Project Report

on

Skin Disease Detection Using CNN

Submitted in partial fulfilment of the requirements

of

the degree of

Bachelor of Technology

by

Aniket Kumar (Roll No. 221230008)
Kshitiz Kanwatia (Roll No. 221230032)
Md. Zamin Alam (Roll No. 221230034)
Nitin Kushwah (Roll No. 221230042)

Supervisor:
Dr. Sachin Singh

Department of Electrical Engineering

NATIONAL INSTITUTE OF TECHNOLOGY DELHI

APPROVAL SHEET

This Project work entitled "Skin Disease Detection Using CNN" by Aniket Kumar, Kshitiz Kanwatia, Md. Zamin Alam and Nitin Kushwah is approved for the degree of Bachelor of Technology

Examiner

…………………………………

Dr. Manoj Kumawat

Supervisor

…………………………………

Dr. Sachin Singh

Chairman

…………………………………

Date: …………………………………

Place: …………………………………

DECLARATION

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources, which have thus not been properly cited, or from whom proper permission has not been taken when needed.

………………………	………………………	………………………	………………………
Aniket Kumar	**Kshitiz Kanwatia**	**Md. Zamin Alam**	**Nitin Kushwah**
221230008	**221230032**	**221230034**	**221230042**

Date:

ACKNOWLEDGEMENTS

We would like to thank Honourable Director Sir, Prof. Ajay K. Sharma, for providing us with excellent facilities and a conducive environment necessary to complete our project work.

We would like to express our deep sense of gratitude to our Supervisor, Dr. Sachin Singh, who has always guided us and provided us with valuable feedback throughout the period of our project work.

Last but not the least, we would like to thank our parents and well-wishers who supported us unconditionally throughout this phase.

………………………	………………………	………………………	………………………
Aniket Kumar	**Kshitiz Kanwatia**	**Md. Zamin Alam**	**Nitin Kushwah**
221230008	**221230032**	**221230034**	**221230042**

Table of Contents

Abstract—Globally, skin diseases affect many people. Treatment and care for these diseases can only be effective if detected early and accurately. Machine learning and computer vision have advanced to the point that convolutional neural networks (CNNs) are now powerful tools for automatically detecting skin diseases. A study on the application of CNNs to skin disease detection is presented, demonstrating the potential of deep learning algorithms to aid dermatologists and healthcare professionals. An image dataset containing images of various skin diseases is used in the proposed method, including dermatitis, eczema, psoriasis, and other common conditions. High-level representations are captured from input images using a pre- trained CNN (Convolutional Neural Network) model. For disease detection, these learned features are fed into a classification layer. The CNN model learns to identify patterns that correspond to the various diseases after being trained on a vast collection of skin photos. An image of skin disease is fed into the model, and the model is trained to extract features that can be used to build a classification model that can accurately identify the disease.

Index Terms - Machine learning, CNN (Convolution neural networks), pattern recognition, feature extraction, and disease identification.

I. INTRODUCTION

Skin disease detection techniques are complex due to the diverse types of skin diseases that may have similar symptoms. Factors such as the appearance of the rash or lesion, its location, and associated symptoms like itching or pain must be considered for a correct diagnosis. In the era of evidence-based medicine, medical professionals must rely on cutting-edge diagnostic techniques like telescopes, laboratories, and radiology to identify skin illnesses. Technological advancements have made diagnosis more logical and practical, but physicians often rely on personal judgment to conclude symptoms, leading to misperceptions and incorrect treatments.

This project aims to aid in the early diagnosis of skin disorders using photos as input. Convolutional neural networks (CNNs) have emerged as powerful tools for computer vision tasks, including skin disease detection and classification. CNNs have revolutionized the field of skin disease detection by enabling accurate and reliable diagnosis of a wide range of skin diseases. However, when trained on certain datasets, models struggle with new and unfamiliar cases, which can compromise accuracy and inclusiveness, particularly in diverse groups of people.

In conclusion, technology and modern science play a vital role in recognizing the symptoms of skin diseases. This project aims to aid in the early diagnosis of skin disorders using photos as input.

II. LITERATURE REVIEW

Skin situations are the maximum common place class of human illnesses. Viruses, bacteria, fungal infections, and different situations are some of the reasons for pores and skin situations. Dermatologists are vital within the traditional approach of figuring out pores and skin situations. A dermatologist begins by researching an affected person and makes use of his information and enjoy to decide the affected person‘s pores and skin condition. Dermatoscopy, an imaging approach, is then done to view the shape of the pores and skin. Due to dermatologists’ subjective visible evaluation, diagnosing pores and skin sicknesses may be difficult. However, a CNN version has been proven to carry out the identical stage as dermatologists in figuring out pores and skin sicknesses, together with cancer and keratinocyte carcinomas . The version will be used as a diagnostic resource for dermatologists in scientific practice.

CNN’s AI (Artificial Intelligence) machine for ailment popularity cannot update dermatologist’s holistic expertise. Further research is wanted to deal with interpretability, trust, and moral concerns in integrating AI into scientific practice. Future studies ought to discover combining dermatologists’ judgment with AI’s performance even as improving AI’s explainability and addressing ability biases.

Dermatoscopy is a non-invasive imaging approach used to study pores and skin lesions for early detection of cancer and different pores and skin situations. Analyzing Dermatoscopy snapshots calls for distinguishing among awesome sorts of lesions, a venture that may be complicated because of the diffused and elaborate visible styles gift inside the snapshots. Dermatologists should create dependable and automatic diagnostic structures due to the fact guide evaluation is timeeating and susceptible to inter-observer variability . By utilising deep ResNet architectures, the version can seize each low- stage and excessive-stage feature, enhancing its capacity to discriminate among visually comparable lesions.

The dermatologic analysis is complicated due to the huge sort of pores and skin sicknesses with overlapping visible features. The venture is in addition complex through the want for early detection and differentiation of diverse pores and skin sicknesses, together with cancer, to make well-timed and powerful remedy decisions. Existing techniques have trouble reaching excessive accuracy and generalization throughout unique pores and skin kinds and situations. Their answer leverages the strengths of deep getting-to-know strategies to enhance accuracy and generalization in diagnosing pores and skin sicknesses from clinical snapshots.

III. EXISTING SYSTEM

Deep learning models have led to significant progress in the classification of skin diseases. Dermatologists are always faced with the challenging task of accurately identifying a variety of skin disorders, one of which is melanoma, which requires fast and exact detection in order to enable effective treatment. Conventional diagnostic techniques frequently display subjectivity and are susceptible to differences in observation between parties. As a result, researchers have investigated the potential applications of deep learning techniques, particularly Convolutional Neural Networks (CNNs), for automating and enhancing the classification of skin conditions.

The objective of this research delve into an existing framework for skin disease classification utilizing Deep Learning. The system harnesses an advanced CNN architecture to categorize Dermatoscopy images of skin lesions into distinct diagnostic classes. Our objective encompasses a comprehensive exploration of the system's blueprint, architecture, and methodology, elucidating its pivotal role in elevating diagnostic accuracy within the domain of dermatology.

The implemented system relies on a meticulously curated dataset of Dermatoscopy images, encompassing an array of skin lesion types and conditions. The dataset is thoughtfully partitioned into training, validation, and testing subsets, ensuring a robust evaluation of the model's efficiency.

Preceding the training phase, the dataset undergoes preparatory procedures encompassing resizing, normalization, and augmentation. Augmentation techniques, including rotation, scaling, and flipping, are thoughtfully employed to enrich the diversity of training instances and bolster the model's proficiency in generalization.

The CNN design, which consists of several convolutional layers, pooling components, and fully connected layers, is strengthened with batch normalization and dropout methods to thwart overfitting tendencies and improve training convergence.

Training involves optimizing the CNN by minimizing a predefined loss function through iterative updates of model weights. Esteemed optimization algorithms like stochastic gradient descent (SGD) or its variants guide this iterative weight refinement process .

By evaluating the training process using the validation dataset, it is continuously observed. Model checkpoints are systematically documented according to the best validation performance, preventing overfitting and confirming the model's ability to handle new data.

Upon completion of training, the model is subjected to rigorous testing against an independent test dataset to gauge its precision, recall, F1 score, accuracy, and other pertinent metrics. System performance is judiciously compared against benchmark standards and, when feasible, against the diagnostic capabilities of human dermatologists.

The experimental outcomes lucidly highlight the system's efficacy in accurately classifying a diverse array of skin lesions. Our discourse further unravels these results, spotlighting the system's strengths, and unveiling avenues for refining and enhancing its performance.

IV. METHODOLOGY

• Step 1: Data collection for skin disease detection using Convolutional Neural Networks (CNNs) is a crucial step that requires collecting a diverse and representative dataset of skin images. We start by resizing the images to 32, 32 for better learn and then add the names and labels, after which the plot parameters are determined. The image pixels are stored as a dependent variable, while the target label is stored as an independent feature.

• Step 2: Using TensorFlow's image preprocessing function, the uploaded image goes through a series of trans- formations. First, it is encoded into a byte stream and then decoded into an image format. It is then resized to a fixed size of 32 x 32 pixels and reshaped to match the input shape provided by the model. By dividing the image's pixel values by 255, the image is then normalized. This process makes sure that the forecasts are consistent, which improves the model's accuracy.

• Step 3: The training data set and the test data set are the two groups into which the data set is split during the classification process. The test set is used to assess the CNN model's performance after it has been trained using the training set. The data is fit to the training set and the new shape is checked, after which it is reshaped into 3 dimensions again to fit the Convolutional neural network.

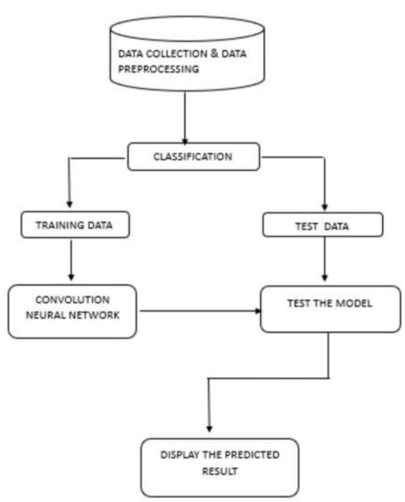


Fig. 1. Methodology

• Step 4: After the model has made its predictions for the test set in the skin disease detection task, the next step is to analyze and interpret these results. This is done by examining the output of the model for each image in the test set and comparing it to the actual labels. The predicted result is displayed using the web application.

V. WORK PROPOSED

1) Integrated Libraries and Modules: Data analysis, visualization, and deep learning are implemented by importing the necessary libraries and modules. For web applications, streamlit is used, NumPy and Pandas are used for data manipulation and analysis, and seaborn and matplotlib. Pyplot is used for data visualization, io is used for input/output operations, and TensorFlow and Keras are used for deep learning.

2) Pre-trained Model and Image Upload with File Uploader: Predictions are made with the loaded model from a pre-trained model file named'Model.h5'. Distinct types of skin diseases are also represented by a dictionary of classes. The file uploader function prompts the user to upload an image to predict skin diseases .

3) Image Preprocessing for Model Input : With TensorFlow's image preprocessing function, the uploaded image is converted into a byte stream, loaded as an image, resized to a specific size of 32x32 pixels, and reshaped to the required input shape of the model. The next step is to normalize the image by dividing its pixel values by 255 to ensure consistency in the prediction and improve accuracy .

4) Predicting Skin Diseases and Displaying Results : After image preprocessing, the model predicts skin diseases.

5) Data Set about Ham10000 : Training of neural networks for automated diagnosis of pigmented skin lesions is hampered by the small size and lack of diversity of available dataset of dermatoscopic images. We tackle this problem by releasing the HAM10000 ("Human Against Machine with 10000 training images") dataset. We collected dermatoscopic images from different populations, acquired and stored by different modalities. The final dataset consists of 10015 dermatoscopic images which can serve as a training set for academic

machine learning purposes. Cases include a representative collection of all important diagnostic categories in the realm of pigmented lesions: Actinic keratoses and intraepithelial carcinoma / Bowen's disease (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv) and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, vasc).

More than 50% of lesions are confirmed through histopathology (histo), the ground truth for the rest of the cases is either follow-up examination (follow_up), expert consensus (consensus), or confirmation by in-vivo confocal microscopy (confocal). The dataset includes lesions with multiple images, which can be tracked by the lesion_id-column within the HAM10000_metadata file.

VI. OVERCOMING CHALLENGES

Our project aims to address the limitations of existing skin disease detection systems by leveraging the advancements in Convolutional Neural Networks (CNNs). Unlike conventional approaches that suffer from limited interpretability and data quality issues, our system harnesses the high accuracy and feature learning capabilities of CNNs to achieve robust and efficient skin disease diagnosis . By incorporating trans- fer learning and adaptability to diverse datasets, our model can generalize well even with limited labeled medical data, ensuring reliable performance across different skin types and imaging conditions.

Automation and efficiency are enhanced through automated skin disease detection, reducing the reliance on manual inspection and enabling early detection, which is critical for timely intervention and improved patient outcomes. Moreover, our system facilitates multiclass classification, scalability, and continuous learning, allowing for the integration of new data and diagnostic modalities to provide a comprehensive under- standing of skin conditions. These advancements, coupled with ongoing research in CNN architectures, contribute to continual improvements in skin disease detection, addressing ethical concerns and enhancing overall diagnostic accuracy and efficiency.

VII. PREDICTION AND CLASSIFICATION

• CNN (Convolution Neural Network: Convolutional Neural Networks (CNNs) are a class of deep learning models specifically designed for tasks involving structured grid-like data such as images and videos. They have demonstrated exceptional efficiency in a range of imageprocessing applications, including picture segmentation, object detection, and classification.

CNNs are especially well adapted for this purpose because of their capacity to learn complicated patterns and features from images. This allows them to discern between different types of skin issues.

Various skin disease images are collected and labeled. There should be a wide range of skin conditions, skin tones, lighting conditions, ages, and anatomical sites included in this dataset. To train an accurate CNN model, a high-quality data set must be available .

• Model Architecture: The CNN architecture was chosen or created specifically for the goal of detecting skin diseases. Transfer learning is frequently used to optimize a pre-trained CNN model for the job of detecting skin diseases, such as VGG, ResNet, or Inception, which was trained on a sizable dataset like ImageNet. The earlier layers of the pre-trained network serve as feature extractors, and the later layers are fine-tuned to learn disease-specific features. The model looks at features within the image (such as color, texture, pattern, etc.) and applies its learned knowledge to decide what skin disease or condition is present .

• Pre-Processing: Skin disease images are pre-processed before feeding onto CNN. Pre-processing approaches frequently used to improve the diversity of training data include scaling the images to a consistent size, standardizing pixel values, and data augmentation methods . The pre-processing technique utilized in this work is essential for getting the input data ready for precise segmentation and classification. Pre-processing entails a set of procedures used to the raw images to enhance their quality and remove pertinent elements for further analysis. In the context of skin lesion analysis, this step is particularly important to ensure accurate segmentation of lesions and extract meaningful information for classification.

• Prediction: Detection of lesion margins is a fundamental step in the analysis of Dermatoscopy images, as it delineates the boundary between normal and potentially abnormal areas of skin. Classification, however, involves dividing detected lesions into specific classes or categories based on their visual characteristics. This step is crucial for diagnosing skin diseases and distinguishing between benign and malignant lesions. The study assesses how well the ANN model performs, reporting metrics like accuracy, sensitivity, and specificity that show how well it can diagnose various skin disorders .Neural networks can be used for both classification and feature extraction. A transfer learning model is developed in which the VGG-16 layered CNN architecture for feature extraction can extract 1000 features from the input image.

• Evaluation Metrics: The performance of the proposed model was measured using the accuracy metric as defined by equation (1) given by

$$Accuracy = \frac{TP + TF}{TP + FP + FF + TF} \times 100$$

Where:

TP = number of correctly predicted Positive.

TF = number of correctly predicted false.

FP = number of incorrectly predicted positive.

FF = number of incorrectly predicted false.

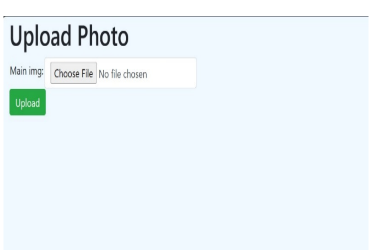


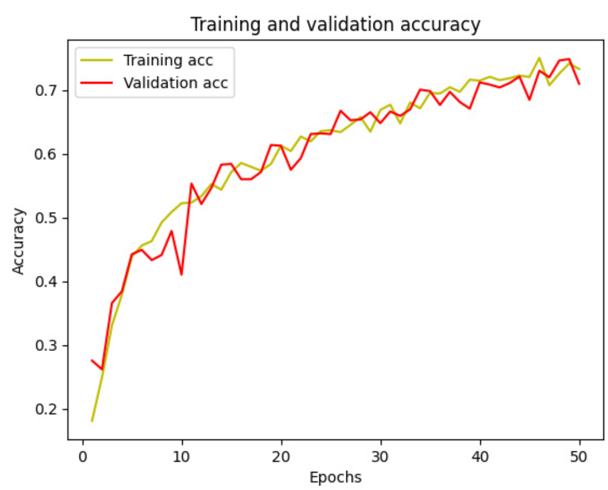
Fig. 3. Image Browser Input

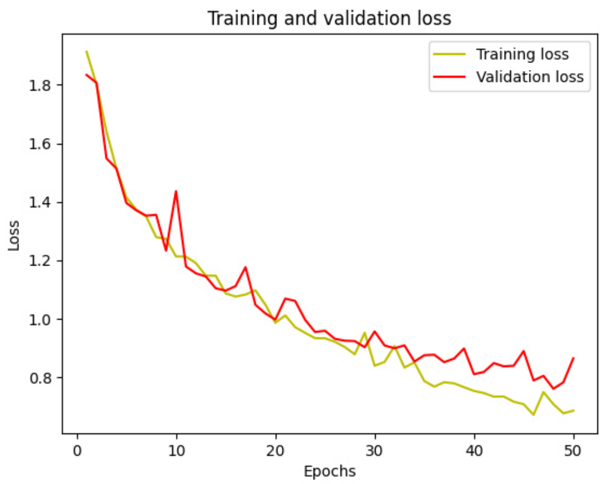
Fig. 4. Predicted Output

This provides insights into how well the model is performing for different classes and overall. Both prediction and classification are critical components in building effective skin disease detection systems because they allow automatic and accurate identification of various skin diseases.

B. SAMPLE IMAGE OF COMMON SKIN DISEASE

C. MODEL STATISTICS






```
Test accuracy: 0.7097142934799194
```

VIII. CONCLUSION

In conclusion, employing Convolutional Neural Networks (CNNs) for skin disease identification has demonstrated promising results in precisely diagnosing a variety of skin efficiency in learning from and extracting information from dermatological images. This technology can revolutionize dermatology by providing a reliable and efficient tool for early diagnosis and treatment of skin diseases. However, it is critical to further refine and validate these models with diverse and representative datasets to ensure their effectiveness in different populations. Furthermore, integrating these CNN-based detection systems into clinical practice can increase the efficiency and accuracy of dermatologic assessments, improving patient outcomes and dermatology healthcare. CNN models can analyze skin images to identify specific disease features such as lesions, patterns, and textures,which helps dermatologists and medical professionals make informed decisions and provide timely treatment. Future research and development efforts may focus on refining and extending CNN architectures, incorporating multimodal data sources (e.g., clinical information, patient history) and improving the interpretability of CNN-based models. Collaboration between dermatologists, data scientists, and engineers is essential to advance the field, improve the accuracy of skin disease detection, and contribute to more effective treatment and management of skin disease.

IX. REFERENCES

[1] Diagnosis of skin diseases using Convolutional Neural Networks Publisher: IEEE Jainesh Rathod; Vishal Wagh mode; Aniruddh Sodha; Praseniit Bhavathankar Jainesh Rathod; Vishal Waghmode; Aniruddh Sodha; Praseniit Bhavathankar.

[2] Z. Ma and J. M. R. S. Tavares, ”A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model”, IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 2, pp. 615- 623, March 2016.

[3] In [2] Megha D. Tijare et al. “Detecting skin disease by accurate skin segmentation using various color spaces”.

[4] In [1] Nawal Soliman, ALKolifi AlEnezi at el. “A method of skin disease detection using Image Processing and machine learning”.

[5] Eldho Paul, RS Sabeenian (2022) “Modified convolutional neural network with pseudo-CNN for removing nonlinear noise in digital images”

[6] Karthik R, Tejas Vaichole and Sanika Kulkarni. Channel Attention based Convolutional Network for skin disease classification. ScienceDirect, August 2021.

[7] CNN with Transfer Learning for Skin Lesion Identification Using Tasmanian Devil Optimization AlgorithmVineet Kumar Dubey, Vandana Dixit Kaushik.

[8] Thurnhofer-Hemsi K, Lo´pez-Rubio E, Dom´ınguez E, Elizondo DA (2021) Skin lesion classification by ensembles of deep convolutional networks and regularly spaced shifting. IEEE, vol 9.