DETECTION OF SKIN CANCER USING DEEP LEARNING MODEL MOBILENET

Aqdus Charolia
MS in Computer Science
University of Missouri Kansas City
Missouri, 64112
aacy5x@umsystem.edu

Affan Charolia
MS in Computer Science
University of Missouri Kansas City
Missouri, 64112
aacbb8@umsystem.edu

Vyoma Desai
MS in Computer Science
University of Missouri Kansas City
Missouri, 64112
vhdmmh@umsystem.edu

Mohammad Shaik
MS in Computer Science
University of Missouri Kansas City
Missouri, 64112
ms6bz@umsystem.edu

***Abstract* — Skin cancers are the most common form of cancers in humans, a physician faces many difficulties for accurate diagnosis of lesions through its characteristics and in the naked eye. Major causes of Skin cancer is in the result of DNA cells which causes the presence of genetic mutations on the skin. Eventually it starts spreading all over the body and hence at earlier stages we need to focus on curing as soon as possible. With more and more research in detecting early stage cancer, there are various skin parameters which have been used in detecting skin cancer. They are Shape, Size, Symmetry and Color. [7] Using deep learning we have a detailed review of detecting skin cancer. The report presents a detailed review of detection of skin cancer using Deep Learning. For that it is necessary to improve the default methods in order to increase diagnostic accuracy.**

In this paper, we evaluate the chance of deep learning algorithms to detect skin cancer classifying Melanoma. For our research we used the Ham10000 dataset. After pre-processing, we got a total of 38568 training images and 938 validation images. We developed our model using Keras and TensorFlow. Our model shows promising results with an accuracy of 89.77%. As a result, we developed a web based application where Skin-Cancer images can be uploaded to get the closest prediction between 7 different classes of skin cancer (Benign, Keratosis, Basal Cell Carcinoma, Actinic Keratosis, Vascular Skin Lesions, Dermatofibroma and Neoplasms) using Deep Learning to perform the prediction.

Keywords— deep learning mobile net skin cancer detection, Melanoma, Benign, Keratosis, Basal Cell Carcinoma, Neoplasms, Ham10000 dataset, pre-processing, prediction, Keras and Tensorflow, web application.

I. Introduction

Most active cancer in the USA is skin cancer with nearly 5.4 million cases. At a glance, by old age, 1 in 5 Americans will develop skin cancer. Every hour a couple of people die because of skin cancer. Since skin is an important part of the body and so it is important to consider skin cancer seriously. There are two types of skin cancer, melanoma and non-melanoma. Melanoma is a type of skin cancer that is dangerous, uncommon, and deadly. According to official sources such as American Cancer society, only 1% of total cases are related to melanoma skin cancer, but it leads to a higher mortality rate. This type of skin cancer grows in cells called melanocytes. It begins when healthy melanocytes begin to grow uncontrollably, creating a cancerous growth and can affect any part of the human body. Areas of the body exposed to sunlight, such as face, neck, hands and lips etc. are much more likely to catch this skin cancer. This type of skin cancer is only treatable if detected in the initial stages else they tend to spread across the body and cause a very painful death. There are different types of melanoma skin cancer such as nodular melanoma, superficial spreading melanoma, acral lentiginous, and lentigo maligna. Majority of skin cancer types are generally under non melanoma stages, such as basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and sebaceous gland carcinoma (SGC). These are formed in the upper layers of the epidermis. Nonmelanoma cancer is relatively easy to treat compared with melanoma cancer. Doctors normally use the lavage techinque for skin cancer detection. This process is painful, slow, and time-consuming. When detected in the late stages its survival rate is less than 14%. If detected in the early stages the chances for survival are about 97%, but for this detection to occur the diagnosis has to be done really fast and luck should be on one's side. With the advancement in AI research & computer vision, software based solutions offer a speedy diagnosis and low cost or inexpensive of skin cancer symptoms. It simply comprises capturing the image of the skin, preprocessing it, segmenting it, feature extraction and training the model using Deep Learning Algorithm and classifying it to 7 different classes of skin cancer.

Deep learning as we know today has evolved a lot over the past few years especially with the CNN. The field of computer vision has grown leaps and bounds with the help of deep learning. As a result the medical field has benefited greatly. Various deep learning approaches have been developed over the years to tackle the diagnosis of skin cancer. Our report focuses on reviewing previously developed models and comparing our model with them. [1]

II. Problem Statement

Most active cancer in the USA is skin cancer, nearly 5.4 million cases.

At a glance, by old age, 1 in 5 Americans will develop skin cancer. Every hour a couple of people die because of skin cancer. When detected in the late stages its survival rate is less than 14%. If detected in the early stages the chances for survival are about 97%, but for this detection to occur the diagnosis has to be done really fast and luck should be on one's side. People are reluctant to get themselves checked when they notice any unusual lesion developing on their skin.

III. Our Solution

Presented a Web based application where Skin-Cancer images can be uploaded to get the closest prediction between 7 different classes of skin cancer using Deep Learning to perform the prediction. Resnet50, MobileNet, VGG16, CNN models have been trained on the HAM10000 Dataset after which the model with best validation accuracy has been employed in the webapp. Classes of Skin Cancer Images are Melanoma, Benign Keratosis, Basal Cell Carcinoma, Actinic Keratosis, Vascular Skin Lesions, Dermatofibroma and Neoplasms The application is designed to be used by medical professionals for a quick result.

IV. Related Work

Skin cancer detection and classification has always been significantly improved over the years. There is a tremendous growth in predictive image analysis of skin cancer ever since visual-based technologies came into the modern world. The latest research case studies have proved significant chances of detecting skin cancer. In 2018, the first ever contest held by ISIC was a far more promising event which has now become a touchstone in this area and development of mobile applications can also be used for real-time cancer detection. More and more research work has gained the best accuracy in diagnosis of early skin cancer detection and provided optimal solutions by various different classification techniques.

In the era of deep learning, various techniques and methods have been developed and implemented on image analysis. CNN (Convolutional Neural Network) is the best model for image classifications. Every improvement of different types of CNN architecture such as AlexNet, GoogleNet, VGGNet, ResNet in performance accuracy and efficiency and reducing the number of parameters made great benefits for deep image classifications.

Xie et al. [8] have proposed a classification system for lesions on skin into benign and malignant. In their proposed system they have implemented 3 phases. The first phase of the system employs a self-generating Neural Network that extracts skin lesions. The second phase of the system uses image segmentation to extract information from the image and shape out borders, textures and colors. This procedure featured out a total of 57 features and 7 novel features related to descriptions of lesion borders. To reduce the dimensionality of the features Principal Component Analysis was implemented. In the third and final step Ensemble Neural Networks were used as this process employs various models of Deep Learning to train the data and give the best accuracy in classification.Ensemble Neural Network employs backpropagation (BP) NN and fuzzy neural networks.

A new method for skin cancer detection was described by Aswin et al.[9] and in his case study. They defined and proposed the generic ANN algorithm to detect cancer and classify images. Preprocessing was focused on removal of hair and medical imaging software anmed dull Rozar. In their research we found out that the Region of Interest (ROI) was extracted using the threshold method. Mainly for the classification of lesion images into cancerous and non cancerous classes, he has implemented GLCM technique and used a hybrid ANN and GA classifier and gained an overall accuracy score as 88%.

All the above authors have proposed deep learning models but as we are moving towards more mobile devices, we need to use a model which is small in size and at the same time computationally quick in providing predictions. Thus we have built our model using MobileNet. MobileNet is one of the first TensorFlow's models which gained tremendous popularity in designing to be used by mobile applications. We have mostly focused on MobileNet architecture for our skin cancer detection. This has helped us in implementing a well, fast and efficient way to give us quick results as compared to ResNet50, VGG16 and custom CNN model.

V. Materials and Methods

Dataset:

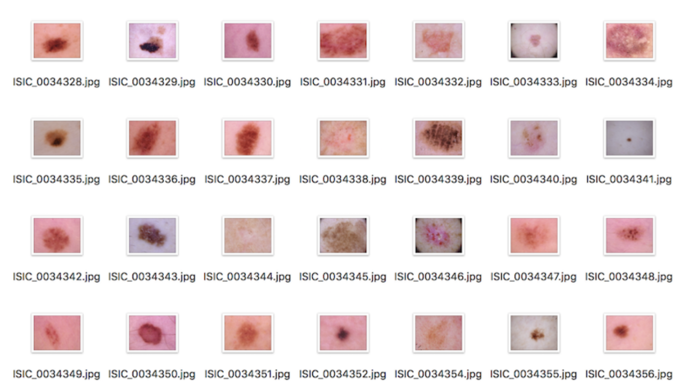

Fig 1 : HAM10000 dataset part 1

We have used the Skin Cancer HAM10000 Dataset which consists of 10015 dermatoscopic images which can serve as a training set for academic machine learning / deep learning purposes. Was part of the 2018 ISIC Challenge where Lesion Segmentation, Lesion Attribute Selection and Disease Classification were the tasks to be performed. The dataset contains 6705 nv, 1113 mel, 1099 bkl, 327 akiec, 142 vasc, and

24 df classes of Skin Cancer images. The dermatoscopic images are collected from different populations, acquired and stored by different modalities by the authors.

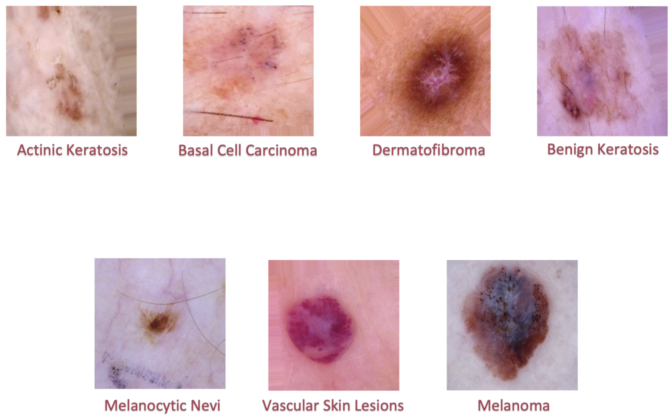

Fig 2 : Skin Cancer Image Classes

<u>Data Preprocessing and Feature Extraction</u>:

Dataset split first into train and test images, each with the further classification based on the skin cancer type (7 classes) using the ham10000.csv file which details which image label belongs to which class. Data Augmentation performed on the training dataset to increase the overall size of the image dataset and improve the accuracy of the learning based model.

Mainly for performing this task we have used Tensorflow's ImageDataGenerator library. Performed an image rotation, width shift, height shift, zoom, horizontal and vertical flip and image normalization (0-255) and reshaping them to (224,224,3). Pre-processed 38568 training images and 938 validation images.

VI. Research Methodology

We have used following Deep Learning Techniques for Skin Cancer Detection :

Model 1 - CNN :

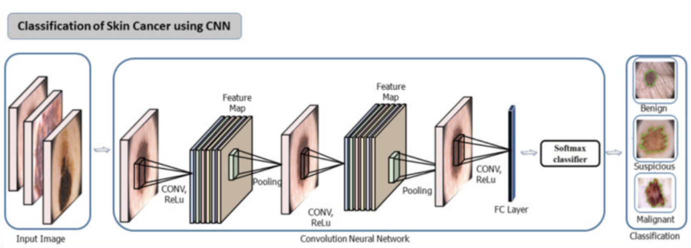

Fig 3: Classification of Skin Cancer using CNN (Source: Internet)

We have built our own custom Convolutional Neural Network layers starting from 16 layers to 256 layers followed by max pooling and dropout layers. Initiated with a batch normalization layer to avoid overfitting. [5] Our CNN consists of 5 conv2d layers, 1 fully connected layer, 5 dropout and maxpool layers and 1 flatten layer. A total of 646,471 params were trained.

input_7 (InputLayer)	[(None, 224, 224, 3)]	0
conv2d_28 (Conv2D)	(None, 222, 222, 32)	896
max_pooling2d_28 (MaxPooling	(None, 111, 111, 32)	0
dropout_34 (Dropout)	(None, 111, 111, 32)	0
conv2d_29 (Conv2D)	(None, 109, 109, 64)	18496
max_pooling2d_29 (MaxPooling	(None, 54, 54, 64)	0
dropout_35 (Dropout)	(None, 54, 54, 64)	0
conv2d_30 (Conv2D)	(None, 52, 52, 128)	73856
max_pooling2d_30 (MaxPooling	(None, 26, 26, 128)	0
dropout_36 (Dropout)	(None, 26, 26, 128)	0
conv2d_31 (Conv2D)	(None, 24, 24, 256)	295168
max_pooling2d_31 (MaxPooling	(None, 12, 12, 256)	0
dropout_37 (Dropout)	(None, 12, 12, 256)	0
flatten_6 (Flatten)	(None, 36864)	0
dense_12 (Dense)	(None, 7)	258055

Fig 4: CNN Model Summary

Model 2 - VGG16 :

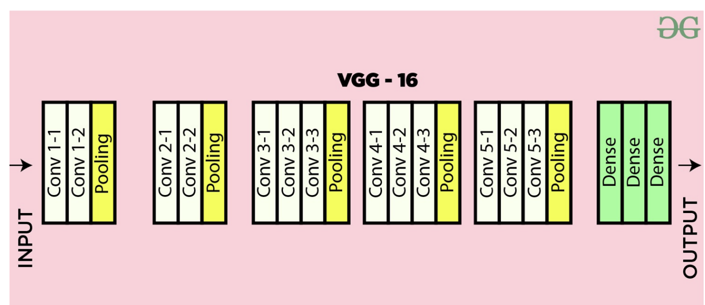

Fig 5: VGG16 Architecture (Source: Internet)

The VGG16 model achieves 92.7% top-5 test accuracy in ImageNet. ImageNet dataset of over 14 million images belonging to 1000 classes. It has multiple 3×3 kernel-sized filters one after another using transfer learning, freezing the previous layers of the model (keep the same trained weights) and performing global_avg_pool followed by adding custom fully connected layers to train our data.

block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
global_average_pooling2d (Gl	(None, 512)	0
dense (Dense)	(None, 128)	65664
dense_1 (Dense)	(None, 64)	8256
dense_2 (Dense)	(None, 32)	2080
dense_3 (Dense)	(None, 7)	231

Total params: 14,790,919
Trainable params: 76,231
Non-trainable params: 14,714,688

Fig 6: VGG Model Summary

Model 3 - ResNet50 :

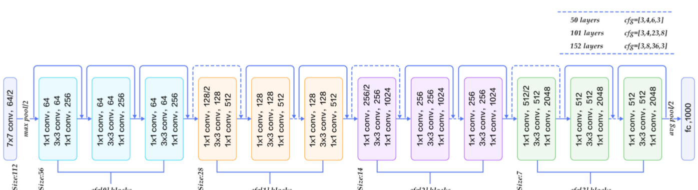

Fig 7: ResNet Architecture (Source: Internet)

ock2_out (Activation)	(None, 7, 7, 2048)	0	conv5_block2_add[0][0]
ock3_1_conv (Conv2D)	(None, 7, 7, 512)	1049088	conv5_block2_out[0][0]
ock3_1_bn (BatchNormali	(None, 7, 7, 512)	2048	conv5_block3_1_conv[0][0]
ock3_1_relu (Activation	(None, 7, 7, 512)	0	conv5_block3_1_bn[0][0]
ock3_2_conv (Conv2D)	(None, 7, 7, 512)	2359808	conv5_block3_1_relu[0][0]
ock3_2_bn (BatchNormali	(None, 7, 7, 512)	2048	conv5_block3_2_conv[0][0]
ock3_2_relu (Activation	(None, 7, 7, 512)	0	conv5_block3_2_bn[0][0]
ock3_3_conv (Conv2D)	(None, 7, 7, 2048)	1050624	conv5_block3_2_relu[0][0]
ock3_3_bn (BatchNormali	(None, 7, 7, 2048)	8192	conv5_block3_3_conv[0][0]
ock3_add (Add)	(None, 7, 7, 2048)	0	conv5_block2_out[0][0] conv5_block3_3_bn[0][0]
ock3_out (Activation)	(None, 7, 7, 2048)	0	conv5_block3_add[0][0]
verage_pooling2d_3 (Glo	(None, 2048)	0	conv5_block3_out[0][0]
(Dense)	(None, 128)	262272	global_average_pooling2d_
(Dense)	(None, 64)	8256	dense_12[0][0]
(Dense)	(None, 32)	2080	dense_13[0][0]
(Dense)	(None, 7)	231	dense_14[0][0]

rams: 23,860,551
e params: 272,839
nable params: 23,587,712

Fig 8: RESNET Model Summary

Deep residual network - ResNet-50. Convolutional neural network (CNN) which is 50 layers deep. This is a very popular model and used in a wide variety of applications. It uses skip connections to add the output from an earlier layer to a later layer which helps tackle the vanishing gradient problem. We have performed transfer learning, by freezing the previous layers, followed by global avg maxpool and then added our custom dense layers to train our data.

Model 4 - MobileNet :

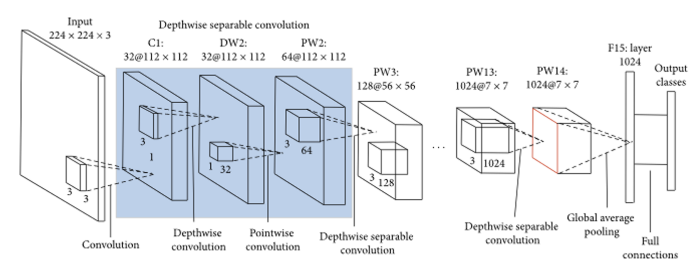

Fig 9: MobileNet Architecture (Source: Internet)

MobileNet was developed with the intention of using them in embedded mobile devices. They are developed by using depth wise separable convolutions which eventually leads to lower latency. [4]

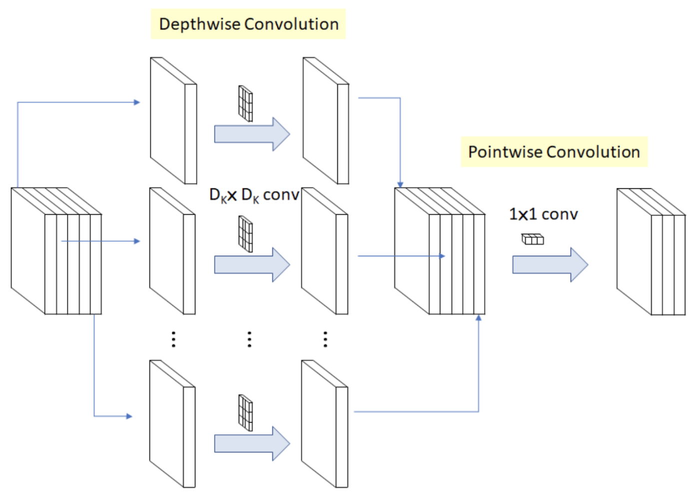

Fig 10: Depthwise followed by pointwise Convolution. (Source: Internet)

The primary layer in MobileNet is the depthwise separable convolution layer to reduce the number of features. The training process was run on Macbook with M1 GPU, 1296 MHz core speed, 32 unified pipelines. We created our own model architecture by excluding the last 5 layers of mobileNet using transfer tuning. Created a new dense layer for predictions and flatten the output into 1 dimension. Our dropout value is 0.25 and made predictions on 7 classes with softmax as the activation function. Keeping model freezing weights into consideration we have freezed the weights of the last 23 layers and retrained into our new model. Owing to this research and experiment we have gained Training accuracy as 96% and validation accuracy as 89.7%. The parameters we used in compilation of this study is with learning rate as 0.01 and with categorical cross-entropy loss.

A depthwise separable convolution is made from two operations. Depthwise convolution is the channel wise spatial convolution. In the image below we will have 5 x 5 spatial convolution. Mainly it is a map of a single input channel on each separately. Whereas pointwise convolution will have 1 x 1 convolution to change the dimension mainly combining the features created by depthwise convolution.

```
conv_pw_12_bn (BatchNormaliz (None, 7, 7, 1024)        4096

conv_pw_12_relu (ReLU)       (None, 7, 7, 1024)        0

conv_dw_13 (DepthwiseConv2D) (None, 7, 7, 1024)        9216

conv_dw_13_bn (BatchNormaliz (None, 7, 7, 1024)        4096

conv_dw_13_relu (ReLU)       (None, 7, 7, 1024)        0

conv_pw_13 (Conv2D)          (None, 7, 7, 1024)        1048576

conv_pw_13_bn (BatchNormaliz (None, 7, 7, 1024)        4096

conv_pw_13_relu (ReLU)       (None, 7, 7, 1024)        0

global_average_pooling2d (Gl (None, 1024)              0

flatten (Flatten)            (None, 1024)              0

dense (Dense)                (None, 1024)              1049600

dropout (Dropout)            (None, 1024)              0

dense_1 (Dense)              (None, 7)                 7175
=================================================================
Total params: 4,285,639
Trainable params: 2,914,823
Non-trainable params: 1,370,816
```

Fig 11: MobileNet Model Summary

VII. Results

The following table below shows the evaluation of the CNN, VGG16, ResNet50 and MobileNet over the computational time to train the model, their training and validation accuracy and their training and validation losses.

Models Characteristic	CNN (10 epochs)	VGG16 (10 epochs)	Resnet50 (20 epochs)	MobileNet (20 epochs)
Timeframe	15.01 Mins	38.83 Mins	44.6 Mins	14.56 Mins
Training Accuracy	66.64%	75.37%	77.60%	95.37%
Training Loss	86.37%	76.80%	57.28%	12.58%
Validation Accuracy	82.4%	83.05%	86.67%	89.77%
Validation Loss	55.93%	47.78%	40.00%	35.20%

Table 1: Comparative results table

As we can see from the table, MobileNet is the best model overall with training accuracy as 95.37% and in entire processing was done in 14.56 mins time frame as compared to other models. ResNet has inference time greater than CNN model since it has many parameters and layers to train. Since we wanted to focus and integrate on devices such as smartphones, mobiles we used the MobileNet architecture model to efficiently train on our HAM10000 dataset. We could easily, accurately detect skin cancer with very less inference time as 14.56 mins.

The best way to compare the models is by plotting graphs and we have implemented our multi-line plot and compared the model accuracy with loss using both models such as ResNet50 and MobileNet respectively.

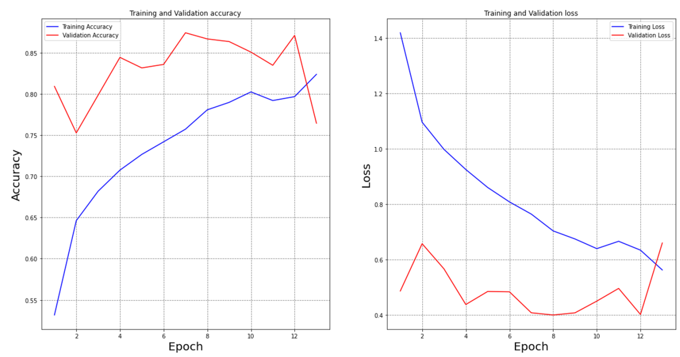

Fig 12: ResNet50 - Model Accuracy and Loss Plot

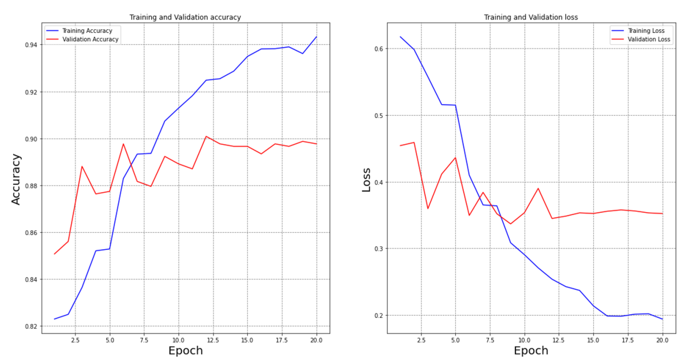

Fig 13: MobileNet - Model Accuracy and Loss Plot

As ResNet50 and MobileNet have the best training and validation accuracy we have implemented the curves for them in Fig12 and Fig13. We notice in ResNet50 that the training loss decreases at about 12 epochs and the validation loss increases at 12 epochs. As for the MobileNet model the validation loss remains constant at 0.35 after 12 epochs.
After the comparisons we have seen that MobileNet accuracy is far much better than ResNet model since in resnet there are many parameters and inference time is quite larger than MobileNet. MobileNet proves to be the best model in detecting and classifying image skin cancer and as the number of parameters have been reduced this results in light-weight deep neural network and much more useful when designing architecture for mobile applications.

VIII. Web Application

Fig 14: Skin Cancer Detection Web Application - Prediction of results using trained MobileNet model

We have presented a Web based application where Skin-Cancer images can be uploaded to get the closest prediction between 7 different classes of skin cancer using Deep Learning to perform the prediction. The application is designed to be used by medical professionals for a quick result.

Elements in Web Applications:
Upload Skin Lesion Image - This button will help to upload the skin image onto the web application.
Predict - This button will predict the image and detect the actual type of skin cancer using backend API.
Displaying the image with the type of skin cancer has been successfully implemented.

IX. Conclusion

Training Accuracy of 95.37% and Validation Accuracy of 89.77% is achieved using MobileNet Deep Learning Model.
MobileNet performed well in classifying all the Skin Cancer Lesions. MobileNet performance is much faster and more accurate compared to ResNet50, VGG16 and cutom CNN model. MobileNet is best suited for mobile applications as it has a small size and faster performance.

X. Future Work

In terms of future work we will want to implement newer versions of MobileNet namely version 2 and 3 with new parameters. We can raise the accuracy to more than 95% hence thereby producing much more accurate results. Also it would be interesting to apply optimization techniques like pruning and quantization on these models to improve accuracy.
Another aspect we would like to work on is increasing the size of the dataset not by augmentation but more images of skin lesions.
The proposed system is implemented in a web application where the image is uploaded and the model will predict and display the classification result. Day by Day we are moving towards mobile applications for completing most of our desired tasks in a quick and efficient way. Hence, we aim at creating a mobile application where the end user can capture the image and detect the skin cancer within seconds

XI. References

1] Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., Alsaiari, S. A., Saeed, A., Alraddadi, M. O., & Mahnashi, M. H. (2021). Skin Cancer Detection: A Review Using Deep Learning Techniques. International journal of environmental research and public health, 18(10), 5479. https://doi.org/10.3390/ijerph18105479
2] Mohammad Ali Kadampur, Sulaiman Al Riyaee, Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images, Informatics in Medicine Unlocked, Volume 18,2020, 100282, ISSN 2352-9148, https://doi.org/10.1016/j.imu.2019.100282. (https://www.sciencedirect.com/science/article/pii/S2352914819302047)
3] MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications
Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam
(https://arxiv.org/abs/1704.04861)
4] D. Sinha and M. El-Sharkawy, "Thin MobileNet: An Enhanced MobileNet Architecture," 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 2019, pp. 0280-0285, doi: 10.1109/UEMCON47517.2019.8993089.
5] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," no. November 2015.
6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, vol. 4, pp. 770– 778
7] K. Steiner, M. Binder, M. Schemper, K. Wolff, and H. Pehamberger, "Statistical evaluation of epiluminescence dermoscopy criteria for melanocytic pigmented lesions," Journal of the American Academy of Dermatology, vol. 29, no. 4, pp. 581–588, 1993
8] Xie F., Fan H., Li Y., Jiang Z., Meng R., Bovik A. Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model. *IEEE Trans. Med. Imaging.* 2017;**36**:849–858. doi: 10.1109/TMI.2016.2633551. [PubMed] [CrossRef] [Google Scholar]
9] Aswin R.B., Jaleel J.A., Salim S. Hybrid Genetic Algorithm: Artificial Neural Network Classifier for Skin Cancer Detection; Proceedings of the 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT); Kanyakumari, India. 10–11 July 2014; pp. 1304–1309. [CrossRef] [Google Scholar]