A report on

"Skin Disease Detection using ML"

Submitted in partial fulfilment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

SHIVAM KASAUDHAN
19BTRCR069

SAURABH MISHRA
19BTRCR067

SUNNY KUMAR YADAV
19BTRCR071

MANISH KUMAR SAH
19BTRCR062

Under the guidance of

Dr M Sivaram

Assistant Professor

Faculty of Engineering & Technology

Jain (Deemed-To-Be University)

B. Tech in Computer Science and Engineering

Jain Global Campus, Kanakapura Taluk - 562112

Ramanagara District, Karnataka, India

2022-2023

B. Tech in Computer Science and Engineering

Jain Global Campus, Kanakapura Taluk - 562112

Ramanagara District, Karnataka, India

CERTIFICATE

This is to certify that the project work titled **"Skin Disease Detection using ML"** is carried out by **Shivam Kasaudhan (19BTRCR069), Saurabh Mishra (19BTRCR067), Sunny Kumar Yadav (19BTRCR071) and Manish Kumar Sah (19BTRCR062)** the bonafide students of Bachelor of Technology at the Faculty of Engineering & Technology, Jain (Deemed-to-be University), Bangalore in partial fulfilment for the award of degree, Bachelor of Technology in Computer Science and Engineering, during the Academic year **2022-2023**.

Dr M SHIVRAM
Assistant Professor
Faculty of Engineering & Technology,
Jain (Deemed-to-be University)
Date:
Signature:

Dr S RAMESH
Assistant Professor and
Dept. of CSE
Faculty of Engineering & Technology,
Jain (Deemed-to-be-University)
Date:
Signature:

Name of the Examiner Signature of Examiner

1.

2.

DECLARATION

We , **Shivam Kasaudhan (19BTRCR069), Saurabh Mishra (19BTRCR067), Sunny Kumar Yadav (19BTRCR071) and Manish Kumar Sah (19BTRCR062)** are students of seventh semester B. Tech (Honours) in **Computer Science (Data Science)**, at Faculty of Engineering & Technology, **Jain (Deemed-To-Be University)**, hereby declare that the project work titled **"Skin Disease Detection using ML"** has been carried out by us and submitted in partial fulfilment for the award of degree in **Bachelor of Technology (Honours) in Computer Science (Data Science)** during the academic year **2022-2023**. Further, the matter presented in the project has not been submitted previously by anybody for the award of any degree or any diploma to any other University, to the best of our knowledge and faith.

Shivam Kasaudhan Signature
USN: 19BTRCR069

Saurabh Mishra Signature
USN: 19BTRCR067

Sunny Kumar Yadav Signature
USN: 19BTRCR071

Manish Kumar Sah Signature
USN: 19BTRCR062

Place: Bengaluru
Date: 12/12/2022

ACKNOWLEDGEMENT

It is a great pleasure for us to acknowledge the assistance and support of a large number of individuals who have been responsible for the successful completion of this project work.

First, we take this opportunity to express our sincere gratitude to the Faculty ***of Engineering & Technology, Jain (Deemed-to-be University),*** *for providing us with a great opportunity to pursue our Bachelor's Degree in this institution.*

In particular we would like to thank ***Dr. Hariprasad S A, Director, Faculty of Engineering & Technology, Jain (Deemed-to-be University),*** *for his constant encouragement and expert advice.*

We place on record, our sincere thank you to ***Dr. Geetha, Dean, Faculty of Engineering & Technology****,* ***Jain (Deemed-to-be University),*** *for the continuous encouragement.*

It is a matter of immense pleasure to express our sincere thanks to ***Dr. S. Ramesh Program Head****,* ***Computer Science & Engineering****,* ***Jain (Deemed- to-be University),*** *for providing the right academic guidance that made our task possible.*

It is a matter of immense pleasure to express our sincere thanks to ***Prof. Mohamed Zabeeulla, Assistant Professor, Dept. of Computer Science & Engineering, Jain (Deemed-to-be University),*** *for providing the right academic guidance that made our task possible.*

We would like to thank our Project Coordinator ***Dr M. John Basha, Assistant Professor, Dept. of Computer Science & Engineering, Jain (Deemed- to-be University),*** *for providing the right academic guidance that made our task possible.*

We would like to thank our guide ***Dr M Shivaram Assistant Professor****,* ***Dept. of Computer Science & Engineering****,* ***Jain (Deemed-to-be University),*** *for sparing his valuable time to extend help in every step of our project work, which paved the way for smooth progress and fruitful culmination of the project.*

We would like to thank all the staff members of ***Dept. of Computer Science & Engineering, Jain (Deemed-to-be University)****, for their support.*

We are also grateful to our family and friends who provided us with every requirement throughout the course.

We would like to thank one and all who directly or indirectly helped us in completing the Project work successfully

Signature of Students

TABLE OF CONTENTS

ABSTRACT

Skin diseases are the most common form of infections occurring in people of all ages. As the costs of dermatologists to monitor every patient is very high, there is a need for a computerized system to evaluate a patient's risk of skin disease using images of their skin lesions. Many researchers have used different pre-processing and classification techniques to determine whether a skin image suffers from diseases or not. Feature extraction is very important for predictive modelling applications. Feature extraction in image processing is a method of capturing visual content of images for indexing and retrieval. Primitive image features can be either general features, such as extraction of colour, texture and shape or domain specific features. Texture based features are widely used in image analysis for medical diagnosis.

Skin disease recognition and observing is a major challenge looked by the medical industry. Because of expanding contamination and utilization of lousy nourishment, the tally of patients experiencing skin related issues is expanding at a quicker rate. Well-being isn't the main concern, however unfortunate skin hurts our certainty. Customary and appropriate skin checking is a significant advance towards early discovery of any destructive or starting changes in skin that may bring about skin disease. Machine learning methods can add to the improvement of capable frameworks which can order various classes of skin illnesses. To identify skin maladies, first, it is required to separate the skin and non-skin. In this paper, five diverse machine learning algorithms have been chosen and executed on skin infection data set to anticipate the exact class of skin disease.

LIST OF FIGURES

NOMENCLATURE USED

S.N.	Symbol	Meaning
1.	DL	Deep Learning
2.	CNN	Convolutional Neural Network
3.	ML	Machine Learning
4.	ANN	Artificial Neural Network

Chapter 1

INTRODUCTION

1.1. Overview

Skin is the largest and most sensitive part of the human body which protects our inner vital parts and organs from the outside environments. Skin diseases occur commonly among humans. They are usually caused by factors like different organism's cells, different diets, and internal and external factors.

Skin disease diagnosis includes a series of pathological laboratory tests for the identification of the correct disease. Skin disease has been a matter of concern as of their sudden arrival and their complexity. Skin diseases should be detected at the early stage and treated properly.

As there is rapid improvement of computer use in the medical sector, computer technology can make it simpler to detect the disease just from the images of the infected skin. We can implement ML models for the classification or detection of skin diseases in the early stage.

1.2. Problem Definition

Skin disease is the major problem nowadays. People need to spend a lot of money in minor cases. When we go to the hospital, doctors have to do a lot of diagnosis to determine the patterns of diseases. To avoid all these processes and expenditures, ML helps in prior recognition of the diseases and it first measures to control the diseases. This ML model helps doctors also to know the diseases without doing a lot of check-ups.

1.3. Objectives

- It gives prior information to the infected person about the diseases.
- Easily detection of the skin diseases before doing numerous tests.
- It assists the doctor to diagnose the different skin diseases (Time Saving).

1.4 Methodology

The proposed work is based on the classification of skin diseases in the human being. Objective of this model is to detect the skin diseases types and give the first information about the diseases to the patient. In order to demonstrate this, we have utilized deep learning techniques. As the model used in this project is trained in ANN, we have used the CNN for the feature extraction from the images of different skin diseases.

1. Get the HAM10000 data set from the Kaggle.
2. The dataset contains the different value of the picture pixel in it.
3. The images are passed through the model CNN through the different layers of CNN for the feature extraction.
4. Different dense layers and conv2D layers are used and activation functions are used to boost the model.
5. Batch sizes differ in different iterations for the model evaluation.
6. Different samples are taken to learn the feature from the picture along with alternative activation and max-pooling layer.
7. At last ANN is used through which the extracted features are passed from ANN for the classification of the images of different skin diseases.
8. After this the model will predict the label of the skin diseases.

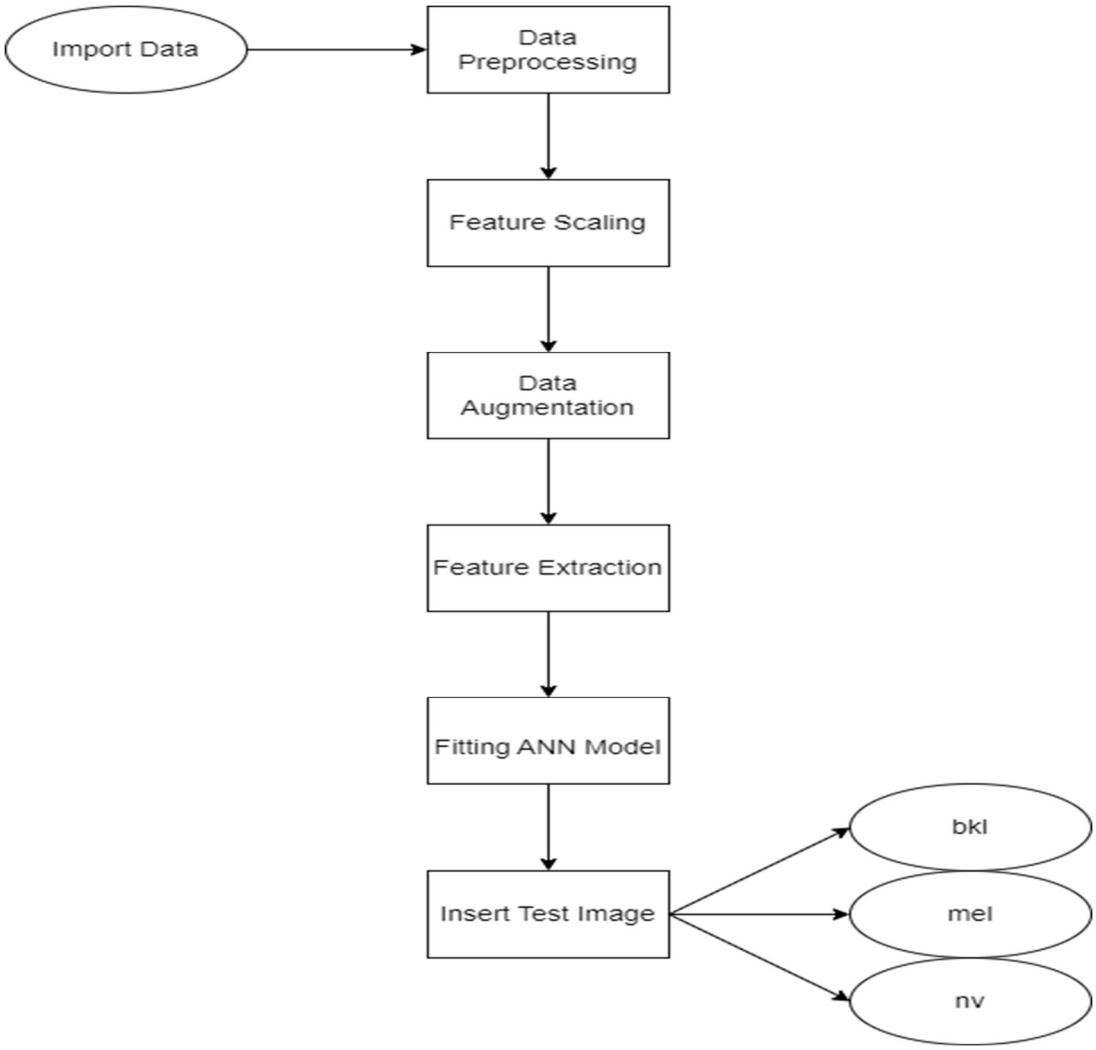

Fig 1.1: Flow Chart of the system used in this project

1.5 Hardware and Software tools used

This section gives a detailed description about the hardware tools and software tools involved in developing this system and how they are used.

Software:

- **Python**: Python is clearly one of the best languages for machine learning. Python does contain special libraries for machine learning namely SciPy, pandas, Scikit-learn and NumPy, which is great for linear algebra, and getting to know the kernel method of machine learning. The language is great to use when working with machine learning algorithms and has easy syntax.

- **Tensorflow Lite**: A deep learning framework for on-device inference. Train and deploy machine learning models on mobile and IoT devices, Android, iOS, Edge TPU, Raspberry Pi.
- **Jupyter Notebook:** The Jupyter Notebook is a web-based interactive computing platform. The notebook combines live code, equations, narrative text, visualizations.

Hardware:

- A simple computer that can run a web browser.

Chapter 2

LITERATURE SURVEY

2.1. Existing System

Methods	Overview	Accuracy	Future Scope
Nasr-Esfahmi et al	• Two layered CNN is used for melanoma classification • Illumination correction mask generation , gaussian filter is applied to remove noise issues. • Cropping and rotation is used for data augmentation.	81%	• Deep layered network may be used for increased results. • Training will get improved by collecting more data
Khryashcheve at al	• CNN u-Net based algorithm is used proposed for medical image segmentation at expert level • AlexNet based algorithm is used for automatic markup of image database.	Not defined	• The algorithm is not used in real time analysis due to more time complexity

Sahu et al	• Classification of benign and malignant skin lesions. • Hybrid deep learning mobile mobileNet , domain specific knowledge and features suggested by dermatologists are used to improve the accuracy of the classifier.	78% to 80%	• Detection of cancer with raspberry pi is not good option because to analyse image high level camera is required

2.2. Limitation of Existing System

- If the appearance of two or more skin diseases are the same, then it will be difficult to distinguish by the model.
- New diseases which have not been discovered yet can't be detected by the model.
- Lack of a dataset for each disease can result in an underfit model.
- Some diseases can't be diagnosed by the model alone without laboratory tests.
- Model training can be expensive due to extensive training of image classification algorithms

2.3. Proposed System

In base paper, they have used SVM as their main classifier. Using SVM, the accuracy we get is around 70-80%.

In this experiment we are replacing the SVM classification part with ANN to classify the data into cancerous cells or non-cancerous cell.ANN, or Artificial Neural Network, is a machine learning technique based on the working of the human brain. It has input layers, hidden layers and output layer. Computation is done in hidden layers. It can be used to solve all the problems that can be expressed as a linear function.

2.4 Advantages of Proposed System

- **Early detection:** Skin disease detection can help to detect any skin disorder at an early stage and prevent it from worsening. This is important as it helps to limit the spread of the disease and minimizes the risk of complications.
- **Accurate diagnosis:** Skin disease detection is usually done through physical examination and laboratory tests. This helps to make sure that the diagnosis is accurate and that the treatment is tailored to the individual's skin condition.
- **Cost-effective:** Skin disease detection is a relatively low cost procedure as compared to other medical treatments. This makes it more affordable for people who cannot afford expensive medical treatments.
- **Improved quality of life:** Early diagnosis of skin diseases can help improve the quality of life for those affected as treatment can be started early to limit the spread of the disease and minimize the risk of complications.
- **Awareness:** Skin disease detection helps to create awareness about different skin conditions and the importance of regular check-ups. This can help people to take better care of their skin and pay attention to any changes.

Chapter 3

System Design

3.1 Architecture

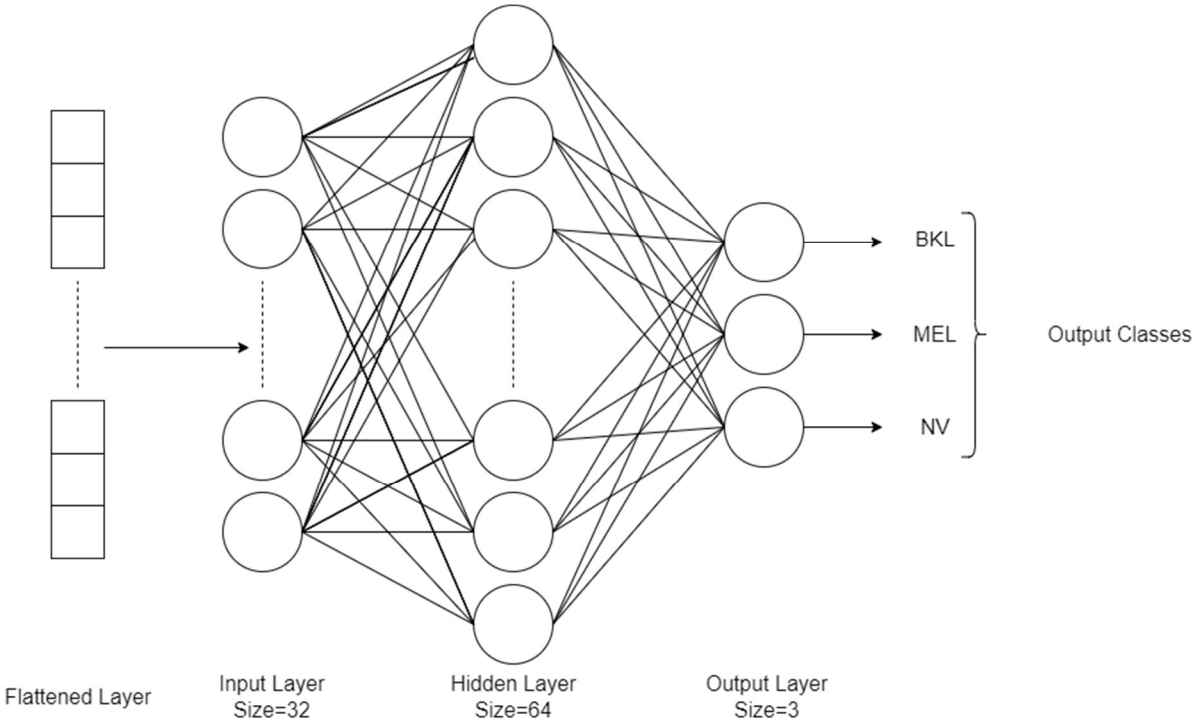

Fig 3.1: Working of ANN

Convolutional Neural Networks (CNNs) are a type of artificial neural network architecture that are particularly well suited for image-based tasks. CNNs use a series of convolutional layers to capture the spatial relationships between pixels in an image, allowing them to efficiently process image data and identify patterns. In addition, CNNs can be trained to recognize objects and classify images, making them a powerful tool for computer vision. The overall system design of a CNN typically involves inputting the image data into a series of convolutional layers, followed by one or more fully connected layers, and a final output layer. Each layer has a set of learnable weights and biases that are adjusted during the training process, allowing the network to learn the underlying relationships in the data. The output layer produces the desired output, such as classifying the input image as a certain object or providing a probability of a certain outcome.

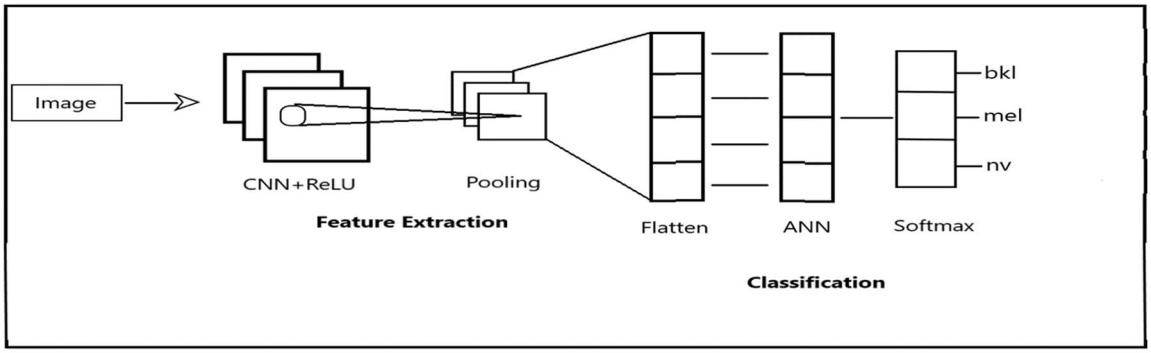

Fig 3.2: Feature Extraction using CNN

CNN is a neural network that extracts input image features and another neural network classifies the image features. The input image is used by the feature extraction network. The extracted feature signals are utilized by the neural network for classification. The neural network classification then works on the basis of the image features and produces the output. The neural network for feature extraction includes convolution layer piles and sets of pooling layers. As its name implies, the convolution layer transforms the image using the process of the convolution. It can be described as a series of digital filters. The layer of pooling transforms the neighboring pixels into a single pixel. The pooling layer then decreases the image dimension. As CNN's primary concern is the image, the convolution and pooling layers' procedures are intuitively in a two-dimensional plane. This is one of CNN's distinctions with other neural networks.

Chapter 4

TOOL DESCRIPTION

This section gives a detailed description about the hardware tools and software tools involved in developing this system and how they are used.

4.1 Software:

- **Python**: Python is clearly one of the best languages for machine learning. Python does contain special libraries for machine learning namely SciPy, pandas, Scikit-learn and NumPy, which is great for linear algebra, and getting to know the kernel method of machine learning. The language is great to use when working with machine learning algorithms and has easy syntax.
- **Tensorflow Lite**: A deep learning framework for on-device inference. Train and deploy machine learning models on mobile and IoT devices, Android, iOS, Edge TPU, Raspberry Pi.
- **Jupyter Notebook:** The Jupyter Notebook is a web-based interactive computing platform. The notebook combines live code, equations, narrative text, visualizations.

4.2 Hardware:

- A simple computer that can run a web browser.

Chapter 5

IMPLEMENTATION

- All the programming is done in Jupyter Notebook and VS Code.
- ANN model file and label file is saved into the assets folder.
- Flask app is used to provide a user interface for prediction.

STEPS TO RUN THE PROGRAM:

- The released app is installed on web browsers.
- Go to localhost to run the web app.
- On the first run, the app asks the user permission to access the camera.
- In the next window, choose one of the sample images.
- The app will then try to predict the class of the image.

Chapter 6

RESULTS AND ANALYSIS

6.1. Result Discussion

The result of different ML algorithms used are:

Algorithm	mIOU
SVM	80%
KNN	82%
ANN	84%

6.2. Analysis

From the above table we can see that the algorithm ANN has produced the best mIOU score. Our experimental results show that the proposed model sets a new state-of-the-art performance on HAM10000 datasets. We have used mean Intersection-over-Union as an accuracy metric. Hence, in the final production, ANN model is used to predict new data.

Chapter 7

CONCLUSION

Our proposed model employs an ANN structure where convolution layers are used to extract the contextual information and then fed into the ANN model to predict the output class. One could also apply classical feature extraction techniques like Gabor filters, etc., depending on the available computation resources. Finally, our experimental results show that the proposed model sets a new state-of-the-art performance on HAM10000 datasets.

REFERENCES

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

Articles on web: qhttps://www.fiercehealthcare.com/tech/industry-voices-how-al-solving-some- healthcare-s-biggest-challenges-during- pandemic#text=A1 has the power to,severe complications962 from COVID 19.

Anurag Kumar Verma, Saurabh Pal, Surjeet Kumar, Comparison of skin disease prediction by feature selection using ensemble data mining techniques, Informatics in Medicine Unlocked, Volume 16,2019, 100202.https://doi.org/10.1016/j.imu.2019.100202

From the report on: http://www.laeng.org/publication/WCECS2013/WCECS2013_pp850-854.pdf https://hal.archives-ouvertes.fr/hal-02141241/document

F.A. Spanhol, L.S. Oliveira, P.R. Cavalin, C. Petitjean, L. Heutte; Deep features for breast cancer histopathological image classification. In 2017 IEEE international conference on systems, MAN and cybernetics, SMC 2017. Banff, AB, Canada, October 5-8, 2017 (2017), pp. 1868-1873.

A. Chon, N. Balachandra, P. Lu, Deep convolutional neural networks for lung cancer detection, Stanford University.

R. Kumar, R. Srivastava, S. Srivastava; Detection and classification of cancer from Microscopic biopsy images using clinically significant and biologically interpretable features, Journal of medical engineering (2015)

M. Grewal, M. M. Srivastava, P. Kumar, and S. Varadarajan, "Radiologist level accuracy using deep learning for haemorrhage detection in CT scans," 2017.

U.-O. Dorj, K.-K. Lee, J.-Y. Choi, M. J. M. T. Lee, and Applications, "The skin cancer classification using deep convolutional neural network," pp. 1-16, 2018.

APPENDIX - I

SOURCE CODE

GitHub: https://github.com/BOT-Hunter/Skin-Disease-Detection-using-ANN

Note: The proposed model is working well then past models and can be made better in future.