INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Skin Lesion Segmentation and Classification Using Deep Learning

Mr. Gokul K S[1], Ms. Pumenitha S T[2], Mr. Hema Vasanth S V[3], Dr. S. A. Pasupathy[4]

B.E (Electronics and Communication Engineering), Kumaraguru College of Technology, Coimbatore, India

Abstract: Classification of skin lesions has recently received a lot of attention. Due to the high degree of similarity between the skin lesions, physicians frequently take a long time to examine them. A deep learning-based automated classification system can help doctors identify the type of skin lesion and improve the patient's health. With the development of deep learning architecture, the classification of skin lesions has emerged as a popular area of research. In this research, we present a method to use segmentation and techniques like Averaging of Deep Learning Architectures and classify skin lesions.
We evaluated the proposal using a large dataset: HAM 10000. Our numerical results using VGGNet and ResNet using good results on the mentioned dataset.

I. INTRODUCTION

One of the most dangerous cancers, skin cancer, can develop from skin lesions, which are changes to the skin that are irregular in comparison to the tissue that is nearby. Skin cancer is divided into two main categories: cancers other than melanoma. The significant rise in mortality and morbidity in recent years is attributable to melanoma lesions; They are the most harmful and risky of the various types of lesions. The cure rate can be increased to 90% if the doctors catch the lesions earlier. Additionally, due to the high degree of similarity between various skin lesion types (such as nonmelanoma and melanoma), visual inspection for skin cancer is challenging, which can result in a false diagnosis. The use of machine learning (ML) for the automatic classification of lesion images is one option for image inspection and healthcare systems.

The contributions of this project are as follows:
1) We propose a method to segment the skin image using the U-Net Model.
2) Averaging-based model employing two deep learning architectures the VGG16 and the ResNet.

II. LITERATURE REVIEW

To study the concepts of image segmentation and classification on Skin Lesions we have studied many papers.

Ali Kadampur, Sulaiman, Al Riyaeeal. [1] in this work, A cloud-based deep learning-based model-driven architecture for classifying images of dermal cells was proposed by the author. The fact that the ROC value is higher indicates that the classifier model frequently diagnoses cancer as cancer and non-cancer as cancer. The higher resource consumption compared to in-house methods is one of the paper's limitations. According to the observations, the DLS-built models produce superior outcomes to many of the related work's findings. DLS makes model building simple and efficient.

Mehwish Dildar, Shumaila Akram, Muhammad Irfan [2] The classification and detection of skin cancer using neural networks has been the subject of numerous authors' discussions. These methods are all non-invasive. Preprocessing and image segmentation are two steps in the skin cancer detection process, followed by feature extraction and classification. The classification of lesion images by ANNs, CNNs, KNNs, and RBFNs was the focus of this review. The study demonstrated that CNN outperforms other kinds of neural networks in terms of output.

Md. Arman Hossin, Farhan Fuad Rupom [3] The authors demonstrated that the ResNet50 architecture produced superior skin cancer classification results. In addition, this research's implementation has revealed the effects and changes in accuracy of using multi-layered CNN in melanoma cancer detection systems. Dermatologists' presentation was not explored to be contrasted and the model precision.

Marriam Nawaz, Zahid Mehmood, Tahira Nazir [4] proposed a novel approach that makes use of a contemporary deep learning-based strategy, faster-RCNN with FKM. Even in the presence of various artifacts, such as variations in illumination, noise, the presence of tiny blood vessels, or hair, the presented method's experimental results demonstrate significant improvements in melanoma detection and segmentation. Because of its shallower network, the presented melanoma detection method outperforms current skin lesion detection methods in terms of accuracy.

Neema M, Arya S Nair, Annette Joy [5] had proposed a deep CNN model that could divide the various types of melanoma into benign and malignant groups. In this work, a simpler model is used, and around 70% accuracy was achieved. In addition, the system incorporates a user-friendly and accountable graphical user interface.

Ameri A [6] proposed using dermoscopy images to identify skin cancer using a deep convolutional neural network. Malignancies other than melanoma were also included in the data. The findings demonstrated the effectiveness of deep learning in the detection of skin cancer. Smartphones can be used to enable self-diagnosis of skin cancer using this method.

Vatsala Anand, Sheifali Gupta, Deepika Koundal [7] proposed the modified U-Net model architecture for the dermoscopy image segmentation of skin lesions in order to accurately classify skin diseases. The dermoscopy images come from the 200-image PH2 dataset. The study demonstrates that the modified U-Net architecture model's Dice Coefficient and Precision could still be improved.

III. BLOCK DIAGRAM

U-Net is an architecture for semantic segmentation. It consists of a contracting path and an expansive path. The contracting path follows the typical architecture of a convolutional network.

It consists of the repeated application of two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. At each downsampling step we double the number of feature channels. Every step in the expansive path consists of an upsampling of the feature map followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers.
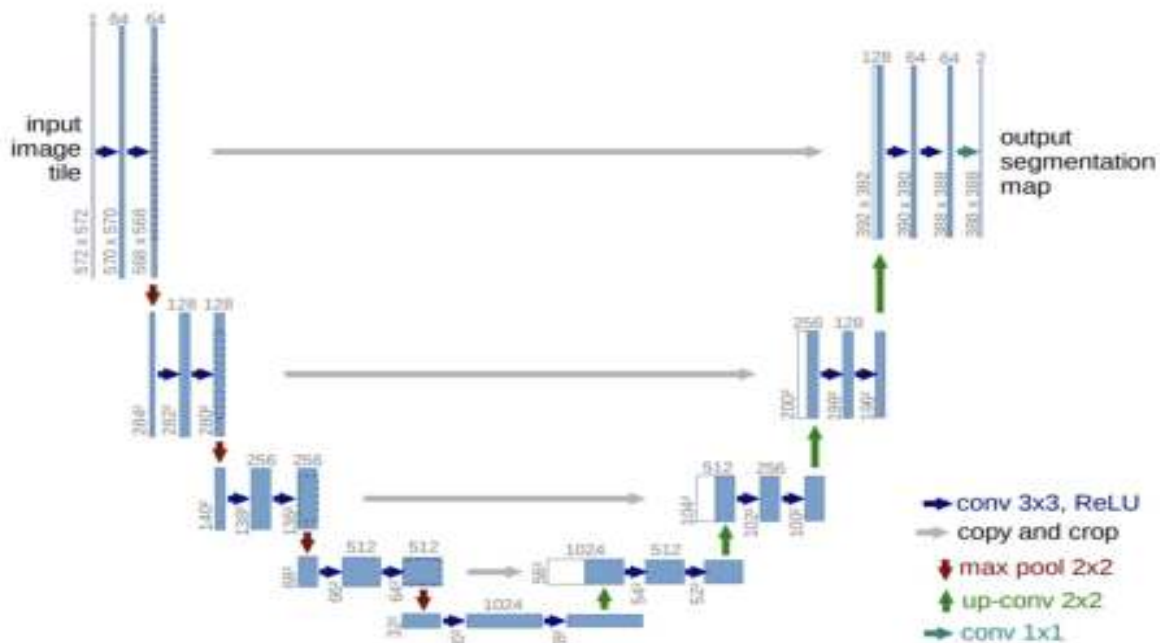

Figure. 3.1 U-Net Architecture

VGG was developed at Oxford by a group known as Visual Geometry Group, it came as an improvement of AlexNet with the aim of achieving higher accuracy. There are three versions of VGG such as VGG11, VGG16, and VGG19. The architecture of the VGG19 consists of nineteen layers, as shown in the figure, the VGG19 has sixteen convolution layers each with 3x3 filter, two fully connected layers each with 4,096 neurons using ReLU activation function, five max pooling layers, and a SoftMax layer with one thousand neurons. The input is an RGB image (224, 224, 3), the kernel size is 3x3 with a stride of one, and the max pooling used a 2x2 filter with a stride of two. In our case, VGG16 was issued for feature extraction and model creation.
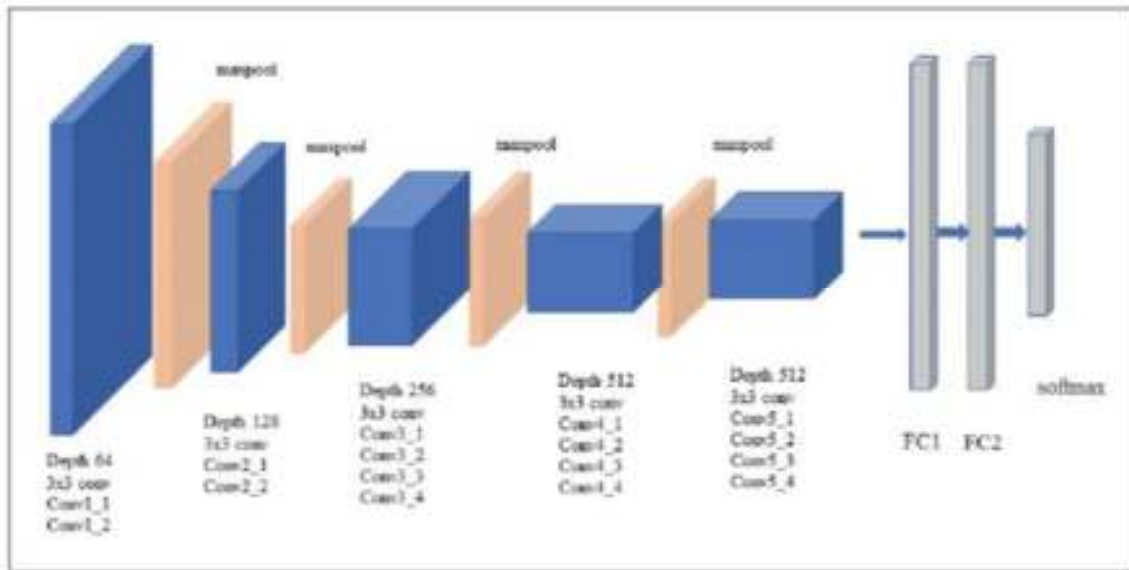
Figure. 3.2 VGG19 Architecture

Some deep neural networks fail to train because of vanishing and exploding gradients. Increasing network depth makes neural networks prone to higher training error, which causes degradation of the training accuracy overtime. To overcome this problem, a residual neural network (ResNet) was introduced. ResNets create skip connections that allows the network to jump over some layers during forward propagation However, the network can still be trained using stochastic gradient descent with back propagation.
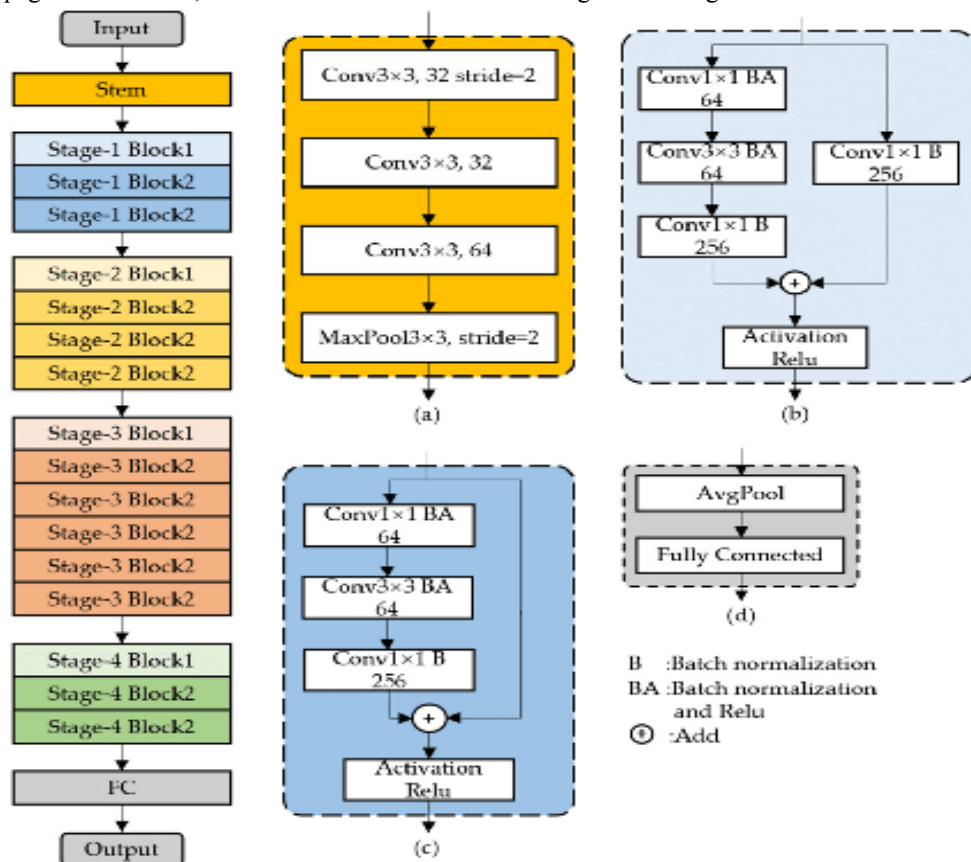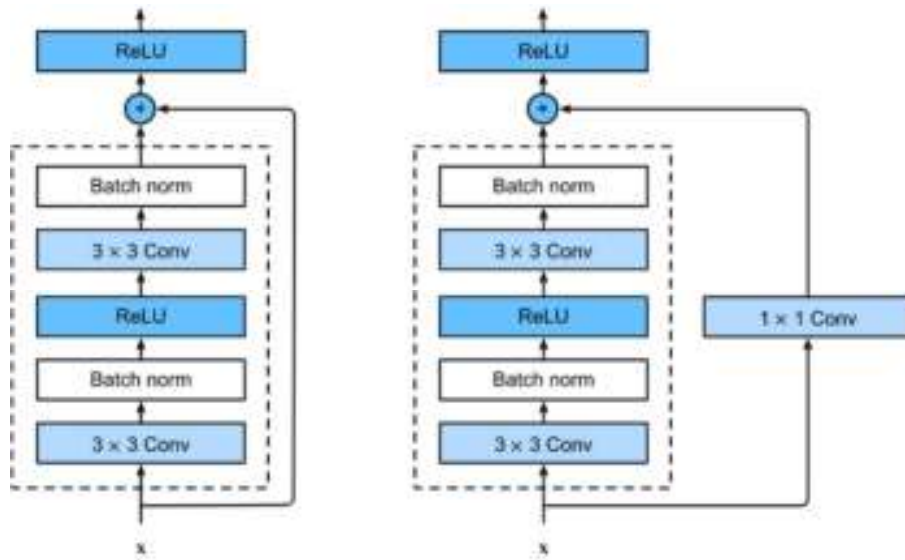
Figure. 3.3 ResNet Architecture

IV.   ALGORITHM

To fully understand the model's algorithm, it is important to understand that the model is divided into two sub-models. This division helps in the easy identification of bugs and better performance of the model.
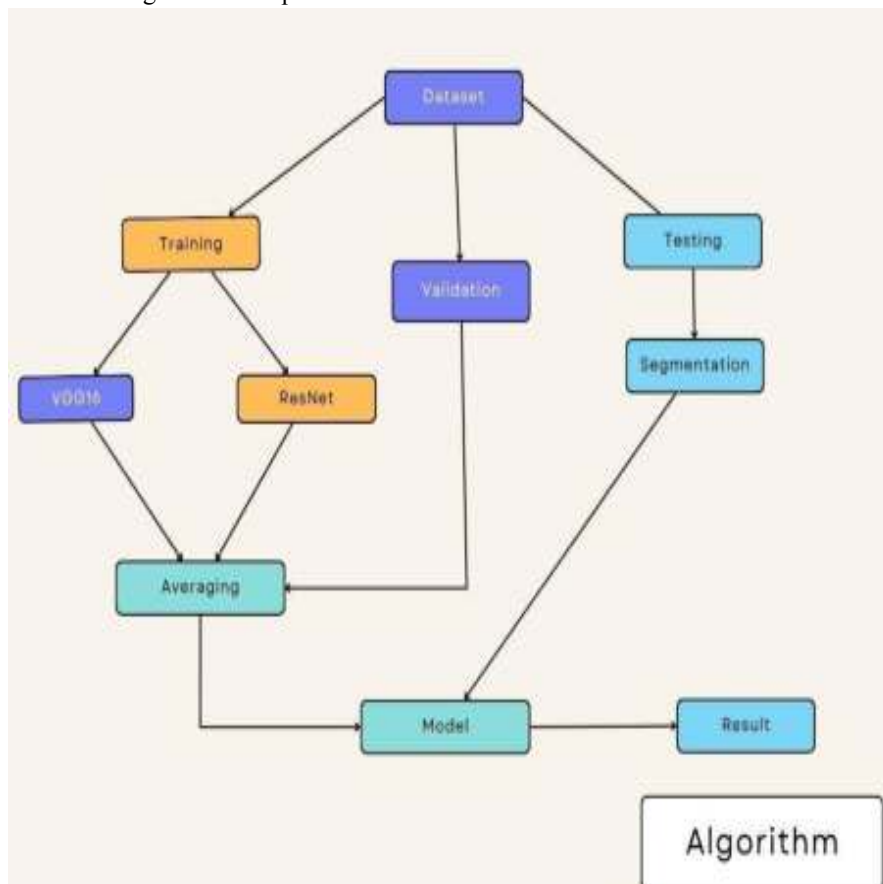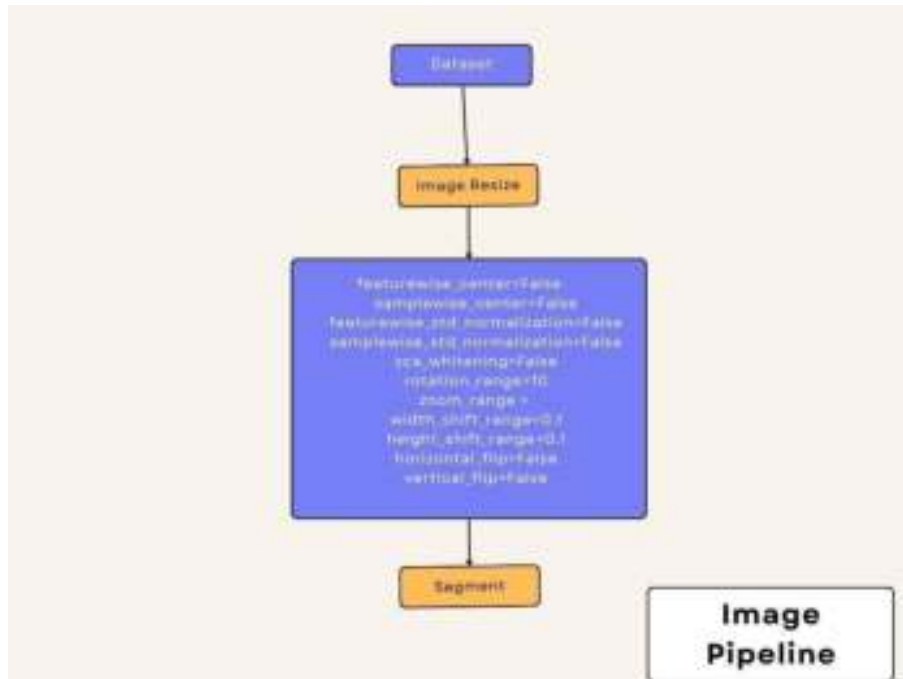
Figure. 4.1 Algorithmic Workflow

A U-shaped architecture consists of a specific encoder-decoder scheme: The encoder reduces the spatial dimensions in every layer and increases the channels. On the other hand, the decoder increases the spatial dims while reducing the channels. The tensor that is passed in the decoder is usually called bottleneck. In the end, the spatial dims are restored to make a prediction for each pixel in the input image.
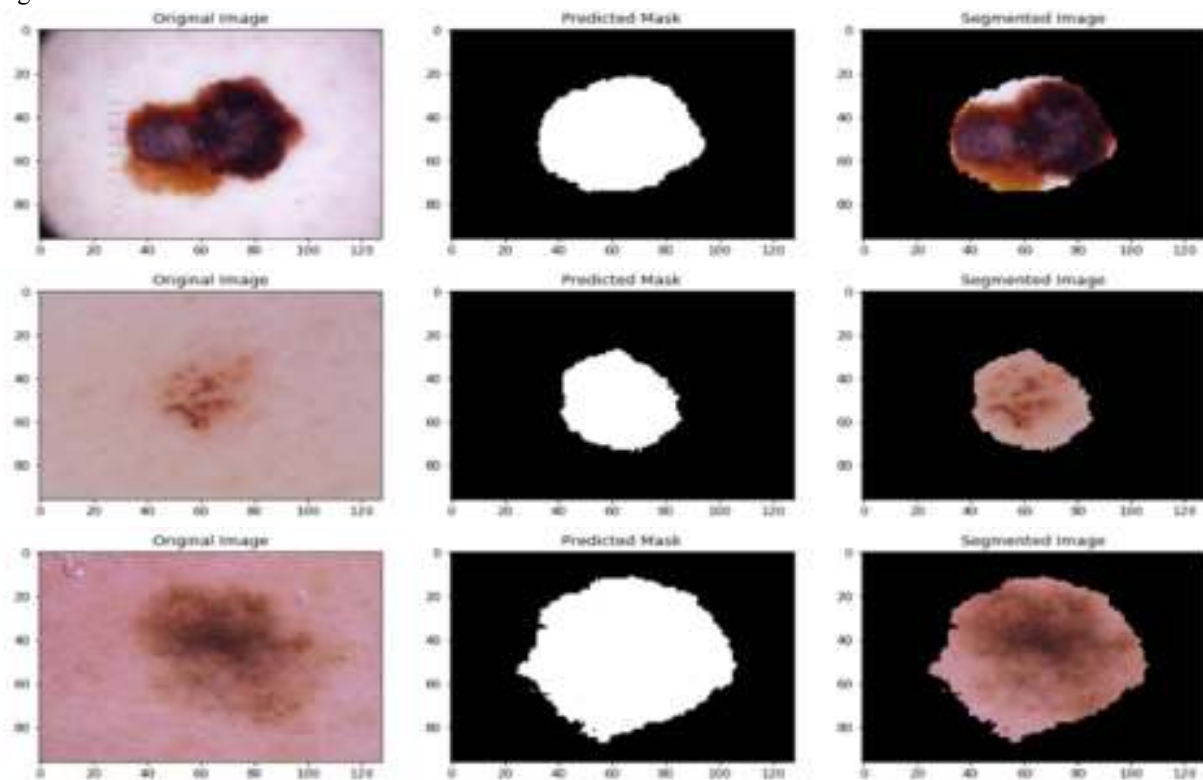

Figure. 4.2 Segmented Images

The data is split into the following:
1) Training Data comprises of 90% of the total dataset
2) Testing Data comprises of 10% of the total dataset
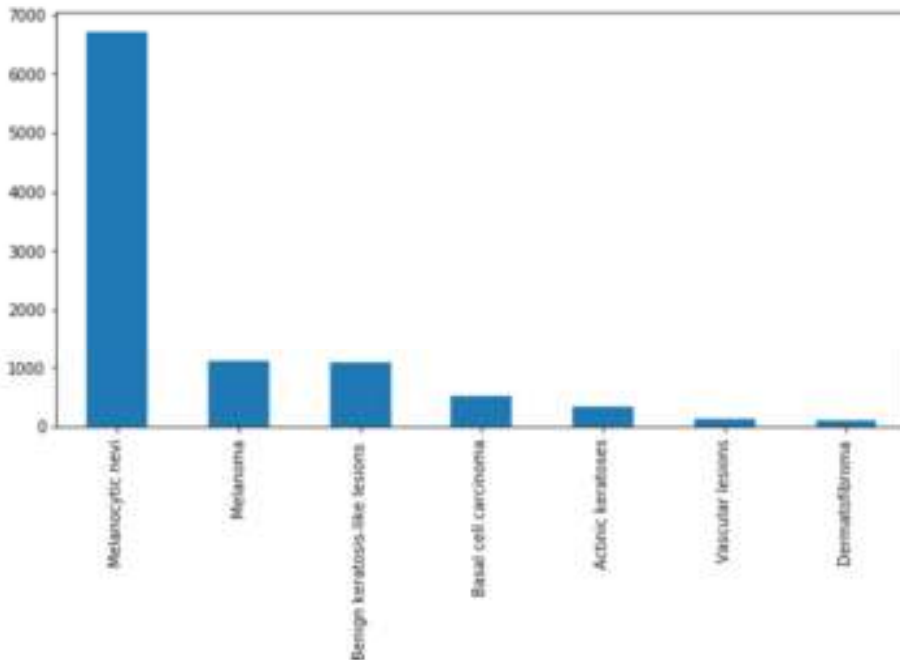3) Validation Data is 10% of the training data

Figure. 4.3 Dataset Distribution

There is a clear skewed distribution in the data, therefore can lead to bias in the model and there is a need for resampling. The resampling of data points is done by upsampling and downsampling of data, which gives a relatively better distribution.
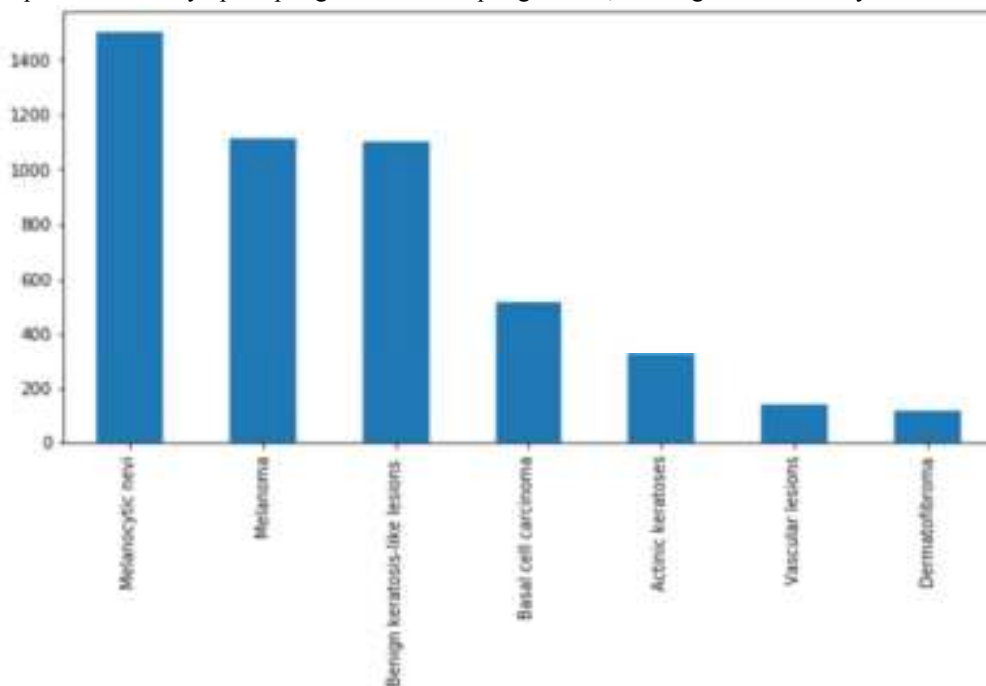

Figure. 4.4 Dataset after Sampling

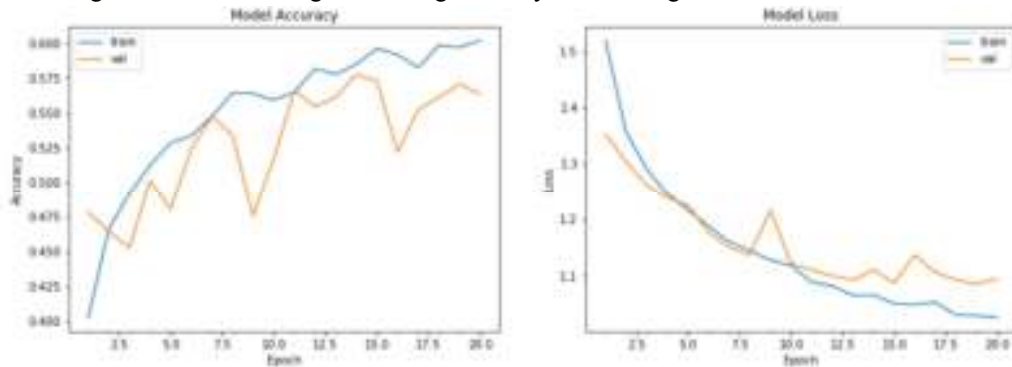The results for training in VGG16 including the training accuracy and loss is given below:

Figure. 4.5 VGG16 Training Metrics

The results show that there needs to be an improvement in the model because of very average results. When it ran for 20 epochs, the model accuracy was 76.25% and validation accuracy was 60.97%. The training and validation loss were of 0.6039 and 1.664, with shows that there is room for improvement

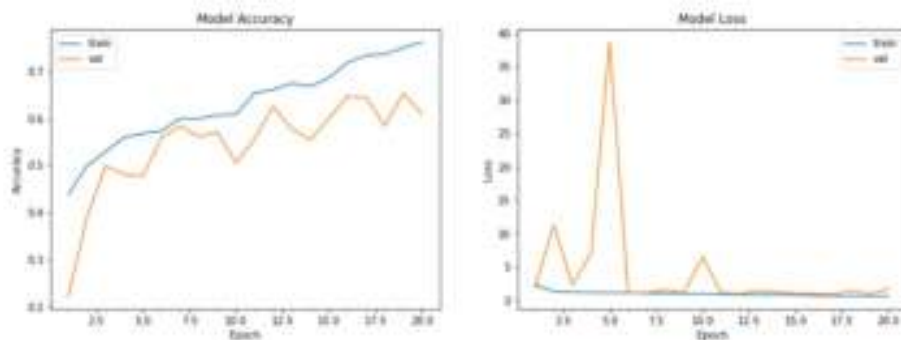The results for training in ResNet50 including the training accuracy and loss is given below:

Figure. 4.6 ResNet Training Metrics

The results gave an indication of it being a relatively good model. When it ran for 20 epochs, the model accuracy was 76.25% and validation accuracy was 60.97%
The training and validation loss were of 0.6039 and 1.664, with shows that there is room for improvement
Layer that averages a list of inputs element-wise. It takes as input a list of tensors, all of the same shape, and returns a single tensor (also of the same shape). Averaging is done to combine the results of multiple models to get better results. For instance in random forest algorithm, the results of multiple trees are used to get the collective result.

Figure. 4.7 Averaging Layer

The result shows an increase in validation accuracy upto 65% for 20 epochs, which is an improvement compared to previous models. Therefore, it shows that ensembling models generalizes well and can extend to better architectures to provide better results. We can also try to train for a higher number of epochs to compare the performance.
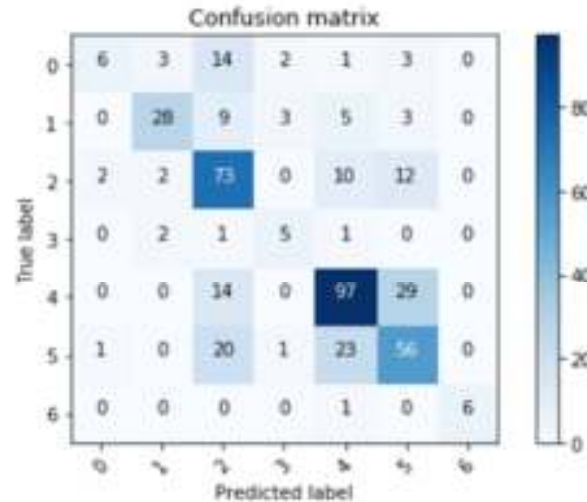

Figure. 4.8 Confusion Matrix

The above picture is the confusion matrix of the ensemble model. A Confusion matrix is an *N x N matrix* used for evaluating the performance of a classification model, where N is the number of *target classes*. The matrix compares the actual target values with those predicted by the machine learning model.

The results show that the ensemble model performs well for certain classes and shows wrong predictions for certain classes.

It is better for cross-validation with the wrong predictions to evaluate the performance of the model on various classes. As we can see in the table below the model does worse for the Melanocytic Nevi class and the best for the Dermatofibroma class
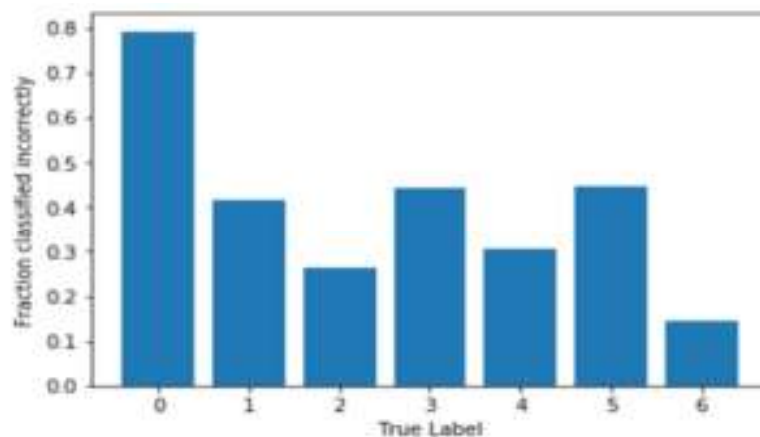

Figure. 4.9. Incorrectly Classified Labels

V. RESULTS

The output results are evaluated after performing segmentation on random sampling on the test data and using 5 images.

The U-Net architecture is applied on the images to obtain the segmented output and those are fed into the ensemble model to get predictions. The output are probabilities of the classes ranging from 0 to 1 and the argmax of the results will give the maximum probability class which is the predicted output.
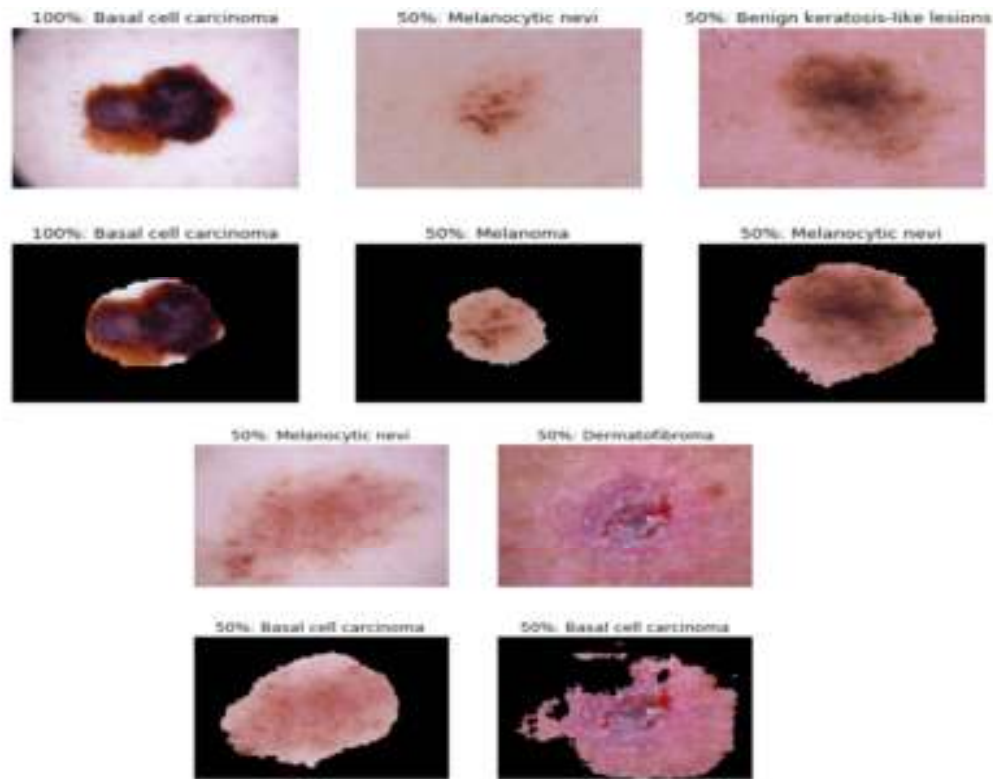
Figure. 5.1. Predicted Images

The output shows that the segmented images perform well in the comparison, as it has predicted 4 out of 5 images correctly with one misclassification as Basal cell carcinoma instead of Melanocytic Nevi. This shows that UNet Segmentation has improved the performance of the pipeline and can be used in prior sections for further improvement.

VI.   CONCLUSION

In conclusion, the pipeline involving both segmentation and classification performed well in the task of skin lesion classification. As hypothesized, the ensemble model performed well that the individual architectures and opened up the space of trying out various architectures. U-Net Segmentation proved effective in improving the metrics and accuracy of the predictions. This shows that this pipeline has a huge potential in the domain of medical diagnosis and can be further improved.

The project implementation also shows that the performance of the pipeline depends on the architecture used and may vary from various architectures available therefore there is a need for proper choice by testing and validating of models with different parameters such as epochs and learning rate.

REFERENCES

[1]   Kadampur, M. A., & Al Riyaee, S. (2020). Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images. Informatics in Medicine Unlocked, 18, 100282.

[2]   Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., ... & Mahnashi, M. H. (2021). Skin cancer detection: a review using deep learning techniques. International journal of environmental research and public health, 18(10), 5479.

[3]   Hossin, M. A., Rupom, F. F., Mahi, H. R., Sarker, A., Ahsan, F., & Warech, S. (2020, October). Melanoma skin cancer detection using deep learning and advanced regularizer. In 2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS) (pp. 89-94). IEEE.

[4]   Nawaz, M., Mehmood, Z., Nazir, T., Naqvi, R. A., Rehman, A., Iqbal, M., & Saba, T. (2022). Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. Microscopy Research and Technique, 85(1), 339-351.

[5]   Neema, M., Nair, A. S., Joy, A., Menon, A. P., & Haris, A. (2020). Skin lesion/cancer detection using deep learning. International Journal of Applied Engineering Research, 15(1).

[6]   Ameri, A. (2020). A deep learning approach to skin cancer detection in dermoscopy images. Journal of Biomedical Physics and Engineering, 10(6), 801-806.

[7]   Anand, V., Gupta, S., Koundal, D., & Singh, K. (2023). Fusion of U-Net and CNN model for segmentation and classification of skin lesion from dermoscopy images. Expert Systems with Applications, 213, 119230.

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⊙ (24*7 Support on Whatsapp)