Skin Lesion Classification Analysis: A Comparative Study

PROJECT REPORT

Submitted in fulfillment for the J-Component of Technical Answers
for Real World Problems (TARP)- ITE3999

B. Tech – Information Technology

By

Reshabh Kumar Jain	17BIT0048
Aradhya Mathur	17BIT0146
Nehal Lalia	17BIT0187
Nikhil Singh	17BIT0192
Shubham Jhamb	17BIT0273

Under the guidance of

Dr. ASWANI KUMAR CHERUKURI

SITE

School of Information Technology and Engineering

Winter Semester 2019-20

Abstract

In today's parlance, there are three standard forms in which skin cancer exists namely melanoma, squamous cell carcinoma (SCC) and basal cell carcinoma (BCC) of which melanoma is the most hazardous due to its unpredictable nature. Hence, early detection of lesions which form the base of skin cancer is required to cure it. To begin with, the approach of dermoscopy has empowered a sensational lift in clinical diagnostic capacity to the point that skin cancer can be identified in the clinic at an early stage using either visual dermoscopy or some other artificial intelligence supported tool. The worldwide acceptance of this technology has permitted usage of enormous collection of dermoscopy pictures of lesions and benevolent sores approved by histopathology. The increment of trends in innovations of the regions of image processing and AI has enabled us to permit differentiation of dangerous lesions from the numerous benign mimics that require no biopsy. In this paper, using the HAM 10000 dataset, we compare six different Transfer Learning nets which provide accurate and more classified identification of lesion images into different form of cancer producing cells. VGG19, InceptionV3, InceptionResNetV2, ResNet50, Xception and MobileNet transfer learning techniques are used for comparison of skin lesion. We added layers such as DropOut, Pooling, Batch Normailzation and Dense in the bottom most layer of transfer learning nets. From the experimental results, we identified that MobileNet is giving better accuracy when compared with VGG19, InceptionV3, InceptionResNetV2, ResNet50, Xception nets.

Keywords*: artificial intelligence, CNN, dermoscopy, image classification, Skin lesion, transfer learning,*

1. Introduction

Skin cancer is the abnormal growth of one of the three types of skin cells which constitute the skin. The most common cause of skin cancer is the over exposure to ultraviolet radiations coming from the sun. The rate of being affected by exposure to UV radiations is more in fair skinned, more sun-sensitive people than in dark skinned, less sun-sensitive people.

There are three major types of skin cancers namely basal cell carcinoma (BCC), Squamous cell carcinoma (SCC) and melanoma. But majority of skin cancer identified is of non-melanoma type. With the advent of deep learning concepts, we can classify skin cancer detection in seven diagnostic categories namely melanocyctic nevi, melanoma, benign keratosis-like lesions, basal cell carcinoma, actinic keratosis, vascular lesions and dermatofibroma. Generally, a dermatologist specialized in skin cancer detection follows a fixed sequence i.e. starting with visual examination of suspected lesion with naked eyes, followed by dermoscopy and finally a biopsy.

In today's era, with the usage of artificial intelligence and deep learning in medical diagnostics, the efficiency of predicting a result increases exponentially as compared to the dependency on visual diagnostic. Convolutional neural network (CNN) is an important artificial intelligence tool in feature selection and object classification. Deep convolutional neural networks help in

classifying skin lesions into 6 different categories with the help of their dermoscopic images covering all the lesions found in skin cancer identification. CNN trains a network of large-scale datasets using high performance GPUs, thus providing better outcome. Deep learning algorithms backed by these high performing GPUs in computing large datasets have shown better performance than humans in skin cancer identification.

2. Background

2.1 Literature Survey

In this section we discuss about the literature survey work which was carried out previously. Previous models built upon this dataset provide substantial results and a higher value of ROC (region of convergence) implies that the classifier models seldom fail in classification of cancer and non-cancer. Chaturvedi, gupta and Prasad [3] makes use of MobileNet in skin lesion detection but accuracy achieved is 83% and also 7 different skin lesion are not concentrated. Codella [8] makes use of ISIC 2017 dataset with conventional machine learning methods in order to predict melanoma precisely. However, the accuracy achieved is pretty low. Goyal [4] uses ensemble deep learning in skin lesion dataset but they worked on segmentation rather than detection. Resnet was used in this work and 19,398 images were used from Asan dataset. This paper focused on ROC-AUC score of each class. Milton [5] presented with transfer learning algorithms that were trained on HAM10000 dataset, they used fine tuning and freezing of two epochs which led to decreased accuracy. Here inception-resnetV2 is used which gives the accuracy of 70%. Another case of skin lesion classification in which VGG net was used, managed to attain the 81%. This paper's sole focus is on using VGG and improving the accuracy and a lot of fine tuning is carried out in that work. The author used VGG19 model to classify diabetic retinopathy from fundus images. Here VGG net is used along with PCA and SVD to improve the classification accuracy.

From the literature survey carried out on skin cancer detection, it was found that most papers had done classification of lesions in the three standard categories i.e. basal cell carcinoma, squamous cell carcinoma and melanoma. The dataset used for classification was not so recent and not sufficient enough to identify all types of lesions. By keeping all this, three objectives are framed. That is to classify the images from HAM 10000 dataset into seven different types of skin cancer and to find their individual count and to use CNN for feature selection and classification so as to identify all types of lesions found in skin cancer and also train the data set with six different transfer learning nets and identify the net that gives highest accuracy along with a comparative analysis of all the transfer nets used. In this paper, HAM 10000 dataset is used to train the model for skin cancer classification. We train our dataset using these six transfer learning nets and also plot their training and validation loss, training and validation accuracy along with their individual confusion matrices. Then we shall perform a comparative analysis of accuracy of all these learning nets and conclude with the net which gives highest accuracy in identifying all the lesions. Section 3 describes about the proposed methodology adopted in six different nets. Section 4 describes about experimental results and section 5 describes analysis. Section 6 ends with conclusion.

2.2 PROBLEM STATEMENT

The cases of skin cancer have been increasing at a very high pace over the past few decades. There were around 96000 new cases of invasive lesions that were diagnosed in 2019 in the United States of America, with a projected number of around 7000 cases that would result in death in case of melanoma alone. According to the United States Centers for Disease Control and Prevention, the rate of new melanoma has doubled over the past 3 decades.

The most dangerous characteristic of skin lesions is that it can spread quickly over the body, hence, early diagnosis of them is of great importance for the prognosis of the disease. Devise a solution for fast, and accurate detection of skin cancer helping dermatologists and patients in achieving a higher chance of survival.

2.3 OBJECTIVES

- Accurately train the model to differentiate between all skin cancer pigment images.
- Efficient and fast determination of melanoma skin cancer to help pathologists and patients.

3. PROPOSED MODEL

In our project we will make and compare different type of highly modularized network architecture for image classification like: MobileNet, ResNeXt, DenseNet, Xception, etc. We will compute the results of all the methodologies and finally come up with a best possible solution which would be better than the technologies currently in the market. We train our dataset using these six transfer learning nets and also plot their training and validation loss, training and validation accuracy along with their individual confusion matrices. Then we shall perform a comparative analysis of accuracy of all these learning nets and conclude with the net which gives highest accuracy in identifying all the lesions. Defining the dataset used for skin lesion classification, we use different transfer learning nets used for skin lesion classification and describes the proposed methodology adopted in six different nets.

Initially, the different models are trained for 50 epochs with batch size of 32. In every epoch training accuracy, training error and validation accuracy, validation error are calculated. We adopt Adaptive Momentum (Adam) optimizer with a learning rate (LR) of 0.001 and loss function as Categorical Cross Entropy. In order to make the optimizer converge faster and closer to the global minimum, annealing method was used with a LR. To keep the advantage of the faster computation time with a high LR, we decreased the LR dynamically every 3 epochs depending on the validation accuracy. With the 'ReduceLROnPlateau' function from 'Keras.callbacks', we choose to reduce the LR by half if the accuracy is not improved after 3 epochs.

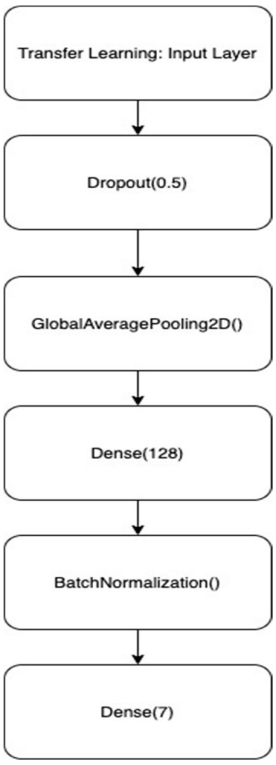


Fig. Model Architecture

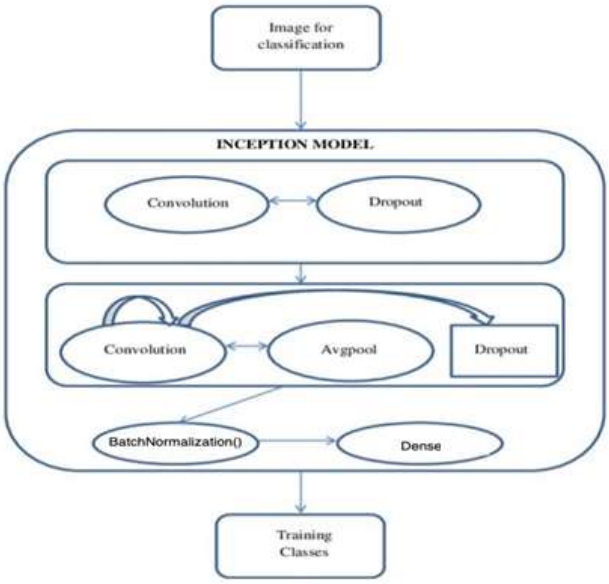


Fig. Architecture of InceptionV3

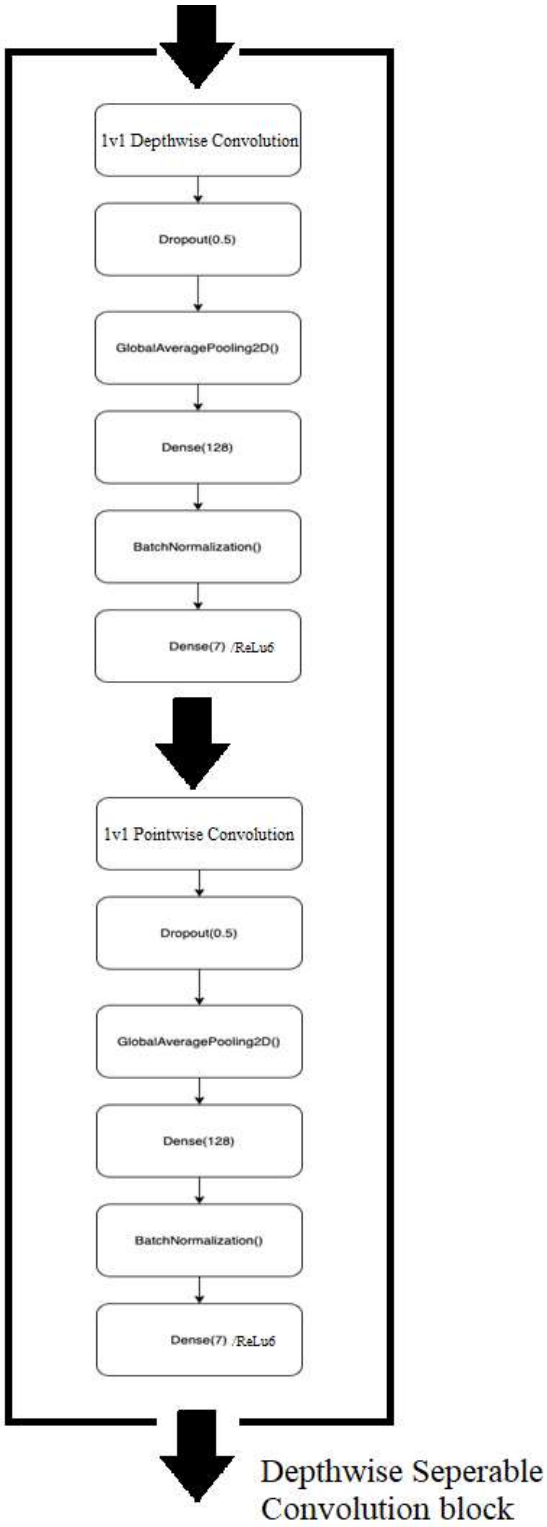


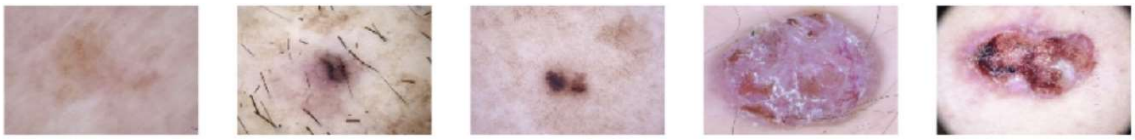
Fig. Architecture of MobileNet

(Other models also follow almost similar)

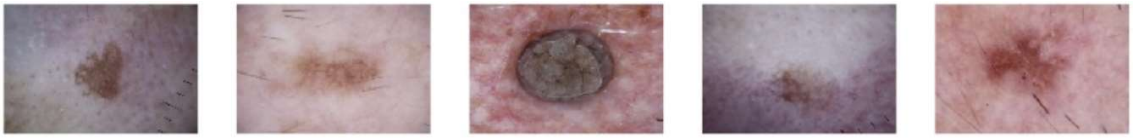
Dataset Description for skin lesion

To carry out the project work, we use HAM 10000 dataset (Human against machine) which has 10015 dermatoscopic images and 7 different classes such as actinic keratosis(327), basal cell carcinoma (541), benign keratosis (1099), dermatofibroma (155), melanocytic nevi (6705), melanoma (1113) and vascular skin lesions (142) . Seven types of lesion are shown in the figure below along with their occurrences later in which the x-axis represents the type of lesion and y-axis represents the corresponding count. This dataset is used for training of models.

a. Actinic keratoses



b. Basal cell carcinoma

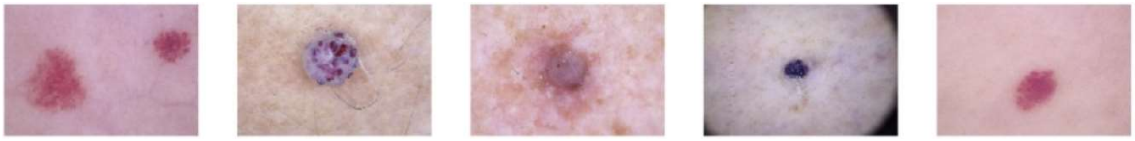


c. Benign keratoses-like lesions

d. Dermatofibroma

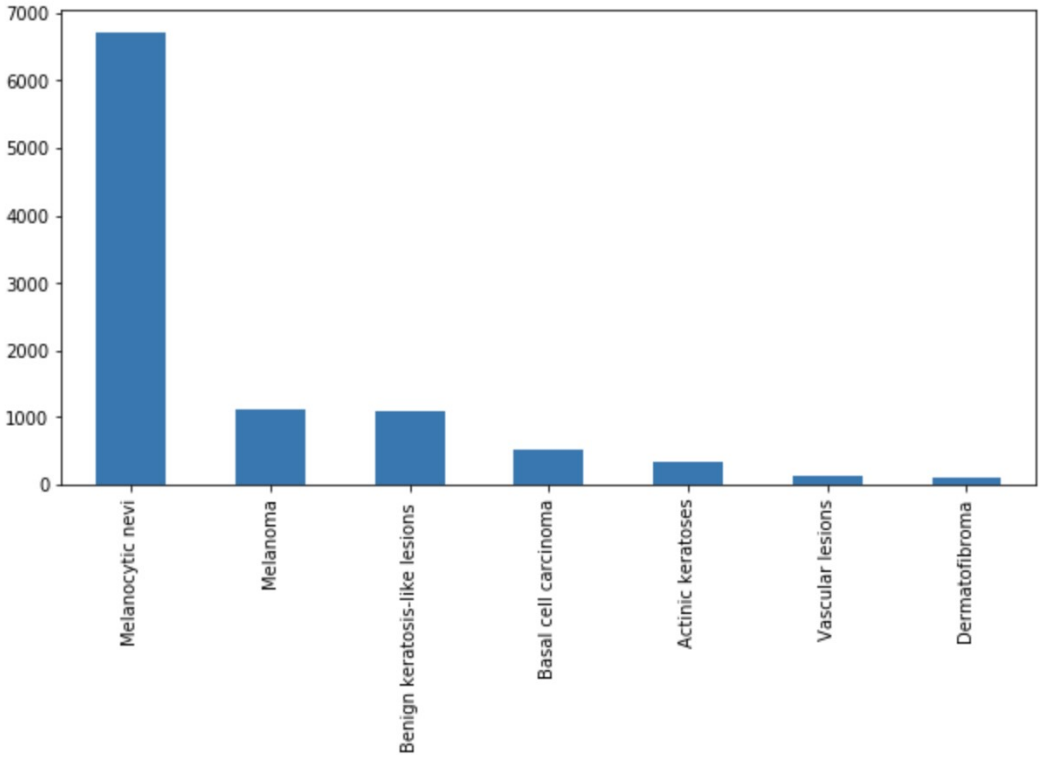
e. Melanocytic nevi

f. Melanoma



g. Vascular lesions

Seven different types of diseases caused from lesions (top to bottom): Actinic keratosis (akiec), Basal cell carcinoma (bcc), Benign keratosis (bkl), Dermatofibroma (df), Melanocytic nevus (nv), Melanoma (mel) and Vascular (vasc).



Occurrence of images of each type of skin cancer

What is the innovation & novelty?

- The older deep learning approach basically uses two different networks to individually perform lesion segmentation and classification, but we are using multi-scale fully-convolutional residual networks.
- Using ADAM optimizer to reduce memory usage and improve computational efficiency
- Using the general features as a starting point to evolve new features, we also obtained better classification accuracy which helped with the discrimination between malignant and benign skin lesions.
- To keep the advantage of the faster computation time with a high LR, we decreased the LR dynamically every 3 epochs depending on the validation accuracy. With the 'ReduceLROnPlateau', we choose to reduce the LR by half if the accuracy is not improved after 3 epochs.

- We decreased LR for every 3 epochs to achieve a global minimum which isn't possible if we reduce LR with every epoch. We didn't use 5 epochs since it was too time taking.

Proposed Methodology

Algorithms (Transfer Learning Nets used for comparison)

In this section, we focus on six different transfer learning nets such as VGG19, InceptionV3, InceptionResNetV2, Resnet50, Xception and MobileNet. We train skin lesion dataset using these six transfer learning nets and analyze the prediction accuracy. In addition to this, we also plot their training and validation loss, training and validation accuracy along with their individual confusion matrices. Then we perform a comparative analysis of accuracy of all these learning nets and conclude with the net which gives highest accuracy in identifying all the lesions.

1 VGG19: This network is characterized by its simplicity. It has 5 blocks each of 3×3 convolutional layers stacked on top of each other. Volume size is reduced by max pooling of 2x2 kernels and stride of 2. Followed by two fully-connected layers, each of 4096 nodes with ReLU activation function. And the final layer has 1000 nodes with softmax as its activation function . VGG19 has about 143 million parameters in total.

2 InceptionV3: InceptionV3 is the refined version of the GoogLeNet architecture. The basic idea of this net is to make this process simpler and more efficient. Inception module acts as a multi-level feature extractor. It computes 1×1, 3×3, and 5×5 convolutions within the same module of the network. The output of these filters are then stacked on each other and then fed into the next layer in the network.

3 InceptionV2Resnet: In this net, residual version of Inception nets are used rather than simple inception modules. As discussed in the inceptionV3 architecture, each inception block is used for scaling up the dimensionality. Inception-ResNet v2 matches the raw cost of the advanced Inception V4. Inception V4. In Inception-Resnet, we use batch normalization in the earlier layers not on summation layers

4 ResNet50: These are the deeper convolutional neural nets which make use of skip connections. These residual blocks greatly resolve gradient degradation and also reduce total parameters. Residual Networks (ResNet) architecture follows two simple design rules Firstly, for the same output map size, layers have same number of filters and secondly, when the feature map size is halved, filters count is doubled. Batch normalization is performed after each convolution layer and before ReLU activation function. If the input and output have same size then shortcut is used. When the dimensions increase, the projection shortcut is used. Final layer has 1000 nodes with softmax activation with 50 layers.

5 Xception: The Xception architecture is an extension of the Inception architecture. It replaces the standard Inception modules with depthwise separable convolutions. It does not perform partitioning on input data and maps the spatial correlations for each output channel separately. Xception net then performs 1x1 depthwise convolution which captures cross-channel correlation. It slightly outperforms Inception V3 in terms of smaller data and vastly on bigger data.

Original depthwise separable convolution has a depthwise convolution first followed by a pointwise convolution, while in the modified version it has pointwise convolution first followed by a depthwise convolution. The model motivation is taken from the Inception module, but the inception model has non-linearity after the first operation while in the modified version there is no such intermediate ReLU non-linearity.

6 MobileNet: This net makes use of Depthwise separable connections, similar to Xception net. For MobileNets, the depthwise convolution applies a single filter to each input channel. The pointwise convolution then applies a 1×1 convolution to combine the outputs the depthwise convolution. A standard convolution both filters and combines inputs into a new set of outputs in one step. The depthwise separable convolution splits this into two layers, a separate layer for filtering and a separate layer for combining. This factorization has the effect of drastically reducing computation and model size. MobileNet is particularly useful for mobile and embedded vision applications. It has fewer number of parameters compared to others and also lesser complexity. This architecture is a concise form of Xception and Inception net.

Methods

In this section, we propose a method to classify the images for skin lesions. We divide the 10015 dataset into training, testing and validation dataset with 70% for training, 10% for validation and 20% for testing. 7 types of skin cancer diseases are segregated and their occurrences are calculated. All images are resized from 450 x 600 to 100 x 150 dimensions for convenience in processing.

1 Data Augmentation: This technique is used to avoid over-fitting problems. Here, we expand the training dataset artificially. We alter the training data with small transformations to reproduce variations. Few techniques like rotation, zooming, shifting vertically and horizontally are implemented. Approximately 10% of the data was used for the above purpose. All the images are a part of HAM 10000 dataset.

2 Preprocessing: After the image acquisition task, we perform image pre-processing. Image pre-processing involves normalization of the images. In normalization, mean and standard deviation of all images in the dataset is calculated. The mean of all the images is subtracted from initial images and then the obtained result is divided by standard deviation. On the other hand, the 7 diseases are one hot encoded. The image has been resized to 150,100.

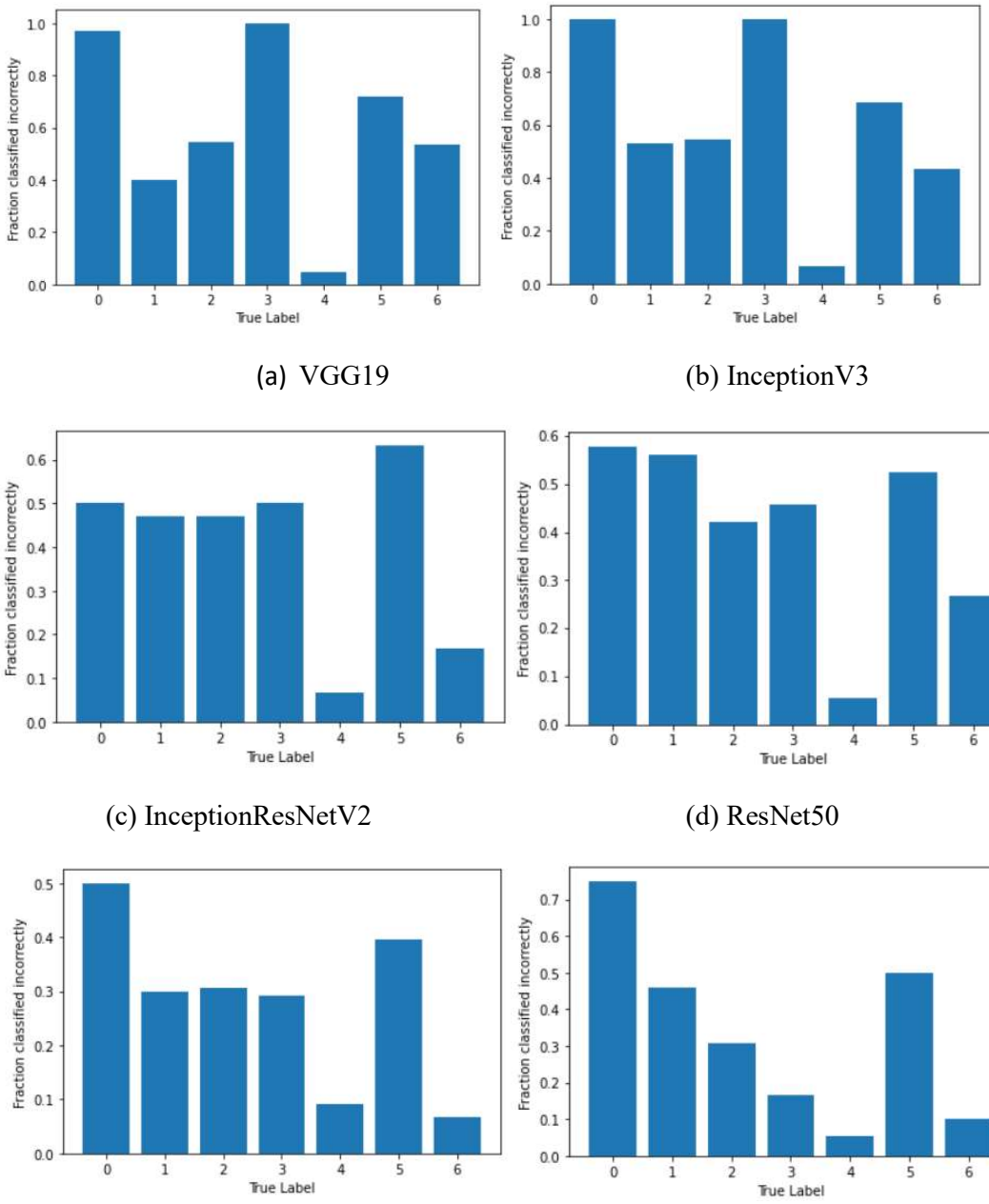
3 Feature Extraction: Feature extraction is the most crucial step in classification. Feature extraction was carried out by pre-trained transfer learning models. This involves looking up for important features in an image and then deriving information from them. Several CNNs are stacked up back to back in order to make a model. Here we use pre-trained models such as VGG19, InceptionV3, Resnet50, Xception, InceptionResNetV2 and MobileNet. we train the top layers and adjust the bottom layers with few more layers, so that it can work for our dataset.

All of the above pre-trained nets used the weights of the Imagenet. The bottom layers, as shown in the figure below, are DropOut layer, GlobalAveragePooling2D, Dense layer, Batch Normailzation layer and dense layer. Dropout layer has a value of 0.5 which helps in reducing computation complexity and also helps to avoid overfitting. Followed by GlobalAveragePooling2D which calculates the average output of each feature map in the previous layer and also prepares for the final layer. Then a dense layer is used which maps the previous layer to 128 output features applying ReLU activation. Followed by a Batch Normalization layer for improving the performance of image classification. And finally, a dense layer with 7 outputs with sigmoid function as its activation function.

4 Classification and Evaluation: The final layer outputs a vector which indicates the class number. The class number is in the correspondence to 7 different skin cancers. The class numbers assigned for different lesions are actinic keratosis (0), basal cell carcinoma (1), benign keratosis like lesions (2), dermatofibroma (3). melanocytic nevi (4), melanoma (5), vascular skin lesions (6) and in the evaluation phase, we use validation dataset for validating the different nets for skin lesion dataset. This phase keeps check on the training of the system.

4. Experimental Results

In this section, we present the experimental results and analysis of our models used on HAM10000 dataset. We applied six different types of transfer learning models- VGG19, InceptionV3, Resnet50, Xception, InceptionResNetV2 and MobileNet to carry out the experiments. Initially, the different models are trained for 50 epochs with batch size of 32. In every epoch training accuracy, training error and validation accuracy, validation error are calculated. We adopt Adaptive Momentum (Adam) optimizer with a learning rate (LR) of 0.001 and loss function as Categorical Cross Entropy. In order to make the optimizer converge faster and closer to the global minimum, annealing method was used with a LR. To keep the advantage of the faster computation time with a high LR, we decreased the LR dynamically every 3 epochs depending on the validation accuracy. With the 'ReduceLROnPlateau' function from 'Keras.callbacks', we choose to reduce the LR by half if the accuracy is not improved after 3 epochs.



(a) VGG19 (b) InceptionV3

(c) InceptionResNetV2 (d) ResNet50

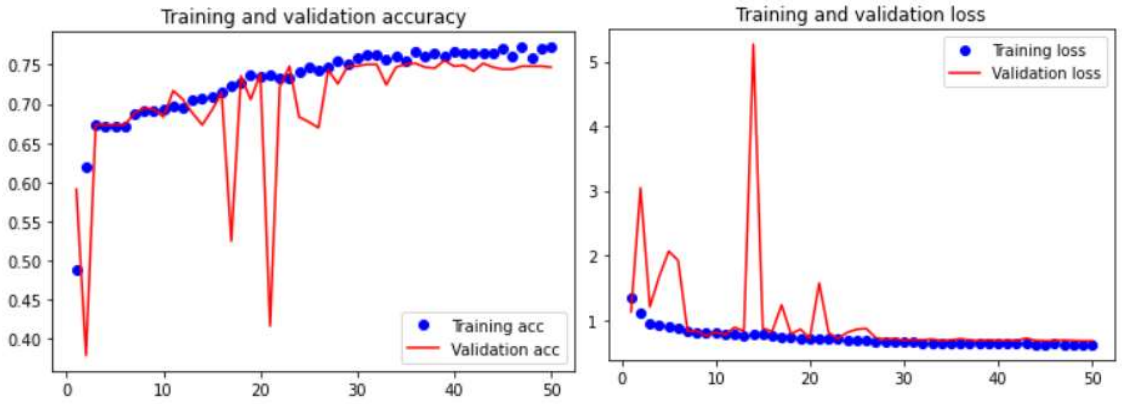
(e) Xception (f) MobileNet

Fig.4. Fraction classified incorrectly for all six models

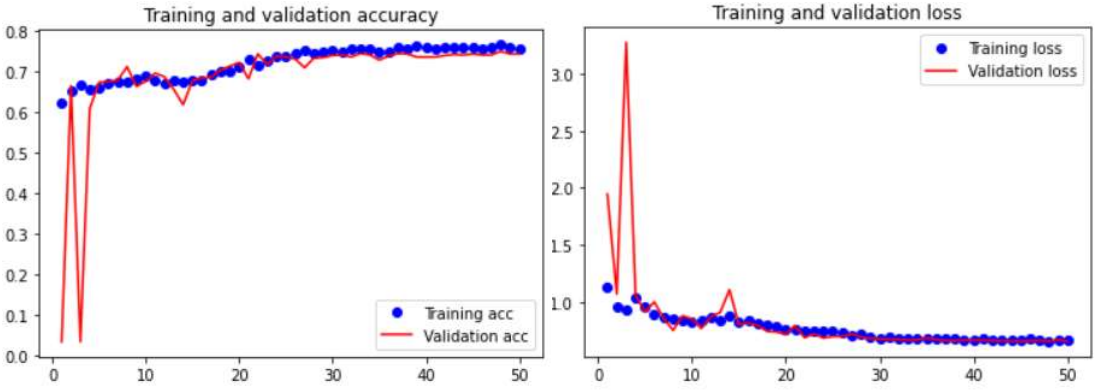
Figure 4 shows the incorrectly classified results for different nets. Incorrectly classified graphs in Fig 4 are derived from the confusion matrices. They can be calculated as percentage of incorrect classification out of the total number of images in each class. We can follow the same process for correctly classified graphs. Figure 4a to 4f shows the fraction that has been classified incorrectly by the different transfer learning nets. From sub-figure (a), we can infer that VGG19 provides the highest incorrect classification for 'Df' thereby making it as the most

difficult lesion to be identified by the system. Also 'Nv' having the least incorrect classification, is most likely to be identified by the system. Similarly, sub-figure (b) represents the incorrect classification by InceptionV3 net where the highest incorrect classification was for 'Df' which makes it the most difficult lesion to be identified by the system and 'Vasc' having the least incorrect classification, is most likely to be identified by the system. Sub-figure (c) represents the incorrect classification by InceptionResNetV2 where the highest incorrect classification was for 'Akiec' which makes it the most difficult lesion to be identified by the system and 'Df' having the least incorrect classification, is most likely to be identified by the system. Sub-figure (d) represents the incorrect classification by ResNet50 where the highest incorrect classification was for 'Akiec' which makes it the most difficult lesion to be identified by the system and 'Df' and 'Vasc' having the least incorrect classification, is most likely to be identified by the system. Sub-figure (e) represents the incorrect classification by Xception where the highest incorrect classification was for 'Akiec' which makes it the most difficult lesion to be identified by the system and 'Df' and 'Vasc' having the least incorrect classification, is most likely to be identified by the system. Sub-figure (f) represents the incorrect classification by MobileNet where the highest incorrect classification was for 'Akiec' which makes it the most difficult lesion to be identified by the system and 'Df' and 'Vasc' having the least incorrect classification, is most likely to be identified by the system. Overall we may infer that 'Akiec' is the most difficult lesion to be identified and 'Vasc' is the most easily identifiable lesion.

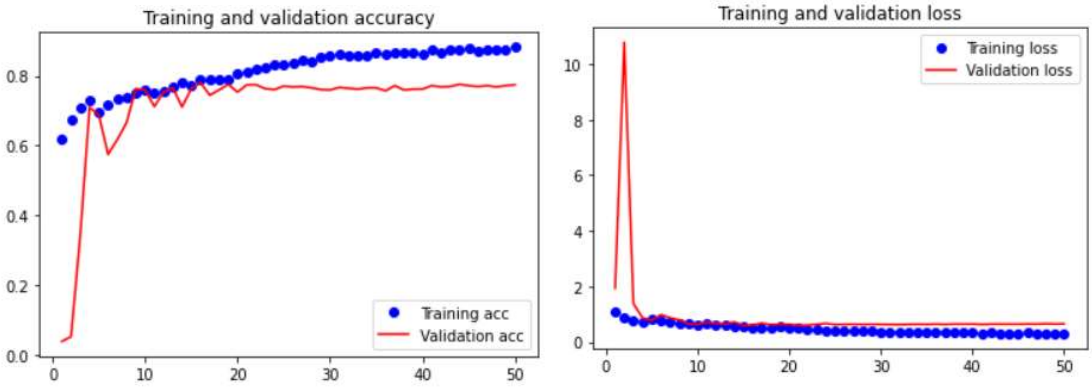
Figures 5a to figure 5f shows the training and validation accuracy and loss for six different nets. Figure 5a shows that, **VGG19 shows the validation loss of 0.67 and accuracy of 74.6%, on skin lesion classification. Figure 5b shows that InceptionV3 has 0.65 and 74.3%, ResNetV2Inception of 0.65 and 77.3%, Resnet50 of 0.63 and 79.4%, Xception of 0.66 and 81.2% and finally MobileNet of 0.76 and 83.1%.**



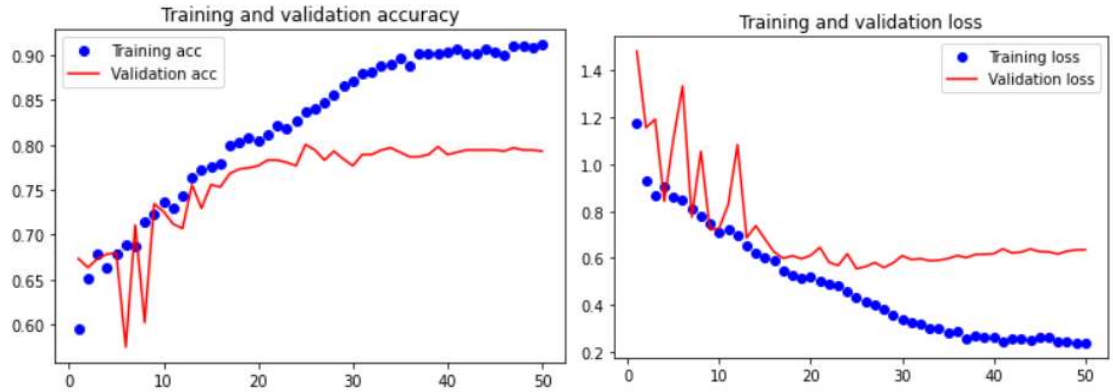
a) VGG19



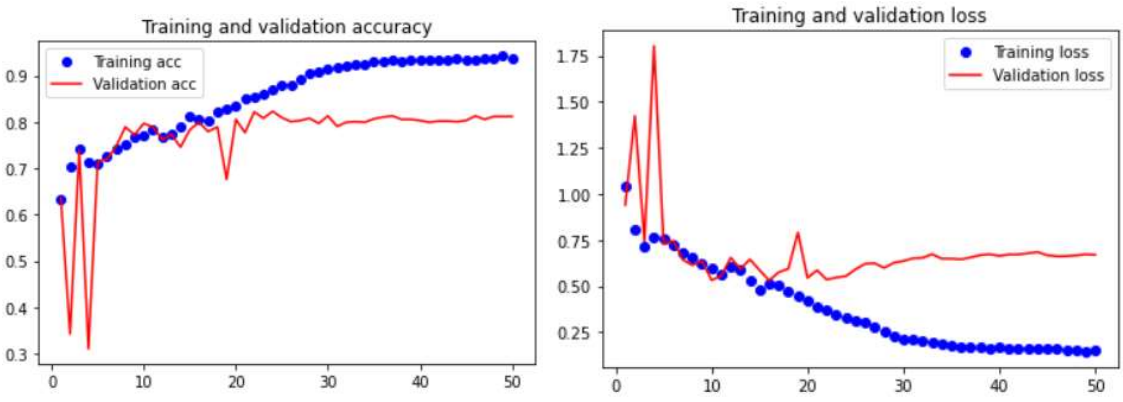
b) InceptionV3



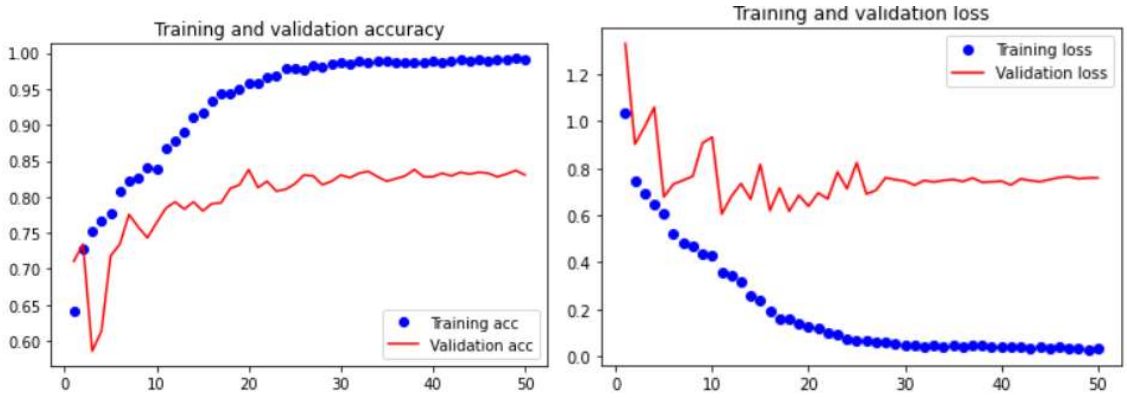
c) Inception ResNetV2



d) Resnet50



e) Xception



f) MobileNet

Figure 5 Training, Validation accuracy and loss results
(Here on x-axis is number of epochs and on y-axis is accuracy)

5. ANALYSIS

Comparative Analysis of Different Transfer Learning Nets

Here we compare the performance of all six nets by using confusion matrix. Confusion matrix is constructed for every network. Performance of different nets is tested by passing validation data which is 10% of the total data. Accuracy is considered as a measure for calculating the performance with skin lesion dataset. Figure 6a to figure 6f shows the confusion matrix results we achieved for different nets.

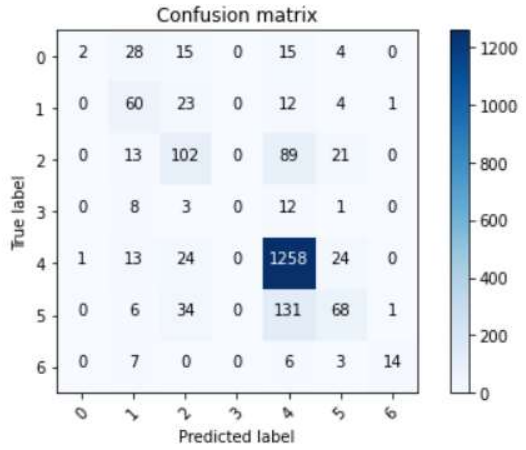
From Fig.6., we can infer true positive (TP), true negative (TN), false positive (FP) and false negative (FN) values for each class. These values can help in calculating Precision, Recall, F1 Score and accuracy.

Precision: TP/(TP+FP)

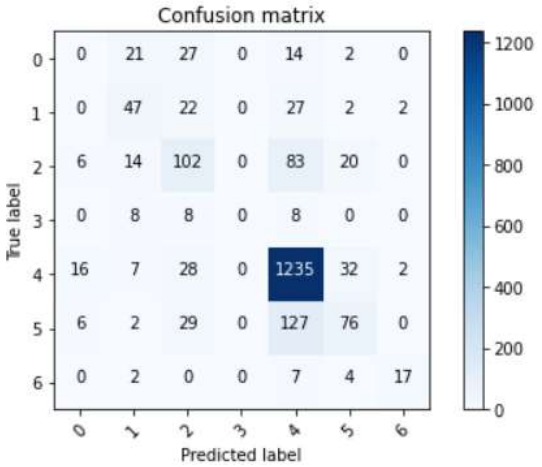
Recall: TP/(TP+FN)

F1 Score: (2*Precision*Recall)/(Precision+Recall)

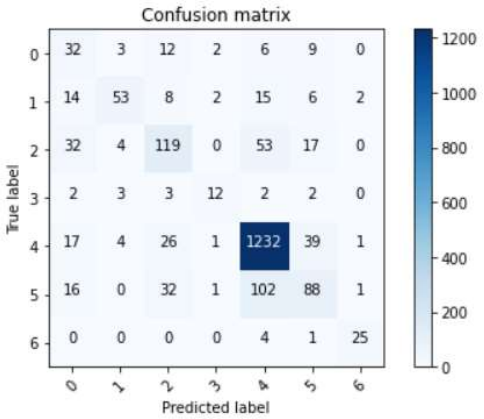
Accuracy = TP + TN/(TP+TN+FP+FN)



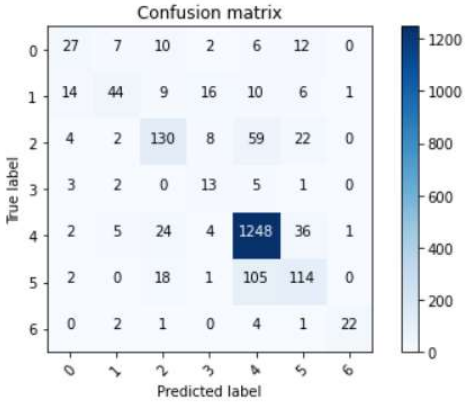
a) VGG19



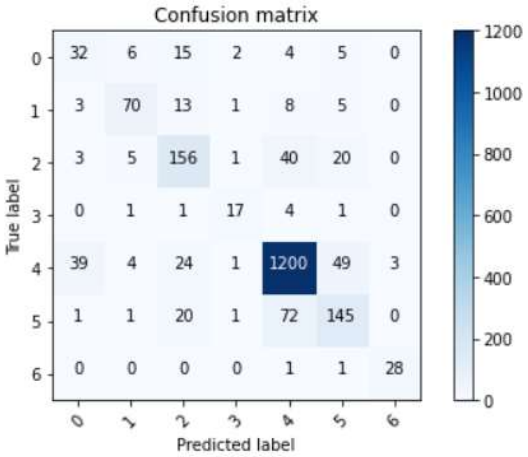
b)InceptionV3



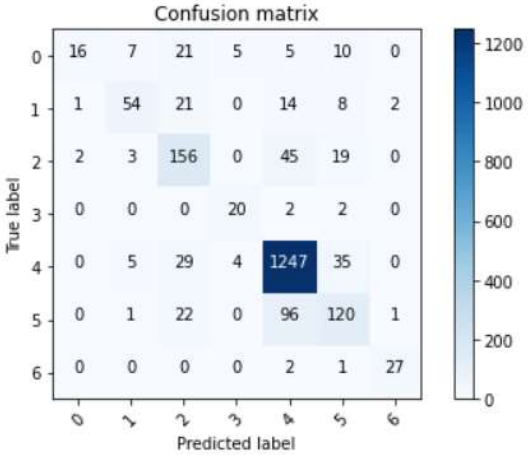
c) InceptionRestNetV2



d)RestNet50



d) Xception



e) MobileNet

Figure 6 Confusion Matrix results of different Nets.

Table 1. Accuracy and loss values of all learning nets used

Transfer Learning Nets	Accuracy	Loss
VGG19	74.6	0.67
InceptionV3	74.3	0.65
InceptionResNetV2	77.3	0.65
Resnet50	79.4	0.63
Xception	81.2	0.66
MobileNet	83.1	0.76

From table 1, we infer the accuracy of the models as 74.6, 74.3, 77.3, 79.4, 81.2 and 83.1 and it is clear that MobileNet showed maximum accuracy of 83.1%.

Table 2. Confusion matrix of MobileNet with highest accuracy of 83.1

Metrics	Akiec	Bcc	Bkl	Df	Nv	Mel	Vasc
Akiec	16	7	21	5	5	10	0
Bcc	1	54	21	0	14	8	2
Bkl	2	3	156	0	45	19	0
Df	0	0	0	20	2	2	0
Nv	0	5	29	4	1247	35	0
Mel	0	1	22	0	96	120	1
Vasc	0	0	0	0	2	1	27

For the testing phase, images were passed through the net and the corresponding confusion matrix of the most accurate model- MobileNet was plotted (as shown in Table 2). Each element

in the confusion matrix depicts the comparison between predicted label and true label. The diagonals represent the predicted labels correctly predicted. This matrix showed that Melanocytic nevi gives best result of 1247 correct prediction out of 1321 and Actinic keratosis showed worst results. Table 3 shows the incorrect and correct predictions of the testing data based on the confusion matrix in Table 2.

Table 3. Test result of different lesions upon training with MobileNet

Disease	Data	Truly classified	False classified
Akiec	64	16	48
Bcc	100	54	46
Bkl	225	156	69
Df	24	20	4
Nv	1321	1247	74
Mel	240	120	120
Vasc	30	27	3

This is a sign of overfitting: Train loss is going down, but validation loss is rising.

Fig. Running model without dropout layer

(which resulted in overfitting.)

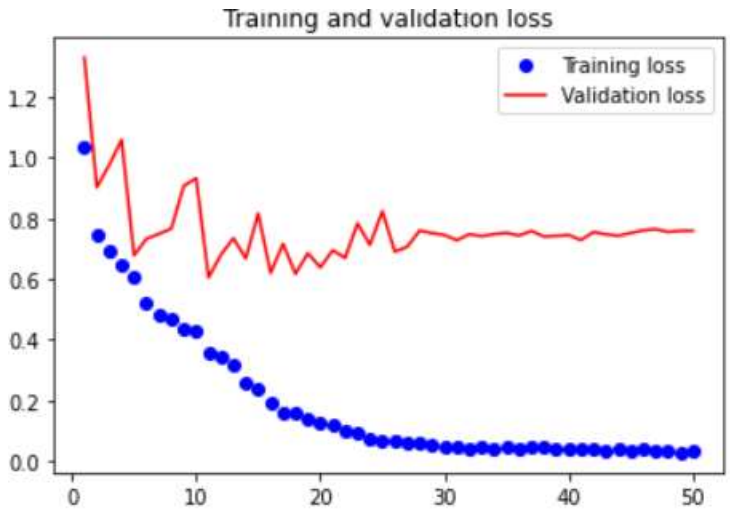


Fig. Running model with dropout layer

(we can clearly see that there is no overfitting)

Comparison with earlier works:

1. The impact of patient clinical information on automated skin cancer detection (2019)

Model used -

Mobile Net - 69.1%

Resnet50 - 69.1%

Our Accuracy: 83.1%

2. Transfer Learning in Medical Image Classification (2020)

Model used -

Mobile Net v2 - 71.3%

Ours Accuracy: 83.1%

3. Vision-Based Classification of Skin Cancer using Deep Learning

Model used -

VGG-16 CNN - 78%

Ours Accuracy: 83.1%

4. Deep Learning for Skin Lesion Classification: Augment, Train, and Ensemble

Model used -

Mobile Net - 82%

Ours Accuracy: 83.1%

6. Conclusion

We demonstrated that by the use of transfer learning algorithms, we can achieve a competitive classification performance. Though our dataset was unbalanced with classes 3 and 6 having very less images and whereas 4 having the greatest number of images, the performance turned out to be 83.1% accurate with MobileNet. We also applied different transfer learning models with their accuracy ranging from 74.3% to 83.1% and out of which MobileNet provided promising results. Transfer learning algorithms different from the ones used in the paper with proper fine tuning and addition of few more layers may result in better performance.

7. 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] Nugroho, A. A., Slamet, I., & Sugiyanto. (2019, December). Skins cancer identification system of HAMl0000 skin cancer dataset using convolutional neural network. In *AIP Conference Proceedings* (Vol. 2202, No. 1, p. 020039). AIP Publishing LLC.

[3] Chaturvedi, S. S., Gupta, K., & Prasad, P. (2019). Skin lesion analyser: An efficient seven-way multi-class skin cancer classification using MobileNet. arXiv preprint arXiv:1907.03220.t.

[4] Goyal, M., Oakley, A., Bansal, P., Dancey, D., & Yap, M. H. (2019). Skin Lesion Segmentation in Dermoscopic Images with Ensemble Deep Learning Methods.IEEE Access.

[5] Milton, M. A. A. (2019). Automated skin lesion classification using ensemble of deep neural networks in isic 2018: Skin lesion analysis towards melanoma detection challenge. *arXiv preprint arXiv:1901.10802*.

[6] Mateen, M., Wen, J., Song, S., & Huang, Z. (2019). Fundus image classification using VGG-19 architecture with PCA and SVD. *Symmetry,* 11(1), 1.

[7] Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., ... & Uhlmann, L. (2018). Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Annals of Oncology*, *29*(8), 1836-1842.

[8] Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... & Halpern, A. (2018, April). Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 168-172). IEEE.

[9] Han, S.S., Kim, M.S., Lim, W., Park, G.H., Park, I., Chang, S.E.: Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. J Invest Dermatol. 138(7), 1529–1538 (2018)

[10] Li, Y., & Shen, L. (2018). Skin lesion analysis towards melanoma detection using deep learning network. Sensors, 18(2), 556.

[11] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, *5*, 180161.

[12] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1251-1258).

[13] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.

[14] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118.

[15] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[16] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

[17] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[18] Hoffman, R., Benz Jr, E. J., Silberstein, L. E., Heslop, H., Anastasi, J., & Weitz, J. (2013). *Hematology: basic principles and practice*. Elsevier Health Sciences

[19] Armstrong, B. K., & Kricker, A. (2001). The epidemiology of UV induced skin cancer. *Journal of photochemistry and photobiology B: Biology*, *63*(1-3), 8-18.

[20] Marks, R. (1995). The epidemiology of non-melanoma skin cancer: who, why and what can we do about it. *The Journal of dermatology*, *22*(11), 853-857.

[21] Kricker, A., Armstrong, B. K., & English, D. R. (1994). Sun exposure and non-melanocytic skin cancer. *Cancer Causes & Control*, *5*(4), 367-392.