifi Alenezi, N. S. (2019). A Method of Skin Disease Detection Using Image Processing and Machine Learning. *Procedia Computer Science*, *163*, 85–92. https://doi.org/10.1016/j.procs.2019.12.090
8. Mohammed, S. S., & Al-Tuwaijari, J. M. (2021). Skin Disease Classification System Based on Machine Learning Technique: A Survey. *IOP Conference Series: Materials Science and Engineering*, *1076*(1), 012045. https://doi.org/10.1088/1757-899x/1076/1/012045
9. Bhadula, S., Sharma, S., Juyal, P., & Kulshrestha, C. (2019). Machine Learning Algorithms based Skin Disease Detection. *International Journal of Innovative Technology and Exploring Engineering*, *9*(2), 4044–4049. https://doi.org/ 10.35940 /ijitee. b7686.129219
10. Jaychandra Reddy, V., & Nagalakshmi, T. J. (2019). Skin disease detection using artificial neural network. *Indian Journal of Public Health Research and Development*, *10*(11), 3829–3832. https://doi.org/10.5958/0976-5506.2019.04183.4

11. Verma, A. K., Pal, S., & Kumar, S. (2019). Comparison of skin disease prediction by feature selection using ensemble data mining techniques. *Informatics in Medicine Unlocked*, *16*(April), 100202.https://doi.org/10.1016/j.imu.2019.100202
12. Hasan, M., Barman, S. Das, Islam, S., & Reza, A. W. (2019). Skin cancer detection using convolutional neural network. *ACM International Conference Proceeding Series*, *March 2020*, 254–258.https://doi.org/10.1145/3330482.3330525
13. Alkolifi Alenezi, N. S. (2019). A Method of Skin Disease Detection Using Image Processing and Machine Learning. Procedia Computer Science, 163, 85–92. https://doi.org/10.1016/j.procs.2019.12.090
14. Leelavathy S, Jaichandran R, Shobana R, Vasudevan, S. S. P. and N. (2020). Skin Disease Detection Using Computer Vision and Machine Learning Technique. European Journal of Molecular & Clinical Medicine, 7(4), 2999–3003.
15. Mohanaiah, P., Sathyanarayana, P., & Gurukumar, L. (2013). Image Texture Feature Extraction Using GLCM Approach. International Journal of Scientific & Research Publication, 3(5), 1–5.