

# Performance Evaluation of a Classification Model for Oral Tumor Diagnosis

Awoyelu I. O.<sup>1</sup>, Ojo B. R.<sup>1</sup>, Aregbesola S. B.<sup>2</sup>, & Soyele O. O.<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

<sup>2</sup>Department of Oral Maxillofacial Surgery and Oral Pathology, Faculty of Dentistry, College of Health Sciences, Obafemi Awolowo University, Ile-Ife, Nigeria

Correspondence: Awoyelu I. O., E-mail: iawoyelu@gmail.com

Received: October 30, 2019

Accepted: November 25, 2019

Online Published: December 20, 2019

doi:10.5539/cis.v13n1p1

URL: <https://doi.org/10.5539/cis.v13n1p1>

## Abstract

This paper extracted features from region of interest of histopathology images, formulated a classification model for diagnosis, simulated the model and evaluated the performance of the model. This is with a view to developing a histopathology image classification model for oral tumor diagnosis. The input for the classification is the oral histopathology images obtained from Obafemi Awolowo University Dental Clinic histopathology archive. The model for oral tumor diagnosis was formulated using the multilayered perceptron type of artificial neural network. Image preprocessing on the images was done using Contrast Limited Adaptive Histogram Equalization (CLAHE), features were extracted using Gray Level Confusion Matrix (GLCM). The important features were identified using Sequential Forward Selection (SFS) algorithm. The model classified oral tumor diagnosis into five classes: Ameloblastoma, Giant Cell Lesions, Pleomorphic Adenoma, Mucoepidermoid Carcinoma and Squamous Cell Carcinoma. The performance of the model was evaluated using specificity and sensitivity. The result obtained showed that the model yielded an average accuracy of 82.14%. The sensitivity and the specificity values of Ameloblastoma were 85.71% and 89.4%, of Giant Cell Lesions were 83.33% and 94.74%, of Pleomorphic Adenoma were 75% and 95.24%, of Mucoepidermoid Carcinoma were 100% and 100%, and of Squamous Cell Carcinoma were 71.43% and 94.74% respectively. The model is capable of assisting pathologists in making consistent and accurate diagnosis. It can be considered as a second opinion to augment a pathologist's diagnostic decision.

**Keywords:** classification, oral tumor diagnosis, multilayer perceptron neural network

## 1. Introduction

A tumor is an abnormal growth of cells. It is a mass of tissue that is grown abnormally, this mass of tissue could be solid or liquid like. A tumor, otherwise, known as a neoplasm or a lesion, is a general term that refers to both benign and malignant growths. It is a kind of growth that can't be controlled which affect any body part. These growths, most of the time, attacks surrounding tissues and migrates to different areas in the body. Oral tumor is the abnormal growth that starts in the mouth (oral cavity). As shown in Figure 1, the oral cavity constitutes the bony roof of the mouth (hard palate), lips, the floor of the mouth below the tongue, the inside linings of the lips and cheeks, the front two-thirds of the tongue, buccal mucosa, the gums and the teeth. The oropharynx and oral cavity helps a person to swallow, breathe, chew, talk and eat. Unimportant salivary glands, which resides in the oropharynx and oral cavity, are responsible for the saliva that makes our mouth moist so as to ease the digestion of food.

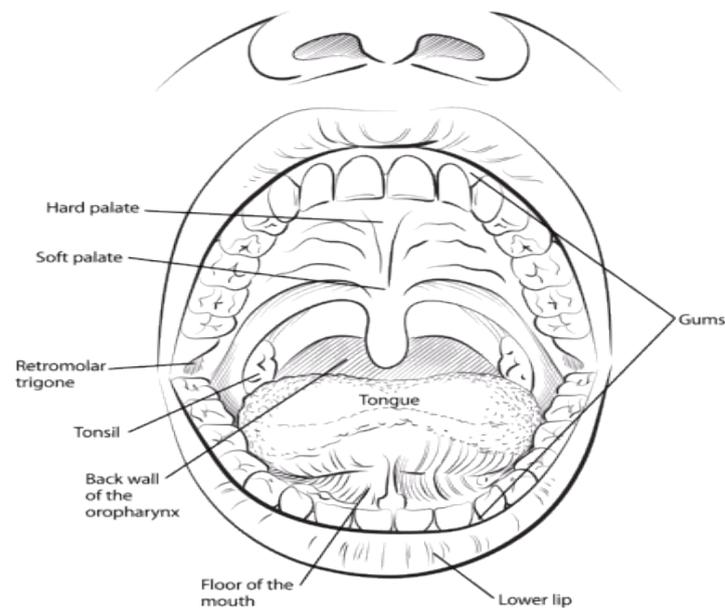


Figure 1. Illustrative Diagram of the Oral Cavity  
(Source: American Cancer Society (2016))

Oral cancer is a class of neck and head cancer, which has risen globally in occurrence and is growing critically in many regions of the world (Werning, 2011). It is caused by a basic lesion that originates from any kind of oral tissues, by metastasizing from an origin that is farther from the site, or by extending it from a neighboring anatomy specification, like nasal cavity. Oral cancer can also evolve from the tissues in the mouth. It is the sixth most common type of cancer that occurs in the world. Histopathology pictures have been used widely in the different ways to diagnose cancer whether it is normal, oral precancerous and cancer lesions. Computer-aided diagnosis (CAD) in histopathology is usually based on quantitative methods that are gotten from natural features of pictures gotten from samples that are histological. CAD allows radiologists/pathologists to use result gotten from computer as an alternative opinion in other to make decisions that will be final.

Data mining helps to analyze unprocessed information using computers to extract their patterns. It can also be described as a way of detecting unknown or potential information from an unstructured data in large volume Health care data mining helps to investigate countless alternatives from hidden patterns in datasets. These identified patterns can be used by health experts to prescribe treatments, diagnoses and prognoses for sick or indisposed patients (Milovic and Milovic, 2012).

In traditional oral cancer diagnosis, the examination and interpretation of an histopathology image is done manually by the pathologist and this depends solely on a labour-intensive process that can be affected by human factors such as fatigue, reader variation and bias to detect disease which leads to subjective diagnosis of image sample that may vary with the level of expertise of examiner (Krishnan *et. al*, 2011; Belsare and Mushrif, 2012, Vandana and Anthony, 2012; Kalkan *et al*, 2012). Computer-aided diagnoses have been developed as quantitative techniques to assist pathologists in the process of tumor diagnosis. Although, existing works have classified tumors as benign and malignant, this is not sufficient for effective definitive diagnosis. Therefore, there is the need for more detailed classification to enhance effective treatment. Hence, this paper proposes a classification model for oral tumor diagnosis that classifies oral tumor into five classes.

## 2. Existing Works

A number of works have stressed the effect of machine learning as regards disease classification. Anuradha and Sankaranarayanan (2013) performed a comparative analysis of three-feature extraction techniques for the classification of oral tumors. They used Intensity Histogram, Gray Level Co-occurrence Matrix (GLCM) and Gray Level Run Length Matrix (GLRLM) to extract features from region of interest of Dental X-rays. 50 dental x-rays were preprocessed, features were extracted and classification into benign and malignant classes were done

using Support Vector Machine. Gray Level Co-occurrence Matrix outperformed other feature extraction methods. Krishnan *et al.* (2011) proposed a model to improve classification accuracy based on textural features in the development of a computer assisted screening of Oral Sub-mucous Fibrosis (OSF). Among the specimens collected from 42 clinically diagnosed OSF patients, 20 revealed OSF with various grades of epithelial dysplasia and another 22 specimens showed OSF without Dysplasia. A hybrid paradigm that included Higher Order Spectra (HOS), Local Binary Pattern (LBP) and Laws Texture Energy (LTE) were used to extract features from the histopathological images. The feature vectors were fed to five different classifiers, namely: Decision Tree, Sugeno Fuzzy, Gaussian Mixture Model, K-Nearest Neighbour and Radial Basis Probabilistic Neural Network to select the best classifier. The combination of Fuzzy classifier and HOS gave the best result. Xu *et al.* (2014) developed an integrated framework that could simultaneously perform image level classification, medical image segmentation and patch level clustering for histopathology images by embedding the concept of clustering into Multiple Instance Learning (MIL). Weakly supervised learning direction, which is similar to semi supervised learning was taken, in which a training set consists of a number of bags. Each bag includes many instances and the goal is to predict both bag and instance level labels while only bag-level labels are given in training. Result obtained showed that the supervised learning method is more accurate in comparison to the weakly supervised method.

Amin *et al.* (2019) worked on brain tumor detection. Weiner filter with different wavelet bands were used for lesion enhancement and different statistical methods for brain tumor segmentation. Global threshold and different mathematical morphology operations were used to isolate the tumor region in Fluid Attenuated Inversion Recovery (Flair) and T2 Magnetic Resonance I (MRI). For classification, Local Binary Pattern (LBP) and Gabor Wavelet Transform (GWT) features were fused. The results obtained showed peak signal to noise ratio value of 76.38, mean squared error of 0.037 and structured similarity index (SSIM) of 0.98 on T2 and 76.2, 0.039 and 0.98 on Flair respectively. Ahmad, Ghuffar and Khurshid (2019) worked on the classification of breast cancer histology images. Transfer learning was used on the pre-trained networks: AlexNet, GoogleNet and ResNet50 architectures using augmented patch dataset. Patch-wise classification accuracy of four classes was calculated by creating twelve non-overlapping patches from the whole slide images of the total size of 512 x 512. Image-wise classification accuracy was calculated by passing the twelve non-overlapping patches extracted from open slide images and image label was decided on the basis of the majority voting rule among the twelve. The accuracy obtained for patch-wise classification were 79.84%, 81.07% and 83.60 % for AlexNet, GoogleNet and ResNet50 respectively while for image-wise classification, the accuracy obtained were 82.30%, 83.6% and 85.0% for AlexNet, GoogleNet and ResNet50 respectively.

Most of the existing methods are limited to a 2-way classification of histopathology images and are more profound in other parts of the body. This paper provides 5-way classification of oral histopathology images.

### 3. Proposed Model

This paper used supervised machine algorithm for the classification of oral tumors. Figure 2 shows a diagram of the methodology framework that was used. The proposed model makes use of oral histopathology images to train artificial neural network model. The images were preprocessed to acquire an even contrast and filter out noise that arose from image acquisition. 15 Gray Level Co-occurrence Matrix (GLCM) features were extracted from histopathology images, sequential feed forward algorithm was used to select the most relevant features from the pool of features. These selected features were fed as inputs into artificial neural network (Multi-layer perceptron) model to classify oral tumor as Ameloblastoma, Mucoepidermoid carcinoma, Pleomorphic adenoma, Giant cell lesion and Squamous cell carcinoma.

The model made use of Contrast Limited Adaptive Histogram Equalization (CLAHE) for the enhancement of histopathology images. Otsu threshold for segmentation of background from the area of interest (foreground), feature extraction was carried out using Gray Level Co-occurrence Matrix (GLCM).

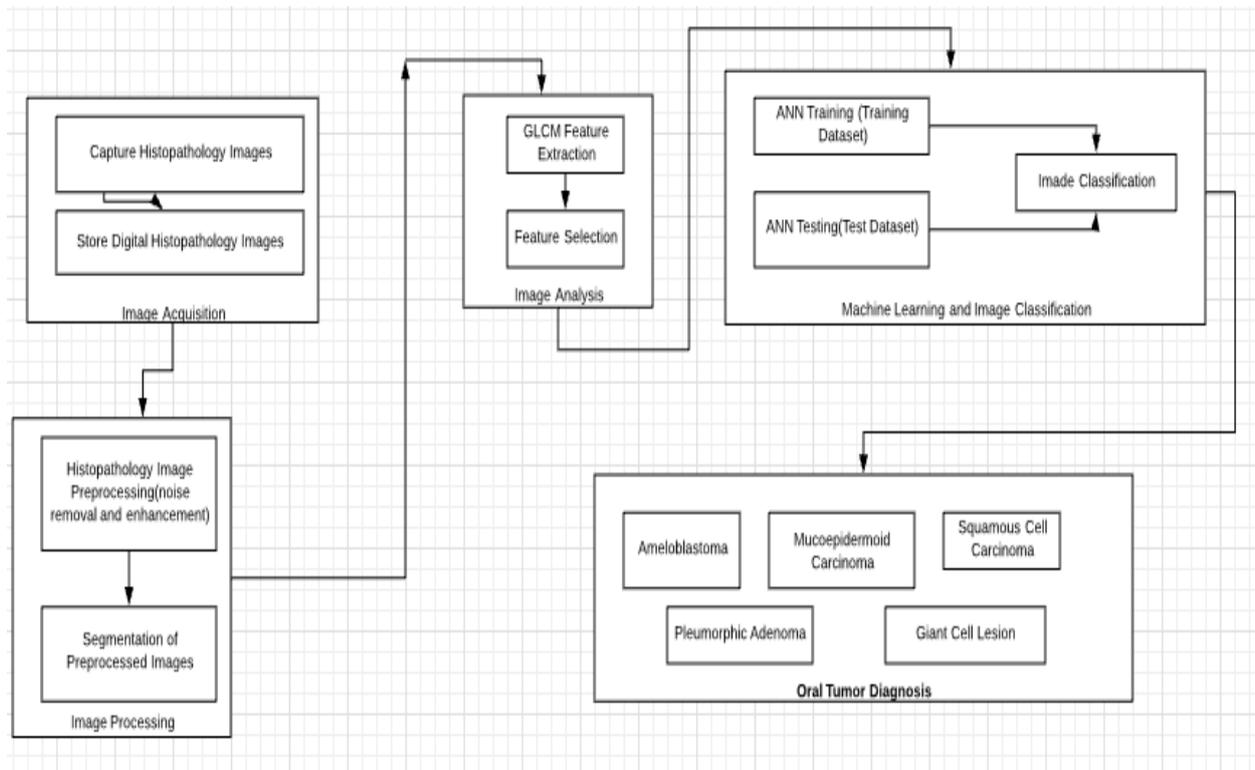


Figure 2. Proposed Framework for Classification of Oral Tumors

### 3.1 Data Acquisition

The input to the proposed model are digital histopathology images. These images were obtained from the histopathology archive of the Dental Clinic, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria. Archived histopathology slides were mounted on Leica dm500 microscope, viewed by an oral pathologist with a magnification of 40x and captures with a 13megapixel Samsung camera. An image sample from the dataset is shown in Figure 3.

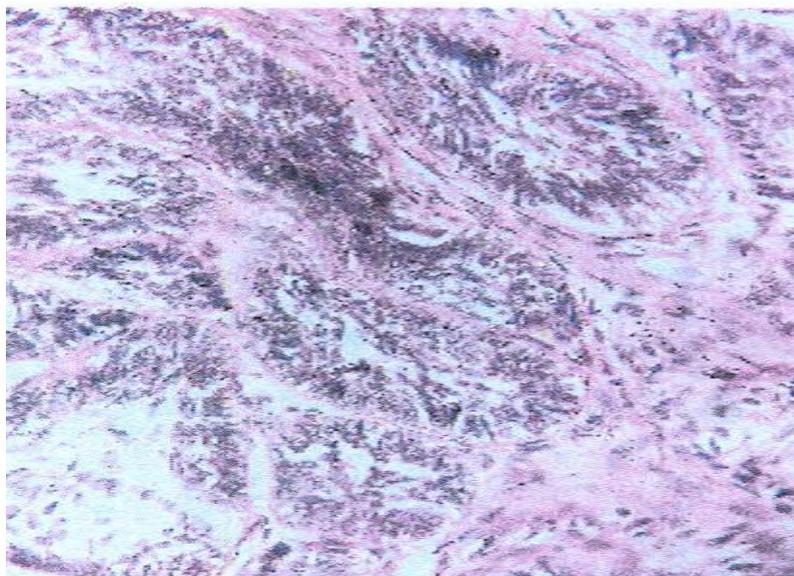


Figure 3. A Sample Oral Histopathology Image from the Dataset (Histomicrograph)

### 3.2 Image Processing and Segmentation

An image enhancement technique was done to process an image so that the resulting image gives more visual information than the original image. The first stage in image processing is the image preprocessing. This was done in MATLAB simulation environment using image processing toolbox. CLAHE algorithm was used. The CLAHE algorithm partitions the images into contextual regions and applies the histogram equalization to each one. This evens out the distribution of used gray values and thus makes hidden features of the image more visible. The initial image from dataset, image after Edge detection and image after CLAHE is as shown in Figure 4. Dealing with images, it is important to differentiate the background from the implicating areas. The methods used for this are segmentation techniques. There are different methods of thresholding used for identification of cell structure from the background area as an object in an image. Otsu threshold method was used.

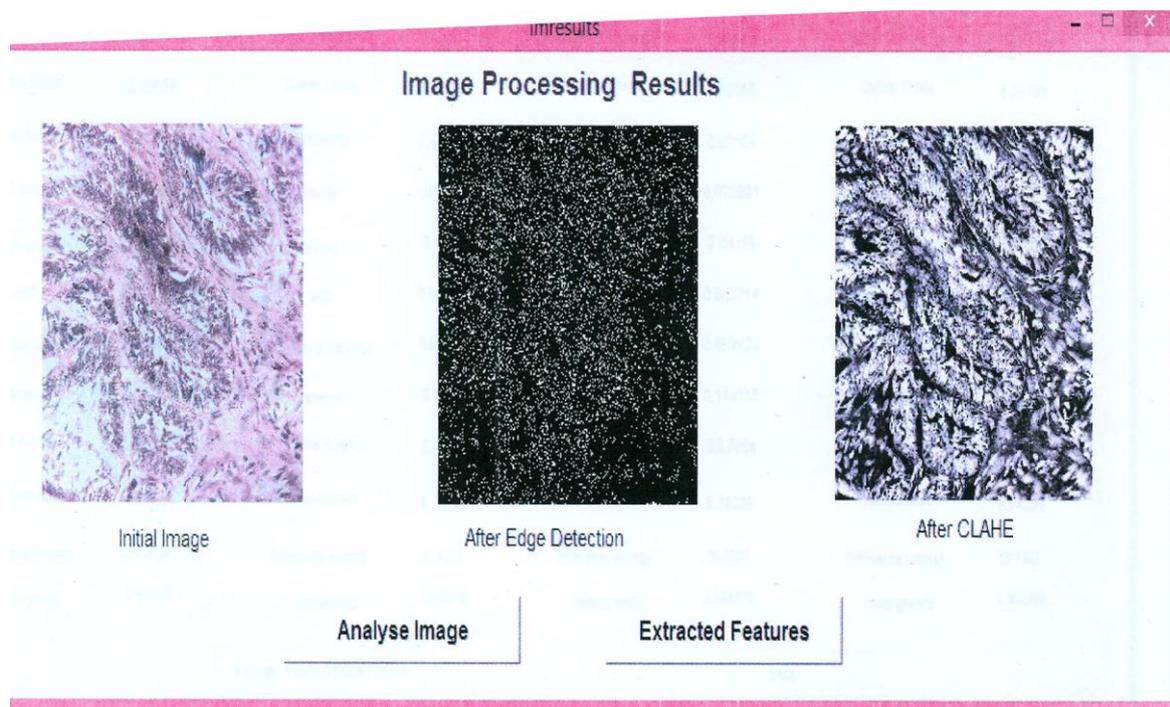


Figure 4. Image Preprocessing

### 3.3 Model Formulation

The dataset used was 100 histopathology images. Data was collected with respect to the specific cases as earlier diagnosed by an oral pathologist. Table 1 indicates the number of images collected with respect to each class.

Table 1. Data Collection Summary

Ameloblastoma	Giant Cell Lesions	Pleomorphic Adenoma	Squamous Carcinoma	Cell	Mucoepidermoid Carcinoma
25	20	15	25		15

### 3.4 Adopted Multilayer Perceptron Approach

Each neuron receives inputs, which are selected features extracted from oral histology images denoted as  $X = x_1, x_2, x_3, \dots, x_n$ , here  $n$  is the number of inputs and it is passed into the neuron on the next layer via its corresponding weights:  $w_1, w_2, w_3, \dots, w_n$ .

The neuron's activation is computed by summing all the products of the inputs and weights by the function in Equation 3.1.

$$O_j = \sum_{k=1}^n w_{kj} x_k + b \quad (3.1)$$

Where  $b$  is the bias that is used to stabilize the node;  $w_{ij}$  is the weight connecting node  $i$  to node  $j$ ; and  $x_i$  refers to the input at node  $i$ .

The output of the neuron  $y_j$  is the outcome of the activation function  $\phi$  on the value of  $O_j$  as denoted in Equation 3.2.

$$y_j = \phi(O_j) = \frac{1}{1+e^{-O_j}} \quad (3.2)$$

The discrepancy between the expected output ( $p$ ) and the actual output ( $a$ ) is deduced using Equation 3.3.

$$e = (p - a) \quad (3.3)$$

The error determines to what degree the weight should be adjusted. If  $e = 1$ ; weight is adjusted back using Equation 3.4 until convergence ( $e=0$ ).

$$\Delta W_{ij} = -\alpha \frac{\partial e}{\partial w_{ij}} \quad (3.4)$$

The Equation 3.4 is interpreted as the change in weight with respect to input rate of change of  $e$  with respect to change in weight with respect to input  $i$  at node  $j$ .

Where  $\Delta W_{ij}$  is the change of weight from node  $i$  to node  $j$ ;  $\alpha$  is the learning rate; and  $\frac{\partial e}{\partial w_{ij}}$  is the rate of error

change with respect to the weight change from node  $i$  to node  $j$ .

Equation 3.4 is used by the back propagation algorithm to adjust the value of the weight attached to the inputs at each neuron.

### 3.5 Model Simulation

The images loading, preprocessing, feature extraction, feature selection as well as the classification were simulated using Matlab R2013a with image processing and statistics toolboxes. The inputs were histopathology images obtained from Obafemi Awolowo University Dental Clinic, Ile-Ife, Nigeria. 72 images were used as the training set while 28 images were used for testing. After an image is loaded, the image was converted to gray scale. Edge detection algorithm was applied to the image to filter out information that are less relevant and reduce the amount of data to be processed while preserving the important structural properties of the image. The image is processed by an image enhancement algorithm called CLAHE that performs both noise reduction and contrast adjustment. The enhanced images were subjected to GLCM - a feature extraction algorithm. The more discriminating features were selected by the feature extraction algorithm. Features extracted by GLCM provide information about the frequency of occurrence of two neighbouring pixel combination in an image. They provide information concerning image texture heterogeneity and coarseness, which is not perceived visually. The textural features at four different angles  $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$  were extracted from the images as shown in Table 2 and Table 3 shows the weighted value from the 4 angles at which the features were computed.

Table 2. GLCM Features Extracted at Different Angles

Feature	GLCM Feature @ 0°	GLCM Feature @ 45°	GLCM Feature @ 90°	GLCM Feature @ 135°
MC1	23.6588	23.5186	23.655	23.5762
Contrast	0.358059	0.641811	0.370189	0.526623
Correlation	0.962173	0.932194	0.96089	0.944363
Cluster Prominence	636.755	619.337	636.178	626.177
Cluster Shade	8.50828	8.20353	8.52005	8.27701
Dissimilarity	0.287587	0.432461	0.29564	0.380943
Energy	0.0739652	0.0596327	0.072891	0.0639741
Entropy	2.87102	3.10198	2.88498	3.02523
MC2	0.866218	0.811024	0.862714	0.829104
Difference Variance	0.863074	0.803538	0.859422	0.823435
Variance	0.115156	0.10222	0.114115	0.105837
Sum average	23.7006	23.7039	23.7054	23.7057
Sum variance	8.74184	8.74226	8.74239	8.74227
Difference entropy	56.4634	55.4316	56.3889	55.7307
Homogeneity	0.994646	0.990696	0.994472	0.992289

### 3.6 ANN Classification

72 histopathology images were used to train the Multilayered Perceptron Neural Network classification machine. 28 images were used for testing. Table 4 gives the values of the training and testing dataset. The classification results are as shown in the Confusion Matrix in Table 5.

Table 3. Weighted GLCM Features

Weighted Average GLCM	Feature
MC1	23.6021
Contrast	0.47417
Correlation	0.949905
Cluster Prominence	629.611
Cluster Shade	8.37722
Dissimilarity	0.349158
Energy	0.0676158
Entropy	2.9708
MC2	0.842265
Difference Variance	0.837367
Variance	0.109589
Sum average	23.7039
Sum variance	8.74219
Difference entropy	56.0036
Homogeneity	0.993021

Table 4. Values of Training and Testing Dataset

	Ameloblastoma	Giant Lesions	Cell	Pleomorphic Adenoma	Squamous Cell Carcinoma	Mucoepidermoid Carcinoma
Training set	18	14		11	18	11
Testing set	7	6		4	7	4

Table 5. Classification Results

	Ameloblastoma	Giant Cell Lesions	Pleomorphic Adenoma	Squamous Cell Carcinoma	Mucoepidermoid Carcinoma
<b>Ameloblastoma</b>	6	0	0	1	0
<b>Giant Cell Lesions</b>	0	5	1	0	0
<b>Pleomorphic Adenoma</b>	1	0	3	0	0
<b>Squamous Cell Carcinoma</b>	1	0	0	5	1
<b>Mucoepidermoid Carcinoma</b>	0	0	0	0	4

From Table 5, there were 6 cases classified correctly as Ameloblastoma, 5 cases correctly classified as Giant Cell Lesions, 3 cases correctly classified as Pleomorphic Adenoma, 5 cases correctly classified as Squamous cell

Carcinoma and 4 cases correctly classified as Mucoepidermoid Carcinoma. Total instances correctly classified were 23.

2 instances were incorrectly classified as Ameloblastoma, no instance of misclassification of Giant Cell Lesions, 1 instance of misclassification of Pleomorphic Adenoma, 1 instance of misclassification of Squamous Cell Carcinoma and 1 instance of misclassification of Mucoepidermoid Carcinoma. Total misclassified instances were 5 instances.

### 3.7 Performance Evaluation

The performance of the model was evaluated to ascertain the model’s accuracy using the metrics Specificity and Sensitivity. Specificity is the probability of a negative diagnosis test among people that do not have the disease and it defined as:

$$Specificity = \frac{\text{Number of True Negatives}}{\text{Number of True Positives} + \text{Number of False Negatives}}$$

while Sensitivity is the probability of a positive diagnosis test among people that have the disease and it is defined as:

$$Sensitivity = \frac{\text{Number of True Positives}}{\text{Number of True Positives} + \text{Number of False Negatives}}$$

The Specificity and Sensitivity values are as shown in Table 6. The model’s accuracy is the overall true positives divided by the whole testing instance. The accuracy of the model is 82.14%. The graph of the Specificity and Sensitivity values for the five classes is as shown in Figure 5.

Table 6. Specificity and Sensitivity values for each Class

	Ameloblastoma	Giant Cell Lesions	Pleomorphic Adenoma	Squamous Cell Carcinoma	Mucoepidermoid Carcinoma
Sensitivity (%)	85.71	83.33	75.00	71.43	100
Specificity (%)	89.47	94.74	95.24	94.74	100

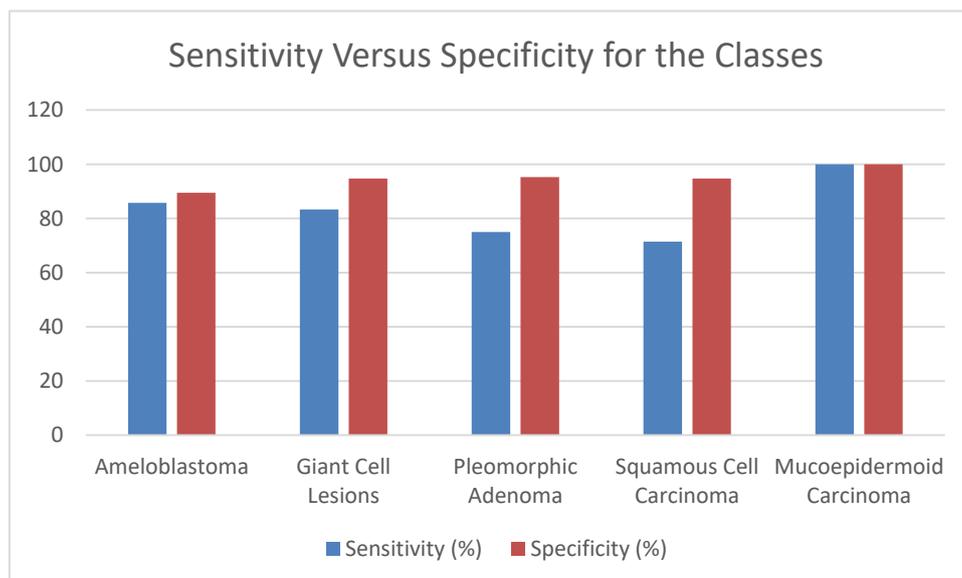


Figure 5. Sensitivity and Specificity Values for Classification of Oral Tumor Diagnosis

#### 4. Conclusion

The proposed model acts as a second opinion to pathologists in decision and diagnostic process. It also improves the diagnostic accuracy and consistency of histopathology image interpretation. In conclusion, the developed model has contributed to the body of knowledge on the classification of images to help reduce subjective interpretation of histopathology images. Histopathology images contain colors, therefore future work should take into consideration the color variations and extract features based on this.

#### References

- Ahmad, H. M., Ghuffar, S., & Khurshid, K. (2019). Classification of Breast Cancer Histology Images Using Transfer Learning. *16<sup>th</sup> IEEE International Bhurban Conference on Applied Sciences and Technology*. <https://doi.org/10.1109/IBCAST.2019.8667221>
- American Cancer Society. (2016). *Oral Cavity and Oropharyngeal Cancers Detailed Guide*. American Cancer Society. Atlanta, Ga.
- Amin, J., Sharif, M., Raza, M., Saba, T., & Anjum, M. A. (2019). Brain Tumor Detection Using Statistical and Machine Learning Method. *Elsevier Journal of Computer Methods and Programs in Biomedicine*, *177*, 69-79. <https://doi.org/10.1016/j.cmpb.2019.05.015>
- Anuradha, K., & Sankaranarayanan, K. (2013). Comparison of Feature Extraction Techniques to Classify Oral Cancers using Image processing. *International Journal of Application on Innovation in Engineering and Management (IJAIEM)*, *2*(6), 456-462.
- Belsare, A. D., & Mushrif, M. M. (2012). Histopathological Image Analysis using Image Processing Techniques: An Overview. *Signal and Image Processing*, *3*(4), 23-30. <https://doi.org/10.5121/sipij.2012.3403>
- Kalkan, H., Nap, M., Duin, R. P., & Loog M. (2012). Automated Classification of Local Patches in Colon Histopathology. *In the International Conference Proceedings of Pattern Recognition (ICPR)*, pp. 61-64.
- Krishnan, M. M. R., Choudhary, A., Chakraborty, C., Ray, A. K., & Paul R. R. (2011). Texture based Segmentation of Epithelial Layer from Histopathological Images. *Micron*, *42*(6), 632-641. <https://doi.org/10.1016/j.micron.2011.03.003>
- Milovic, B., & Milovic, M. (2012). Prediction and Decision Making in Health Care using data Mining. *Kuwait Chapter of the Arabian Journal of Business and Management Review*, *1*(12), 126-136. <https://doi.org/10.11591/ijphs.v1i2.1380>
- Vandana, B. S., & Anthony, P. J. (2012). Automated Segmentation using Histopathology Images as a Diagnostic Confirmatory Tool in Detection of Bone Cancer. *International Journal of Computer Applications*, *57*(11), 45-55.
- Werning, J. W. (2007). *Oral Cancer: Diagnosis, Management and Rehabilitation*. Thieme Medical Publishers, New York. <https://doi.org/10.1055/b-002-85470>
- Xu, Y., Zhu, J. Y., Eric, I., Chang, C., Lai, M., & Tu, Z. (2014). Weakly Supervised Histopathology Cancer Image Segmentation and Classification. *Medical Image Analysis*, *18*(3), 591-604. <https://doi.org/10.1016/j.media.2014.01.010>

#### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).