Analysis of Oral Squamous Cell Carcinoma into Various Stages using Pre-Trained Convolutional Neural Networks

A. Ramalingam\textsuperscript{1}, P. Aurchana\textsuperscript{2}, P. Dhanalakshmi\textsuperscript{3}, K. Vivekananadan\textsuperscript{4} and V. S. K. Venkatachalapathy\textsuperscript{5}

\textsuperscript{1}Professor, Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Puducherry, India
\textsuperscript{2}Research Scholar, Department of Computer Science and Engineering, Annamalai University, Chidambaram, India
\textsuperscript{3}Professor, Department of Computer Science and Engineering, Annamalai University, Chidambaram, India
\textsuperscript{4}Professor, Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry, India
\textsuperscript{5}Professor, Department of Mechanical Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India

Abstract. Oral cancer is a commonly prevalent disease in the world. Cancer begins when alterations in healthy cells take place and grow gradually. Finally, it forms a mass called cancer. In general cancer can be classified into malignant and benign. In malignant, the cells can grow and multiply affects other parts of the body but, benign does not spread. Among the cancers in the South East Asia, the cancers in oral cavity ranks among the third most common types of cancer. Oral Squamous Cell Carcinoma is a highly prevalent oral cancer affecting the head and neck more than 90\% than other parts of the body. Until now, classification of Oral Squamous Cell Carcinoma classification into various stages is based on the cytological and architectural change which relies on the pathologist. Every pathologist while assessing the lesions of Oral Squamous Cell Carcinoma into various stages led to mistakes. To overcome this, Computer Aided Diagnosis gives the exact stages of Oral Squamous Cell Carcinoma into poorly differentiated, moderately differentiated and well differentiated. In this work, two Convolutional Neural Network Architectures Inception-v3 and Resnet50 are used as feature extractors. Then, the derived features are given to Multi-class Support Vector Machine and Random Forest. Random Forest and Multi-class Support Vector Machine classifies the Oral Squamous Cell Carcinoma into various stages namely poorly differentiated, moderately differentiated and well differentiated. The features obtained from Resnet50 when given to Random Forest gives the satisfactory performance of 92.08\%.

Keywords: Oral Squamous Cell Carcinoma (OSCC), Convolutional Neural Network (CNN), Support Vector Machine (SVM), Fuzzy Cognitive Map (FCM), Gray Level Co-occurrence Matrix (GLCM), Deep learning (DL), Reinforcement Learning (RL), Random Forest (RF).
1. Introduction
Oral squamous cell carcinoma is the most common malignant epithelial neoplasm affecting the oral cavity which consists of collection of neoplasms. This will affect any area of oral cavity, salivary glands and pharyngeal areas. This term tends to replace with Oral Squamous Cell Carcinoma. The most important danger of neoplasm is that it could not be noticed in earlier stage. Generally, at earlier stage, there is no pain but has burning sensation or pain when it develops. The most common affected areas of OSCC are lips, tongue and floor of the mouth. Some OSCC seems as normal mucosa but others are premalignant lesions particularly erythroplakia and leukoplaikia.

- The most beneficial menace aspect for oral squamous cell carcinoma are
- Smoking (especially >2packs/day)
- Use of Alcohol
- Other factors also play a role. These include:
- An impaired facility to renovate DNA injured by mutagens
- Adamagein metabolize carcinogens
- Lack of vitamins A, E or C or trace elements
- Defects in Immune

The widespread type of cancer found in oral epithelium is the Oral Squamous Cell Carcinoma (OSCC). Despite sufficient screening methods and absence of accuracy, Oral Squamous Cell Carcinoma (OSCC) is noted at the final stage [1]. More than 90% of oral cancer, particularly, OSCC is caused mainly due to the consumption of alcohol and chewing of tobacco. It is mainly prevalent in south Asian countries. Like other diagnoses of cancer, OSCC is also diagnosed by taking the tissue samples and evaluating the affected region by using the microscope [2]. In this way, the pathologist grades the stages of OSCC.

At present, the pathologist finds it difficult to identify the stages of OSCC and this consumes more time. To overcome this issue, Computer Aided Diagnosis (CAD) model is introduced which helps the pathologist in making decisions. In this work, various algorithms are utilized to categorize the OSCC into various stages, namely, poorly differentiated, well differentiated, and moderately differentiated. In this proposed work, two CNN architectures (Inception-v3 and Resnet50) are used as feature extractors. Then, the derived features are given to (Multi-class SVM) MSVM and (Random Forest) RF for classification of OSCC into various stages namely poorly differentiated, moderately differentiated and well differentiated. Fig. 1 shows the overall framework of OSCC classification into various stages.

![Diagram of OSCC classification](image)

**Figure. 1** Proposed method for categorizing the OSCC affected images into poorly differentiated, moderately differentiated and well differentiated.
2. Literature Review

In [3] the author proposed various diagnosis methods to find the oral cancer such as biopsy method in which a small sample of tissues is removed from a part of body and tested using the microscope and some screening methods. But the drawback is that it cannot actually clearly detect the tumor of cancer cells as well as they couldn’t classify how much cells are affected by cancer so this paper discerns and categorize the ostentatious cancerous cell in the oral sector by digital Image processing procedures component extraction delegates clear visualization of cancer affected areas. Here they used firefly algorithm to detect the cancer tumor in the MRI image. And Expectation maximization (EM) algorithm to categorize the cancer cells precisely the project is proficient utilizing mat lab program.

In [4] the author proposed work uses Fuzzy Cognitive Map (FCM) to classify the lesions. The map was formulated utilizing the five Gray Level Co-occurrence Matrix (GLCM) constituents extracted from Dental Radiographs. The sequel from this tactic were resembled with other tactics. 50 cases were classified by experts as benign and malignant. The categorization precision prevailed for vicious cases is 87.50% and benign cases was 84.61%. The operation of the system is scrutinized utilizing Accuracy, Sensitivity and Specificity. The results were and given accuracy of 87.21%.

In [5] the author utilized histological constituents to categorize oral cancers from 123 cases. FCM was developed and Active Hebbian Learning algorithm was used to train the FCM. It achieved an accuracy of 90.58% for normal cases and 89.47% for abnormal cases.

In [6] the author deployed detected cysts from dental images. They used image processing techniques and neural network methods for the measurement of severity detection. The dubious vesicle provinces are interprets utilizing Radial Basis Function Network. The ferocity of the vesicles is deliberated utilizing circularity appraises and the sequels unveils the part of the vesicles distilled. In [7] the author survey under the topic of Genetic Programming for Oral Cancer Detection and Secured Image Restoration, this work presents the detection of oral cancers using Image Processing. Dental X – Rays and RGB facsimiles are utilized as the input image for perception of cancer. At first, Gabor filter is utilized to detach noise from the images. This is utilized for image intensification in image preprocessing tactic. Genetic Algorithm is utilized to distilled the traits of tumors from the inflated images. The ferocity of the vesicles is deliberated utilizing circularity appraises and the sequels unveils the part of the vesicles distilled. In [8] the author made the research to design a lung cancer detection system based on analysis of microscopic image of biopsy using digital image processing. Microscopic facsimiles of biopsy are constituents distilled with the Gray Level Co-occurrence Matrix (GLCM) design and categorize utilizing back proliferation tactile network. This method is implemented to detect both normal and cancerous lung of biopsy samples. In the phase of tutoring, 20 biopsy facsimile specimens were scrutinized utilizing back proliferation sensational network with 95% precision. On the other hand, 16 biopsy specimens were scrutinized in the course of appraising, with an precisioin of 81.25%. These sequels show that microscopic biopsy facsimile processing can be carried out in a system of lung cancer detection.

In [9] the author used dental X – ray images to classify normal and abnormal cases. The hybrid fuzzy based classification was used for predictions. Accuracy, Specificity and Sensitivity were calculated. Oral cancers were detected using Android Application. In [10] authors experimented the Squamous Cell Carcinoma Detection and Measurement in Image Processing. Mainly this mechanism detects the disease as soon as possible and also finds out the exact point of disorder and calculates the growth of this disease, especially in Squamous cell carcinoma in lower lip. Actinic keratosis which is 1/4 inch in diameter, is a pink or flesh colored rough spot which is one of the most important causes of squamous cell carcinoma, which is mainly grown in sun-exposed area. It is usually grown slowly and affects epidermal layer to dermis layer. Our proposed strategy pinpoints on five disparate modules. These methods are included in Image Acquisition module and respective Preprocessing (Image Clipping, Smoothing and Enhancement), Segmentation (Thresholding, Histogram Analysis), Filtering Phase and Edge Detection modules.

In [11] the authors took a tour on Unsupervised Deep Learning. They said, both supervised and
unsupervised deep learning achieved promising results in the area of medical imaging and image analysis. Contrasting supervised learning which is prejudiced concerning how it is being administered and manual enterprises to erect class label for the algorithm, unsupervised learning procures discernment precisely from the data itself, categorize the data and assist to make data-driven firmness without any external oblique. This appraise consistently presents assorted unsupervised paradigms appealed to medical facsimile scrutiny, comprising auto encoders and its assorted variants, Restricted Boltzmann machines, Deep belief networks, Deep Boltzmann machine and Generative adversarial network.

In [12] authors did scrutinize in Deep learning (DL), Reinforcement learning (RL), and their amalgamation (Deep RL) assurance to recast Artificial Intelligence. The expansion in computational power assisted by brisk and expanded data storage and diminishing computing costs have heretofore permitted scientists in several fields to appeal these procedures on datasets that were hitherto awkward for their extent and intricacy. This scrutiny tract bestows an exhaustive prospect on the application of DL, RL, and Deep RL approaches in mining Biological data. In inclusion, we emulate interpretations of DL procedures when solicit to distinct datasets across several application realms. Eventually, we contour open circumstances in this stimulating scrutiny area and dispute subsequent evolution aspect.

2.1 PRE-TRAINED CONVOLUTION NEURAL NETWORKS

The feature extraction performance in microscopic images of this chapter presents two re-training techniques known as Transfer-learning and Fine tuning. In this work, transfer learning has been used on the Inception-v3 and Resnet50. Inception-v3 and Resnet50 use the weights of its network to perform the task of classification of microscopic image dataset which is classified into three stages namely, poorly differentiated, moderately differentiated and well differentiated.

Transfer Learning

Transfer learning is a technique used to deal with small or large amounts of labelled information from the source domain to a destination domain, in order to build an efficient prediction model. The Inception-v3 CNN model has been pre-trained by utilizing the ImageNet dataset, to freeze its layers rendering them non-trainable. Then the last layer is removed from the fully connected or SoftMax. A new layer is included entirely linked, taking the respite of the network for constituent coercion and for guiding the model Inception-v3model

Inception network is an up-to-date deep learning model. It is mainly used to solving image identification and detecting problems. The Inception deep convolutional architecture was launched by Google Net in 2015 and it was named as Inception-v1 [13]. Next, the inception was refined by batch normalization and then Inception-v2 came into existence. Now, in Inception-v3, more factorization is introduced.

Based on the earlier versions, the factorization of 3 × 3 convolutions takes place instead of standard 7 × 7 convolution. A set of 3 standard inception models are incorporated for the network Inception part at 35×35 besides 288 filters each. It is minimized to 17 × 17 grid with 768 filters with grid reduction. It is proceeded by 5 recurrence of factorized inception modules. The Inception modules consists a set of 8 × 8 level with linked output filter bank size of 2048 for each tile.
However, the network quality is relatively stable towards modifications. The Inception-v3 is 42 layers deep, which works more efficiently than VGGNet and it performs concatenation of many various sized convolutional filters into a new filter [14]. This model reduces the number of parameters which undergo the training and thereby minimizes the computation complexity. Fig. 2 illustrates the overall architecture of Inception-v3 model. The architecture consists of factorization into smaller Convolutions, Factorization into Asymmetric Convolution, Grid Size Reduction, Auxiliary Classifier and Regularization via Label Smoothing.

**Figure. 2** Basic architecture of Inception-v3

**Figure. 3** Architecture of Inception-v3 model

*Factorization into smaller convolutions*
In this stage, the dimensions/parameters are reduced without decreasing the efficiency of the factorizing convolutions.

**Grid Size Reduction**

At this point, when more convolutional layers are applied, the Grid size decrease is realized by pooling, followed by the convolution operation.

**Auxiliary Classifiers**

Auxiliary classifiers improve the convergence of very profound networks. The main intention is to push the essential gradients to the lower layers during the training by combating the disappearing gradient issue in very deep networks.

For one facsimile, we distilled a 2048-dimensional constituent from the endmost fully-connected logits layer. The layers before the fully connected layer of the pre-trained networks perform the feature extraction for the images and the fully-connected layers are used for classification. In this work, the fully connected layers are replaced by MSVM to perform classification. This model is referred as a single transfer learning network. The information related to the features produced from the pre-trained deep CNN is given here. The overall architecture of Inception-v3 is depicted in Fig. 3.

Resnet50 is otherwise called the Residual network used to identify mapping by shortcuts. It is the commonly used model in CNN [15]. Residual Networks comprises of various subsequent residual modules, which are the basic foundations block of Resent. As the network goes deeper and deeper, the training is more difficult. Generally, the input feature map will be followed by the convolutional filter, non-linear activation function and a pooling operation and finally the output is the next layer. Here, back propagation algorithm is implemented. As the network goes deeper and deeper, it is hard to converge.

[Figure 4: Architecture of ResNet50 model]

The architecture of ResNet50 is depicted in Fig. 4. The construction of ResNet50 has 4 stages as shown in Fig 4. The input size of the image is 224 x 224 x 3. Every ResNet structure makes the first convolution and max pooling using 7 x 7 and 3 x 3 kernel sizes distinctively. Next, first stage of the
network commences and it comprises of 3 Residual blocks containing with 3 layers each. The size of the kernels utilized to perform the convolution operation with all 3 layers of the block of the first stage is 64, 64 and 128 distinctively. The undulated arrows cite to the discerning interrelation. The dashed interrelated arrow indicates that the contortion process in the Residual block is carried out with stride 2, therefore, the extent of input will be lowered to half in regard to height and width but the channel width will be dual. As we shift from one instant to the succeeding, the channel width is raised dual as much and the extent of the input is lowered to half. For extensive networks like Resnet 50, Resnet152, etc, constriction pattern is utilized. For each surplus process F, 3 layers are drifted one over the other. The three layers are 1x1, 3x3, 1x1 convolution. The 1x1 convolution layers are responsible for decreasing and then replacing the dimensions. The 3x3 layer remains as a bottleneck with less input/output dimensions. Finally, the network has an average pooling layer by a connected layer with 1000 neurons.

2.2 MODELING THE FEATURES FOR OSCC

Multi-SVM
In this work, MSVM is applied to allot labels to the occurrence in which the labels are taken from definite set of various factors. Each training point is linked to one of N different classes. The goal is to construct a function which gives a new data point, that will exactly confirm the class to which the new point refers to [16]. The result to predict various classes using MSVM is to alleviate the only multiclass issue into various binary classification problems. In this chapter, the MSVM model is applied to classify the given image into three classes namely poorly Differentiated, moderately differentiated and well Differentiated.

Random Forest
Random Forest (RF) classifier is otherwise called the random decision forests [17]. RF is a collection of tree predictors which is called as forest. The RF classifier is mainly used for categorizing and regression issues. For implementing out the classification process, a trained RF model is applied and the steps are given below:

- Take the test features and employ the rules of every arbitrarily generated decision tree for predicting the output and save the expected result(target)
- Determine the votes for every predicted target.
- Assume the high voted predicted target as the final prediction from the RF model.
- To carry out the classification process, the trained RF model is needed to render the test features using the rules of every arbitrarily generated tree. The comprehensive pattern is given in Fig. 5.
During training, all the trees are trained with the same parameters. Here, the bootstrap procedure is followed for each training set, then we randomly select the usual number of vectors and the vector is replaced randomly. During training at each node, a new subset is produced. All the variables are not utilized to divide the node; a subset is selected randomly to generate a new subset in a node. But the size is affirmed for every node and tree. During training, the current tree is drawn by replacement, while some vectors are left out. This is called as out-of-bag.

2.2 PERFORMANCE MEASURES
To evaluate the performance of MSVM classifier utilizing HOG and LBP feature extraction techniques, a set of evaluation parameters namely Accuracy, Precision, recall and F-score are utilized.

3. EXPERIMENTAL RESULTS
In this work, the architecture of the Inception-v3 and Resnet50 is provided in the keras package. By analyzing the effectiveness of the network and by utilizing the Pre-Trained network as a feature extractor, the weights are transferred to the proposed system. The optimal setup values vary while fine tuning the deep networks. In the Inception-v3 and Resnet50 the fully connected layer followed by an output layer was selected and was replaced by the MSVM and RF classifier. Implementing the microscopic images Resnet50 with the Random Forest gives the satisfactory results.

Datasets
The datasets were collected form Rajah Muthiah Dental College and Hospital (RMDC & H). The images were collected with Haemotoxylin and eosin staining with 10x magnification. A total of 350 images were accumulated from different patients. Here, 75 images were used for testing and 275 images were utilized for training.

Pre-Trained Convolutional Neural Network used for Feature Extraction

Pre-Trained Inception-v3 as a Feature Extractor
The input layer takes an image in the size of 299 x 299 x 3 and the output layer is the softmax prediction on 1000 classes. From the input layer to the last is the max pooling layer by 8x8x2048 which is referred as the feature extraction of the model, while the rest of the network is regarded as the classification of the model. In this work, inception-v3 is used to extract the features hence we arrive at 2048 feature vector for an individual image.
Pre-Trained Resnet50 as a Feature Extractor

The input layer adopts an image in the size of 224 x 224 x 3 and the output layer is the softmax prediction on 1000 classes. From the input layer to the end is the max pooling layer by 7x7x2048 which is regarded as the feature extraction of the model, while the remnant of the network is presented as the classification of the model. In this work, inception-v3 is utilized to extract the features and we conclude at 2048 feature vector for a single image.

Classification using MSVM with Inception-v3

The training process analyzes OSCC training data to find an optimal way to classify OSCC affected images in their distinctive classes. A nonlinear support vector classifier is used to discriminate the various stages. The N class classification predicament can be elucidated using N SVMs. Every individual SVM differentiates a particular class from the remaining classes (one-vs-rest approach). Support vector machine is trained to identify OSCC features of a stage from the remaining stages. Three SVMs are created for each stage. For training 275 feature vectors, each of 2048-dimension are extracted from the images. The training process analyzes the OSCC training data to find an optimal way to classify OSCC affected images into its respective stages namely, poorly differentiated (Category 0), moderately differentiated (Category 1) and well differentiated (Category 2). The derived support vector is utilized to categorize the images. For testing the 75 feature vectors, each of 2048 dimensions are given as input to the SVM model and the distance between each of the feature vectors and the SVM hyperplane is calculated. The standard interval is enumerated for each template. The average distance gives a better performance than using distance for each feature vector. The stages of OSCC are decided based on the maximum distance. The performance of OSCC stratification for Polynomial, Gaussian and Sigmoidal kernels is contrived. From the analysis, Gaussian kernel gives the best performance using MSVM classification. Table 1 shows the classification performance of OSCC images using Inception v3.

Table 1 Performance of SVM Kernel for OSCC images using Inception-v3

| OSCC Stages          | Gaussian Kernel (in %) | Polynomial Kernel (in %) | Sigmoidal Kernel (in %) |
|----------------------|------------------------|--------------------------|-------------------------|
| Poorly Differentiated| 82.67                  | 81.12                    | 80.54                   |
| Moderately Differentiated | 85.32                  | 83.21                    | 82.42                   |
| Well Differentiated  | 86.71                  | 84.11                    | 83.40                   |

From the above table, it is clear that Gaussian kernel achieves better when compared to polynomial and Sigmoidal kernel.

Classification using MSVM with ResNet50

The training process analyzes OSCC training data to find an optimal way to classify OSCC affected images into their distinctive classes. A nonlinear support vector classifier is utilized to differentiate the various stages. The N class stratification predicament can be elucidated using N SVMs. Each SVM separates a single class from all the classes (one-vs-rest approach). Support vector machine is trained to discriminate OSCC features of a stage from all the rest. Three SVMs are created for each stage. For training 275 feature vectors each of 2048-dimension are extracted from the images. The training process analyzes the OSCC training data to find an optimal way to categorize OSCC affected images into its distinctive stages, namely, poorly differentiated (Category 0), moderately differentiated...
(Category 1) and well differentiated (category 2). The derived support vector used to classify the images. For testing 75 features, vectors each of 2048 dimensions are given as input to the SVM model and the distance between each of the feature vectors and the SVM hyperplane is procured. The standard interval is deliberated for each design. The average distance gives better a performance than using distance for each feature vector. The stages of OSCC are decided based on the maximum distance. The performance of OSCC classification for Polynomial, Gaussian and Sigmoidal kernels is studied. From the analysis Gaussian kernel gives the best performance using MSVM classification. The OSCC belonging to three abnormal categories namely poorly differentiated, moderately differentiated and well differentiated. Table 2. Show the classification performance of OSCC images using Resnet50.

| OSCC Stages          | Gaussian Kernel (in%) | Polynomial Kernel (in%) | Sigmoidal Kernel (in%) |
|----------------------|-----------------------|-------------------------|------------------------|
| Poorly Differentiated | 80.00                 | 79.12                   | 77.54                  |
| Moderately Differentiated | 80.71               | 78.21                   | 75.42                  |
| Well Differentiated  | 88.00                 | 86.11                   | 83.40                  |

Classification using Random Forest with Inception v3

The training process analyses OSCC training data to find the decision tree to classify OSCC affected images into their relevant stages. Random forest is structured for a complete learning procedure for categorizing with a set of decision trees that grow randomly by selecting the sample data. A nonlinear decision tree is applied to discriminate the various stages. Random Forest is trained to ascertain OSCC features. The bootstrap procedure is followed for each training set; the samples are selected randomly. For training 275 feature vectors, each of 2048-dimensional are extracted from the images. At each point, a new subset is generated; current tree is drawn by replacement of vectors. This is called as out of bag. The training process analyses the OSCC data to categorize the OSCC affected images into its distinctive stages, namely, poorly differentiated, moderately differentiated and well differentiated. For testing 75 feature vectors each of 2048 dimensions are given as input to the Random Forest model. While testing, predictions are arrived by finding the average of the study of each decision tree.

| OSCC Stages          | Precision (in %) | Recall (in %) | F-Score (in %) | Accuracy (in %) |
|----------------------|------------------|---------------|----------------|-----------------|
| Poorly Differentiated | 92.00            | 82.00         | 86.7           | 90.00           |
| Moderately Differentiated | 60.1              | 83.00         | 69.6           | 82.00           |
| Well Differentiated  | 92.00            | 79.00         | 83.5           | 89.00           |
Classification using Random Forest with ResNet50

The training process analyses OSCC training data to identify the decision tree to categorize OSCC affected images into their distinctive classes. Random forest is structured to find a total procedure for categorization with a set of decision trees that grow randomly by selecting the sample data. A nonlinear decision tree is used to discriminate the various stages. Random Forest is trained upon to differentiate OSCC features of a stage. The bootstrap procedure is followed for each training set; the samples are selected randomly. For training 275 feature vectors, each of 2048-dimension are extracted from the images. At each node a new subset is generated and the current tree is drawn by replacement of vectors. This is called as out-of-bag. The training process analyses the OSCC data to find the OSCCC affected images into its respective stages, namely, poorly differentiated, moderately differentiated and well differentiated. For testing 75 feature vectors, each of 2048 dimensions are given as input to the Random Forest model. While testing, predictions are made by averaging the predictions of each decision tree. Based on the greater number of votes the stages are classified.

Table 4 Performance of Resnet50 with Random Forest

| OSCC Stages             | Precision (in %) | Recall (in %) | F-Score (in %) | Accuracy (in %) |
|------------------------|------------------|---------------|----------------|-----------------|
| Poorly Differentiated  | 88.00            | 95.61         | 91.60          | 94.62           |
| Moderately Differentiated | 92.15           | 88.42         | 90.00          | 93.31           |
| Well Differentiated    | 96.10            | 92.31         | 96.12          | 96.00           |

3.1 Comparative Results analysis

This section made a comparative study of two different feature extractions namely, Inception-v3 as well as Resnet50 on the Random Forest and MSVM. The results are provided in Table 5 and also in Fig. 6. From the Fig. 6, the results show that Resnet50 with Random Forest achieved comparably better results when compared with other techniques. It is noted that the highest accuracy of 92.08 % is obtained with Random Forest using Resnet50 features. Fig. 7 shows the categorized model of Resnet50 with MSVM and Random Forest. Fig 8 shows the classified model of Inception-v3 with MSVM and Random Forest.

Table 5 Overall performance of the Pre-trained model features with MSVM and Random Forest

| Pre-trained Models | Classifier | Accuracy (in %) |
|--------------------|------------|-----------------|
| Inception-v3       | MSVM       | 84.90           |
| Inception-v3       | Random Forest | 81.37          |
| ResNet50           | MSVM       | 82.90           |
| ResNet50           | Random Forest | 92.08          |
**Figure 6** Overall performances of the Pre-trained model features with MSVM and Random Forest

**Figure 7** Classified image by ResNet50 model with MSVM and RF classifier


4. CONCLUSION

In this work, a computer-aided diagnosis model is introduced which helps the pathologist in a decision-making process. A feature extraction-based classification model for the identification of various OSCC classification into various stages namely, poorly differentiated, well differentiated, and moderately differentiated is presented in detail. Here, a Pre-Trained CNN is used for extracting features using Inception V3 and ResNet50 models. Then, the extracted features are used on MSVM and Random Forest classifiers for proper identification of the OSCC classification. The above experimental analysis is verified to prove that the Resnet50 with Random Forest model gives the highest accuracy of 92.08%.

References

[1] Aubreville M, Knipfer C, Oetter N, Jaremenko C, Rodner E, Denzler J and Maier A 2017 Automatic classification of Cancerous tissue in laser End microscopy images of the Oral Cavity using deep learning, Scientific Reports, 7(1) 1-10.
[2] Cho N, Shaw J E, Karuranga S, Huang Y, da Rocha Fernandes J D, Ohno-Machado L and Malone B 2018 IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045, Diabetes research and clinical practice, 138 271-281.
[3] Ahmed CM, Narayanan M, Kalaivanesan S, Narayanan K S and deshmukh K 2017 To detect and classify oral cancer in MRI image using firefly algorithm and Expectation Maximization algorithm, International Journal of Pure and Applied Mathematics, 116(21) 149-154.
[4] Anuradha K 2018 Efficient Oral Cancer Classification Using GLCM Feature Extraction and Fuzzy Cognitive Map from Dental Radiographs, International Journal of Pure and Applied Mathematics, 15 651-655.
[5] Anuradha K and Uma K P 2017 Histological grading of Oral Tumours using Fuzzy Cognitive Map, Biomedical and Pharmacology Journal, 10(4) 1695–1700.
[6] Banumathi A, Mallika J P, Raju S and Kumar V A 2009 Automated Diagnosis and Severity Measurement of cyst in dental X – ray images using Neural Network, International Journal of Biomedical Soft Computing and Human Sciences, 14(2) 103 –107.
[7] Rasika Joat and Thakare A P 2018 Genetic Programming for Oral Cancer Detection and Secured Image Restoration: A Survey, International Journal of Advanced Innovative Technology in Engineering (IJAIITE), 3 24-27.
[8] Adi K, Widodo C E, Widodo A P, Gernowo R, Pamungkas A and Syifa R A 2018 Detection Lung Cancer Using Gray Level Co-Occurrence Matrix (GLCM) and Back Propagation Neural Network Classification, Journal of Engineering Science & Technology Review, 11(2) 8-12.
[9] Buhari M P A and Mohideen S K 2018 A combination of fuzzy positioned dental X – ray analysis model to presume the peculiar images, International Journal of Computers and Applications, 1 –6.
[10] Himadri Nath Moulick, ManasKarmakar and Poulami Das 2015 Squamous Cell Carcinoma Detection and Measurement by Image Processing, International Journal on Applications of Information and Communication Engineering, 1 50-54.
[11] Raza K and Singh N K 2018 A Tour of Unsupervised Deep Learning for Medical Image Analysis, IEEE Transactions in Medical Imaging, 35 119 –130.
[12] Mahmud M, Kaiser M S, Hussain A and Vassanelli S 2018 Applications of Deep Learning and Reinforcement Learning to Biological Data, IEEE Transaction on Neural Networks and Learning System, 29(6) 2063–2079.
[13] Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z 2016 Rethinking the inception architecture for computer vision, Proceedings in IEEE conference on computer vision and pattern recognition, 2018:2812-2821.
[14] Xia X, Xu C and Nan B 2017 Inception-v3 for flower classification, Proceedings in Second International Conference on Image, Vision and Computing (ICIVC)783-787.
[15] He K, Zhang X, Ren S and Sun J 2016 Deep Residual Learning for Image Recognition, Proceedings in IEEE Conference on Computer Vision and Pattern Recognition, 770-778.
[16] Van Den Burg G J and Groenen P J 2016 GenSVM: A Generalized Multiclass Support Vector Machine, The Journal of Machine Learning Research, 17(1) 7964-8005.
[17] Breiman L 2001 Random Forests, Machine learning, 45(1) 5-32.