International Journal of Advanced Research in Engineering Technology & Science ISSN: 2349-2819
Impact Factor: 7.10
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Email: editor@ijarets.org Volume-12, Issue-11 November – 2025 www.ijarets.org

ML based Skin disease identification: A Deep learning significantly to improves diagnostic accuracy

Dr. Girish Padhan
Associate professor, VIT,Bargarh Under
Biju Patnaik University,Odisha

Abstract

Machine learning algorithms were used along with image processing techniques for the detection of skin diseases. Compared to machine learning, the dermatological technique for identifying the type of skin illness is quite expensive. Skin conditions can be surprising and difficult to diagnose. The colour of healthy skin differs from that of diseased skin. Aim of the project is to classify the type of disease using machine learning. In order to decide, these algorithms employ feature values from effected images as input. Three parts make up the process of determining the type of skin disease: feature extraction, training, and testing.

The objective is to increase the accuracy of diagnosis for various types of skin diseases. Here we used herpes, eczema, and acne as the four diseases in this study. The process makes use of machine learning technology to train itself with the various skin images. Three important features in image classification are texture, color, shape, and combination of these. In this project work, features such as entropy, variance, contrast, and energy are used to build machine learning algorithm such as logistic regression, Support Vector Machine, and Decision Tree. Accuracy is used to test the performance of the chosen algorithms.

Skin diseases are common health problems around the world. The perils of the infections are invisible, which cause physical health distress as well as initiate mental depression. In addition, it sometimes leads to skin cancer in severe cases. Subsequently, diagnosing skin diseases from clinical images is one of the foremost challenging tasks in medical image analysis. As a result, both patients and dermatologists require automatic skin disease prediction, which makes the treatments plan faster.

Keywords:-CNN,DT,LDA,k means,RGB,HSB,XA

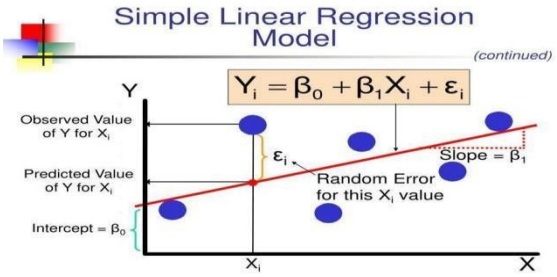
1.1 Introduction

Skin diseases are among the most common health conditions affecting millions of individuals worldwide, with a wide range of types, from benign rashes and infections to more serious conditions such as melanoma. Early detection and timely diagnosis of skin disorders are critical to ensuring effective treatment, minimizing the risk of complications, and improving patient outcomes. However, in many parts of the world, access to dermatological expertise is limited. The integration of technology

into healthcare, particularly in the field of dermatology, has opened up new avenues to address these challenges through automated systems capable of identifying and classifying various skin conditions.This project report presents the design, development, and implementation of a Skin Disease Detection System, an intelligent software tool that leverages advancements in machine learning and computer vision to aid in the accurate detection of skin diseases from images. The core objective of the system is to bridge the gap between the availability of dermatological services and the increasing demand for efficient, accurate, and accessible diagnostic support tools. By utilizing image processing techniques and powerful deep learning algorithms, the system is capable of analyzing images of skin lesions and predicting the presence and type of skin disease with a high degree of accuracy.The proposed system employs a Convolutional Neural Network (CNN) model, trained on a diverse dataset comprising thousands of skin images labeled with various dermatological conditions. These include common disorders such as eczema, psoriasis, acne, as well as more critical conditions like basal cell carcinoma and melanoma. Through feature extraction, pattern recognition, and classification, the model learns to distinguish between different disease types based on visual cues such as texture, color, shape, and border irregularities.The significance of this project lies not only in its potential to support clinical decision-making but also in its ability to provide diagnostic assistance in remote or under resourced regions where specialist care may not be readily available. Furthermore, the integration of this system into mobile or web-based platforms could enable broader accessibility and usability by both healthcare providers and the general public encouraging proactive skin health monitoring.

1.2 Linear regression:

One of the most widely used and fundamental machine learning techniques is linear regression. A predictive analysis is conducted using this statistical technique.The dependent and independent variables are shown to have a linear relationship via the linear regression technique. The link is predicted by linear regression as a slope or straight line. It is a prediction algorithm to predict or estimate one variable using another. The Prediction of Ratings using Opinions is an example of linear regression.



1.3 Logistic regression:

Logistic Regression is a type of supervised learning algorithm. By combining a number of known independent variables, it is utilized to forecast the definite dependent variable. When the dependent variable is categorical, the output is foreseen using logistic regression. With this approach, you may determine if a result is 1 or 0.

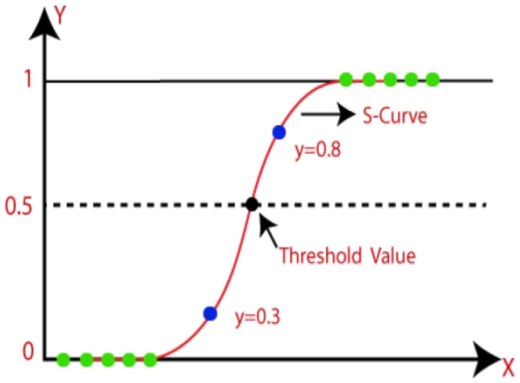


Fig 1.2: Logistic regression

Fig - 1.1: Linear Regression

1.4 Decision Tree (DT):

It is a supervised learning algorithm with a structure like a tree. It has roots and nodes. Nodes of the decision tree have various types like Decision nodes and leaf nodes shown in fig. 2.4

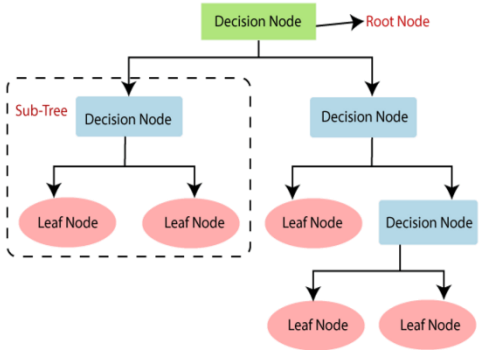


Fig. - 1.3: Decision tree

In a decision tree, the algorithm starts at the Root node and moves upward to identify the dataset's class. This method compares the results of the base attribute and the data (real data) attributes, then follows the branch to move to the next location. Then the value of the next number is compared with the value of the other child nodes, and the process is completed.

1.5 K-Means Clustering Algorithm:

In machine learning or data science, clustering problems uses unsupervised learning method. Clustering is an unsupervised learning method which is used to solve problems in data science or machine learning. K-means clustering is used to solve this kind of problems. Uses an iterative process to split unsigned data into k unique groups, each a group in common with the others. No training needed, this is a simple way to independently identify groups in anonymous data and allows us to aggregate data into different groups. Clustering based on centroid as shown in Fig 2.6.

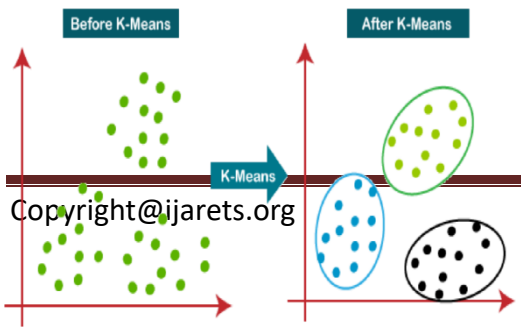


Fig – 1.4K means Clustering Algorithm

Since this method is center-based, each cluster has a center point assigned to it. The main purpose of this method is to decrease the distance between each data source and its related groups. The algorithm considers unlabeled data as input, divides it into K groups, and repeats the process until there are no groups left. In this method, the value of K must first be determined.

1.6 Feature extraction:

A good feature set contains discrete information that can distinguish one item from other items. It should be as strong as possible against generating different feature values for objects of same class. The selection process should be a small enough that it is useful to distinguish between different models, but similar models in same class. Features can be of four types: Texture features, Color features, Geometric features, and Statistical features.

1.6.1 Types of features:

Energy: Energy is used to measure uniformity. The results are same for all combinations resulting in small energy profiles; on the contrary, high energy might be required where results are not balanced. It provides information on image homogeneity. Energy has a low value when the probabilities of the grey level pairs are similar and high values otherwise. The term energy describes the local changes of a certain quality of the image.

S.no	Entropy	Variance	Contrast	Energy
1	6.855	3426.488	234.843	22.783
2	6.674	1551.767	129.207	21.876
3	6.942	1396.821	97.726	20.316
4	6.546	1544.117	108.891	14.938
5	7.325	1247.297	102.318	37.179
6	7.437	2008.435	147.77	34.494
7	7.24	2025.067	166.076	35.602

1.7 four energy we shown

8	6.591	936.089	81.235	13.895
...	...	...	...	...
313	5.79	185.873	14.886	6.264

1.8 The the which

S.no	Entropy	Variance	Contrast	Energy
1	6.514	740.004	12.926	62.28
2	7.181	1019.697	11.811	73.8
3	5.713	157.438	11.896	12.826
4	7.119	431.661	16.434	38.748
5	6.865	336.584	16.69	29.808
6	7.135	342.671	18.095	29.969
7	6.746	278.737	11.408	28.219
8	6.789	288.45	14.997	30.065
...	...	...	...	...
312	7.207	401.525	13.821	37.587

Extracted Features: In our work, we have selected features named entropy, variance, contrast, and for the better classification of skin diseases. Next, extracted the selected features of the required diseases. The results of the feature extraction were below for required disease separately.

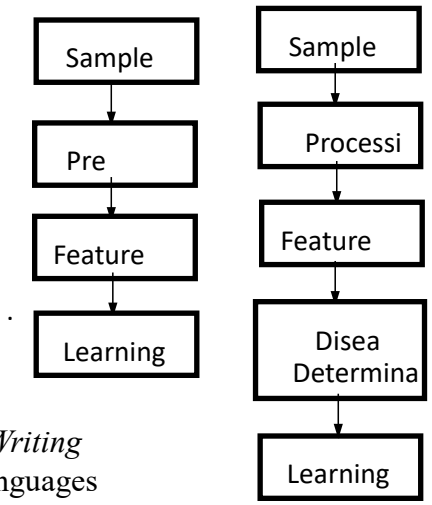
TABLE.1:Extracted features of keratosis

TABLE.2:Extracted features of Eczema

Methodology

methodology of Classification of skin diseases consists of two phases namely the Training phase and Testing phase. Training Phase is a phase where several images of disease classes are obtained, from image features obtained are stored. The extracted features of different disease classes are used for creating the required learning model using different machine learning algorithms for classifying the skin diseases. In the Testing Phase, few images of disease classes are obtained, from which image features are obtained and are compared with respect to a determined model created using the Training phase.Based on the model given image input is classified accordingly.

METHODOLOGICAL BLOCK



TrainingPhase Testing Phase

[Fig- 1.5]

1.9 Features Of Vscode

VS Code is packed with powerful features that make it super popular for development — lightweight, fast, and crazy customizable. Here's a breakdown of its top features: Visual Studio Code (VS Code) is a versatile code editor used for various development tasks. Here are some of its primary uses:

Writing and Editing Code: Supports multiple programming languages like Python, JavaScript, C++, and more.

Debugging: Integrated debugging tools help developers identify and fix errors efficiently.

Version Control: Built-in Git support allows seamless collaboration and code management.

Customization: Users can install extensions, change themes, and configure settings to enhance productivity.

Remote Development: Enables coding on remote machines, containers, and cloud environments.

Web Development: Ideal for front-end and back-end development with support for frameworks like React, Angular, and Node.js.

2.0 Result :

Skin disease images Dataset:

The proposed algorithms were trained and tested from the features of images which were obtained from Kaggle website. This anonymous dataset was collected from patients suffering from different skin diseases and categorized the images accordingly. These images are obtained from the public access "Dermnet", one of the biggest online dermatology resource designed for providing online medical education. The image is in JPEG format and has three channels RGB. Resolution from image to image and category to category varies but most of the images do not have very high-quality images.

Fig 5.1. Acne

Fig 5.2. Herpes

Fig5.3Eczema

Fig 5.4. Keratosis

Table.5.1:Count of images considered for each disease set.

The obtained dataset has many disease image collections like Acne, Exanthems, Lupus, Melanoma and such 23 disease sets with total of 19,500 images around, from which we have selected the diseases that occur more often like Acne and Herpes. After downloading the dataset, we have filtered the images which have images with disease shown clearly so that it gives better feature values during extraction. Table 5.1 indicates the count of images considered for different disease classes.

Disease	Count
Acne	290
Eczema	312
Herpes	297
Keratosis	314

Normal Skin	87

2.1 Image Processing:

Image processing is a technique used to perform few particular operations on a particular image to obtain an enhanced image or for extracting some useful data from it. Image classification was done was based on the extracted features from the images of the dataset of required diseases. For this we used image processing, based methods for purpose of extracting required features from the images.

Image processing methods for image classification can be of Texture features and color features. Features we considered are Entropy, Variance, Energy and Contrast; in which Contrast is a color feature and Entropy, Variance, Energy are Texture Features.

2.2 Future Work for Skin disease:

1.Expand Dataset Diversity
•Include more diverse skin tones, age groups, and rare diseases.
•Use datasets from different geographical regions to improve model generalization.

2. Real-Time Detection via Mobile/Web App
•Deploy the model in a mobile app or web interface.
•Real-time image capture + instant feedback for users.

3. Multi-Disease Classification
•Expand beyond binary classification (e.g., healthy vs diseased).
•Include classification for multiple skin conditions like eczema, psoriasis, vitiligo, etc.

2.3 Conclusion :-

Machine learning algorithms combined with image processing techniques were utilized to detect and classify skin diseases from digital images. The work specifically focused on four common skin conditions: herpes, keratosis, eczema, and acne. Accurate and timely detection of these skin diseases is vital for effective treatment and improving patient outcomes. Traditional methods often rely on manual diagnosis, which can be time-consuming and subjective. Therefore, this project explored the potential of automated skin disease classification systems using computational techniques.

The primary features used in this study were entropy, variance, contrast, and energy. These features represent the statistical properties of the images and provide a quantitative basis for differentiation among the skin diseases. Once the feature extraction phase was completed, three machine learning algorithms were applied for classification: Logistic Regression (LR), Support Vector Machine (SVM), and Decision Tree (DT).

In conclusion, this project demonstrated that machine learning models, when equipped with carefully selected image features, can provide meaningful assistance in the diagnosis of skin diseases. Although improvements are needed in data diversity and algorithm optimization, the initial results are promising and suggest a strong foundation for further research. Future work may focus on incorporating deep learning methods, enhancing real-time classification capabilities, and expanding the range of detectable skin conditions, ultimately aiming to create a robust, automated diagnostic tool for dermatological applications.

REFERENCES

[1] K. V. Swamy, B. Divya, "Skin Disease Classification using Machine Learning Algorithms", 2021 2nd International Conference Communication Computing and Industry 4.0 (C2I4), 2022, doi: 10.1109/C2I454156.2021.9689338 .

[2] V. B. Kumar, S. S. Kumar and V. Saboo, "Dermatological disease detection using image processing and machine learning", 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR), doi: 10.1109/ICAIPR.2016.7585217

[3] P. R. Hegde, M. M. Shenoy and B. H. Shekar, "Comparison of Machine Learning Algorithms for Skin Disease Classification Using Colour and Texture Features", 2018 International Conference on Advances in Computing Communications and Informatics(ICACCI), doi: 10.1109/ICACCI.2018.8554512

[4] D.M.T. Rasanga, G.K.A.A. Tharushika, Ishara Weeruthunge, P. Bandara, "Machine Learning-Based Skin and Heart Disease Diagnose Mobile App", 2022, doi: 10.1109/ECAI52376.2021.9515126

[5] N.V Kumar, P.V. Kumar, K. Promodh, y> Karuna. "Classification of Skin Diseases using Image Processing and SVM", 2019, DOI: 10.1109/ViTECoN.2019.8899449

[6] Honey Janoria, Jasmine Minj and Pooja Patre, "Classification of Skin Disease from Skin images using Transfer Learning Technique", 2020 4th International Conference on Electronics,Communication and Aerospace Technology (ICECA), doi: 10.1109/ICECA49313.2020.9297567.

[7] T. Chauhan, S. Rawat, S. Malik and P. Singh, "Supervised and unsupervised machine learning based review on diabetes care", 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 10.1109/ICACCS51430.2021.9442021.

[8] Khushi Kumari Jha, Roshan Jha, Ankita Kumari Jha, Md Al Mahedi Hassan, Saurav Kumar Yadav, Tr Mahesh, "A Brief Comparison On Machine Learning Algorithms Based On Various Applications: A Comprehensive Survey", 2021 IEEE International Conference on Computation System and Information Technology for SustainableSolutions(CSITSS)doi: 10.1109/CSITSS54238.2021.9683524.

[9] P. Singh, S. P. Singh and D. S. Singh, "An introduction and review on machine learning applications in medicine and healthcare", 2019 IEEE Conference on Information and Communication Technology, doi: 10.1109/CICT48419.2019.9066250.

[10] Saibee Alam , Pankaj Mohrut,A Review on "Breast Cancer Stroma Maturity Classification Using DtG Filter and SVM", 2019 International Conference on Intelligent Sustainable Systems (ICISS), doi: 10.1109/ISS1.2019.8908047.

[11] Srividhya E, Muthukumaravel A , Feature Extraction of Tongue Diseases Diagnosis Using SVM Classifier , 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), doi: 10.1109/ICCIKE47802.2019.9004326.