International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Survey Paper: Skin Disease Prediction

D. Jena Catherine Bel[1], Bharath P[2], Edwin Mathews[3], Nidharshana B[4], Monika R[5]

[1,2,3,4,5]Artificial Intelligence and Data Science, Sri Sairam Engineering College, Chennai, India

Abstract

Machine Learning (ML) is an emerging technology and a discipline to study how to use the machine to simulate human learning activities. It focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. Prediction in ML involves using a trained model to make predictions on forecasts, about new unseen data based on patterns learned from past data. On the contrary healthcare plays an important role in human life. Healthcare presents a range of complex and challenging problems that require sophisticated solutions. For example, accurately diagnosing diseases based on medical images can be difficult, but ML models can be trained to identify subtle patterns that might be missed by human experts. Skin diseases are very common nowadays and spreading widely among people in present time. Therefore, development of data mining techniques can efficiently distinguish classes of skin disease is important. Due to lack of medical facilities available in the remote areas, patients usually ignore early symptoms which may worsen the situation as time progresses. Hence, there is a rising need for automatic skin disease detection system with high accuracy. Therefore, an application is developed which uses Convolutional Neural Network (CNN) as it is the most efficient algorithm for image classification. Then a machine learning model is built for better accuracy. Then Softmax is used for multi classification. Deep Learning is used to train the model, Deep Learning is a part of Machine Learning in which unlike Machine Learning it uses large dataset and hence the number of classifiers is reduced substantially. The final model will be able to identify and classify the type of disease which has affected a particular area of skin.

Keywords: Machine Learning, Convolutional Neural Network, Skin disease prediction, Softmax.Tensorflow.

I. INTRODUCTION

Machine learning is a subset of artificial intelligence (AI) that involves using algorithms and statistical models to analyze data and learn patterns in order to make predictions or decisions without being explicitly programmed to do so. Machine learning has many real-world applications, including image and speech recognition, natural language processing, recommendation systems, fraud detection, and more. It is used in a variety of industries, including finance, healthcare, e-commerce, and marketing. The development of machine learning has been driven by advances in computing power, data storage, and data processing techniques. Today, machine learning is a rapidly evolving field, with new algorithms, techniques, and applications being developed all the time.

The healthcare sector in India is a critical area that has undergone significant development and transformation in recent years. India's healthcare system faces numerous challenges, including inadequate infrastructure, limited access to healthcare services, and a shortage of healthcare professionals.

The healthcare industry is facing many challenges, including rising costs and a shortage of healthcare professionals. Machine learning and artificial intelligence (AI) have the potential to address many of these challenges by improving the accuracy and efficiency of medical diagnoses, identifying early warning signs for diseases, predicting treatment outcomes, and more.

Skin diseases are diverse and can range from mild rashes to severe conditions like skin cancer. The early detection and diagnosis of skin diseases are crucial for effective treatment and management. However, skin diseases are often difficult to diagnose due to their varying symptoms and visual appearance. Machine learning models can aid in the early detection and diagnosis of skin diseases. By analyzing a patient's medical and lifestyle data, as well as visual data such as images of the skin, a model can predict the likelihood of a person developing a skin disease and identify the specific condition. This can enable healthcare professionals to provide more accurate diagnoses and personalized treatment plans, ultimately improving patient outcomes.

A hand-held device for skin disease prediction could be a portable diagnostic tool that utilizes machine learning algorithms and visual data analysis to predict skin diseases in real-time. The device could be equipped with a high-resolution camera to capture images of the patient's skin and send them to a machine learning model for analysis.

The CNN, machine learning model could be trained on a large dataset of skin images and associated diagnostic information. By analyzing the image data such as skin color and texture. Then comparing it to the dataset, the model could accurately predict the likelihood of a person developing a skin disease and identify the specific condition.

A hand-held skin disease prediction device could provide healthcare professionals with a more accurate and efficient tool for diagnosing skin conditions. Patients could also benefit from earlier detection and treatment, resulting in improved outcomes and quality of life. Additionally, this device guides the users to nearby dermatologists using Global Positioning Device (GPS) for further treatment if the condition is found to be critical.

II. RELATED WORKS

Skin disease prediction is a popular research area with many studies focused on developing machine learning models for accurate diagnosis and prognosis of skin conditions. Here are a few recent literature surveys on skin disease prediction.

Rimi, T. A., Sultana, N., & Ahmed Foysal, M. F. (2020). “Derm-NN: Skin Diseases Detection Using Convolutional Neural Network”. 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). In this research, they have tried to develop a prototype to detect skin diseases using CNN. Earlier detection works have been done using DNN which is a deep neural network. Right now, have classes to identify a typical skin malady called dermatitis hand, eczema hand, eczema subcute, lichen simplex, statis dermatitis and ulcers. This paper is a sandwich between picture handling strategies and machine learning [1].

Waweru, A. K., Ahmed, K., Miao, Y., & Kawan, P. (2020), “Deep Learning in Skin Lesion Analysis Towards Cancer Detection”. 2020 24th International Conference Information Visualisation (IV). In this

paper, it proposes a method for skin lesion classification and Melanoma detection based on the DenseNet201, a Deep Convolutional Neutral Network (DCNN), using transfer learning. The experiments with the DenseNet201 neural network model utilizing the HAM10000 public dataset has achieved a 78% balanced accuracy. HAM10000 dataset was used as a benchmark in several studies [2].

Prakriti Dwivedi, Akbar Ali Khan, Amit Gawade, Subodh Deolekar,(2020), " A deep learning based approach for automated skin disease detection using Fast R-CNN", This paper is an attempt to develop a system using deep learning technology to detect skin diseases accurately. Using the Fast R-CNN architecture of deep learning, appropriate annotation technique and proper selection of parameters, the results were obtained. We are able to detect the specified skin disease from the given classes with an overall accuracy of 90% and the loss of 0.3 which shows the effectiveness of the model [3].

Roy, K., Chaudhuri, S. S., Ghosh, S., Dutta, S. K., Chakraborty, P., & Sarkar, R. (2019). "Skin Disease detection based on different Segmentation Techniques". 2019 International Conference on Opto-Electronics and Applied Optics (Optronix). In this paper, image processing techniques like adaptive thresholding, edge detection, K-means clustering and morphology-based image segmentation have been used to identify the skin diseases from the given image set [4].

Chakraborty, S., Mali, K., Chatterjee, S., Banerjee, S., Mazumdar, K. G., Debnath, M., Roy, K. (2017). "Detection of skin disease using metaheuristic supported artificial neural networks". 2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON). In this work, a technique based on meta-heuristic supported artificial neural network has been proposed to classify images. Here 3 common skin diseases have been considered namely angioma, basal cell carcinoma and lentigo simplex [5].

Kolkur, S., & Kalbande, D. R. (2016). "Survey of texture based feature extraction for skin disease detection". 2016 International Conference on ICT in Business Industry & Government (ICTBIG). This paper presents a comprehensive survey of texture-based feature extraction for detection of skin diseases and proposes a system based on the findings [6].

Navarro, M. C. R., Bustillos, E., & Barfeh, D. P. Y. (2019). "Skin Disease Analysis using Digital Image processing". 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE). This research study is focused on the detection and classification of skin diseases with the use of the Improved Bag of Features Algorithm. This study used the combined Speed-Up Robust Features (SURF) algorithm for features extraction and k-means clustering [7].

Anthal, J., Upadhyay, A., & Gupta, A., (2017). "Detection of Vitiligo Skin Disease using LVQ Neural Network". 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC). This paper provides the idea which is used to detect the affected area of the Vitiligo disease with help of image captured by camera and classified the affected area from non-affected area in image using LVQ neural network [8].

Rathod, J., Wazhmode, V., Sodha, A., & Bhavathankar, P. (2018). "Diagnosis of skin diseases using Convolutional Neural Networks". 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA). In this study, feature extraction is done using complex techniques such as Convolutional Neural Network (CNN), classify the image based on the algorithm of softmax classifier and obtain the diagnosis report as an output [9].

Suganiya Murugan, S.R. Srividhya, S. Pradeep Kumar, & B. Rubini, (2023), "A Machine Learning Approach to Predict Skin Diseases and Treatment Recommendation System". 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT). This research work predicts the type

of disease by using the clinical parameter of a patient, which is achieved by using various machine learning classification algorithms. Maximum accuracy of 98.4% is achieved by the Naive Bayes algorithm. The Graphical User Interface (GUI) was developed for easy utilization of this prediction module for both doctors and dermatologists [10].

Suparyati, Ema Utami, Alva Hendi Muhammad, (2022), "Lumpy Skin Disease Prediction Based on Meteorological and Geospatial Features using Random Forest Algorithm with Hyperparameter Tuning" , 2022 5th International Conference on Information and Communications Technology (ICOIACT). The purpose of this study is to optimize the Random Forest classification in predicting LSD with Genetic Algorithm (GA) as a hyperparameter tuning and using SMOTE as a resampling technique for unbalanced datasets.

III. METHODOLOGIES

Convolutional Layer: The first layer of a CNN is typically a convolutional layer. This layer applies a set of learnable filters to the input image, resulting in a set of output feature maps. Each filter looks for specific features in the image, such as edges, lines, or textures.

Activation Function: An activation function is applied to the output of each filter in the convolutional layer to introduce non-linearity into the network. The most commonly used activation function is the ReLU (Rectified Linear Unit) function.

Pooling Layer: The pooling layer downsamples the output feature maps by taking the maximum or average value within a sliding window. This reduces the size of the feature maps while preserving the most important information.

Fully Connected Layer: The fully connected layer takes the output of the previous layers and applies a set of weights to generate the final output. This layer performs the classification by computing the probabilities of the input image belonging to each class.

Softmax Function: The softmax function is applied to the output of the fully connected layer to convert the final output into a probability distribution over the different classes. The class with the highest probability is then selected as the predicted class for the input image.

Training: The CNN is trained on a large dataset of labeled images. The training process involves iteratively feeding batches of images through the network and updating the weights using backpropagation. The goal is to optimize the network's performance on the training set so that it can generalize well to new, unseen images.

Testing: Once the CNN has been trained, it can be used to classify new, unseen images. The input image is fed through the network, and the predicted class is obtained using the softmax function. The performance of the network is evaluated on a separate test set to assess its accuracy and generalization ability.

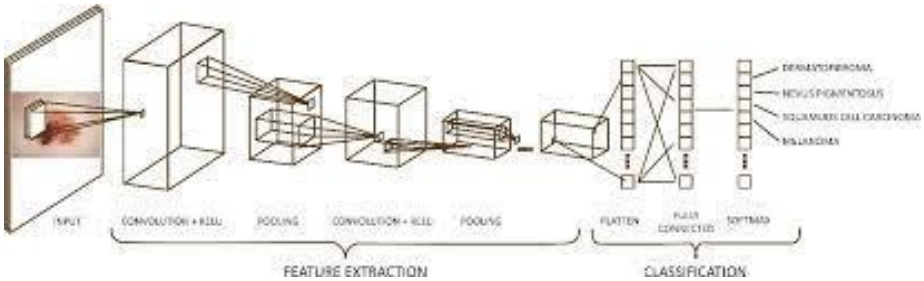

Figure.1

IV. ALGORITHM USED

CNN: Convolutional Neural Networks (CNNs) are a type of deep learning model that is primarily used for image and video analysis tasks. CNNs are inspired by the structure and function of the human brain's visual cortex, which contains neurons that are highly specialized for detecting specific visual patterns in the input signals.

The most important types of layers in a CNN are:

Convolutional Layers: These layers apply a set of filters to the input data and produce a set of feature maps. Each filter is a small matrix that slides over the input data and performs a dot product at each location, resulting in a new value in the output feature map.

Pooling Layers: These layers reduce the spatial dimensions of the feature maps by taking the maximum or average value within a small window. This helps to make the network more robust to small variations in the input data.

Fully Connected Layers: These layers connect all the nodes in the previous layer to all the nodes in the current layer, allowing the network to learn complex, non-linear relationships between the input and output data.

SOFTMAX CLASSIFICATION

FORMULA USED:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Softmax classification is a type of multiclass classification algorithm used in machine learning and deep learning. It is commonly used for problems where there are multiple classes, and the output of the model needs to be a probability distribution over all possible classes.

The softmax function is used in the final layer of a neural network to convert the output of the network into a probability distribution over the classes. The softmax function takes the raw output values from the previous layer and normalizes them so that they sum to 1, representing a probability distribution.

Softmax classification is widely used in deep learning models for image classification, natural language processing, and other multiclass classification problems. It is computationally efficient and can handle large numbers of classes, making it suitable for many real-world applications.

V. FUTURE SCOPE

As machine learning algorithms continue to advance, we can expect skin disease prediction models to become even more accurate and reliable. This could lead to earlier detection of skin diseases, which could ultimately save lives. Skin disease prediction models may be able to provide personalized treatment recommendations based on a patient's skin type, medical history, and other factors. This could lead to more effective treatment outcomes and improved patient satisfaction.

VI. CONCLUSION

Skin disease prediction using machine learning algorithms is a rapidly growing field that has the potential to revolutionize dermatological diagnosis and treatment. With the increasing availability of large datasets

and advances in computer vision and artificial intelligence technologies, skin disease prediction models are becoming more accurate and reliable.

However, there are still challenges to overcome, including the need for standardized datasets and the difficulty in differentiating between similar skin diseases. Despite these challenges, skin disease prediction using machine learning is a promising area of research that has the potential to improve patient outcomes and increase access to care. It is likely that we will see continued progress and innovation in this field in the coming years.

VII. REFERENCES

1. Rimi, T. A., Sultana, N., & Ahmed Foysal, M. F. (2020). "Derm-NN: Skin Diseases Detection Using Convolutional Neural Network". 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS).
2. Waweru, A. K., Ahmed, K., Miao, Y., & Kawan, P. (2020), "Deep Learning in Skin Lesion Analysis Towards Cancer Detection". 2020 24th International Conference Information Visualisation (IV).
3. Prakriti Dwivedi, Akbar Ali Khan, Amit Gawade, Subodh Deolekar,(2020), " A deep learning based approach for automated skin disease detection using Fast R-CNN" , 2021 Sixth International Conference on Image Information Processing (ICIIP)
4. Roy, K., Chaudhuri, S. S., Ghosh, S., Dutta, S. K., Chakraborty, P., & Sarkar, R. (2019). "Skin Disease detection based on different Segmentation Techniques". 2019 International Conference on Opto-Electronics and Applied Optics (Optronix).
5. Chakraborty, S., Mali, K., Chatterjee, S., Banerjee, S., Mazumdar, K. G., Debnath, M., Roy, K. (2017). "Detection of skin disease using metaheuristic supported artificial neural networks". 2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON).
6. Kolkur, S., & Kalbande, D. R. (2016). "Survey of texture-based feature extraction for skin disease detection". 2016 International Conference on ICT in Business Industry & Government (ICTBIG). This paper presents a comprehensive survey of texture-based feature extraction for detection of skin diseases and proposes a system based on the findings.
7. Navarro, M. C. R., Bustillos, E., & Barfeh, D. P. Y. (2019). "Skin Disease Analysis using Digital Image processing". 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE). This research study is focused on the detection and classification of skin diseases with the use of the Improved Bag of Features Algorithm. This study used the combined Speed-Up Robust Features (SURF) algorithm for features extraction and k-means clustering.
8. Anthal, J., Upadhyay, A., & Gupta, A., (2017). "Detection of Vitiligo Skin Disease using LVQ Neural Network". 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC). This paper provides the idea which is used to detect the affected area of the Vitiligo disease with help of image captured by camera and classified the affected area from non-affected area in image using LVQ neural network.
9. Rathod, J., Wazhmode, V., Sodha, A., & Bhavathankar, P. (2018). "Diagnosis of skin diseases using Convolutional Neural Networks". 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA).
10. Suganiya Murugan, S.R. Srividhya, S. Pradeep Kumar, & B. Rubini, (2023), "A Machine Learning Approach to Predict Skin Diseases and Treatment Recommendation System". 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT). This research work predicts the

type of disease by using the clinical parameter of a patient, which is achieved by using various machine learning classification algorithms. Maximum accuracy of 98.4% is achieved by the Naive Bayes algorithm. The Graphical User Interface (GUI) was developed for easy utilization of this prediction module for both doctors and dermatologists.

11. Suparyati, Ema Utami, Alva Hendi Muhammad, (2022), "Lumpy Skin Disease Prediction Based on Meteorological and Geospatial Features using Random Forest Algorithm with Hyperparameter Tuning" , 2022 5th International Conference on Information and Communications Technology (ICOIACT). The purpose of this study is to optimize the Random Forest classification in predicting LSD with Genetic Algorithm (GA) as a hyperparameter tuning and using SMOTE as a resampling technique for unbalanced datasets.