Skin Disease Prediction Using Convolutional Neural Networks

A PROJECT REPORT

Submitted by

Avinash, 18BCE0363

Neetigya chahar, 18BCE0794

Shubham Kumar, 18BCE0356

Course Code: CSE4019

Course Title: Image Processing

Under the guidance of

Prof. Swathi J N
Associate Professor Grade 2

SCOPE, VIT, Vellore

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

ACKNOWLEDGEMENT

I express gratitude to my supervisor, **Prof. Swathi JN**, for guidance and suggestions that helped me to complete the project on time. Words are inadequate to express my gratitude to the faculty and peers who encouraged and supported me during the course. Finally, I would like to thank my ever-loving parents for their blessings and my friends for their timely help and support.

Place : Vellore

Date : 03-06-2021

Shubham Kumar (18BCE0356)
Avinash (18BCE0363)
Neetigya Chahar (18BCE0794)

EXECUTIVE SUMMARY

With ever growing pollution and climate disruptions, our way of living and lifestyle is giving rise to many skin related diseases. For some of which early and accurate detection becomes utmost necessary in order to save the patient. Given the availability of skin doctors in hospitals and huge expenditure associated with appointments, it becomes almost impossible for a poor to afford even a diagnosis related to a skin disease. The project aims at providing skin disease prediction technology made using deep learning. The project aims to preprocess the images, perform data augmentation and finally make use of deep learning technology to train itself with the various skin images. Then it is tested with the test set to check for accuracy. Finally a web implementation is developed to deploy our model.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviations

CNN	Convolution neural network
NN	Neural Network
AI	Artificial Intelligence
SD	Skin Disease
TECH	Technology

1. INTRODUCTION

1.1. Objective

The main objective of this project is to detect skin diseases from their visual images using deep learning and neural networks.

1.2. Motivation

The world is growing fast with growing industries and manufacturing processes. But in turn this is affecting the climate of the world we all are living in. Moreover, with changing times, our lifestyle is also getting worse day by day. These factors give birth to millions of dermatological diseases in the world. Ranging from as small as a pimple to as big as cancer. For ailment of these SDs, an accurate detection of the disease is mandatory at an early stage for most of the diseases. For doing so, we don't only need sufficient skin doctors for a country as big as India but also proficient ones. Given prices for an appointment with a private dermatologist, it appears that only well off people can afford SDs. For the sake of argument there are public hospitals having good dermatologists, but still with most of them operating their own private clinics where they charge heavily. Also due to this they are rarely present at the government hospitals. Most people who are not economically well off, most of the time either ignore SDs or hesitate due to expense. Our project aims at providing a way to tackle this. The project is designed to detect SDs with the image itself using NNs. This doesn't only save time and money of the population but also will be helpful for the students who are looking to become dermatologists in the future. They can utilise it to test and learn about SDs, giving us better professionals later. The project uses convolutional NNs for classifying the SDs based on the feed of images provided. The model is first trained on training samples and later tested on various test cases for the judgement of its accuracy. The objective is to give a tool which can be accurate enough to assist dermatologists and warn people who can't afford to have frequent appointments with the dermatologist.

1.3. Background

Artificial intelligence was first introduced at a famous Dartmouth College conference in 1956. AI is gradually interrelated with all disciplines, and also satisfies all aspects of the medical field. In the early 1970s, medical researchers discovered the applicability of AI in life sciences. AI can play a role in many aspects, such as medical image recognition and auxiliary diagnosis, bioTECH, drug research and development etc. Currently, medical image recognition is the most widely used. Dermatology is a subject that relies on morphological features, and the majority of diagnoses are based on visual pattern recognition. Dermatology is exceedingly suitable for applying AI image recognition capabilities for assisted diagnosis. At present, skin imaging TECH is represented by dermoscopy, very high-frequency (VHF) ultrasound, and reflectance confocal microscopy (RCM). Each method of skin imaging equipment has its own advantages and limitations. Dermatologists need to choose different imaging methods according to different conditions of skin lesions. Skin imaging TECH has become a vitally important tool for clinical diagnosis of SDs, and widely accepted and applied in the world.

2. Literature Survey

No.	Title	Author(s)	Year	Dataset used	Methodology	Pros & Cons	Future Work
1	Skin Disease detection based on different Segmentation Techniques [1]	Kyamelia Roy, Sheli Sinha Chaudhuri, Sanjana Ghosh, Swarna Kamal Dutta, Proggya Chakraborty, Rudradeep Sarkar	2019	Four skin diseases images -chicken pox, eczema, psoriasis and ringworm are used.	The images of four diseases were preprocessed with noise removal using an average filter, then were processed against 4 segmentation techniques in openCV (python) - Adaptive thresholding, edge detection, k means clustering and morphological segmentation	Four different segmentation techniques are used. Noise reduction using an average filter. No use of a learning based model.	Accompany segmentation with classifiers to identify diseases.
2	Skin Disease Classificatio	Prof. Jyotsna Gharat,	2020	MNIST: HAM10000 dataset	The dataset images were resized and cleaned before	Use of augmentation to enhance	Improving accuracy of the model.

	n using CNN[2]	, Anjali Bhatt, Maitreyee Nath, Pranali Yamgar		of dermatoscopic images.	augmenting them for the purpose of enriching the training set using keras deep learning. Then the classifier model using convolutional neural networks is formed.	the training set. Mobile based application for practical use.	
3	Detecting Skin Disease by Accurate Skin Segmentation Using Various Color Spaces[3]	Megha D. Tijare, Dr. V. T. Gaikwad	2018	Personal dataset of 5 classified diseases with 726 samples from 141 images.	Proposal of skin disease detection in sequence of these steps - preprocessing using median filtering, choosing color spaces, segmenting regions based on adaptive k-means clustering, then feature extraction using color	Practical use of color spaces such as RGB, HSV etc in segmentation to extract skin regions. F-score of 61% using both SVM	Exploring results using other distance measures such as hamming. Developing practical and user friendly systems from the proposed algorithm.

					histogram technique followed by the final step of building a classifier using KNN.	and KNN.	
4	Implementation of Nearest Neighbor using HSV to Identify Skin Disease[4]	Y A Gerhana, W B Zulfikar , A H Ramdani and M A Ramdhani	2018	200 image samples collected using an android camera.	Hue saturation value is used to process the image where hue is an actual color, saturation is purity of colors and value is light received. The Nearest Neighbour algorithm was used to create the classifier which calculates the distance between test samples and all training samples.	About 80% of accuracy. Stiver size of dataset used. No weightage to samples. Nearest neighbour method is computationally complex.	Comparison to other algorithms such as naive bayes, support vector machines etc. Use of more image training samples.
5	Skin Disease Identificatio	Dr. M. Sunil Babu , M. Sai	2020	Six diseases	Discrete cosine transformation	Achieved upto 80% accuracy.	Inclusion of more skin

	n using MATLAB[5]	Manikanta, P. Ganesh, M. Dharma Rakshak , P. Prudvi Teja		consist of three malignant and three benign tumors. https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign	is applied to images which is a 2D transformation technique. This was used to decompose images using Discrete wavelet transform.Finally levenberg marquardt algorithm is used which matches images saved in databases against input ones.	Limited number of skin diseases used.	colors and skin diseases. Further deployment of algorithms as phone applications.
6.	Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing	A.Gola Isasi B.García Zapirain A Méndez Zorrilla	2018	160 500×500-pixel RGB images (20 images per pattern) catalogue	The system is based on the standard ABCD Rule and dermatological Pattern Recognition protocols. On the one hand,	Data augmentation is still not implemented in the best way possible. Achieved	A better and more reliable implementation of data augmentation in the project

	algorithms [6]			d by dermatol ogists	a complete stack of algorithms for the asymmetry, border, color, and diameter parameterizati on were developed.	upto 85% accuracy. Still continues to help dermatol ogists.	can be noted for a future enhance ment.
7.	Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image [7]	Anushikha Singh, Malay Kishore Dutta, M. Partha, Sarathi Vaclav, Uher Radim	2018	500 image samples collected using an android mobile phone camera.	Feature extraction from the segmented and blood vessel removed the optic disc to improve the accuracy of identification. More significant in comparison to features of the whole or sub fundus image in the detection of glaucoma from fundus image. Several	Genetic algorithm s are used to reduce the dimensio nality of feature vectors. Accuracy of glaucoma identificat ion achieved in this work is 94.7%.	User interface can be improved so as to make the technolog y more user friendly and easily accessibl e to needed people.

					machine learning algorithms are used for prominent feature selection.		
8.	An Intelligent System for Monitoring Skin Diseases [8]	Dawid Połap, Alicja Winnicka, Kalina Serwata, Karolina Kęsik and Marcin Woźniak	2018	Dataset from ISIC2018 Challenge Disease Classification dataset	In this work, we present a smart home system which is using in-built sensors and proposed artificial intelligence methods to diagnose the skin health condition of the residents of the house. The proposed solution has been tested and discussed due to potential use in practice.	Experimental research has shown high efficiency and gives a good start for further development. This kind of support shall be installed in a place where we feel good and	It is possible to extend the system's ability to examine not only the skin but other features of our bodies. It is also possible to develop a portable version of this solution which we

						where the evaluation of the symptoms can be the most efficient.	can take for a holiday or a business trip.
9.	Diagnosis of skin diseases using Convolutional Neural Networks [9]	Jainesh Rathod, Vishal Waghmode, Aniruddh Sodha, Prasenjit Bhavthankar	2018	JNMIT: PUV10001 dataset of dermatoscopic images.	Skin images are filtered to remove unwanted noise and also process it for enhancement of the image. Feature extraction using complex techniques such as Convolutional Neural Network (CNN), classify the image based on the algorithm of softmax classifier and	We propose an automated image based system for recognition of skin diseases using machine learning classification. 95.7% accuracy is achieved through the proposed	This can also be used as a reliable real time teaching tool for medical students in the dermatology stream.

					obtain the diagnosis report as an output.	architectu re system.	
10.	Use of Neural Network-Ba sed Deep Learning Techniques for the Diagnostics of Skin Diseases [10]	E. I. Zakirov, N. N. Shchelkunov a, A. V. Melerzanov, D. A. Gavrilov	2019	Personal dataset of 16 classifie d diseases with 900 samples from 199 images.	Current development of image processing and machine learning technologies allows systems based on artificial neural convolutional networks to be created, these being better than humans in object classification tasks, including the diagnostics of malignant skin neoplasms.	This algorithm can discrimin ate benign and malignant skin tumors with an accuracy of at least 91% by examinati on of dermosco py images.	Presente d here is an algorithm for the early diagnosti cs of melanom a based on artificial deep convoluti onal neural networks that can be further develope d to be used by professor s to teach students.

11.	A Method Of Skin Disease Detection Using Image Processing And Machine Learning .[11]	Nawal SolimanALKolifi ALEnezi	2019	80 training and 20 testing images for 3 diseases. Diseases are Eczema, Melanoma, Psoriasis. Images are from related websites and dermatologists.	The system takes images as input and extracts features using pre-trained CNN. Then the features are classified using Multiclass SVM.	Pros - The system is 100% accurate for tested images. It can also reveal the spread, and severity of the disease. Cons - Model only works for 3 particular diseases.	Development of mobile applications. Detection of skin lesions in the Dermis layer. Finally, detect all the skin diseases in the world.
12.	Image based skin disease detection using hybrid neural network coupled bag-of-features[12]	Shouvik Chakraborty, Kalyani Mali, Sankhadeep Chatterjee, Sumit Anand, Aavery Basu, Soumen	2017	International Skin Imaging Collaboration (ISIC) dataset	First key points are detected using SIFT, then feature descriptor for each key point is computed, then bag-of-feature	Pros - This model can detect Basel Cell Carcinoma	Many more metaheuristic methods can be integrated to train the artificial

		Banerjee, Mitali Das, Abhishek Bhattachary a			s is computed, then training of the Neural Network with Meta-heuristic algorithms are done. Finally, ANN is tested with test images for accuracy.	(cancer) and Skin Angioma. Cons - The accuracy of this model is well below 90%.	neural network.
13.	Deep CNN and Data Augmentati on for Skin Lesion Classificatio n[13]	Tri-Cong Pham Chi-Mai Luong Muriel Visani Van-Dung Hoang	2018	A total of 6,162 train images and 600 testing images from ISBI challeng e and ISIC archive.	The authors proposed image classification based on Deep CNN models. They have also applied data augmentation for more accuracy. Secondly, they have compared data augmentation results with traditional	Pros - Data augment ation reduces model overfitting and works better for less data. Cons - Sensitivit y is average as	Sensitivit y needs improvem ent. Combinat ion of lower levels of CNN and data augment ation has to be improved. Last layer can still be fine

					ways.	compared to top results. Data augmentation is still not implemented in the best way possible.	tuned to reuse the weight of the network trained by 1.2 million images.
14.	Anomaly Detection for Skin Disease Images Using Variational Autoencoder [14]	Yuchen Lu, Peng Xu	2018	Dataset from ISIC2018 Challenge Disease Classification dataset (Task 3)	They have used Variational Autoencoder (VAE) for anomaly detection. The model is trained with normal skin images so that a skin disease can be detected as an anomaly in the image.	Pros - It is an unsupervised ML model letting any and most of the skin diseases as anomalies. Cons - It is hard to label the name or type of disease.	More methods of VAE for anomaly detection have to be explored. For example, Gamma distribution.

15.	Transfer learning with class-weighted and focal loss function for automatic skin cancer classification[15]	Duyen N.T. Le, Hieu X. Le, Lua T. Ngo and Hoan T. Ngo	2020	The HAM10000 dataset. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T	The techniques of end-to-end deep learning process, transfer learning technique, utilizing multiple pre-trained models, combining with class-weighted and focal loss were applied for the classification process.	Pros - Due to the use of transfer learning the issue of domain transfer is reduced. The accuracy is high. Cons- There are artifacts in the model which create biases.	The dataset can be expanded to dermoscopy images and clinical images. SOTA noisy student and finding a lottery ticket model to reduce the size of the model.

3. Project Description and Goals

Our project involves training an image dataset collected from HAM1001 and achieving significant accuracy. We are expecting an accuracy of around 90%. Studies we come across relating to this topic have shown that it is difficult to obtain an accuracy of above

90%. This fact can be attributed to the lack of ample amount of skin disease photos online. There are not enough sources for the dataset to train on.

In the preprocessing step, we aim to use the metadata file given in the HAM1001 dataset to sort the images and make a hierarchical directory structure useful for training image learning classification models using keras.

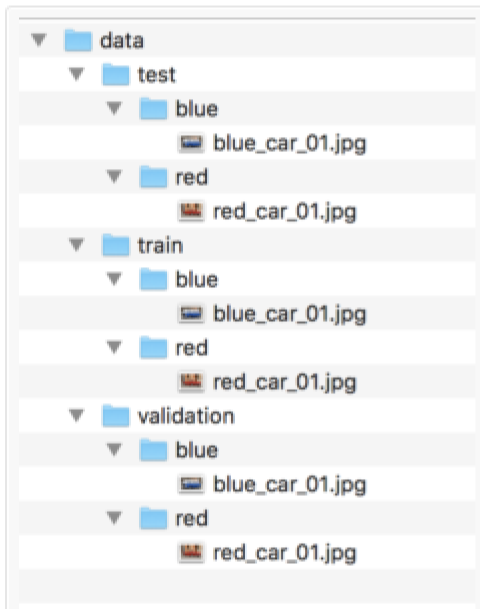


Fig 1. Image dataset directory structure by keras.

Instead of colors shown in fig 1, we have to make folders for skin diseases and put respective images inside them using the shutil library in python. The Shutil module in Python provides many functions for advanced operations on files and file collections. It belongs to the standard Python utility module. This module helps to automate the process of copying and deleting files and directories. Once done with creating the directory structure, the next step is to augment the data to enhance the dataset. Data augmentation was carried out using the Imagedatagenerator() function in the keras package. We used techniques such as random rotation, random zooming, brightness augmentation etc. to increase the dataset using existing images. The files created were kept in a temporary folder from which they were sent to their respective folder label. Next step in the sequence is the most important step. It is training the model. We set the epoch value neither too high nor too low. An epoch value too low would give us unsatisfactory accuracy, on the other hand an epoch value too high will require a huge amount of time and computation power which may not be available to complete the project in a given time frame. Once a model is compiled and we have the accuracy we aim to develop a web implementation for the same which can show the top 3 predictions for a given image.

4. Technical Specifications

Language - Python3

IDE - Spyder 4

Libraries and their functions used-

- Keras.preprocessing.image - Used for implementing various preprocessing techniques for deep learning models.
  ImageDataGenerator() - Generate batches of tensor image data with real-time data augmentation.
- Sklearn.model_selection-
  train_test_split() - Facilitates random partition of the entire dataset into train and test according to given proportion.
- Sklearn.metrics - This includes a function to evaluate classification performance. We will be using it for generating confusion matrices and accuracy.
- Matplotlib.pyplot - Used for creating visualisation.
  imshow() - for plotting pixels of image.
- Keras.models -Keras is a deep learning API written in Python, running on top of the machine learning platform TensorFlow.
  For generating CNN models with given inputs ready for compilation.
  Model() function is used for this.
- Shutil - High level operations on files and collection of files. This library is used to open, change the directory and save the edited image in the file system.

5. Design Approach

5.1. Materials and Methods

The dataset used for training the model is HAM1001. From this We will obtain dermoscopic images to train our model. This is a collection of 10,000 images labelled with 7 Skin diseases.

- Actinic Keratoses
- Basal Cell Carcinoma
- Benign Keratosis
- Dermatofibroma
- Malignant Melanoma
- Melanocytic Nevi
- Vascular Lesions

Augmentation

A process we generate new images from existing images by doing minor Alterations to enrich the training set to avoid overfitting and obtain better results.

Range of operations -

1. Horizontal and vertical shift- Moving the pixels of image in a given direction keeping the dimensions of image constant.
2. Horizontal and vertical flip - Reversing the rows and columns of pixels of an image.
3. Random rotation - Randomly rotating images clockwise in a given range of degrees.
4. Random zoom - zooms image randomly adding the pixel values in or interpolating from neighbours.
5. Random brightness - Randomly darkening or brightening images.

CNN

The solution is to do image classification with deep learning. For this project we are going to use CNNs. The model starts off with a very accurate and random prediction in the beginning. Then with the help of a method called Backtracking, it learns to find and recognize correlations in the input data. The correlations which are further used to predict the results, an approach similar to the neural system of our body.

The step by step approach is as follows:

- **Convolutional operation -** The first building block in our plan is a convolution operation. In this step, we touch on feature detectors, which basically serve as the NN's filters.
- **ReLU layer** - The second part of this step involves the Rectified Linear Unit or ReLU.
- **Pooling -** It is a process of reducing down the number of dimensions of the image so that the model compiles in finite time.
- **Flattening -** The two dimensional layer is converted into one dimension so that it can be fed to a fully connected NN classifier.
- **Full connection -** In this part, everything that we did so far is merged together.

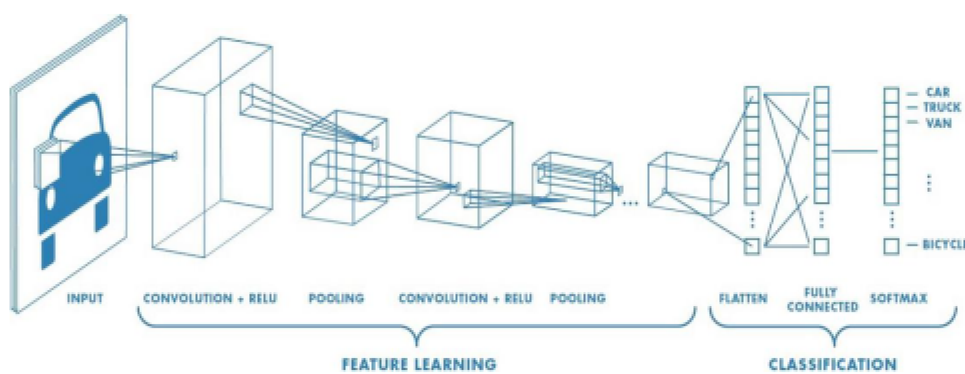


Fig 2. Steps in CNN

The mode of learning used here was transfer learning, which corresponds to taking a pre-trained *Mobilenet* model to train the images. This was done to achieve higher accuracy as the available dataset was less. Moreover, MobileNet is favored due to its high speed.

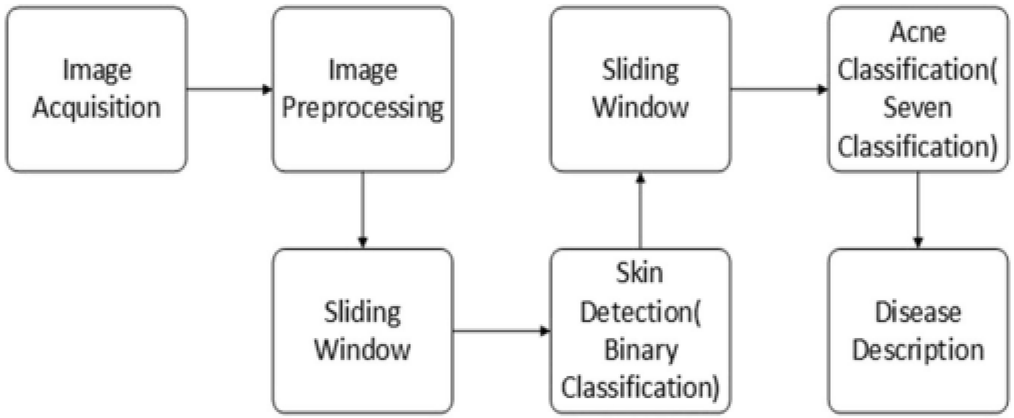


Fig 3. Stepwise process.

5.2. Codes and Standards

The size of images on which the model was trained has to be uniform. If the size of images in the dataset varies, it can be rectified in the preprocessing step. Moreover, the deep learning model will perform better only if there is ample amount of data given as input. To make sure this we used image augmentation and transfer flow models. There are other ways to do it as well which will lead to different results. Once the model is trained and tested, it is always necessary to save the compiled model. It will not only reduce time as one doesn't have to compile the model again and again. Further the same model can be used to deploy the project as well.

5.3. Constraints and Alternatives

Biggest constraint in any neural network project is the lack of a good number of dataset to train the model on. This was the issue found as one of the most common limitations across the related works we covered. On the internet there aren't many sources available with skin disease dataset. Even if there are, the number of images is too less to train a deep learning model and obtain significant results. To fix this constraint we adopted image augmentation. Image augmentation is a known technique in machine learning to enhance the data set. It alters the parameters of the original images in the datasets to form altered clones of them. These images combined with old images solve the problem of less images in the dataset. Studies have shown that by doing this the overall accuracy of deep learning models increases by a noticeable amount as well as help in overcoming overfitting of the model.

Secondly, we used a pre-trained mobile net model to train our deep learning model. In a pre-trained model, few features are already identified, which compensates for lack of dataset. Further the model took a long time to compile on CPU. To overcome this a TPU can be used.

6. Tasks and Milestones

1. Preprocessing

As we are using keras, it is necessary for us to preprocess the directory structure like the one shown in the image (Fig.)

2. Data augmentation

To enhance the dataset we have to perform data augmentation, as this will be really beneficial for us because by augmentation we can achieve the following :

a. A larger dataset, hence a higher accuracy
b. Avoiding overfitting

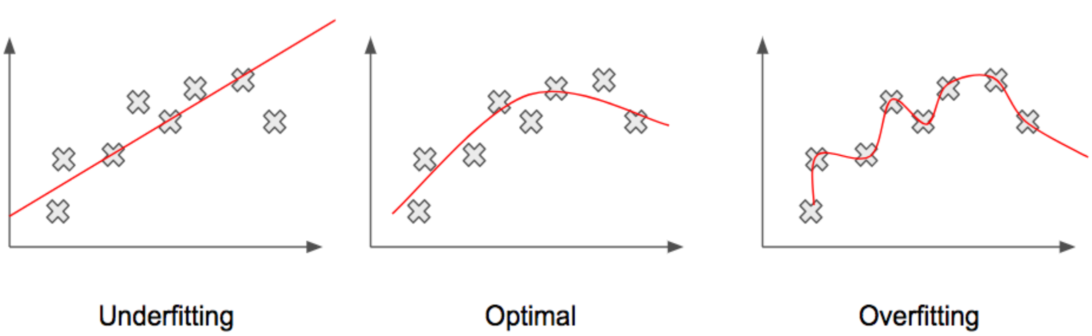


Fig 4. Fitting data on the dataset.

3. Training the model

Last 23 layers of the mobilenet Transfer model were used to train our dataset. As we know it is a pretrained model, this will increase the accuracy of our results further.

4. Testing the model

The model was compiled and generated on the test set. After that, a test set was used to measure the accuracy of our trained model. Test set was generated by segregation of 20% of images randomly. Upon testing, we achieved the top3 accuracy of above 90%.

5. Web implementation

For the implementation of the model in a user-friendly way, we have created a web app where using JavaScript, we read the image uploaded by the user, then we load the model and use the Tensorflow library to match the uploaded image. Backend is simply a web server in Node.js serving the required files.

7. Project Demonstration

Demonstration video link - **Review3Final_imageProcessingJcomp.mkv**

Source code GitHub Link - **https://github.com/neetigyachahar/Skin-disease-classifier**

Preprocessing the directory structure like the one shown in the image below to perform network training on keras package.

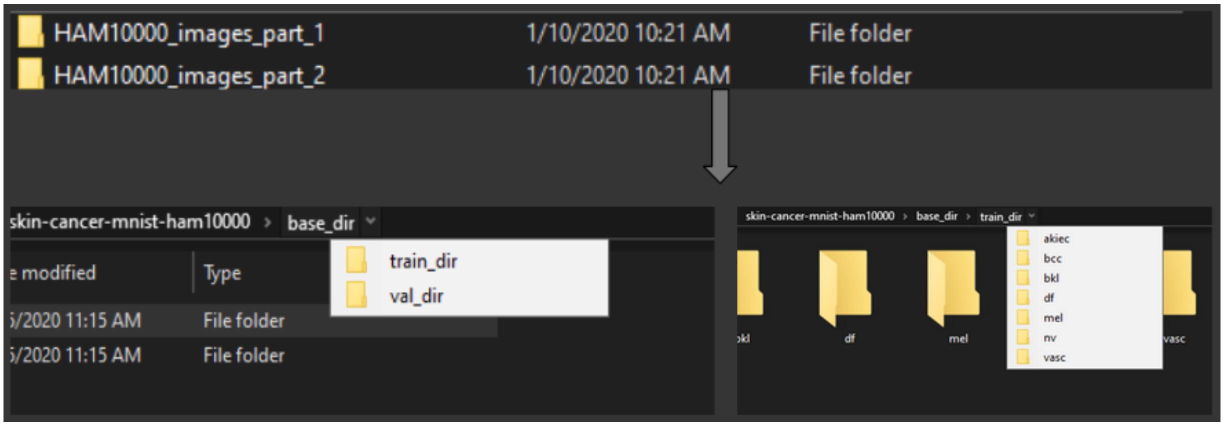


Fig 5. File directory

Data augmentation - Series of operations shown on a melanoma disease image as follows.

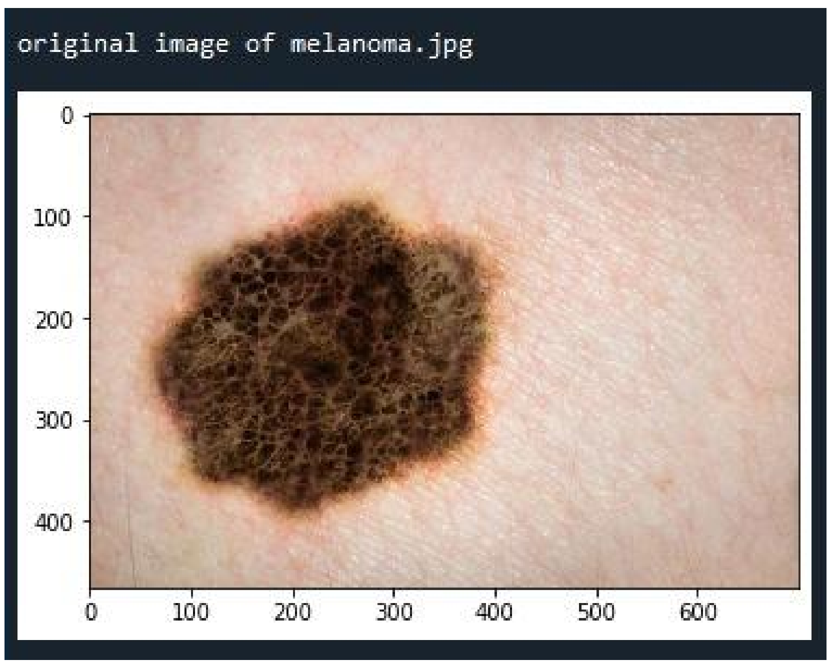


Fig 6. Original image of melanoma

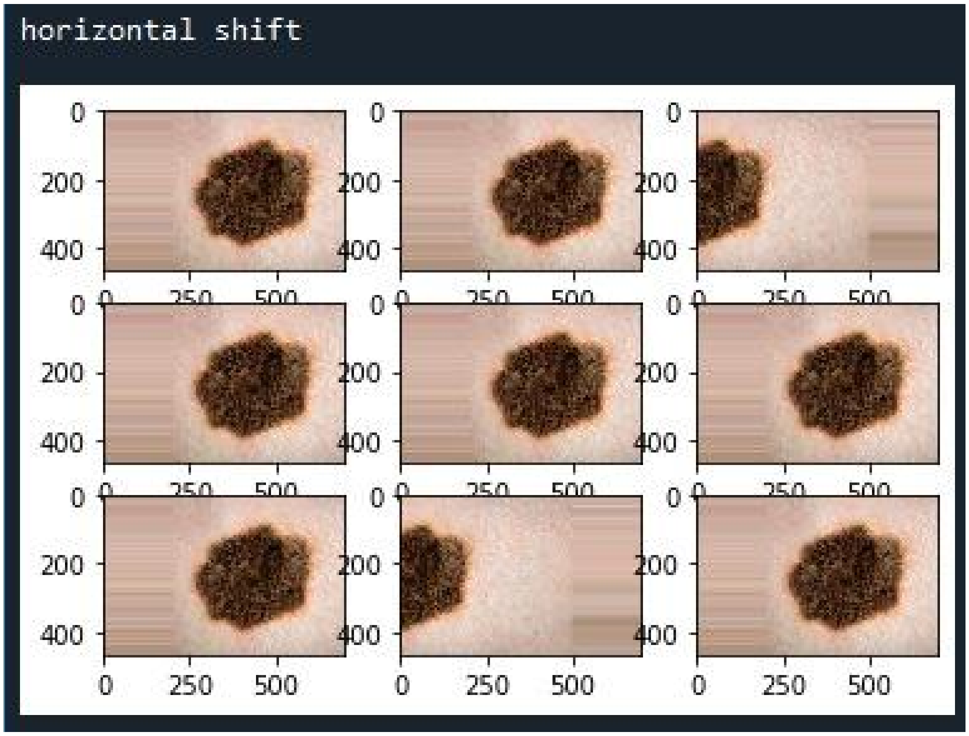


Fig 7. Horizontal shift

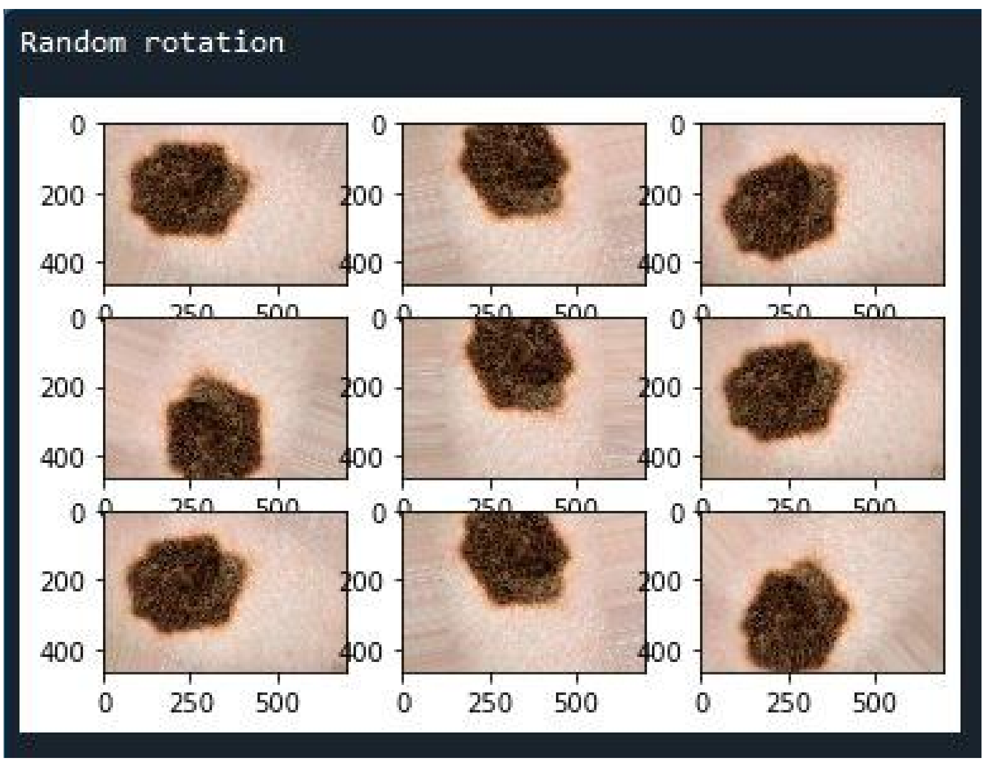


Fig 8. Random Rotation

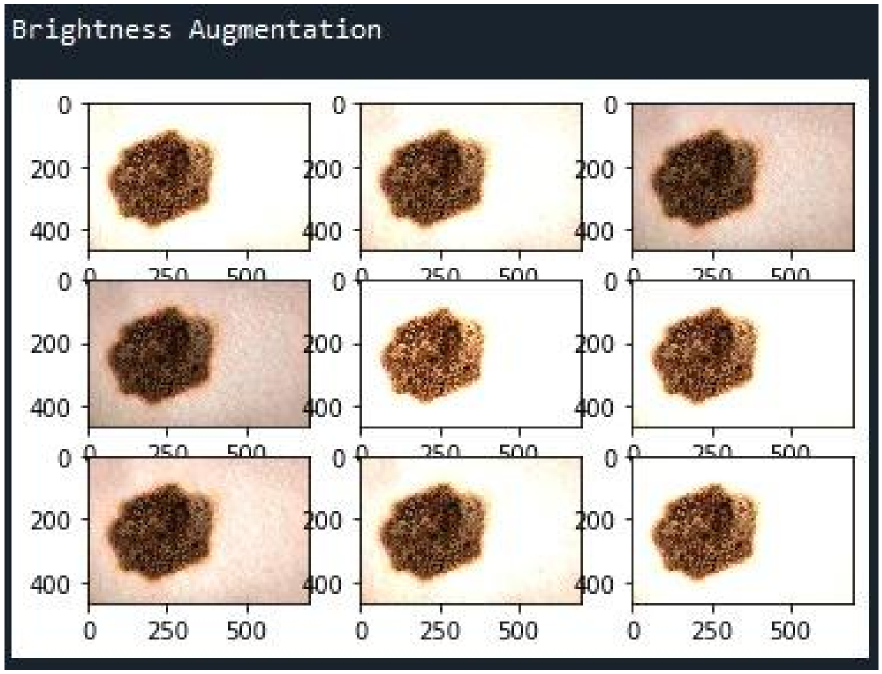


Fig 9. Brightness augmentation

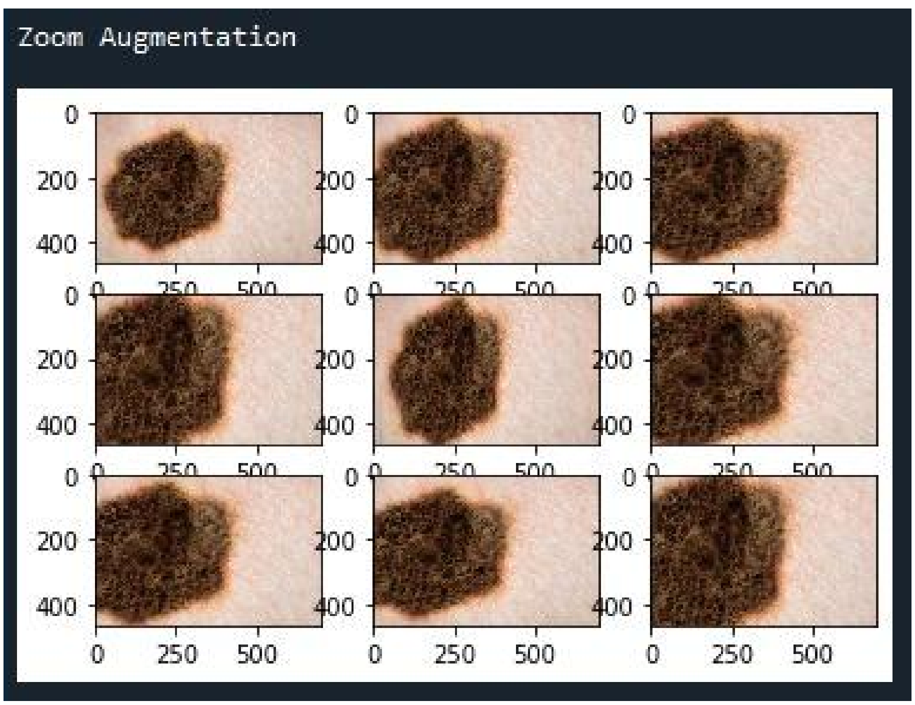


Fig 10. Zoom augmentation

Using the last 23 layers of mobilenet transfer flow model.

```
Using TensorFlow backend.
Found 38704 images belonging to 7 classes.
Found 1002 images belonging to 7 classes.
Found 1002 images belonging to 7 classes.
Model: "mobilenet_1.00_224"
```

Fig 11. Tensor flow backend

Now, training the model using the given dataset.

```
Epoch 1/1
103/902 [==>...........................] - ETA: 24:50 - loss: 3.0643 -
categorical_accuracy: 0.3553 - top_2_accuracy: 0.5466 - top_3_accuracy: 0.7019
```

```
# Fit the model
history = model.fit_generator(train_batches,
                              steps_per_epoch=train_steps,
                              class_weight=class_weights,
                              validation_data=valid_batches,
                              validation_steps=val_steps,
                              epochs=1,
                              verbose=1,
                              callbacks=callbacks_list)
```

Now, testing the model and evaluating the eventual top3 accuracy of the program.

```
val_top_3_accuracy: 0.9082
```

Now finally, implementing the same program as a website, hence completing the deployment portion.

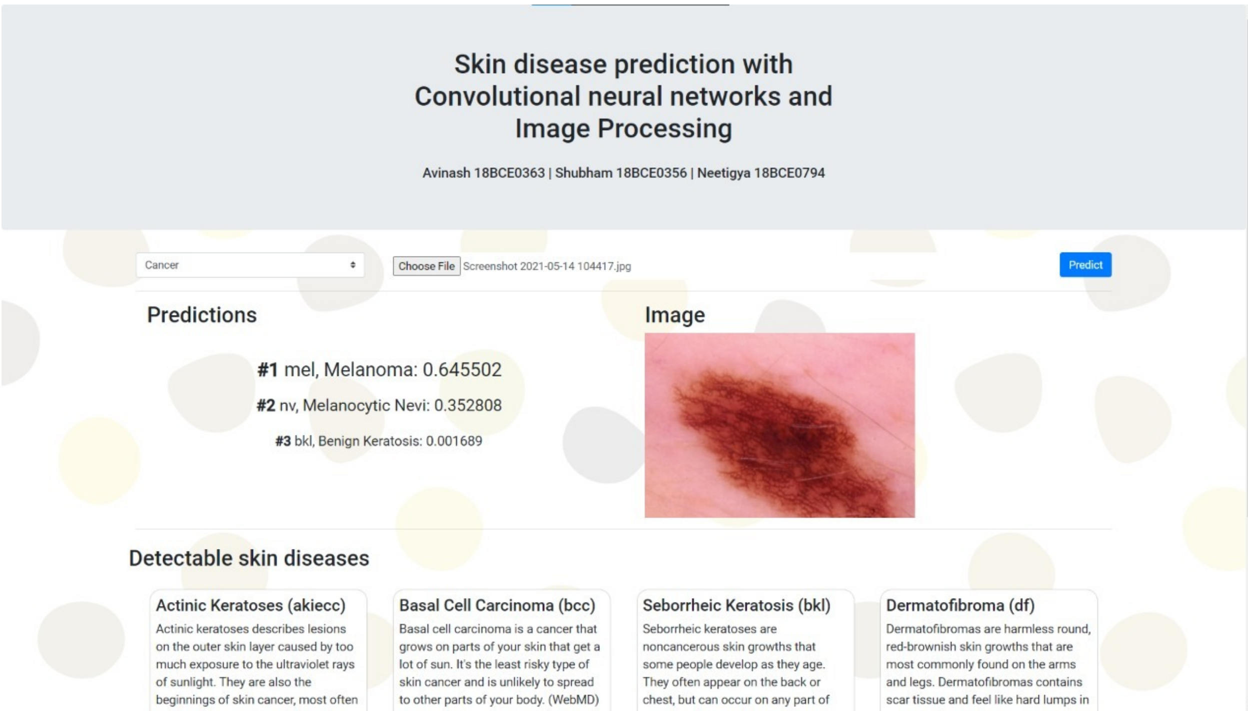


Fig 11. Web deployment

8. Result and Discussion

Before moving on to the results it was necessary to analyze and discuss the dataset and the model. Learning curves are a widely used diagnostic tool in machine learning for algorithms that learn from a training dataset incrementally.

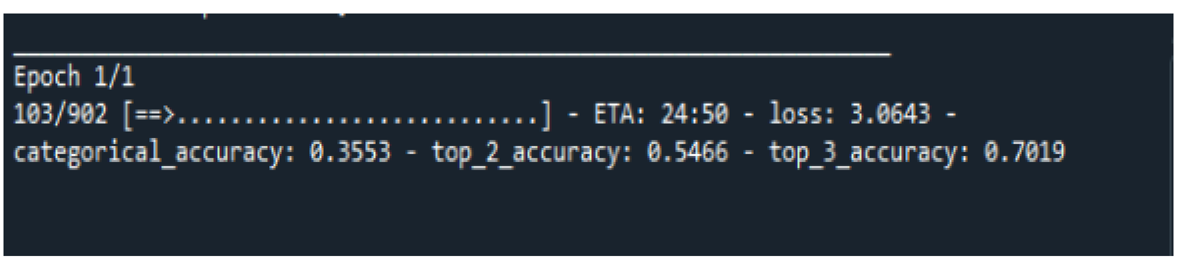


Fig 12. Training the dataset

They can be analyzed by two methods:

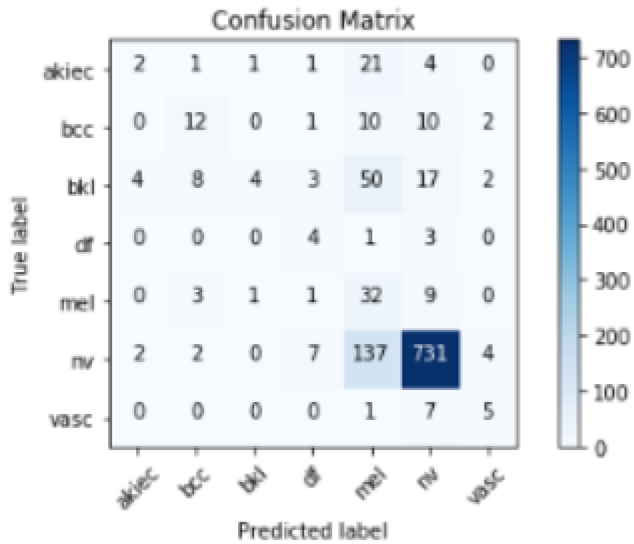
- Loss curves - This is also called a cost function. It illustrates how much the prediction varies from true value.
- Accuracy curves - This is used to measure the classification model's performance. It is basically a percentage that tells you how close the model fitted the actual dataset.

Plotting curves for our model gives the following results -

- Loss Curve

- Accuracy Curve

Now moving on to the result, The model performed very well in fact as the model obtained 90 % top3 accuracy. The model significantly worked well and it was able to predict and classify the SDs. The confusion matrix came out pretty well as well.



The model's full report can be viewed here.

	precision	recall	f1-score	support
akiec	0.25	0.07	0.11	30
bcc	0.46	0.34	0.39	35
bkl	0.67	0.05	0.09	88
df	0.24	0.50	0.32	8
mel	0.13	0.70	0.21	46
nv	0.94	0.83	0.88	883
vasc	0.38	0.38	0.38	13
accuracy			0.72	1103
macro avg	0.44	0.41	0.34	1103
weighted avg	0.84	0.72	0.74	1103

To summarize the model performed very well and was able to fit our aim of predicting SDs pretty proficiently giving all round good result values. The model performed very well in fact as the model obtained 90 % top 3 accuracy. We further investigated the performance of our model on the second dataset and we obtained the following results.

Name	**Total**	**Match**	**Mismatch**	**Percentage**
Vascular skin lesion	30	22	8	73.33%
Melanoma	72	56	16	78.8%
Benign Keratosis	80	60	20	75%
Melanocytic Nevi	80	69	11	86.25%
Actinic Keratoses	80	69	11	86.25%

Table 1. Performance on secondary dataset.

9. Conclusion and Future Work

After all the preprocessing and analysis, the model was trained on top of the transfer model MobileNet, which gave out to us brilliant results with accuracy of around 90% simply outshining other models similar to this league.This can be easily looked upon and improved for being used in real-life scenarios. The applications of these kinds of healthcare technologies are ginormous. The more data and support that we have for this, the more it can grow and become more accurate. With only 10,000 images, we achieved an accuracy of around 90%, but it can be improved if we had millions of images. We could potentially eradicate misdiagnosis in clinics by using AI technology and deep learning methods such as this. AI is already blooming in the healthcare scene. As we continue to adopt it into mainstream medicine, the quality of treatments, as well as diagnosis will improve tremendously. Through the use of AI for diagnosis, we can eliminate preventable diseases, and create a future where we no longer need to worry about getting an accurate diagnosis. We’ve already seen AI replace many manufacturing and clerical jobs, but many don’t consider the impact AI has on a field as important as healthcare.

In future, this project can be elevated to a higher level if the model can be converted using TensorFlowJs to deploy on the web. By doing this it will reach a wider audience and help millions of people suffering from skin diseases. Moreover, due to memory and time constraints we chose only the first 23 layers of the MobileNet model. To increase its accuracy furthermore this can be increased depending on the scenario. Also if we need to increase accuracy, we can try the training model again with different epoch values and verify. Furthermore, the model can be developed to give out a prediction value for each attribute. By doing this the user will get to know if there is a conflict or the system fails to identify due to unfortunate issues.

10. References

1. Roy, K., Chaudhuri, S. S., Ghosh, S., Dutta, S. K., Chakraborty, P., & Sarkar, R. (2019, March). Skin disease detection based on different segmentation techniques. In 2019 International Conference on Opto-Electronics and Applied Optics (Optronix) (pp. 1-5). IEEE.
2. Gharat, J., Bhatt, A., Nath, M., & Yamgar, P. Skin Disease Classification using CNN.
3. Tijare, M. D., & Gaikwad, V. T. (2018). Detecting Skin Disease by Accurate Skin Segmentation Using Various Color Spaces. Journal of Engineering Research and Application, 8(12).
4. Gerhana, Y. A., Zulfikar, W. B., Ramdani, A. H., & Ramdhani, M. A. (2018). Implementation of Nearest Neighbor using HSV to Identify Skin Disease. In IOP Conference Series: Materials Science and Engineering (Vol. 288, No. 1, p. 012153). IOP Publishing.
5. BABU, D. M. S., MANIKANTA, M., GANESH, P., RAKSHAK, M. D., & TEJA, P. P. (2020). Skin Disease Identification using MATLAB.
6. Isasi, A. G., Zapirain, B. G., & Zorrilla, A. M. (2011). Melanomas non-invasive diagnosis application based on the ABCD rule and pattern recognition image processing algorithms. Computers in Biology and Medicine, 41(9), 742-755.
7. Singh, A., Dutta, M. K., ParthaSarathi, M., Uher, V., & Burget, R. (2016). Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image. Computer methods and programs in biomedicine, 124, 108-120.
8. Połap, D., Winnicka, A., Serwata, K., Kęsik, K., & Woźniak, M. (2018). An intelligent system for monitoring skin diseases. Sensors, 18(8), 2552.
9. Rathod, J., Waghmode, V., Sodha, A., & Bhavathankar, P. (2018, March). Diagnosis of skin diseases using Convolutional Neural Networks. In 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1048-1051). IEEE.

10. Gavrilov, D. A., Melerzanov, A. V., Shchelkunov, N. N., & Zakirov, E. I. (2019). Use of neural network-based deep learning techniques for the diagnostics of skin diseases. Biomedical Engineering, 52(5), 348-352.
11. ALEnezi, N. S. A. (2019). A method of skin disease detection using image processing and machine learning. Procedia Computer Science, 163, 85-92.
12. Chakraborty, S., Mali, K., Chatterjee, S., Anand, S., Basu, A., Banerjee, S., ... & Bhattacharya, A. (2017, October). Image based skin disease detection using hybrid neural network coupled bag-of-features. In 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) (pp. 242-246). IEEE.
13. Pham, T. C., Luong, C. M., Visani, M., & Hoang, V. D. (2018, March). Deep CNN and data augmentation for skin lesion classification. In Asian Conference on Intelligent Information and Database Systems (pp. 573-582). Springer, Cham.
14. Lu, Y., & Xu, P. (2018). Anomaly detection for skin disease images using variational autoencoder. arXiv preprint arXiv:1807.01349.
15. Le, D. N., Le, H. X., Ngo, L. T., & Ngo, H. T. (2020). Transfer learning with class-weighted and focal loss function for automatic skin cancer classification. arXiv preprint arXiv:2009.05977.