Seven Class Classification of Skin Lesions by using Incremental Convolutional Neural Network in Python

Ankit Prashant Chopade*
Shri Guru Gobind Singhji Institute of
Engineering and Technology, Nanded
2015bcs060@sggs.ac.in

***Abstract*—Classification of skin lesions in various cancerous type plays a crucial role in diagnosing various, local and gene related, medical conditions in the field of medical science. Classification of these lesions in several cancerous types i.e Melanoma(MEL), Melanomic Neves(NV), Basal Cell Carcinoma(BCC), Actinic Keratosis(AKIEC), Benign Keratosis(BKL), Dermatofibroma(DF) and Vascular Lesion(VASC) gives some insight about disease. Skin cancer are most lethal type of cancer but if these disease are identified in early stages then patients can have high frequency of recovery. Several approaches to automatic classification have been explored by many authors, using various techniques and approaches but this paper proposed novel Incremental approach for Convolution Neural Network on dermoscopy images for classification of skin lesions in various skin cancers. This is generalized approach hence can be implemented in various algorithms for achieving higher accuracy. International Skin Imaging Collaboration (ISIC) 2018 challenge dataset is used in in this paper. The procedure used in this paper yields an accuracy of 90.26%.**

***Index Terms*—Dermoscopy image, Incremental Convolutional Neural Network(ICNN), Lesion classification, Skin lesions**

I. INTRODUCTION

Skin disease is the uncontrolled development of anomalous skin cells. Skin diseases like basal cell carcinoma, and squamous cell carcinoma, melanoma - regularly start to change your skin. Skin malignancy can be relieved if it's found and treated early. Skin malignancy essentially grows on sun uncovered places like legs, arms, etc. Skin malignant growth influences individuals of all skin tones, incorporating those with darker compositions. At the point when melanoma happens in individuals with dim skin tones, it's bound to happen in zones not regularly presented to the sun, for example, the palms of the hands and bottoms of the feet.

As indicated by Indian Cancer Society(ICS), it has been announced that the skin malignant growth rates in India was higher when contrasted with different nations, for example, Canada, the US and the UK. It has been accounted for that about 125,693 new malignant growth cases are spotted yet it was higher than 45,395 individuals are foreseen to death from disease. Out of numerous malignant patients just some got treatment.

The rest of the paper is organized as follows. Section 2 of paper constitute of Description of Dataset, Section 3 contains Convolutional Neural Network Fundamentals, Section 4 is Incremental learning approach for CNN, Section 5 is Architecture of CNN proposed model, Section 6 is Experimental Results and Section 7 is Conclusion.

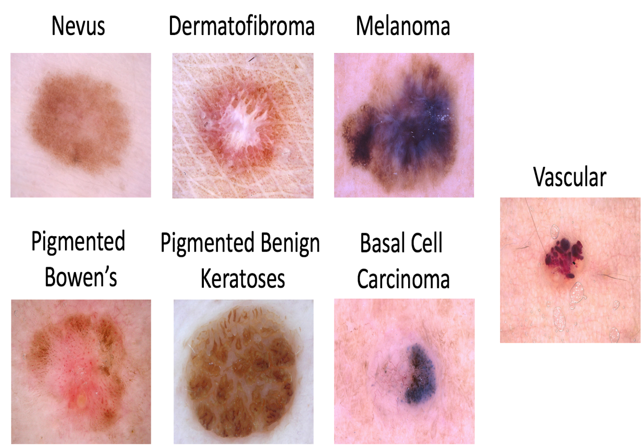

Fig. 1. Different types of skin lesions. [7]

II. DATASET DESCRIPTION

[7]The information is in dermoscopic sore pictures in JPEG design. All sore pictures are named utilizing the plan ISIC.jpg, where is a 7-digit one of a kind identifier. EXIF labels in the pictures have been expelled; any remaining EXIF labels ought not be depended upon to give precise metadata. The sore pictures originate from the HAM10000 Dataset, and were procured with an assortment of dermatoscope types, from every single anatomic site(barring mucosa and nails), from a recorded example of patients displayed for skin malignancy screening, from a few unique establishments. Pictures were gathered with endorsement of the Ethics Review Committee of University of Queensland(Protocol-No. 2017001223) and Medical University of Vienna(Protocol-No. 1804/2017).

The Ground Truth is single CSV record(comma-isolated esteem) document of binary classifications for every one of the 7 illness states, demonstrating the conclusion of each info injury picture.

1) MEL: Melanoma diagnosis confidence
2) NV: Melanocytic nevus diagnosis confidence
3) BCC: Basal cell carcinoma diagnosis confidence
4) AKIEC: Actinic keratosis / Bowens disease(intraepithelial carcinoma) diagnosis confidence

5) BKL: Benign keratosis(solar lentigo / seborrheic keratosis / lichen planus-like keratosis) diagnosis confidence
6) DF: Dermatofibroma diagnosis confidence
7) VASC: Vascular lesion diagnosis confidence

III. Convolutional Neural Network-Fundamentals

[1], [2], [3] A Convolutional Neural Network(CNN) is included at least one convolutional layers(frequently with a sub-sampling step) and afterward pursued by at least one fully connected layers as in a standard multilayer neural system. The engineering of a CNN is intended to exploit the 2D structure of an info picture(or other 2D information, for example, a discourse flag). Another advantage of CNN 's is that they are simpler to prepare and have numerous less parameters than completely associated systems with a similar number of concealed units.

A. Convolutional Layer

Each neuron in the first layer is associated to a region of the input neurons. The region could be big or small. That region is called as local receptive field for that neuron. Then we slide the local receptive field over by stride length to the right to connect to a second neuron, and goes doing, building up the first layer. Here same weights and bias for each of the neurons is used. It means each neuron has shared weights an biases. So it can be infer that all the neurons in the first layer detect the same feature. But at different places in the images. For example the weights and bias are adjusted such that they can detect edge in its field. This weights and biases can be used at other places in image to detect edges. An advantage is it drastically reduces the total number of parameters. It means it reduces the processing time and use of resources.[10]

$$(t) = (x \times w)(t) \tag{1}$$

B. Pooling Layer

CNNs also contain pooling layers. These layers usually used immediately after convolutional layers but not necessary.A pooling layer takes each feature map as input and outputs the reduced feature map. Condensed feature map means the dimension of convolutional layer is reduced. As a example max-pooling, avg-pooling, etc. For instance, if the input layer is a NN layer then pooling layer yields output as $N/k \times N/k$ layer, here $k \times k$ is size of pooling layer.

C. Normalization Layer

We usually normalize the layers by adjusting and scaling the activations i.e when we have features from 0 to 1 and some from 1 to 100, we should normalize them between some range say 0 to 1 to accelerate training. This technique drastically increases the training speed because this technique eliminates many ouliers. Batch normalization provides some liberty to layers i.e by allowing each layer of to learn by itself a little bit more independently of other layers.

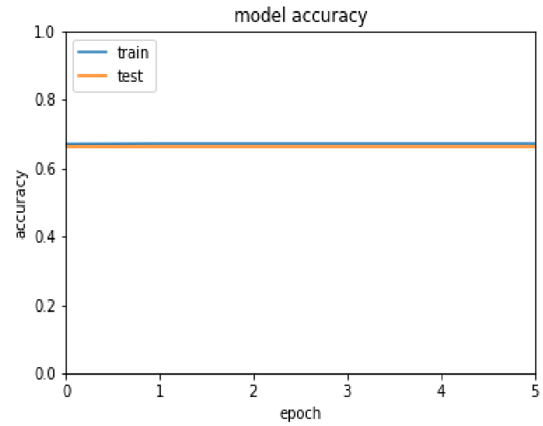

Fig. 2. Accuracy variation in CNN model without using the Incremental approach

D. Dropout Layers

The dropouts are specifically used for preventing the overfitting problem. In this technique neurons are randomly removed during the training. Parameter p is used to controlled the dropouts.It means it determines removing the neuron probability. These layers are used during training only.

E. Fully Connected Layers

The final layer in the CNN network is a fully- connected layer. Pooling layer is connected to every output layer by fully connected layer. This fully-connected architecture is same as in multilayer perceptrons.

IV. Incremental learning approach for Convolutional Neural Network

We proposed this technique for complex, biased, etc. dataset. i.e images in dataset are differ by less probability distribution(less variations). If this type of condition persists in dataset then simple CNN can't able to detect it. If we tried to use simple CNN model then this model cant able to generalize the the result properly and efficiently. But in Incremental training approach CNN can able to classify biased, complex dataset properly and efficiently.

A. Without using Incremental approach

In this traditional approach we fed all dataset to model at once any to increase the accuracy of model. But this will work well with only simpler dataset conversely in complex dataset like ISIC 2018 will reduce the loss but the accuracy will remain constant throughout the training. The main reason for this is the batch size provided during the training is completely biased. And so CNN failed to classify them properly [3], [4].

The main disadvantage of this method is we can't achieve high accuracy though the loss is very low. If we increase number of training cycles then also the accuracy remain same. So for achieving high accuracy Incremental approach is suggested for complex datasets. From Figure 2 and Figure 3 we can infer that loss is decreasing but accuracy is constant

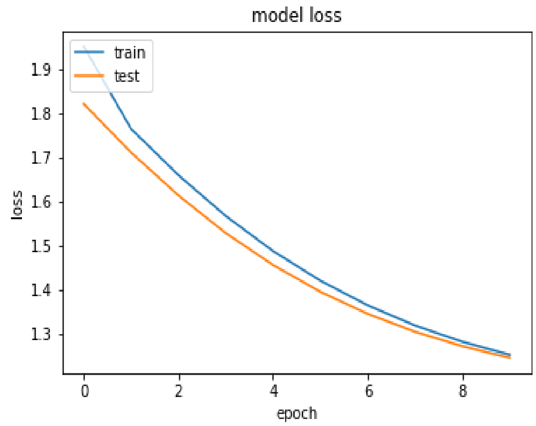

Fig. 3. Loss variation in CNN model without using the Incremental approach

throughout the training. This can be eliminated in Incremental approach.

B. With Incremental approach

Algorithm 1: Incremental CNN Learning

Input: feature set α,
$L_i \in \alpha$, $L_i \in$ [feature]^{n*d} $\forall$i=0,...6,
$\xi \in variation\ function, \gamma \in variation\ value.$
Output: Weight Matrix $W^{\eta * \kappa}$.
$\tau \in \phi, \mu \in \phi$.
while $\alpha \neq \phi$ **do**
  if $L_i = \phi$ **then**
    return W;
  else
    $\tau \subset L_i \ldots \xi(\tau) \geq \gamma$;
    $\mu \in \mu + \tau$;
    W=$Train\_CNN(\mu)$;
    $\tau = \phi$;

Steps for implementing Incremental approach for CNN:

1) We segregate the dataset into its respective class.
2) By observation we select random number of images of all classes which showing large variation i.e they are easily classified. The intuition behind choosing high variant images is due to high variation in dataset CNN can classify them easily.
3) After selecting start the training of model
4) Where to stop training - If we got peak point i.e the condition where variation in accuracy and loss is negligible then this condition is called as saturation point and here we have to stop training.
5) Now modify dataset by adding some more images of all class arbitrarily(do not remove the previous images just keep adding new images from dataset with each increment) which having less variation than previous images. And go to step 3.
6) Continue this procedure until total dataset get exhausted.
7) As we move closer to more complex dataset consider small increment because as complexity increases loss increases accuracy decreases so by using small batch size we led to proper training of model.

TABLE I
COMPARISON TABLE OF SIMPLE AND INCREMENTAL CNN.

Approach	Accuracy	Loss
Simple CNN	64.6%	1.24
Incremental CNN	90.26%	0.35

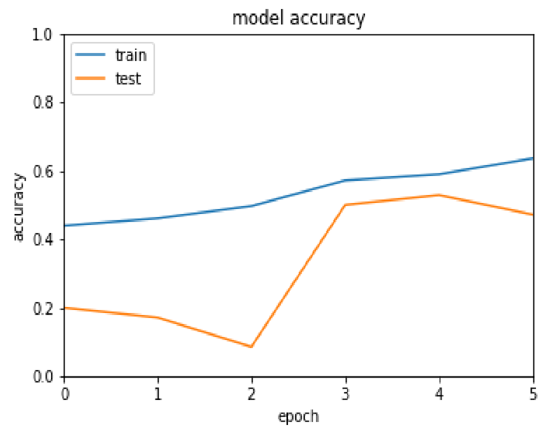

Fig. 4. Accuracy of Model in First Increment

As seen in the Table I Simple CNN model yields only 64.6% accuracy. But incremental model yields 90.26% of accuracy.

Figure 4 shows how training accuracy and testing accuracy varies. The training accuracy increases regularly but testing accuracy varies it show some random trend in its variation. The training accuracy is less than testing accuracy it is due to less data in first increment. As we started adding less variant data the testing accuracy increases.

As shown in Figure 5 we trained model in Second Increment i.e with added less variant data to First Increment. As it is shown the training accuracy and testing accuracy in increased than in First Increment. The difference between training and testing accuracy also reduced. As we go on Increments the accuracy start increasing ans loss starts decreasing. The accuracy is directly proportional to number of Increments.

$$Accuracy \propto Increments \tag{2}$$

Figure 6 shows Third Increment where the model get trained but cant able to generalize it properly. This problem can be solved by increasing data or moving to further Increment. The Second Increment consist of large variant images. As we move further in Increments we have to fed most complex images dataset to model so for achieving best accuracy for most complex dataset add very few images in previous increments. Figure 7 shows training of most complex images in dataset. It can be seen that the training and testing accuracy is nearly same.

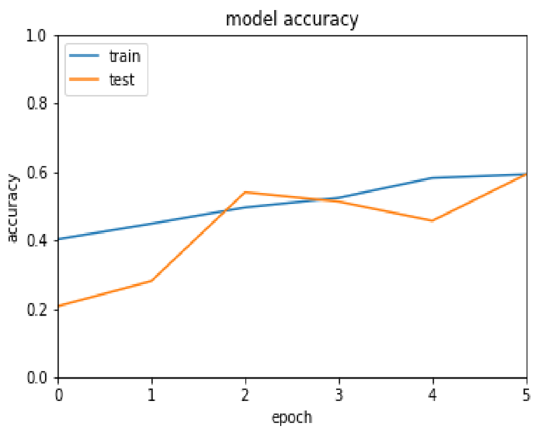

Fig. 5. Accuracy of Model in Second Increment

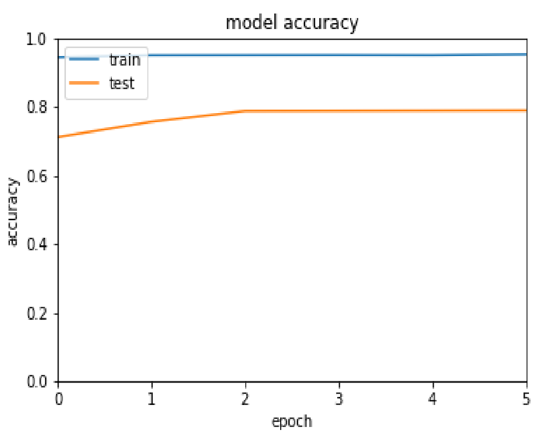

Fig. 6. Accuracy of Model in Third Increment

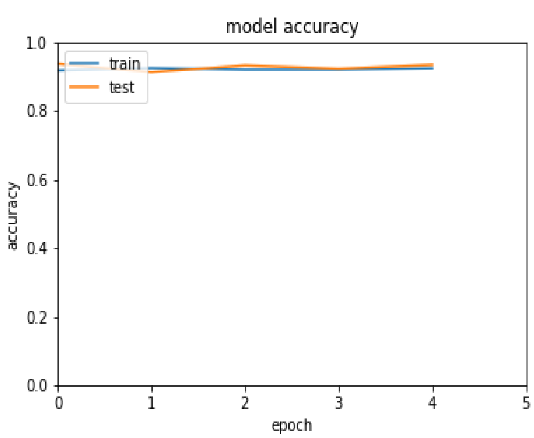

Fig. 7. Accuracy of Model in Fourth Increment

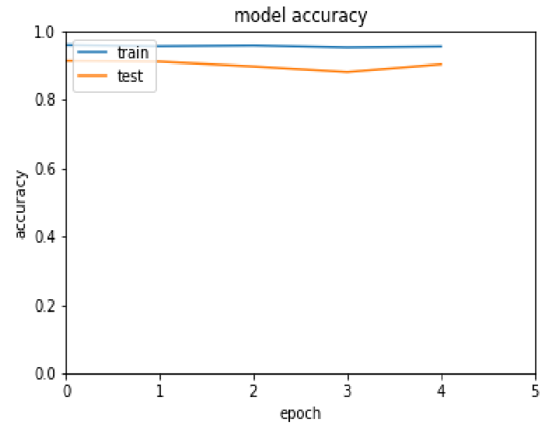

Fig. 8. Accuracy of Final Model

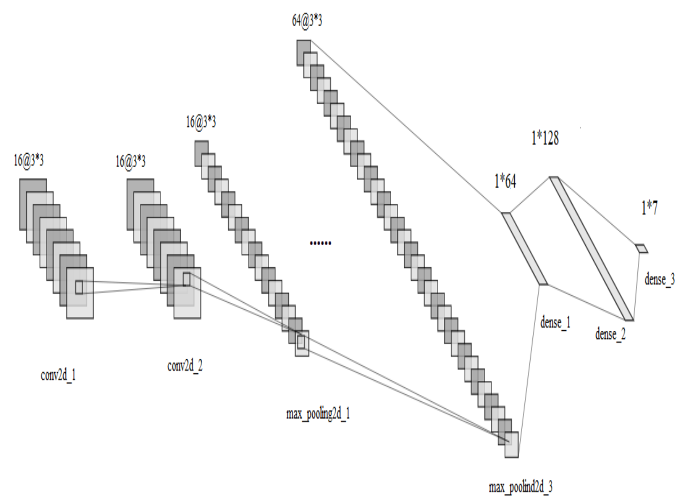

Fig. 9. Proposed Architecture

Figure 8 shows the training and testing accuracy of final model. It can be seen that variation of both accuracy is negligible which is stopping condition.

V. Convolutional Neural Network- Proposed Architecture

[7], [8], [9] The model type that we will be using is 'Sequential'. Sequential is the easiest way to build a model in Keras. It allows you to build a model layer by layer. We use the 'add' function to add layers to our Model. Our layers are Conv2D layers. These are convolution layers that will deal with our input images, which are seen as 2-dimensional matrices. First two consecutive Convolutional layer of kernels size 16 is followed by Max Pooling layer of size 3 is followed by Batch Normalization layer. Second two consecutive Convolutional layer of kernels size 32 is followed by Max Pooling layer of

```
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 300, 300, 16)      160
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 298, 298, 16)      2320
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 99, 99, 16)        0
_________________________________________________________________
batch_normalization_1 (Batch (None, 99, 99, 16)        64
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 99, 99, 32)        4640
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 97, 97, 32)        9248
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 32, 32, 32)        0
_________________________________________________________________
batch_normalization_2 (Batch (None, 32, 32, 32)        128
_________________________________________________________________
dropout_1 (Dropout)          (None, 32, 32, 32)        0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 32, 32, 64)        18496
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 30, 30, 64)        36928
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 10, 10, 64)        0
_________________________________________________________________
batch_normalization_3 (Batch (None, 10, 10, 64)        256
_________________________________________________________________
flatten_1 (Flatten)          (None, 6400)              0
_________________________________________________________________
dense_1 (Dense)              (None, 64)                409664
_________________________________________________________________
dropout_2 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_2 (Dense)              (None, 128)               8320
_________________________________________________________________
dropout_3 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_3 (Dense)              (None, 7)                 903
=================================================================
Total params: 491,127
Trainable params: 490,903
Non-trainable params: 224
_________________________________________________________________
```

Fig. 10. proposed model summary

size 3 is followed by Batch Normalization layer and Dropout layer of size 0.25. Last two consecutive Convolutional layer of kernels size 64 is followed by Max Pooling layer of size 3 is followed by Batch Normalization layer. The number of filters can be changed according to need. This paper uses totally 6 Conv2D layers. Kernel size is the size of the filter matrix for our convolution. So a kernel size of 3 means we will have a 3x3 filter matrix. Activation is the activation function for the layer. The activation function we will be using for Conv2D layers is the ReLU, or Rectified Linear Activation. This activation function has been proven to work well in neural networks. The Max pooling Layer is used. In between the Conv2D layers and the dense layer, there is a Flatten layer. Flatten serves as a connection between the convolution and dense layers. Dense is the layer type we will use in for our output layer. Dense is a standard layer type that is used in many cases for neural networks. Here we have three dense layer having filter size of 64 and 128. The first dense layer is followed by Drop out of size 0.25 and second Dense layer is followed by Drop out layer of size 0.5. We will have 10 nodes in our output layer, one for each possible outcome(0 to 9). The activation is softmax. Softmax makes the output sum up to 1 so the output can be interpreted as probabilities. The model will then make its prediction based on which option has the highest probability. We use drop out between Fully Connected layer to avoid overfitting.

A. *Prepossessing*

Before feeding data to Convolutional Neural Network we have provide some preprocessing it includes image re sizing, data cleaning, etc. To prepare images for training of Convolutional Neural Network each of training images was re sized to 300*300 form 600*400

B. *Implementation*

There three types of data set available in ISIC 2018 challenge first one is Training dataset, second is Teasing dataset and third one is validation dataset. There also a Ground truth also called as labels.

The implementation is based on the technique of Incremental learning for same probability distribution. For the first instance we have select small volume of data having the large variations for training. We trained the model up to maximum extent of accuracy i.e up to peak where accuracy variation is low and accuracy value is large. The weight matrix of this instance is saved for future use. For Second instance using same model but for less variant data we tried to train model by making modifications in model. As soon as we reached same condition in first instance we stopped and save the weights. And continue this Step by Step Learning until our dataset get exhausted. This technique helped our model to classify more complex dataset efficiently. This technique works efficiently if dataset is biased. This technique drops the loss vigorously.

VI. Experimental Result

Motivation behind this work is this is the latest dataset on skin cancer i.e release in 2018 and this dataset contains lesions of 7 type of skin cancers. This algorithm has highest accuracy on this dataset as compared to ISIC2017 dataset which is released in 2017[5]. Comparison of both results is depicted in Table III.

This method is implemented in Python 3. During the implementation we need platform of Intel Xeon and Titan X GPU with 64GB RAM. The performance is calculated by confusion matrix. Firstly input is preprocessed and the classification is carried out deep convolutional neural network. The System can efficiently classified seven types of skin lesions. The convolutional layer acts as feature extractor.

Table II shows comparison of precision, recall, f1-score and support. Here support means number of images taken for tasting of individual class. Class 0 and Class 4 has lower precision compare to the other Classes because of there is large variation in class. Their Symmetry, Border, Color are so variant that the human eye cannot predict it properly. All these the parameter varies randomly so the convolutional neural cant able to predict it. Other than Class 0 and Class 4 all are predicted accurately. Overall precision of network is more than 89%.

We find TP, TN, FP, FN which refer to true positive, true negative, false positive, false negative at the pixel level respectively. Using above terms Precision, Recall and F1-Score are evaluated as follows.

TABLE II
COMPARISON OF PRECISION, RECALL, F1-SCORE FOR INDIVIDUAL CLASS
.

Classes	precision	recall	f1-score	support
Class 0	0.72	0.73	0.72	215
Class 1	0.95	0.92	0.94	1303
Class 2	0.95	0.82	0.88	108
Class 3	0.89	0.86	0.87	64
Class 4	0.70	0.88	0.78	257
Class 5	1.00	0.92	0.96	24
Class 6	1.00	0.96	0.81	32
avg/total	0.89	0.88	0.89	2003

TABLE III
COMPARISON OF RESULTS OF SKIN LESIONS CLASSIFICATION ON ISIC 2017 AND ISIC 2018.

Method	Accuracy	Precision
[6] CSUJT	0.816	0.748
[6] MPG-UCIIIM1	0.849	0.747
[6] RECOD Titans	0.883	0.752
[6] USYD-BMIT	0.888	0.732
[6] IHPC-NSC	0.873	0.665
Proposed Algorithm	**0.9026**	**0.89**

$$precision = \frac{TP}{TP + FP} \tag{3}$$

$$recall = \frac{TP}{TP + FN} \tag{4}$$

$$f1 - score = \frac{precision * recall}{precision - recall} \tag{5}$$

Top 5 result of ISIC 2017 competition combined with result of proposed algorithm is mentioned above in Table III.

Figure 11 and Figure 12 shows the sample of classified and miss-classified images from dataset. As it can be infer from Figure 12 the variation in miss-classified images are more and even they cannot be detected properly by human eye. The variation shown by these miss-classified image on basis of symmetry, border, color are extremely random. Figure 13 shows the confusion matrix of proposed model. It is the comparison between predicted label and true label. As the confusion matrix representing count of classified and miss-classified images. The diagonal of confusion matrix is representing the correctly classified images and non diagonal values represent the miss-classified images. The model correctly classified 153, 1201, 89, 55, 227, 22, 22 of respective class. As class 1 has highest number of samples so it is shown in white to represent the test dataset contains large number of class 1 images. The example images and their representation after convolutional layer is shown in Figure 15.

Figure 14 shows the filters image of last convolutional layer. It looks like most of these have encoded some type of direction or color. Lighter blocks are indicating the smaller weight, as a result feature map responds less to these input pixels. Darker blocks are indicating larger weight, as a result feature map responds more to these input pixels. Figure 15 shows how the feature are extracted in convolutional neural network. These are the some sample images of output after each convolutional layer of trained model. As it can be seen that it get more complex and deeper as we move deeper and

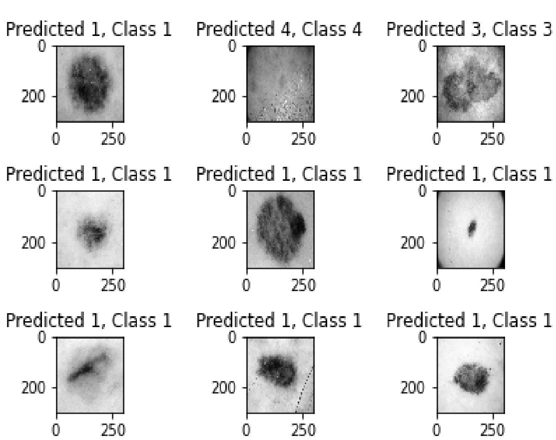

Fig. 11. Correctly Classified class samples

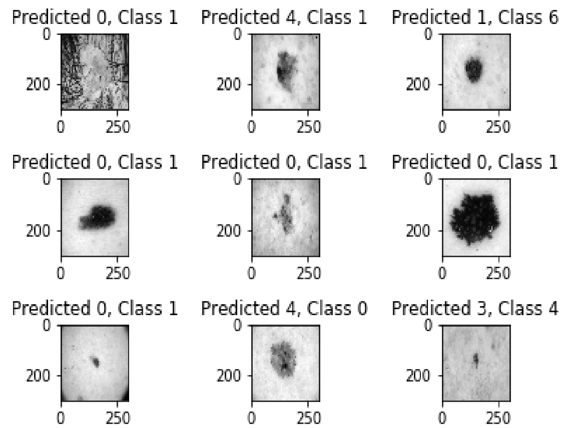

Fig. 12. Miss-Classified class samples

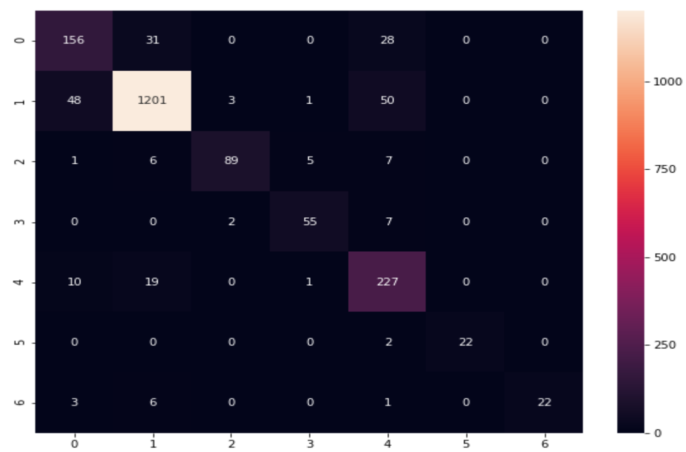

Fig. 13. Confusion Matrix

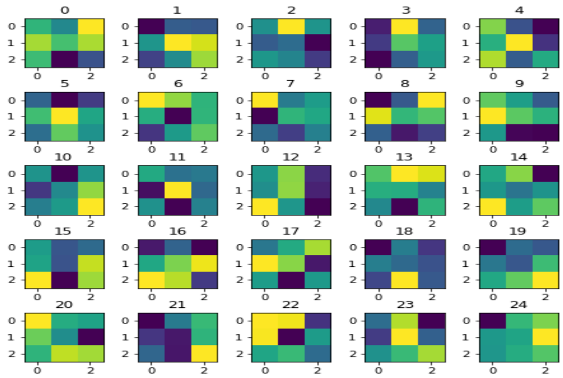

Fig. 14. Filters of last convolutional layer

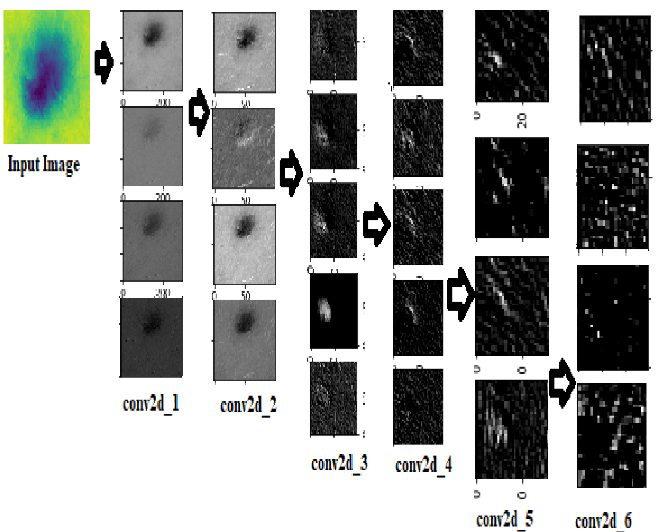

Fig. 15. Sample Image with its Six activation's.

deeper in network.

VII. Conclusion

In this paper we proposed an Incremental approach for convolutional neural network for cancerous image classification. In this paper we have analyzed how incremental model is useful when there is most complex dataset. We compare incremental approach with simple approach and hence concluded that incremental approach gives better result. Further we are aiming to increase the model accuracy by using some diverse methods in combination with incremental approach.

References

[1] Mane, Deepak & Kulkarni, U.V.: Visualizing and Understanding Customized Convolutional Neural Network for Recognition of Handwritten Marathi Numerals. Procedia Computer Science. 132. 1123-1137. 10.1016/j.procs.2018.05.027,(2018).

[2] Tianmei Guo, Jiwen Dong, Henjian Li, Yunxing Gao :Simple convolutional neural network on image classification,pp 721-123,(2017).10.1109/ICBDA.2017.8078730

[3] Thu Thao, Le & Hong Quang, Nguyen, Automatic skin lesion analysis towards melanoma detection. 106-111. 10.1109/IESYS.2017.8233570,pp 106-107,(2017).

[4] Codella, Noel & Gutman, David & Celebi, M. Emre Helba, Brian & Marchetti, Michael & Dusza, Stephen & Kalloo, Aadi & Liopyris, Konstantinos & Mishra, Nabin Kumar & Kittler, Harald & Halpern, Allan : 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), pp 168-171,(2017).

[5] Matt Berseth. : ISIC 2017Skin Lesion Analysis Towards Melanoma Detection. : A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC) https://arxiv.org/pdf/1703.00523.pdf

[6] Yuexiang Li and Linlin Shen : Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC) https://arxiv.org/pdf/1703.00577.pdf

[7] ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection, https://challenge2018.isic-archive.com/. Last accessed 12 Dec 2018

[8] Introduction to CNN with keras, https://www.kaggle.com/amarjeet007/visualize-cnn-with-keras. Last accessed 12 Dec 2018

[9] Understanding of Convolutional Neural Network (CNN)Deep Learning, https://www.medium.com. Last accessed 12 Dec 2018

[10] Chapter 9- Convolutional Networks https://www.deeplearningbook.org/contents/convnets.html. Last accessed 15 Dec 2018