Web based Skin Lesion Detection and Diagnosis using Deep Learning

[1]G.Jemilda, [2]K. Manak, [3]M. Manoj Kumar, [4]M. Mathan, [5]C.Yovel

[1]Faculty, [2,3,4,5]Student
[1]Computer Science and Engineering,
[1]Jayaraj Annapackiam CSI College of Engineering, Nazareth, India

Abstract - Skin cancer is one of the most common types of cancer, and early detection is key for successful treatment. In our work, we propose a skin lesion detection system that utilizes machine learning techniques to identify potential skin cancer lesions. The proposed system includes various stages, such as image segmentation, feature extraction using transfer learning with a pre-trained Mobilenet Convolutional Neural Networks (CNN), data augmentation, and classification using Support Vector Machine (SVM). The system is trained and tested on a dataset, which consists of images of ten different types of skin lesions. The experimental results demonstrate the effectiveness of the proposed system in accurately detecting skin lesions and achieving high classification accuracy. The proposed system can potentially assist dermatologists in the early detection of skin lesion, leading to better patient outcomes.

IndexTerms - Skin Lesions, Machine Learning, Convolutional Neural Networks (CNN), Support Vector Machine (SVM)

I. INTRODUCTION

Skin cancer is a widespread type of cancer that requires early detection for effective treatment. Skin lesions are a common symptom of skin cancer, but accurately identifying and classifying them can be difficult for even experienced dermatologists. To assist in this task, computer vision and machine learning techniques can automatically detect and classify skin lesions using digital images. The goal is to develop a skin lesion detection system that can accurately classify skin lesions, which can aid medical professionals in providing accurate diagnoses and treatment plans. The proposed system consists of several stages, starting with data collection and pre-processing to ensure that all images are of consistent size and quality. Augmentation techniques are applied to increase the size of the dataset and improve the CNN's robustness. In the segmentation stage, the lesion is isolated from the image background using image processing techniques. Then, a pre-trained CNN based on MobileNet architecture is used for feature extraction. Finally, the extracted features are used for classification in the final stage. In conclusion, the proposed skin lesion detection system based on a CNN using MobileNet architecture has the potential to improve the accuracy and efficiency of skin lesion diagnosis. This project highlights the usefulness of computer vision and machine learning techniques in assisting medical professionals and improving patient outcomes in the detection and treatment of skin cancer.

II. LITERATURE SURVEY

This literature survey reviews some of the recent research papers that have used machine learning techniques to identify skin diseases.
Shivam Pandey et al [3] proposed a model-fusion-based CNN for the identification of skin diseases. They added modules to the central blocks to help the network and increase the network's capacity to harvest visual information. The proposed model achieved high accuracy of classification by combining different techniques like up-sampling, parameter tuning, and modelling.
A. Kalaivani. et al [2] proposed a predictive heterogeneous ensemble model for the multiclass dermatitis problem using data mining techniques.. The proposed ensemble model uses Random Forest Deep CNN (RF-DCNN) Classifier, which achieved the highest accuracy of any of these approaches, at 96.1 percent. Furthermore, they used a multi-model ensemble approach to combine different data mining techniques to get the best accuracy.
MWP Maduranga. et al [1] proposed an android mobile application that implements a CNN Model for the identification of skin diseases. They used transfer learning to train the model and found that MobileNet with transfer learning yields an accuracy of more than 85%, making it the most suitable model for automatic skin disease identification. They also highlighted that the identification accuracy depends on the camera quality of the mobile device and lighting conditions.
Sriwong Kittipat. et al [4] proposed a deep learning transfer learning adoption strategy to classify medical data, including image data of skin diseases. They used three kinds of learning models: transfer learning with the existing Alexnet architecture, SVM modeling from feature extraction data of the image (FESVM), and SVM modeling from feature extraction data of the other patient data (PESVM). They found that the transfer learning with the existing Alexnet architecture (Alexnet-TL) performed the best among the three models. The proposed models have the potential to provide faster, cheaper, and more accurate diagnoses of skin diseases, which can lead to better treatment outcome for patients.

III. RESEARCH METHODOLOGY

In skin lesion detection, various modules present are Image Preprocessing, Segmentation, Data Augmentation, Feature Extraction, Classification and Diagnosis. The below figures 1 and 2 shows the flow diagram for training and testing.

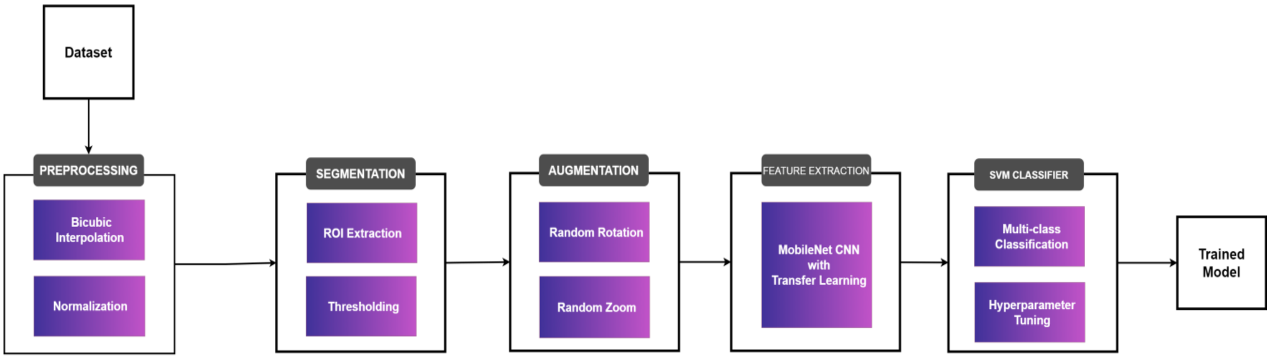

Fig.1 Training

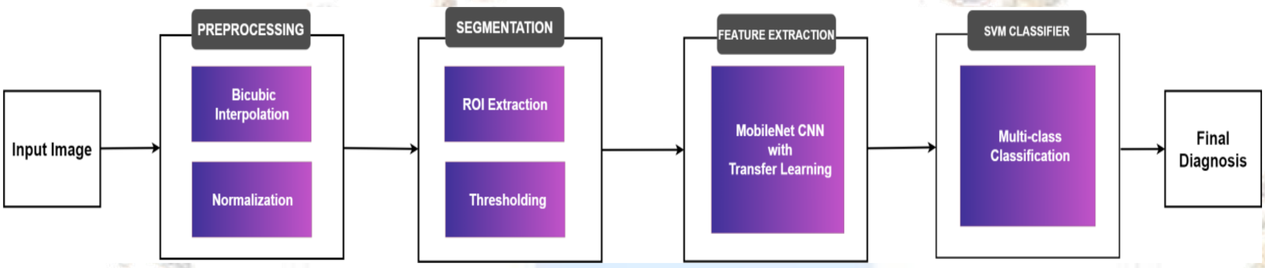

Fig.2 Testing

(1) Image Preprocessing

Image preprocessing is a crucial step in the analysis of skin lesion images. It involves various techniques that are applied to enhance the quality and extract relevant features from the image which is shown in Figure 3.

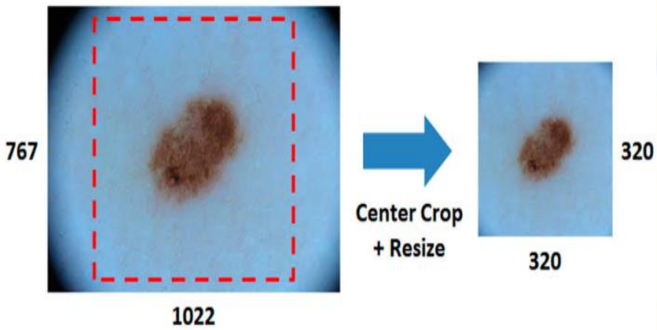

Fig.3 Resizing and normalization of input images

This preprocessing also contains two sub modules as,

(i) Bicubic Interpolation
(ii) Normalization

(i) Bicubic Interpolation

Bicubic interpolation is a commonly used technique in medical image processing and analysis, particularly in skin lesion imaging. It is a method for resampling images, where a new image is created with a different number of pixels than the original image. In skin lesion imaging, this can be useful for increasing or decreasing the resolution of images, or for changing the aspect ratio of an image.

(ii) Normalization

When working with images of skin lesions, normalization is an essential step to ensure that the images are consistent and comparable. Normalization adjusts the pixel values in an image so that they fall within a specific range as in Equation (1). In skin lesion analysis, normalized images have a mean pixel value of 0 and a standard deviation of 1. This removes any differences in brightness or contrast between images, making it easier to compare them.

$$I_N = (\mathrm{I} - \mathrm{Min})\frac{newMax - newMin}{Max - Min} + newMin \quad \text{-----} \ (1)$$

(2) Segmentation

Segmentation is a crucial step in skin lesion analysis, where the objective is to isolate the lesion region from the surrounding healthy skin tissue in an image. The aim of segmentation is to accurately identify and separate the lesion region from the rest of the image as per figure 4, which can be used for further analysis such as feature extraction or classification. Segmentation can be done manually or automatically, with the latter being the most common approach.

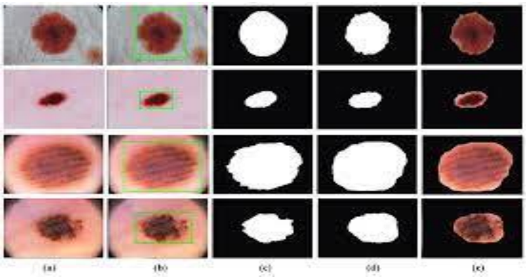

Fig.4 Segmentation of Preprocessed Images

This Segmentation also contains two sub modules.

(i) Extraction of Region of Interest (ROI)

(ii) Thresholding for image segmentation

(i) ROI Extraction

The ROI is the specific area of the lesion that contains the most relevant and diagnostically useful information. By extracting the ROI, focus can be on the most important part of the lesion and reduce the amount of irrelevant data that need to process. The ROI extraction process typically involves identifying the boundary of the lesion and then selecting a smaller area within the boundary that contains the most significant features of the lesion.

(ii) Thresholding

When segmenting skin lesion images, thresholding is a common technique used to separate the foreground (lesion) from the background. It involves selecting a threshold value that separates the image into two parts: pixels with intensity values above the threshold (foreground) and pixels with values below the threshold (background). In adaptive thresholding, the threshold value is determined based on the local pixel intensity of the image, rather than using a fixed threshold value for the entire image. This is important because skin lesion images can have varying lighting conditions, which can make it challenging to use a fixed threshold value. Here, the image is divided into smaller regions, and a threshold value is calculated for each region based on its local pixel intensity. This helps to account for any variations in lighting within the image and results in more accurate segmentation of the lesion.

(3) Data Augmentation

Data augmentation is a technique used to increase the size of a dataset by creating new variations of existing data. It is used to improve the accuracy and robustness of machine learning models by introducing new images with different orientations, scales, and lighting conditions. One common method is random zooming and rotating. This involves randomly selecting an image and applying a random rotation and scaling factor to it. The rotation can be in any direction and the scaling can be either uniform or non-uniform. Figure 5 shows the augmentation of segmented images.

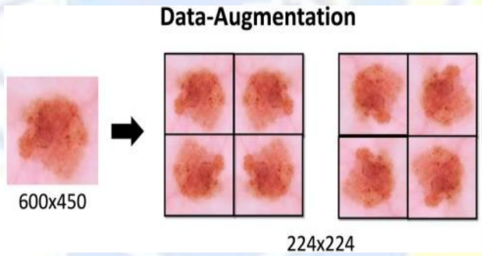

Fig.5 Augmentation of Segmented Images

(4) Feature Extraction

When working with skin lesion images, one of the key steps in the analysis process is feature extraction. This involves identifying and extracting important features of the lesion that can be used for classification and diagnosis. In our project, we are using a pre-trained MobileNet CNN model for feature extraction.

Fig.6 Various Features of Lesions

Some of the features include asymmetry, border irregularity, color, and diameter is shown in Figure 6. These features are known to be important indicators of skin lesions and can help to distinguish between benign and malignant lesions. The MobileNet CNN model is well-suited for this task because it has been trained on a large dataset of images and can accurately identify and extract these features.

(5) Classification

Skin lesion classification is a critical task in dermatology. One of the popular approaches for classification is SVM, which is a supervised learning algorithm used to classify data by finding the best hyperplane that separates different classes in the feature space as per Figure 7. SVM is used for classifying skin lesions into 10 different categories, namely, acne, contact dermatitis, corn cutaneous horn, eczema, herpes zoster, melanoma, psoriasis, rosacea, and urticaria. SVM takes the features extracted from the skin lesion images and learns a decision boundary that separates them into different categories. Figure 8 shows the pseudo code of SVM algorithm. SVM has been widely used for skin lesion classification due to its high accuracy and ability to handle high-dimensional feature spaces.

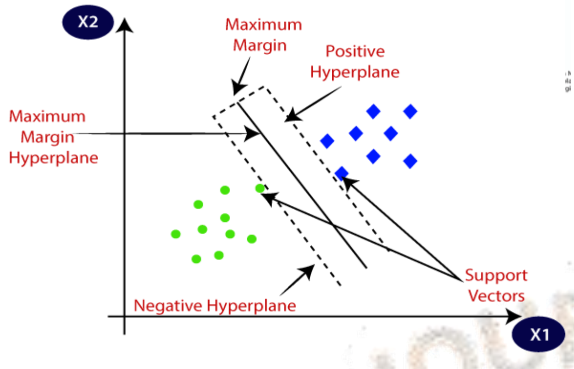

Fig.7 Classification using SVM

```
Define number of features+1 as F and SVs+1 as SV
FOR each SV
        FOR each feature of the SV
        Read streamed data
        Convert it to float
        Store into array_SV's [SV][F]
       END FOR
END FOR
Read streamed data
Convert it to float
Store into array_ay [0] (b value)
FOR cach SV
        Read streamed data
        Convert it to float
        Store into array_ay [SV]
END FOR
FOR each feature
        Read streamed data
        Convert it to float
        Store into array_test [1]
END FOR
FOR each feature
        Clear array AC [F]
END FOR
FOR each SV
        FOR each feature of the SV
        array AC [F] += array ay [SV] array SVS [SV][F]
        END FOR
END FOR
FOR each feature
        Distance value + array_AC [F] * array_test [F]

END FOR
Distance value -= b
IF (Distance value - th) THEN
        RETURN 1
ELSE
        RETURN-1
END IF
```

Fig.8 Pseudo code of SVM Algorithm

(6) Diagnosis

In the final diagnosis section, the SVM classifier outputs the predicted diagnosis for each skin lesion image, along with the corresponding accuracy of the prediction. The predicted diagnosis is the name of the disease or condition that the model has identified from the image. For example, it could be acne, contact dermatitis, corn, cutaneous horn, eczema, herpes zoster, melanoma, psoriasis, rosacea, or urticaria, depending on the specific skin lesion being analyzed. The accuracy is a measure of how confident the model is in its prediction, expressed as a percentage.

IV. CONCLUSIONS

Our work is aimed to develop a skin lesion classification system using machine learning techniques. The dataset was preprocessed by normalizing and augmenting the images to improve accuracy. Segmentation was performed to extract ROI using thresholding techniques. The features of asymmetry, border irregularity, color, and diameter were extracted using a pre-trained MobileNet CNN model. The SVM classifier was used for the classification of ten different skin lesion types. The final diagnosis section displays the name of the disease along with the accuracy. In overall, our work demonstrated the potential of using machine learning techniques to accurately classify skin lesions, which could greatly benefit the field of dermatology. Figures 9 to 14 represent the sample screenshots.

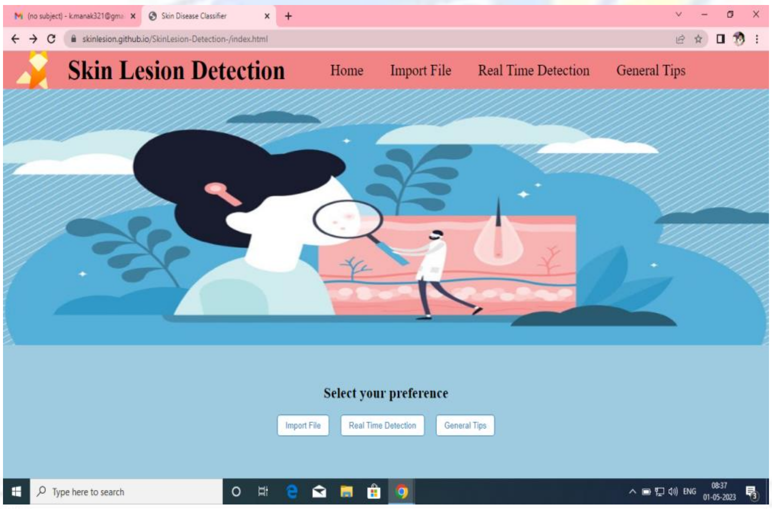

Fig.9 Web Page

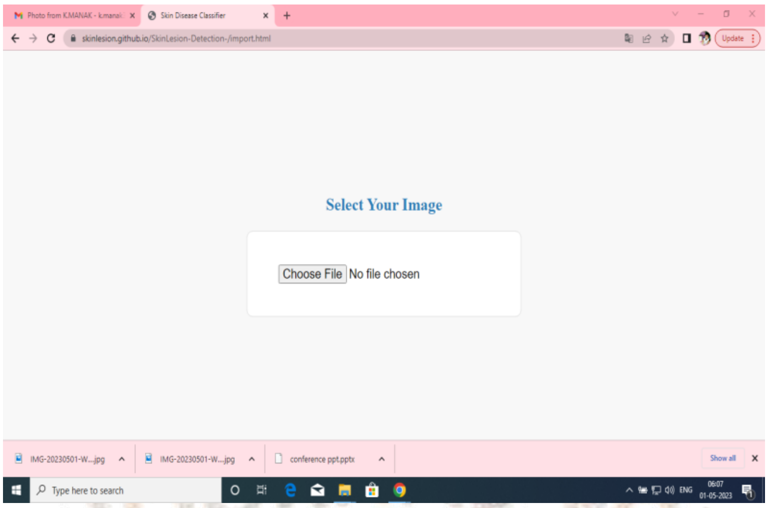

Fig.10 Image Selection

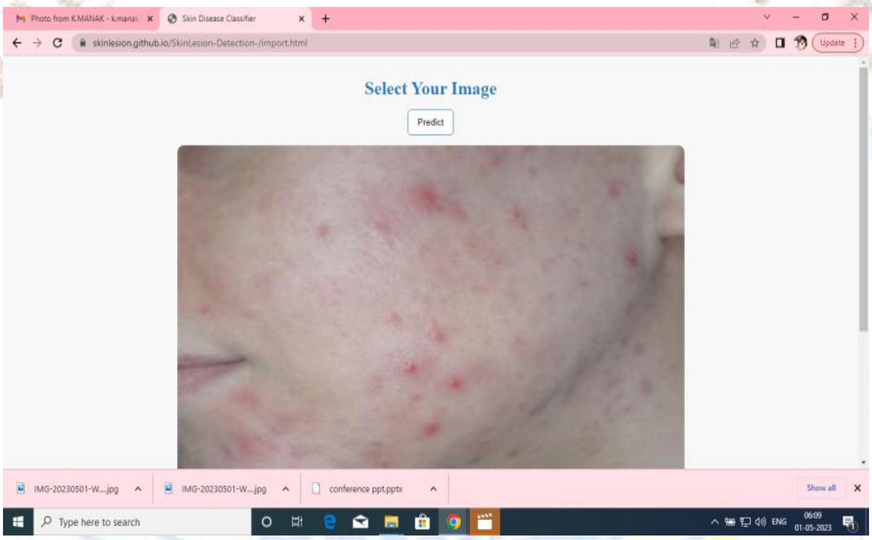

Fig.11 Prediction

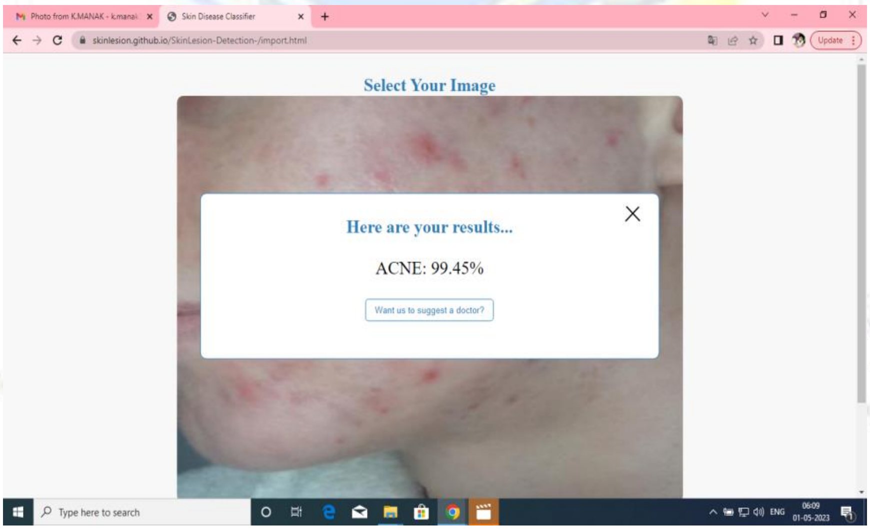

Fig.12 Diagnosis

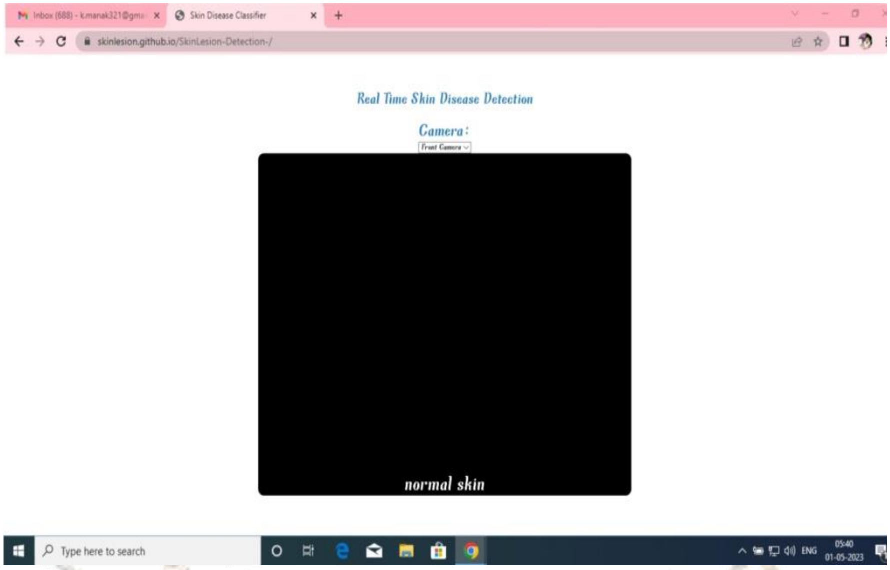

Fig.13 Real time Detection

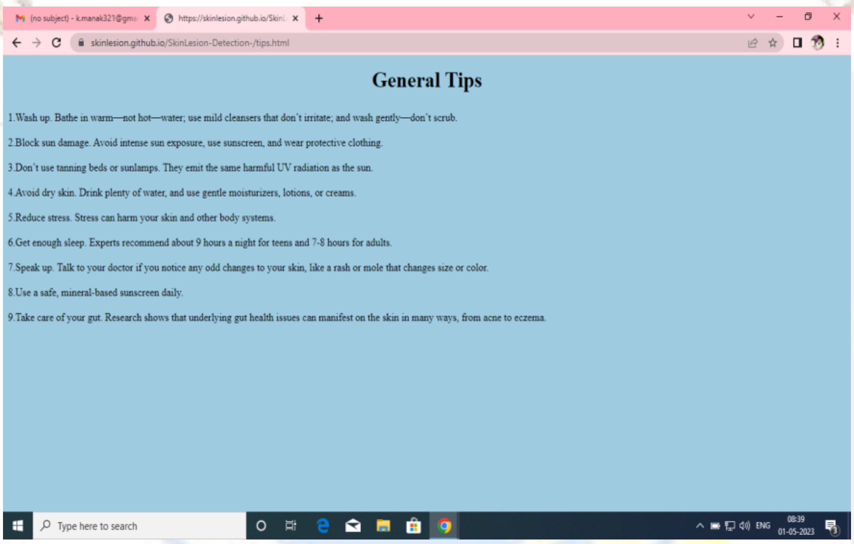

Fig.14 General Tips

V. REFERENCES

[1] MWP Maduranga, Dilshan Nandasena, "Mobile-Based Skin Disease Diagnosis System using Convolutional Neural Networks (CNN)", International Journal of Image, Graphics and Signal Processing (IJIGSP), Vol.14, No.3, pp. 47-57, 2022.

[2] Kalaivani, A., Karpagavalli, S., "Detection and Classification of Skin Diseases with Ensembles of Deep Learning Networks in Medical Imaging", International Journal of Health Sciences, 6(S1), 13624 – 13637, 2022.

[3] Shivam Pandey, Sanchary Nandy, Shivani Bansal, "Skin Disease Detection Based on Deep Learning", International Journal of Scientific Reasearch in Science, Engineering and Technology (IJSRSET), Online ISSN:2394-4099, Print ISSN:2395-1990, Volume 10, Issue 1, pp.120 - 127, January - February 2023.

[4] Sriwong Kittipat, Bunnrit Supaporn, Kerdprasop, Kittisak, Nittaya, "Dermatological Classification using Deep Learning of Skin Image and Patient Background Knowledge", International Journal of Machine Learning and Computing, Vol. 9, 2019.