Skin Cancer Detection Using Convolutional Neural Networks

Haripriya.S[1], Hemalatha.N[2],S.R.Senthil Kumar[3], R. Reena[4]
[1, 2] Dept of Computer Science and Engineering
[3]Assistant Professor, Dept of Computer Science and Engineering
[4]Associate Professor, Dept of Computer Science and Engineering
[1, 2, 3, 4] Prince Shri Venkateshwara Padmavathy Engineering College, Ponmar, Chennai

Abstract- *Melanoma is considered to be the least common Skin Cancer but it is the deadliest one. It spreads to other body parts in no time and must be diagnosed as early as possible. This project aims to reduce the death caused by Skin Cancer. Our objective is to build a Deep Learning based model that could classify benign and malignant skin cancer from outlier lesion images. The approach uses Convolutional Neural Networks (CNN) to classify the images. The Deep Learning models built here are trained and tested on standard datasets and the metric area under the curve was observed 84%.*

Keywords- skin cancer, convolutional neural networks, lesion classification, deep learning, malignant classification, benign classification ,efficient net model.

I. INTRODUCTION

Based on the statistical information of 2019, skin cancer fourth most common disease with 15000 cases and 1900 deaths. The common method for skin cancer detection used by physicians are sampling and testing. The diagnostic procedure for these type of skin cancers starts with clinical screening and then analysis and examination. Dermatologists usually cut a small part of the affected area and examine it under the microscope. This process takes a week or more. This model aims to shorten the time gap from more than a week to just couple of days. Deep Learning has revolutionized the future as it can solve any complex problem. The primary goal is to shorten the time between diagnosing and treatment.

II. LITERATURE SURVEY

i.The Performance of Deep and Conventional Machine Learning Techniques for Skin Lesion Classification

Skin Lesion is any abnormalities occurring to skin tissues which endangers the lives of patients. It occurs in terms of size, colour, texture and shape. This paper helped us to know how Machine Learning helps to detect Skin Lesion based on the images in the data set.

ii.An efficient machine learning approach for the detection of melanoma using dermoscopic images

Initially different types of colour and texture features are extracted from the dermoscopic images based on the structure. Then the features are fed to classifier to know whether it is melanoma or not.

III. EXISTING APPLICATION

In the existing system, Biopsy method is used for detecting the Skin Cancer. It is done by removing the skin cells and sample goes to the various laboratory testing. Few years back, the problem of classifying Melanoma came into the focus of Machine Learning. But research mostly followed pre-processing, feature extraction and classification.

IV. FLOW DIAGRAM

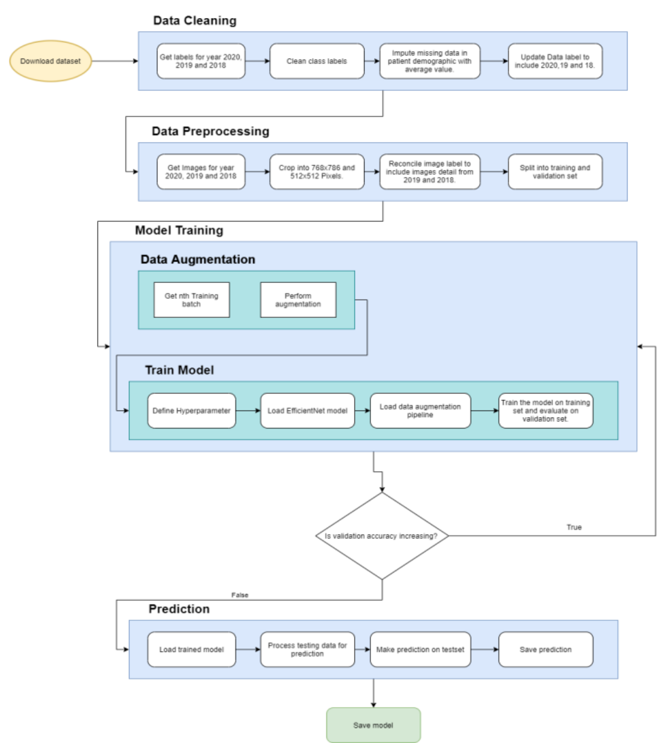


Figure 1 : Flow Diagram

The flow diagram consists of six components which are discussed below:

I. **Dataset:**

The data set consists of around forty-four thousand images. Along with the images the data set consists of a CSV file along with patient details containing patient-id, name, age, sex.

II. **Pre-processing:**

The data set has small to medium sized images and evaluation metric is unstable. However, we have only 1.76% of positive samples (Malignant). We tried to use both 2018- and 2019-year's data along with 2020 year which possibly increased the positive sample ratio to 17.85% (10 times higher).

III. **Data Augmentation:**

If the data set is very small, then image augmentation is recommended to avoid over-fitting the training dataset. After data processing, we have nearly **46k**images in the training data. The dataset contains significant class imbalance, in which most of the classes have a category called "UNKNOWN". We have defined an augmentation pipeline that deals with the class imbalance. The augmentation helps to improve the prediction accuracy of the model.

The selected augmentation are as follows:

Transpose: We are transposing the input images which swaps the rows and columns.
Flip: We are flipping the input images randomly in either direction (horizontally or vertically).This makes the model more robust.
Rotate: Input images are rotated with a random degree chosen from the range .This augmentation makes the model invariant to object orientation.
Random Brightness: The brightness of the input images is either randomly darkened or brightened . This makes the model robust against real world images with different lighting.
Motion Blur: We are applying motion blur to the input images. Images comes from various sources and the quality of images will not be same .Hence, blurring the images improves the model to even predict images with low quality.
Median Blur: In median blur, the central element of the input images is replaced by the median of all pixels. This makes the model robust against blurred images.
Gaussian Blur: We are applying Gaussian blur to the images which is used to remove random noise from the images.
Gauss Noise: We are applying Gauss noise to the input images which generates some random noise. This is used to smooth the sharpness of the images.

V. ALGORITHM USED

The classification of Skin Cancer is done using skin lesion images. In order to achieve high accuracy various Efficient Net models have been used. We have explored other pre trained models, but the problem with them is that they are depth scaled. And adding more layers practically will result in vanishing gradient problem. With Efficient Net we can use depth, width and resolution scaling altogether to improve network accuracy. Efficient Net B4, B5, B7 are being used as the model achieved start-of-the-art 84%.

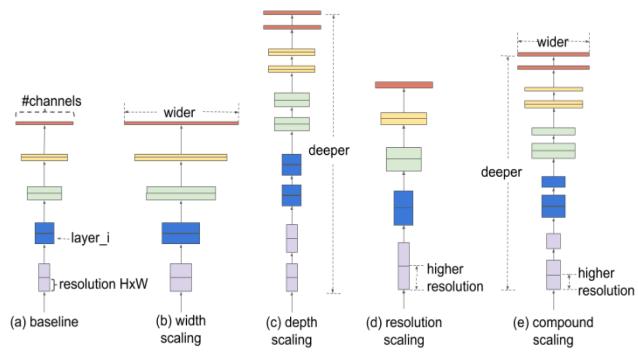


Figure 2: Efficient Net architecture diagram

VI. OUTPUT

We have predicted the images using final output layer. After the prediction of the test images, we evaluated our system with accuracy and loss against the epoch.

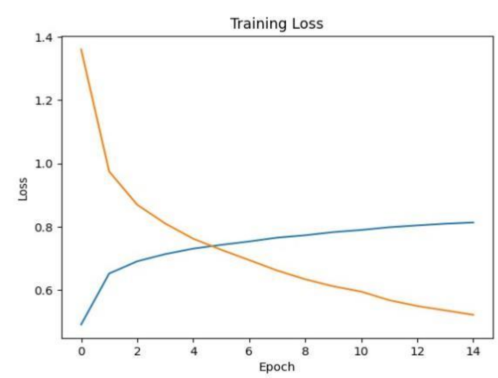


Figure 3: B4 Efficient Net Model

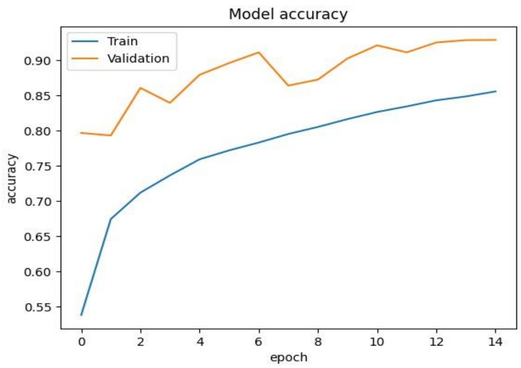


Figure 4: B5 Efficient Net Model

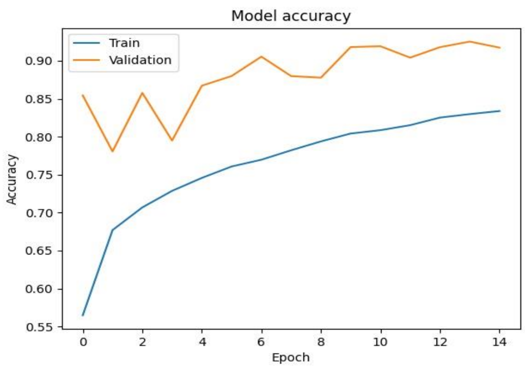
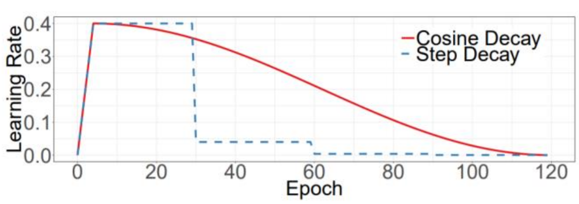


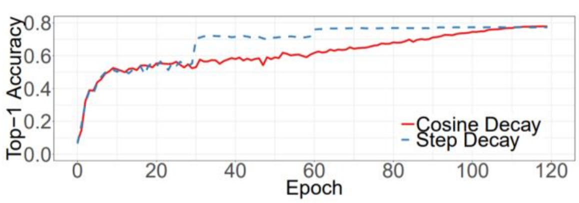
Figure 5: B7 Efficient Net Model

VII. RESULTS

We have chosen B4, B5 and B7 variant of the Efficient Net over B0 because they have achieved higher accuracy in ImageNet competition. The validation loss fluctuates in the first epoch, and it becomes stable at the end of training. The training and the validation loss is decreasing continuously, which shows that training for a larger number of epochs can help us achieve better results.



(a) Learning Rate Schedule



(b) Validation Accuracy

Figure 6 : Cosine Decay

VIII. CONCLUSION AND FUTURE WORK

In this system, a Convolutional Neural Networks - based approach have been proposed for melanoma classification. The proposed method includes images preprocessing for extracting the region of interest in the image itself and then augmenting some images to produce a bigger dataset. The dataset has been applied to CNN model which comprise of several layers. The Efficient Net model is proved to be a better network for the skin cancer detection. It is painless and timeless process and is more advantageous to patients.

The accuracy we have been achieved could be improved if trained the B7 with image resolution of 640. We are using an ensemble methodology, and it is expected that the bigger the ensemble, accuracy score will be more stable. To get a more stable accuracy score, we have proposed 18 different configurations, which can be used to achieve a better result on the skin lesion dataset. The configuration is chosen while considering the model diversity that included backbone architecture and different targets.

REFERENCES

[1] ISIC. (2018). Skin Lesion Analysis Towards Melanoma Detection. Retrieved March 20, 2021.

[2] ISIC. (2019). Skin Lesion Analysis Towards Melanoma Detection. Retrieved March 20, 2021.

[3] Mingxing T., & Quoc, L. (2019). Efficient Net: Rethinking Model Scaling for Convolutional Neural Networks.

[4] SIIM-ISIC Melanoma Classification. (2020). Identify melanoma in lesion images. Retrieved March 20, 2021.

[5] Tong, H., Zi, Z., Hang, Z., Zhong Yue, Z., Jun yuan, X., Mu, L. (2018). Bag of Tricks for Image Classification with CNN.

[6] Qishan, H., Bo, L., Fuxu L. (2020). Identifying Melanoma Images using Efficient Net Ensemble: Winning Solution to the SIIM-ISIC Melanoma Classification Challenge.

[7] S Jinnai, N Yamazaki, Y Hirano, Y Sugawara, Y ohe, Ryuji Hama Mato (2020). The development of Skin Cancer Classification systems for pigmented Skin Lesions using Deep Learning.

[8] L. Bi, J. Kim, E. Ahn and D. Feng (2019). Automatic Skin Lesion Analysis using Large-scale Dermoscopic Images and Deep Residual Networks.