SKIN CANCER ANALYSIS USING CNN

Sharmila Gaikwad[1], Aparna Agnihotri[2], Amreen Khan[3], Vaishnav Kanekar[4]

[1]Sharmila Gaikwad (Professor, Computer Engineering Dept, Rajiv Gandhi Institute of Technology)
[2]Aparna Agnihotri (Student, Computer Engineering Dept, Rajiv Gandhi Institute of Technology)
[3]Amreen Khan (Student, Computer Engineering Dept, Rajiv Gandhi Institute of Technology)
[4]Vaishnav Kanekar (Student, Computer Engineering Dept, Rajiv Gandhi Institute of Technology)

---***---

Abstract - *Skin Cancer is an uncontrollable growth of abnormal cells in the epidermis, which is the outer layer of the skin. It occurs when the DNA is altered and can't correctly control skin cell growth. Skin cancer is also one of the most critical forms of cancer.*

There are four main types of skin cancer: Basal cell carcinoma, Basal cell carcinoma, Merkel cell cancer, and Melanoma. Detection of Skin cancer in the early stage is essential to cure it. The dermatological way to diagnose skin cancer is with the dermoscopic assessment of the lesion followed by biopsy and histopathologic evaluation. This process is very long, leading the patient to critical stages of cancer. Many technologies have been developed to increase the accuracy of detecting skin cancer as early as possible. Computer Vision can play a vital role in medical Image diagnosis, which the existing systems have proved.

In this paper, we have presented a model to detect skin cancer from seven types. The types are Melanoma (MEL), Melanomic Neves (NV), Basal Cell Carcinoma (BCC), Actinic Keratosis (AKIEC), Benign Keratosis (BKL), Dermatofibroma (DF), Vascular Lesion (VASC). We have built the CNN (Convolutional Neural Network) model to perform image processing on various image datasets of skin cancer to analyze and detect its type.

Key Words: CNN, image processing, cells, Melanoma (MEL), Melanomic Neves (NV), Basal Cell Carcinoma (BCC), Actinic Keratosis (AKIEC), Benign Keratosis (BKL), Dermatofibroma (DF), Vascular Lesion (VASC).

1.INTRODUCTION

Skin is the most exposed part of the body because it is the outermost part hence UV rays, dust, micro-organisms. It can lead to several diseases, and in the worst-case scenario, to skin cancer. As stated by World Health Organization (WHO), it is a depressing fact that one in every three cancers identified is skin cancer. As it correctly said, absolute power corrupts absolutely. Because of humans' exponential increase in control over nature and subsequent global warming, the ozone layer is depleted, inversely leading to an increase in the harmful UV radiations reaching Earth's surface because of loss of the protective filter. An additional 4,500 melanoma and 3,00,00 nonmelanoma skin cancer cases will result in a 10% decrease in ozone levels. Other causes of skin cancer include modernizing diets, smoking, and alcohol. Abnormal changes in skin color, appearance, shape, size, itching, bleeding are the cautionary signs of skin cancer. Due to negligence, lack of facilities, or simply unavailability of crucial time, the disease rushes to become fatal by spreading to other body parts.

In this paper we have presented an Artificial intelligence take for developing a skin cancer classifier model. There are 3 steps in this automated classifier:

1.1 DATA COLLECTION

One has to tackle a significant obstacle: collecting datasets for skin cancer classification. Automated analysis and prediction of pigmented lesions on the skin often meet roadblocks such as limited datasets or diverse and unique skin lesion images. It is necessary that the classifier model is exposed to varied ideas during its learning time and doesn't suddenly get petrified but somewhat habituated and confident while dealing with on-the-spot images in routine practice.

1.2 FEATURE EXTRACTION

Feature extraction means to be able to take important notes by taking into consideration only specific essential properties that will be different from other inputs. A CNN model is considered appropriate for processing the images as it works on pixel values. Layers such as Conv2D help by sliding a filter or kernel over input data in 2D and upon performing element-wise multiplication and concluding with summing up of results into a single output pixel. This kernel slides over every location serves the same operation and transforms the input 2D matrix into a new 2D matrix. The ReLU or Rectified Linear Unit is an activation function which outputs zero if the input is not positive and if it is positive. Max pooling assists in calculating the maximum value of every patch in all feature maps.

1.3 CLASSIFICATION

Classification is an essential step because it is in charge of outputting the interpretations about the information acquired from previous actions to classify the input image into Seven types or random images. These were the general steps in building our model for organizing skin cancer and its detection types.

2. SYSTEM MODEL

2.1 INTRODUCTION TO DATASET COLLECTION:

The dataset we have collected is from Kaggle, “Skin Cancer MNIST: 10000”, an extensive collection of skin cancer images, which was the foundational need of the project. It consists of more than 10,000 images belonging to different skin cancer classes such as Bkl, Mel, Nv, Bcc, Df, Vasc, Akiec. It comes along with a .csv file that links a large number of images to their lesion ids, furthermore providing information such as localization of skin cancer, age of the patient, gender of the patient, and the method of concluding type skin cancer.

2.2 BUILDING ALGORITHM FOR MODEL:

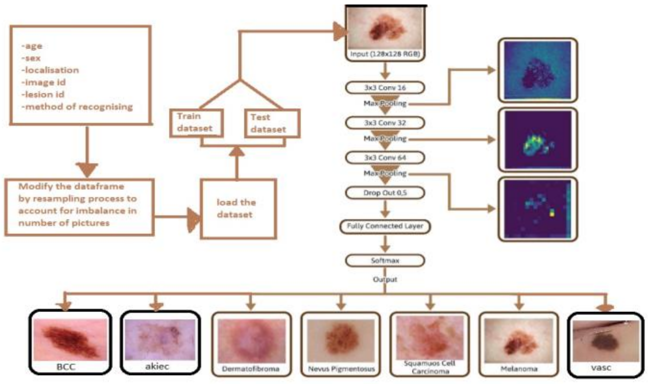

Fig -1: Proposed System

The proposed system flow begins with loading the .csv file. The data frame obtained from the .csv file is modified to suit the programmer’s ease.

Modifications first include label encoding, which converts textual classes of the data frame to numerical types, as data handling is more accessible with numbers than text.

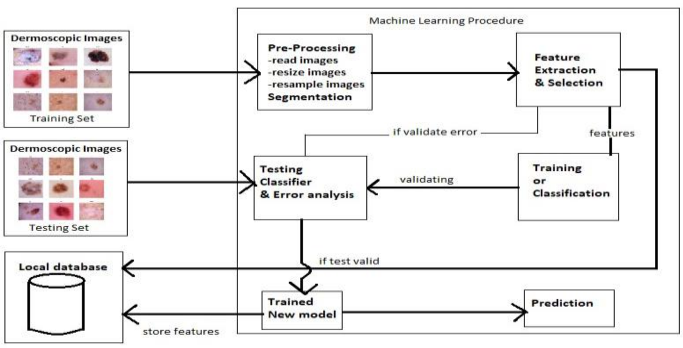

Fig -2: System Architecture

From the System Architecture diagram, it is visualized that the whole dataset is divided into training and testing datasets. The training dataset is preprocessed; for instance, images read are resized and resampled. Resampling is performed to account for an imbalance in the number of pictures for seven types of skin cancer. Now the preprocessed images are fed for feature extraction and selection. At this time, the CNN model comes into the picture, the Conv2D, max pooling, with activation function ReLU, flatten layer, dense layer finally helps build a complete model. The testing dataset is now tested, and the model gives a prediction.

2.3. MODEL AND IT’S METRICES:

```
Model: "sequential_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_9 (Conv2D)            (None, 30, 30, 256)       7168
_________________________________________________________________
max_pooling2d_9 (MaxPooling2 (None, 15, 15, 256)       0
_________________________________________________________________
dropout_9 (Dropout)          (None, 15, 15, 256)       0
_________________________________________________________________
conv2d_10 (Conv2D)           (None, 13, 13, 128)       295040
_________________________________________________________________
max_pooling2d_10 (MaxPooling (None, 6, 6, 128)         0
_________________________________________________________________
dropout_10 (Dropout)         (None, 6, 6, 128)         0
_________________________________________________________________
conv2d_11 (Conv2D)           (None, 4, 4, 64)          73792
_________________________________________________________________
max_pooling2d_11 (MaxPooling (None, 2, 2, 64)          0
_________________________________________________________________
dropout_11 (Dropout)         (None, 2, 2, 64)          0
_________________________________________________________________
flatten_3 (Flatten)          (None, 256)               0
_________________________________________________________________
dense_6 (Dense)              (None, 32)                8224
_________________________________________________________________
dense_7 (Dense)              (None, 8)                 264
=================================================================
Total params: 384,488
Trainable params: 384,488
Non-trainable params: 0
```

Fig -3: Model and Metrices

Model consists of layers such as:

Conv2d Layer: In a Conv2d layer, a filter or kernel slides over the 2D input data, for example, an image array, after which an elementwise multiplication takes place, resulting in a single output pixel. The kernel will repeat the same operation by sliding on every location of the input image array, transforming a 2D matrix of features into another 2D matrix of elements.

ReLU: It is an activation function that goes by the complete form of Rectified Linear Unit. It is a piecewise linear function that outputs the input as positive directly if it is non-negative or outputs it as zero.

Max Pooling: An operation that performs pooling, meaning that it calculates the maximum value in each area of the feature map. But unlike average pooling, the results are down-sampled feature maps that highlight the most critical/present feature from that area. In contrast, average pooling is the average of all the extracted features.

Metrics: Metrics are used to monitor and measure the performance of a model. Metrices such as accuracy, precision, recall, confusion matrix and F-1 Score are used in our model.

Fig -4: Training and Validation Loss

Training and Validation Loss:

In the above (Fig-4) the curve in yellow indicates how well the model is fitting the training data, while the red curve indicates how well the model fits new data.

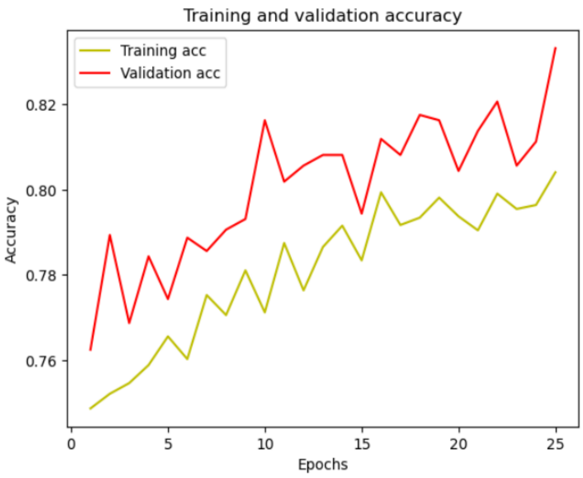

Fig -5: Training and Validation Accuracy

Training and Validation Accuracy:

In the above (Fig-5) the curve in yellow indicates the model classifying images while training the dataset whereas the curve in red indicates the accuracy we calculate for dataset and used for validating the model.

Accuracy of our model achieved stands at 94.09% while recall and precision stand at 94.09.% and 93.99% respectively.

3. FRONTEND WEBSITE:

We have successfully implemented our backend and linked it with our frontend. The website provides functionalities such as finding a dermatologist, information about types of cancer, representation of cancer in pie charts according to age, sex & body parts. Our website will be helpful in the real world and will ease our doctors and patients in determining cancer and treating it at early stages.

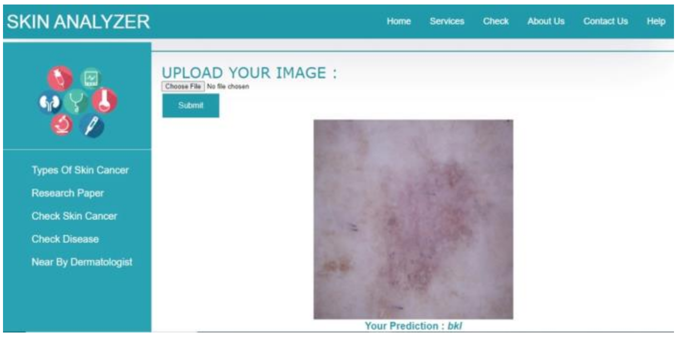

Fig -6: Skin cancer detection.

When any doctor or patient enters any cancerous image, the model will do analysis and predict if the infected part is cancerous or not. If the infected part is cancerous, it will indicate skin cancer from seven types. If the user uploads any random image other than the infected part, the classifier will classify it as a random image.

For example, in the above (Fig-6), the classification is given as Bkl, one of the seven types of skin cancer.

Other Functionalities of Website:

Types of Skin Cancer: The types of skin cancer will give the user information about the seven types of skin cancer, their differentiation, and the pictorial representation of the classes.

Research Paper: This tab will provide the research work we have carried out during the execution of the project and relevant papers and publications.

Check Skin Cancer: This feature allows the user to enter the infection region image and check if it's cancerous or not. It also gives predictions about the seven types of skin cancer.

Fig -7: Nearby Dermatologist.

Near By Dermatologist: This feature will allow the user to search the nearby dermatologist to consult and go with the treatment ahead. Thus, it will help the patients search for the doctor's clinic by entering their locality. It will save the patients time and help them in early action towards treating cancer and trying to cure it.

4. CONCLUSION:

Even if it has occurred in only 4% population, Malignant melanoma accounts for 75% of deaths caused due to skin cancer. In this paper, we have built a CNN-based model and website to help dermatologists detect skin cancer. Doing this will give them more time to perfect their future steps in curing the patients because time plays a vital role in curing this deadly disease. The paper also discusses how the classifier must admit potential faults if it cannot classify images correctly and define them as random rather than give confident predictions. It is a sensitive issue and can harm somebody's life. Detection and curing of cancer play a vital role, and our project will undoubtedly provide a helping hand to patients and doctors and will be helpful in the real-time world. The challenges ahead of us are more and more collection of datasets and more awareness in patients to get themselves tested without being ignorant as it spreads quickly to other body parts by beginning as a tiny lesion.

REFERENCES

[1] Gaikwad, Sharmila. "Study on Artificial Intelligence in Healthcare." 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS). Vol. 1. IEEE, 2021

[2] Gaikwad, Sharmila, Saarah Patel, and Ajinkya Shetty."BrainTumor Detection: An Application Based on Machine Learning." 2021 2nd International Conference for Emerging Technology (INCET). IEEE, 2021

[3] Y. Li and L. Shen, "Skin Lesion Analysis towards Melaoma Detection Using Deep Learning Network"," Sensors, vol. 18, no. 2, p. 556, 2018.

[4] V. M, "Melanoma Skin Cancer Detection using Image Processing and Machine Learning", International Journal of Trend in Scientific Research and Development, vol. -3, no. -4, pp. 780-784, 2019.

[5] Kamalpreet Kaur, et al. "Melanoma Skin Cancer Detection Using Image Processing and Machine Learning.Vol 1 IEEE,2019

[6] M. Dildar et al., "Skin Cancer Detection: A Review Using Deep Learning Techniques", International Journal of Environmental Research and Public Health, vol. 18, no. 10, p. 5479, 2021.

[7] H. Zunair and A. Ben Hamza, "Melanoma detection using adversarial training and deep transfer learning", Physics in Medicine & Biology, vol. 65, no. 13, p. 135005, 2020.

[8] M. A. Hossin, F. F. Rupom, H. R. Mahi, A. Sarker, F. Ahsan, and S. Warech, "Melanoma skin cancer detection using Deep Learning and advanced regularizer," 2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2020.