International Journal of Advanced Dental Sciences and Technology (IJADST)
ISSN: 2582-757X (Online), Volume-1 Issue-1, December2020

Deep Learning on Histopathological Images: Automated Classification of Oral Squamous Cell Carcinoma Stages Detection using Pre-trained Convolutional Neural Networks

P. Aurchana, P. Dhanalakshmi

Abstract: ***Deep neural networks have gained immense popularity in the recent years by achieving very good results in medical analysis. This work aims at finding the stages of Oral squamous cell carcinoma using the Pre-trained convolution neural network models like Alexnet, Googlenet and Resnet50. Every pathologist while evaluating the photomicrograph finds it difficult in analyzing the stages of oral squamous cell carcinoma into poorly differentiated, moderately differentiated and well differentiated. To overcome, this deep convolution neural network model has been implemented. In the proposed work, Deep learning needs huge amount of data to achieve good performance so image augmentation has been implemented to boost the performance in deep networks. Later segmentation has been implemented and it is given to the Pre-trained convolutional neural networks which gives the satisfactory result of above80 %.GoogleNet gives the highest results than Alexnet and Resnet50.***

Keywords: ***Oral Squamous Cell Carcinoma (OSCC), Convention Neural Network (CNN), Alexnet, Google Net, Resnet50.***

I. INTRODUCTION

Oral squamous cell carcinoma is the sixth most common cancer in the world [1]. The incidence of oral cancer is highest worldwide and the prevalence is lower in women than men [2]. The common risk factors of Oral squamous cell carcinoma were tobacco habits of chewing with pan, smoking and alcohol. Over 95 percent of oral cavity cancers are squamous Cell Carcinoma and these cancers are further subdivided and by how closely they resemble normal lining cells: well differentiated, moderately differentiated and poorly differentiated. Figure 1(a): shows the photomicrograph of oral squamous cell carcinoma.

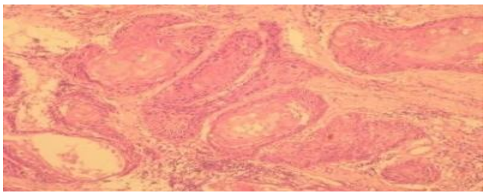


Figure 1: Photomicrograph of Oral Squamous Cell Carcinoma

Manuscript received on 14 November 2020 | Revised Manuscript received on 28 November 2020 | Manuscript Accepted on 15 December 2020 | Manuscript published on 30 December 2020.
***** Correspondence Author
P. Aurchana, Research Scholar, Professor, Department of CSE, Annamalai University, Tamil Nadu, India. E-mail:aurchana85@gmail.com
P. Dhanalakshmi, Professor, Department of CSE, Annamalai University, Tamil Nadu, India.

Histological grade of the tumors revealed that 27%, 40 % and 21% of the tumors were, respectively classified as well, moderately and poorly differentiated OSCC.

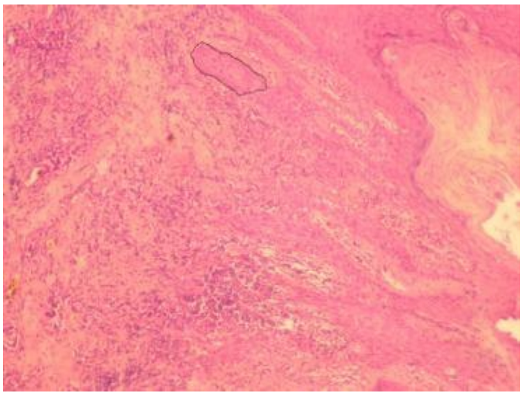


Fig 2: Poorly Differentiated

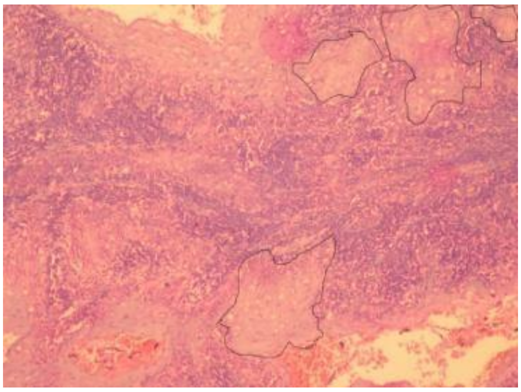


Fig 3: Moderately Differentiated

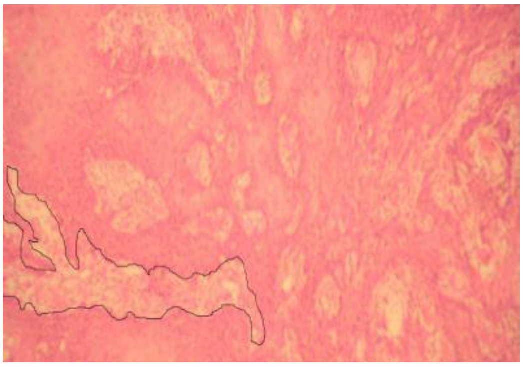


Fig 4: Well Differentiated

Figure 2: shows the image of poorly differentiated OSCC, Figure 3: Shows the image of moderately differentiated, Figure 4: Shows the image of well differentiated.In males it is mostlyaffectedinthe border of tongue, alveolar mucosa/gingival and buccal mucosa/ buccal sulcus in females. Figure 5: shows the overall architecture of the proposed classification system.

Pretrained Convolutional Neural Network

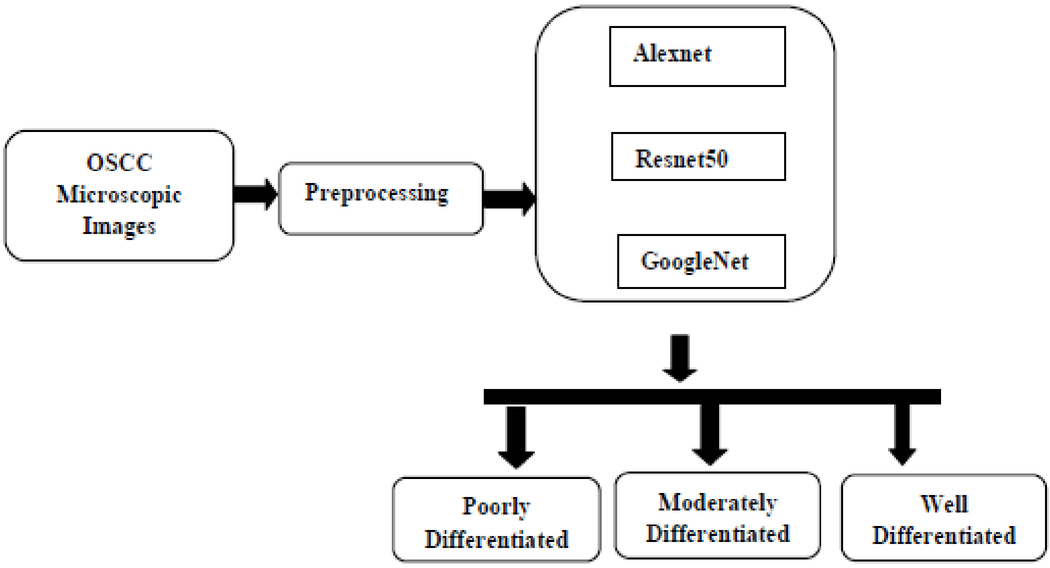


Figure 5: Block Diagram of OSCC classification System

II. RELATEDWORK

In [Siyuan Lu et al] the authors assess the brain detection using Alexnet were the last three layers were replaced by the random weights and the rest of the parameters served as the initial values. In [Merl James Macawile] proposed a work in calculation the WBC count in leukocytes through the pretrained convolutional neural network using Alexnet, GoogleNet, Resnet -101where Alexnet is the winner classifier. In [Paras Lakhani et al] evaluated the efficiency by detecting the Tuberculosis through two different DCNNs namely Alexnet and GoogleNet. In this work Pretrained Augmentation and untrained Augmentation results has been compared through ensembling both the results in which GoogleNet gives the bestperformance.

III. DEEPLEARNING

Deep learning is a new challenge for all well-known applications such speech, image processing and NLP. Deep learning has the greater potential impact in the changing world. In medical image analysis lot deep learning architectures have been applied. Deep learning uses the neural network to learn useful representations and features directly from the data. One of the most important neural networks in deep learning is the convolutional neural network. Convolutional Neural Network uses different pretrained models which work more effectively than the existing techniques. In the proposed work, we concentrate on pre-trained Convolutional models namely Alexnet,Resnet50 and GoogleNet.

A. Pre-Trained Convention Neural NetworkModels

The pretrained networks are trained on more than a million images and can classify images into 100 object categories. The pretrained networks are trained on more than a million ImageNet databases, which is used in ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). Using a pretrained network with transfer learning is typically much faster and easier than training a network from scratch. The Three different pre-trained CNN models have been evaluated in this paper namely Alexnet, Resnet50,GoogleNet.

a. Alexnet

Alexnet is the one of the unique architecture in the area of deep learning. It started from traditional machine learning computer vision and later stepped into deep learning. It was the first neural network in the 21st century.Alexnet is a deep learning network defined from the ImageNet .It consists of eight learning layers in which five convolution layers and three fully connected layers.The convolution layers are responsible for the feature extraction while the fully connected layers are the regular neural networks. The output of the final fully connected layers is the softmax regression which is used for the classification. The main aspects of the softmax are to take results and fit it into the forwarded distribution of 1000 classes. Figure 1:shows the architecture of alexnet.

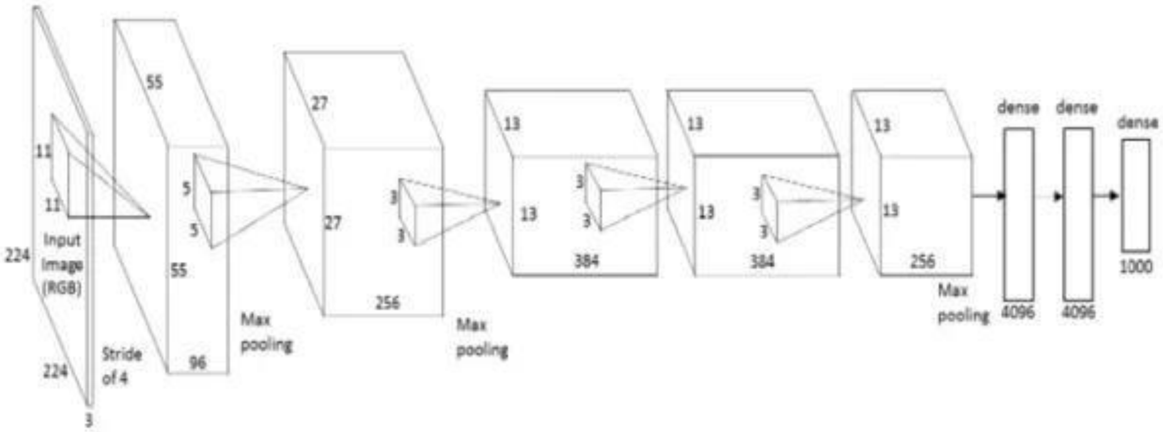


Figure 6: Architecture of Alexnet

b. MaxPooling

Max Pooling is one of the most popular pooling methods, pooling is a form of non-linear down sampling. Max pooling divides the image into non-overlapping sub regions for which the output of each region is the maximum value. Max Pooling is very useful as it eliminates a non-maximal value which in turn reducesthe computation of the upper layerss.This technique is a robust approach to reducing the dimensionality of intermediate responses.

c. Dropout

Dropout is a interesting technique used for normalization. The idea is that the neural network consists of multiple nodes. In order to provide regularization and we have to make sure the neural network is not biased to a particular result we have to perform the random dropping of the neurons in the network. This performs the dropout andwehaveto specify the percentage of neurons to dropout and the network randomly chooses the neurons and dropout.

d. Resnet50

It is the commonly used model in convolutional neural network. It is a very deep networks used in residual connections. It was first implemented by He et al., 2015.Residual Networks consists of multiple subsequent residual modules, which are the basic building block of Resnet architecture. Figure 2: shows the Resnet architecture. It has 152-layer model for ImageNet and it was ILSVRC 2015 winner in (3.57% top 5 error) [3].The main advantage of Resnet is that hundreds, even thousands of these residual layers can be used to create a network and then trained. This is a bit different from usual sequential networks, where we see that there are reduced performance upgrades as we increase the number of layers.

e. GoogleNet

GoogleNet(or Inception Network) is a class of architecture designed by researchers at Google. Googlenet has 22 layers and introduce for the first time the efficient “Inception “module.Also there are fully connected layers with only 5 million parameters , which is 12 times less than Alexnet. Figure 7: The building block of GoogleNet

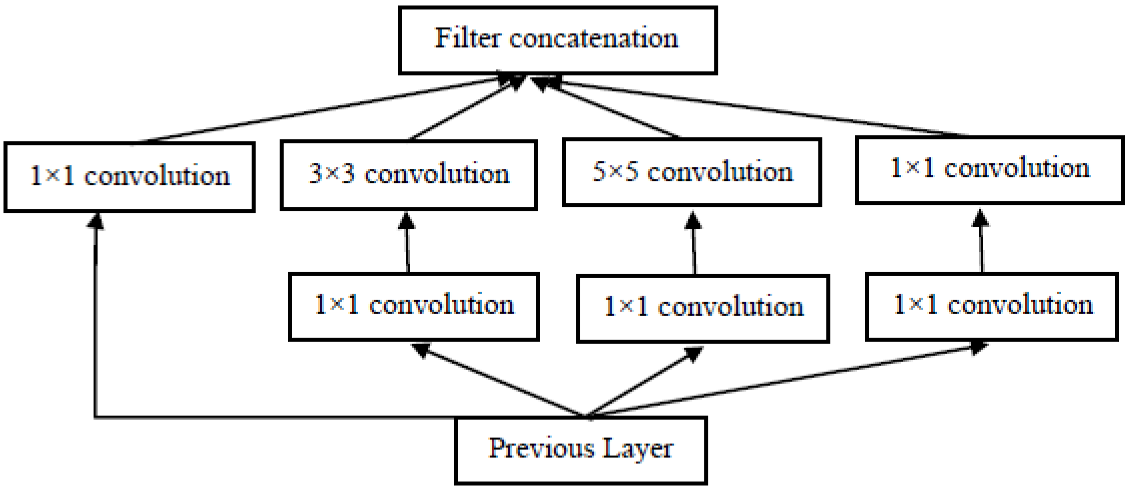


Figure 7: Building block of GoogleNet.

In the above figure 7 it is based on “Inception module” which meant designing a good local network topology (network within a network) and then stacks these modules on top of each other. In this apply parallel filter operations on the input from previous layer: Multiple receptive field sizes for convolution (1x1, 3x3, 5x5) – Pooling operation (3x3) Concatenate all filter outputs together depth-wise. Apply parallel filter operations on the input fromprevious layer: - Multiple receptive field sizes for convolution (1x1, 3x3, 5x5) – Pooling operation depth-wise.

IV. PROPOSEDMETHODOLOGY

a. Preprocessing

In the preprocessing stage, the original image that is obtainedfromtheelectronmicroscopicisofthesize256x 256. Next the region of interest is selected manually for disease identification and the region cropped.

b. Region of InterestCropping

For volumetric examinations, automatic freehand ROI technique, is presented in the system of microscopic image which is used. Using cell differentiation of OSCC our proposed system categories automatically into well-differentiated, moderately differentiated and Poorly Differentiated. On this image mild pleomorphism good amount of intracellular and extracellular keratinization in the form of keratin pearls, few keratin pearl formulation moderate nuclear pleomorphism, moderately marked nuclear pleomorphism, absence of intracellular and extracellular keratinization. Figure 8: Shows the disease affected area

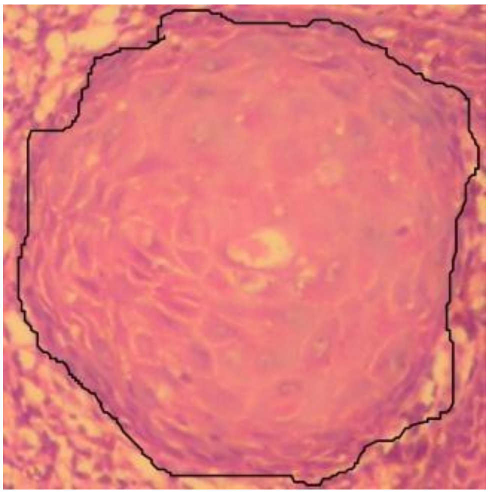


Figure 8: Cropped disease affected image

V. EXPERIMENTALRESULTS

a. Datasets

The photomicrograph of OSCC shows the different stages that were collected from raja Muthiah Dental College and Hospital (RMDC & H). Hematoxylin and eosin stained section of OSCC showing well differentiated, poorly differentiated and moderately differentiated with (10 x magnification) higher magnification were taken.

b. Image Augmentation for DeepLearning

Deep networks need large amount of training data to achieve good performance. Due to the scarcity of images and to build a powerful image classifier using very little training data, image augmentation is usually required to boost the performance of deep networks [5]. The augmented image data store transforms the batches of training, validation, test and prediction data with preprocessing such as resizing, rotation and refection. Augment training image data with randomized image preprocessing operations helps to prevent the network from over fitting and memorizing the exact values of the trainingimages.

To train the network using augmented images, we use augmented image data store to train network.In the proposed work, when we use the augmented data store as a source of training images, the data store randomly perturbs the training datafor each epoch, so that each epoch uses the slightly different data store. The actual training images at each epoch does not change.Here the images input layer normalizes images using the mean of the augmented images, not the mean of the original dataset. The mean is calculated once for the first augmented epoch. All other epoch use the same mean, so that the average image does not change during training.

c. Recognition of Disease using Pre-trained Alexnet Convolutional NeuralNetwork

The Alexnet consists of eight layers in which the first Five is the convolutional layer and the remaining three is the fully connected layers. It uses the different activation function called the ReLU (Rectified Linear Unit) after every convolution layers and it also has new types of processing called the dropout after fully connected layers 1 and 2.The input features are reduced from 154587 to 1024 before sending them to the fully connected layers by the arrangement of convolution and pooling layers. The detailed mapping of the proposed work of Alexnet is shown below:

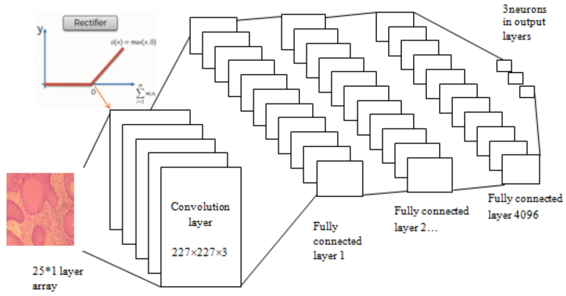


Figure 9: Proposed work of the Pre-trained Alexnet Convolutional Neural Network.

In the above figure the input size of the image is fixed to 227 x 227x3 in the input layer. In the input layer it contains the three featuremap. The first is the convolution layer contains the activation function with 96 feature map and the size of the image is 55x55 where the filter size is 11x11 and the stride is 4 x 4.In thefirst max pooling the feature map remains as 96, the size of the image is 27 x 27, filter size is 3 x 3 and the stride changes to 2 x2.

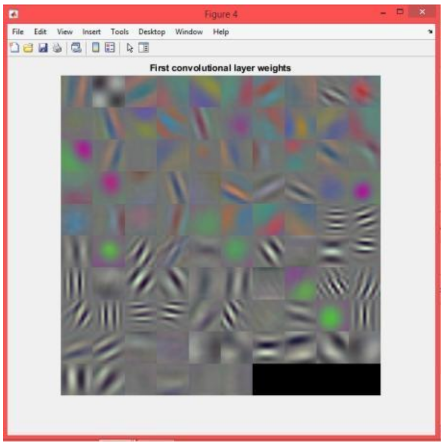


Figure 10: Alexnet First Convolutional Layer Weights

In the second convolution layer the 256 featuremap has been obtained. The size of the image is 23 x 23, the filter size is 5 x 5, stride is 1 x 1.In the second pooling the feature map remains unchanged. The size of the feature map is 11 x 11, the filter size reduces to 3 x 3, and the stride is 2 x 2. In the third convolution layer the feature map is increased to 384, the size of the image is 9 x 9,the filter size is 3 x 3 and stride is 1 x 1. In the fourth convolution layer the feature remains unchanged; the size of the image is 7 x 7 ,filter size 3 x 3 with the stride of 1 x 1. The fifth convolution layer the feature map is 256, the image size is 5 x 5 and filter size is 3 x 3 with the stride of 1 x 1. The third max pooling the feature map remains unchanged, the size of 2 x 2, the feature size is 3 x 3, the stride 2, 2.The output of the final max pooling layer is given 4096 fully connected nodes. The softmax classifier is used to categorize into threeclasses.

d. Recognition of disease using Pre-trained Resnet50 Convolutional NeuralNetwork

Resnet50 has 177 layers in which 50 layers forms the residual layers. The input image of 224x 224 is taken as the input. Here the first convolution layersconsists of 64 filters with 7 x 7 filter size with the stride 2 x 2 is processed.

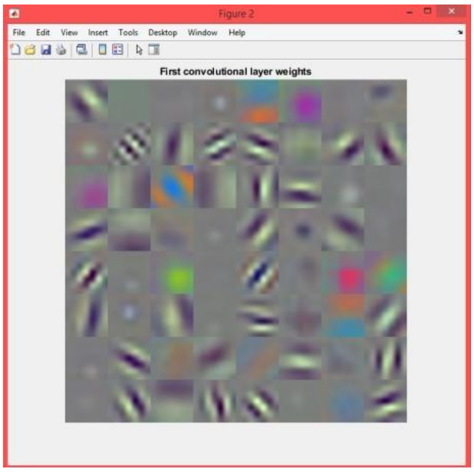


Figure 11: Resnet50 First Layer Weights

Further the batch normalization is processed with 64 channels with Rectifier Linear unit as the activation function and with the max pooling of 3 x 3.Hence this steps is processed for the remaining all the resnet-50 layers. The final convolution yields 2048 parameters with the batch normalization of 2048 channels, with the average pooling of 1000 fully connected layer. Later softmax classifier is used to categorize the three classes into poorly differentiated, moderaly differentiated and welldifferentiated.

e. Recognition of disease using Pre-trained GoogleNet Convolutional NeuralNetwork

The GoogleNet consists of 144 layered arrayswith 22 layers deep. Figure : Proposed work of Pretrained GoogleNet Convolutional Neural Network.

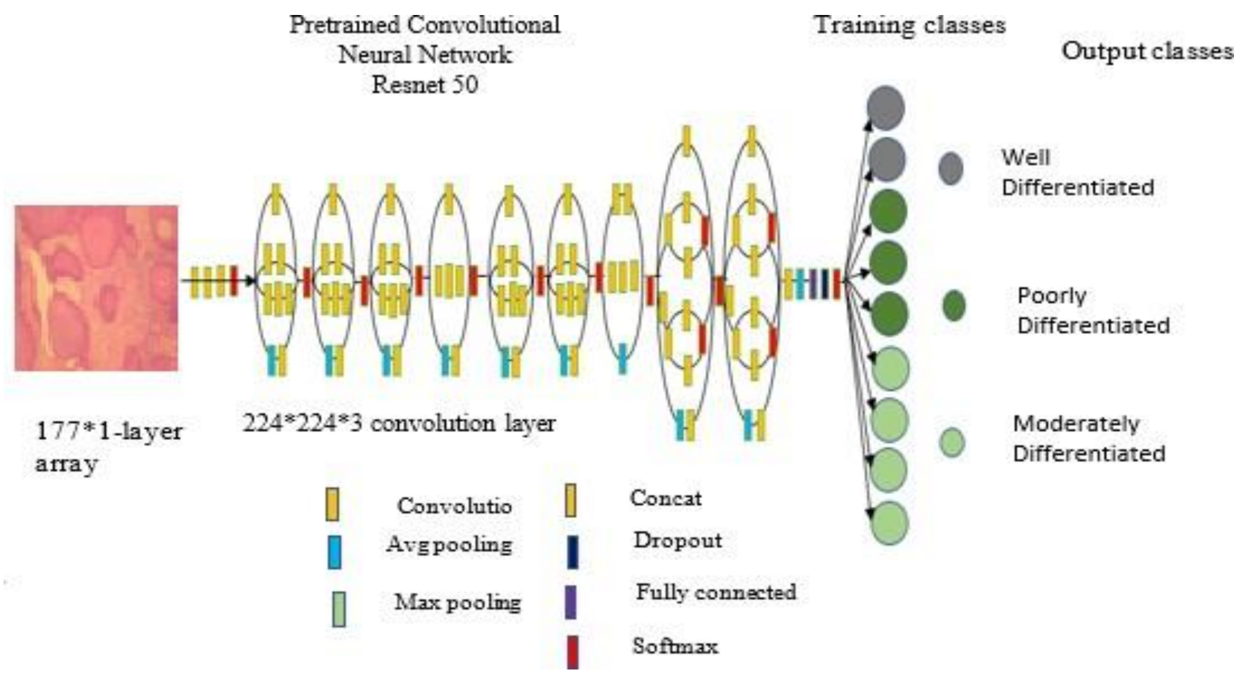


Figure 11: Proposed work of the Pretrained Google Net Convolutional Neural Network.

Here the input is resized into 224 x 224. The first covolutional layer consists of 64 filters with the size of 7 x7 x 3convolutions with stride [2 2] with which it is activated using the Rectified Linear Unit with the max pooling and the channel is normalized with 5 channels per element and the process in continued for the next second convolution. Figure :shows the first convolutional layerweights

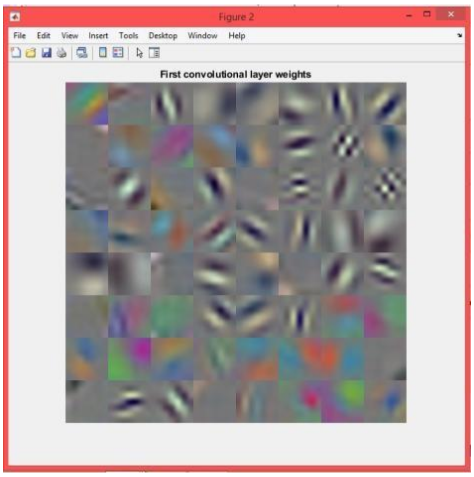


Figure 12: GoogleNet First Convolutional Layer Weights.

Next step it enters into the inception block with the convolution of 64 filters of 1 x 1 size of 192 convolution with the stride of [1 1] and this activated using Rectified Linear unit. This process is continued. The final convolutional layers is processed with depth concatenation with the average pooling of 40 % droout ratio and it gets flattern with the dense layer further it is categorized into three classes namely well differentiated, moderally differentiated and poorly differentiated.

Testing

For testing the histopathologicalimages of different stages were taken for prediction. The prediction of classes was computed by pre-trained neural network in MATLAB 2018b prediction analysis. The prediction matrix is given into three classes namely well differentiated, poorly differentiated, moderly differentiated.

f. PerformanceMeasures

The performance accuracy is calculated using confusion matrix based on predicted samples. In confusion matrix processes factors such as true positive, true negative, false positive, false negative for finding performance metrics, with these factor's accuracy, recall, precision and f-score of the predicated samples haveanalyzed.

The accuracy can be defined as the percentage of correctly classified instances

$$\text{Accuracy}= \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{1}$$

where TP, FN, FP and TN represent the number of true positives, false negatives, false positive and true negatives respectively also you can use standard performance measures:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{2}$$

$$\text{Specificity}= \frac{TP}{TP+FN} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4}$$

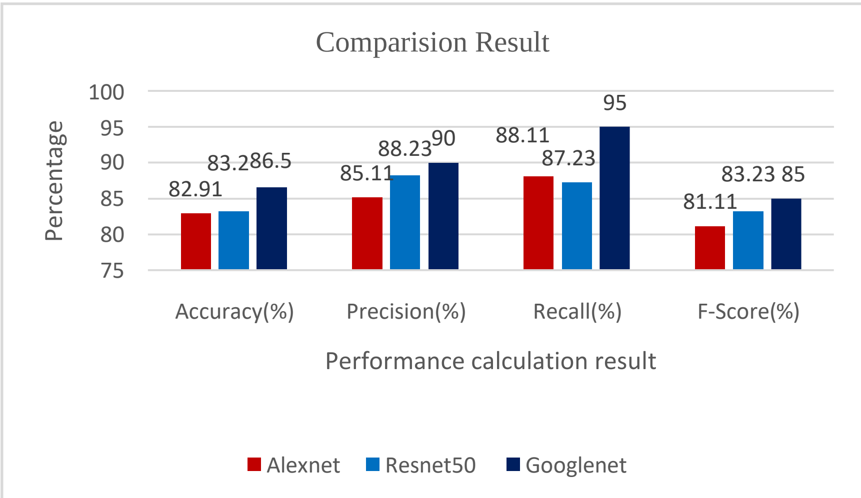
$$\text{F-Measure} = \frac{2PR}{P+R} \tag{5}$$

where P and R are precision and Recall . The below table represents the performance evaluation of oral squamous cell carcinoma :

Table 1: Performance evaluation on classification of OSCC stages using Pre-trained Convolution Neural Network

Proposed Method ofPre-trained Neural Network	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
Alexnet	82.91	85.11	88.11	81.11
Resnet 50	83.20	88.23	87.23	83.23
Googlenet	86.50	90.00	95.00	85.00

The below chart illustrates the comparison chart of thethree Pre-trained Convolutional NeuralNetwork.



VI. CONCLUSION

This research was able to classify the OSCC stages on microscopic images using pretrained convolutional neural networks. Moreoverupon comparison of the three CNN models (Alexnet, GoogleNet and Resnet-50) it was observed that GoogleNet performed best on the task of the classification of OSCC stages sample images as Compared to Alexnet and GoogleNet generating an overall accuracy of 86.50%.

REFERENCES

1. Ramasamy Padma, AmitkumarKalaivani, SivapathamSundaresan and PaulrajSathish,Journal of Oral Maxillofacial Pathology,The relationship between the histological differentiation and disease recurrence of primary squamous cell carcinoma,Volume 21, Issue 3,2017.
2. Balaram, Sridhar H, Rajkumar T, Vaccarella S, Herrero R, Nandakumar A, Ravichandran K, Ramdas K, SankaranarayananR,Gajalakshmi V, Munoz N, FranceschiS,International Journal of Cancer,Oral Cancer in southern India: the influence of smoking, drinking and paan-chewing and oral hygiene, Volume 98, issue 3,March2002.
3. Popa Didi Liliana, FaiqBaji, Popa Radu Teodoru, Overview of Deep Learning in Medical Imaging,Annals of the University of Craiova, Series:Automation, Computers, Electronics and Mechatronics, Volume 14, Issue1,2017.
4. Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun Microsoft Research,10 December2015.
5. Ganesh B, Kumar C, National Conference on Emerging Trends in Computing Technologies, Deep Learning Techniques in image processing, ISSN 2395 – 1303,2018.
6. Merl James Macawile, Vonn Vincent Quinones, Alejandro Ballado Jr, Jennifer Dela Cruz, Meo Vincent caya, “White Blood Cell Classification Using Convolutional Neural Network”, 3rd International Conference on control and Robotics Engineering,2018.
7. Pras Lakhani, Baskaran sundaram, Thoracic Imaging: Deep Learning at Chest radiography: Automated Classification of Pulmonary Tuberculosis by using Convolutional Neural Networks,2017.
8. SiyuanLu,Zhihai Lu, Yu-Dong Zhang, “Pathological brain detection based on Alexnet and transfer learning”, Journal of computional Science,2018.
9. P.Aurchana, P.Dhanalakshmi, “Grading of Oral Epithelial Dysplasia Using Multi-SVM”, Journal of advanced Research in Dynamical and Control Systems, Volume 10, Issue2018.
10. P.Aurchana, P.Dhanalakshmi, SVM Based classification of epithelial Dysplasia Using SURF and SIFT Features, International Journal of Pure and Applied Mathematics, Volume 117, No 15,2017.