International Journal of Future Generation Communication and Networking
Vol. 14, No. 1, (2021), pp. 1330–1340

Detection of Oral Cancer Using Probabilistic Neural Network

*1. **Ramya.R**, Assistant Professor, Sri Ramakrishna Engineering College,Coimbatore-641022, ramya.ece@srec.ac.in*

*2. **Vijay.S**, UG Student, Sri Ramakrishna Engineering College,Coimbatore-641022, vijay.1702245@srec.ac.in*

*3. **Surya Prasath**.M, UG Student, Sri Ramakrishna Engineering College,Coimbatore-641022, suryaprasath.1702234@srec.ac.in*

*4. **Vigneshwaran.K**, UG Student, Sri Ramakrishna Engineering College,Coimbatore-641022, vigneshwaran.1702244@srec.ac.in*

Abstract:-

The number of people being diagnosed positive with oral cancer and its death rate is rapidly increasing every day. This paper aims at enhancing the efficiency of oral cancer diagnosis. The proposed method will detect and classify the affected cancer cell in the oral region using digital Image processing techniques. Initially, the data set images are pre-processed and its contrast is enhanced thus making the images a lot more effective for the other processes involved in detecting oral cancer. Discrete Wavelet Transform (DWT) and Gray Level Co-occurrence Matrix (GLCM) are used for extraction of higher order gradient features for the particular Region Of Interest(ROI) which is determined using the Extraction Maximization Image segmentation process. Maximum Gaussian Mixture model(MGMM) is choose as the pixel classifier. Probabilistic Neural Network (PNN) is used as a classifier here to predict and classify whether the input image is diagnosed with Oral cancer or not.

Keywords: Image processing, Image segmentation, DWT, Feature extraction, PNN classifier, Gradient Features & GLCM.

I. INTRODUCTION

Oral cancer comment on the tumor that fall out in the mouth (oral cavity)region which is a major threat if it is not diagnosed and left untreated. Globally, the cases of oral cancer in India reckon to 86% which is found irrespective of gender. This cancer ruins the life of thousands of people every year also those who have the habit of smoking and chewing tobacco are more prone to this. The oral cavity encompass of gums, teeth, lip, tongue, cheeks, roof and floor of the mouth. Detecting oral cancer becomes onerous and has the least perpetuity rate in compared to all other cancers. Diagnosis of this remains to be provocation to dentist.

Oral cancer cells are classified using SVM,Naive Bayes and CNN. The mean, intensity,centroid, area and perimeter of different blobs in image are calculated [1].Six different SVM classifiers are used to find the breast cancer cells from the image[2]. Various methods of oral

ISSN: 2233-7857 IJFGCN

cancer diagnosis and its growth, also machine learning methods available to predict are given [3,4]. A review of different image segmentation methods are given in [5]. Colour image segmentation based on region, edge and contour is provided in [6,7]

Alfonso Rojas Dominguez used restricted area rising segmentation methods in [9], and Casico used threshold for diagnosing breast cancer in [11]. The number of ROIs increases in this case as the density of breast tissue increases. Hadhoud(2005) improved the image using mathematical morphology methods involving the Top hat algorithm and structural elements such as low pass Gaussian filters (LPGF), which allowed for the extraction of even very fine information [10] . Brijesh used Region Based Segmentation to extract margin, abnormality assessment rank, and other data from the area extraction point [12].

The grey level co-occurrences matrix (GLCM) is a widely used texture-based feature extraction algorithm that assigns the textural relationship between pixels images. The GLCM's key concept is to conduct operations in the picture based on second-order statistics [13, 14].Neural network to let on the oral lesions were done classifying initially with ResNet-101 and lesions are detected with the Faster R-CNN [15]. In study, A. Copeland et al. found a clear link between synthetically derived textures and how different textures are interpreted by human observers [16]. Different feature extraction techniques like Gray Level Co-occurrence Matrix (GLCM), Intensity Histogram and Gray Level Run Length Matrix (GLRLM) is comared in [17]. 2D-Discrete Wavelet Transform is exploited to bring out the image features [18].

We introduces the categorization of typical and atypical portions in the oral images. The image input used is pre- treated applying isotropic Gaussian filter. Once the image enhancement is done, the tumor portion is fragmented and the feature extraction is performed by Grey Level Co-occurrence Matrix (GLCM). Probabilistic Neural Network is used in identifying the cancer cells accurately from large set of input images. Maximum Gaussian Mixture model(MGMM) is choose as the pixel classifier, consecutively solved by Expectation Maximization algorithm.

II.PROBABILISTIC NEURAL NETWORK FOR ORAL CANCER DETECTION.

The multiple features necessary to discriminate benign from malignant tumor cells is investigated in the proposed model. This research employs a Gaussian filter in pre-processing to eliminate image noise while simultaneously improving image quality. Unnecessary noise can be pull out employing the proposed algorithm it and enhances image consistency by ensuring that all edges are visible.

ISSN: 2233-7857 IJFGCN

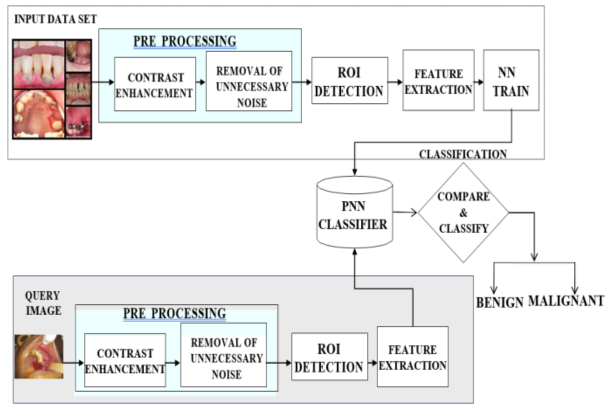

Fig.1 Proposed model for Oral cancer detection

A. Image Pre-processing:

The images are often contaminated with noise whatever be the application we use. The diagnosis accuracy would be harmed by these noises. Each image is requisite to be pre- processed in the view of decreasing the likelihood of errors and increase processing speed. Contrast enhancement, re-sizing such as cropping and scaling, and other operations are performed during the pre-processing stage. The pre-processing 2 – D isotropic Gaussian filter is represented as follows.

$$G(x, y) = (0.5\prod \sigma^2)e^{-(\frac{x^2+y^2}{2\sigma^2})} \qquad (1)$$

where G(x,y) is the Gaussian distribution's standard deviation, x and y are the horizontal and vertical distance from origin along the axis. A convolution matrix is constructed from the Gaussian distribution values and added to the original image.

B. Image segmentation:

Image segmentation helps in grouping pixels into clusters based on pixel similarity. Maximum Gaussian Mixture model(MGMM) [19] is used a a classifier. Expectation Maximization (EM) algorithm plays the role of segmentation in this paper. It helps to solve MGMM. It identifies the Regions of Interest in the pre-processed image. The algorithm is implemented in two steps. The E-step is expectation step which gives the expectation of the unknown distribution based on the observed data. The next one is M-step (Maximization step) which maximizes the E-step parameters. The MGMM is defined as,

$$p(x/\theta) = f(x) = \frac{1}{s}\max(N(x \mid \mu_k, \Sigma_k)) \qquad (1)$$
$$I = \{1,2,....K\}$$

Where N,μ_k, Σ_k represents the multivariate Gaussian distribution, mean and covariance matrix respectively. The likelihood function (X;θ) or its logarithm (X; θ), has the following parameters.

The approximate estimates $p_k^{(i)}, m_k^{(i)}, \sigma_k^{(i)}$ are available, where $p_k^{(i)}, m_k^{(i)}, \sigma_k^{(i)}$ represent probability, mean, and standard deviation at the i^{th} iteration, respectively. Then, by first using the old estimates to create a lower bound bi() for the probability function, and then maximizing the bound with reference to p_k, m_k, σ_k, better estimates $p_k^{(i+1)}, m_k^{(i+1)}, \sigma_k^{(i+1)}$ can be calculated. Expectation Maximization (EM) begins with the parameters' initial values of $p_k^{(0)}, m_k^{(0)} \& \sigma_k^{(0)}$. and performs the 'E' and 'M' steps iteratively before convergence. The logarithm expectation is considered to be the bound, the construction of bi() is known as the "E step." The "M step" refers to the bi() maximization that results in the latest approximations $p_k^{(i+1)}, m_k^{(i+1)}, \sigma_k^{(i+1)}$.

EM iterates the following computations until the likelihood function convergence to a confined limit, given an initial estimate $p_k^{(0)}, m_k^{(0)}, \sigma_k^{(0)}$. The Gaussian distribution function is represented by g(.).

Expectation step

$$p^{(i)}(k|n) = \frac{p_k^{(i)} g(x_n; m_k^{(i)}, \sigma_k^{(i)})}{\sum_{m=1}^{K} p_k^{(i)} g(x_n; m_k^{(i)}, \sigma_k^{(i)})} \quad (2)$$

Maximization step

$$m_k^{(i+1)} = \frac{\sum_{n=1}^{N} p^{(i)}(k|n) x_n}{\sum_{n=1}^{N} p^{(i)}(k|n)} \quad (3)$$

$$\sigma_k^{(i+1)} = \sqrt{\frac{1}{D} \frac{\sum_{n=1}^{N} p^{(i)}(k|n) \left\| x_n - m_k^{(i+1)} \right\|^2}{\sum_{n=1}^{N} p^{(i)}(k|n)}} \quad (4)$$

$$p_k^{(i+1)} = \frac{1}{N} \sum_{n=1}^{N} p^{(i)}(k|n) \quad (5)$$

C. Feature extraction: The process of extracting quantitative information from an image, such as color features, texture, shape, and contrast, is known as feature extraction. In this paper, the Discrete Wavelet Transform (DWT) was used to extract wavelet coefficients, and the gray-level co-occurrence matrix (GLCM) was used to extract statistical features.

1. Using DWT:

To analyze the different frequencies available in an image wavelet is used as a tool. It gives the result in time-frequency domain representation. It uses different windows for low and high frequencies. It breaks the image frequencies into its sub-band using a set of filter bank.

A set of shifted and dilated wavelet functions $\varphi^{LH}, \varphi^{HL}, \varphi^{HH}$and scaling functions φ LL form an orthonormal basis for $L^2(R^2)$ in the 2D-DWT. An image x(s, t) of NxN is decomposed as{ LH,UL,HH}, and $N_j = N/2^j$given a J-scale DWT. Wavelet or DWT sub-bands are LH, HL, and HH in this case.

ISSN: 2233-7857 IJFGCN

$$x(s,t) = \sum_{k,i=0}^{N_j-1} e_{J,k,i} \emptyset^{LL} J,k,i(s,t) + \sum_{B\in B} \sum_{j=1} \sum_{k,i=0}^{N-1} w^B j,k,i \varphi^B j,k,i(s,t)$$

$$\emptyset^{LL} j,k,i(s,t) \qquad (6)$$

$$= 2^{-j/2} \emptyset\left(2^{-j}s - k, 2^{-j}t - i\right), \varphi^B j,k,i(s,t), \varphi^B j,k,i(s,t)(2)$$

$$= 2^{-j/2} \varphi^B\left(2^{-j}s - k, 2^{-j}t - i\right) B \in B, B \qquad (7)$$

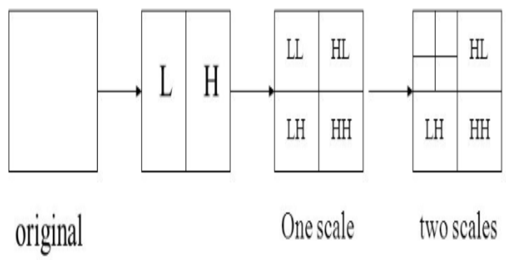

Fig.2: Discrete Wavelet Transform sub-bands

2. GLCM:

A gray level co-occurrence matrix (GLCM), $P(i,j)$ confers the statistical texture features like energy, contrast, entropy, correlation and homogeneity in the image. Energy denotes the textual uniformity among the image pixels. The matrix size will be same as the gray levels in the image. GLCM gives the association between two nearby pixel, of those one act as the reference and the other becomes the adjoining pixel. The GLCM matrix P, gives the distance the pixel of interest and the adjoining pixel. The formula for different statistical texture features is given below.

Energy,

$$\mu = (1/MN) * \sum_{i=1} \sum_{j=1} P(i,j) \qquad (8)$$

Contrast,

$$f_{2=} \sum_{n=0}^{N_9-1} n^2 \left\{ \sum_{i=0}^{N_9-1} \sum_{j=0}^{N_9-1} Pd,\theta(i,j) \right\}, where\ n = |i-j| \qquad (9)$$

Entropy,

$$f_{3=} \sum_{i=0}^{N_9-1} \sum_{j=0}^{N_{9-1}} Pd,\theta(i,j) \log(Pd,\theta(i,j)) \qquad (10)$$

ISSN: 2233-7857 IJFGCN

Correlation,

$$f_{5=}\ \sum_{i=0}^{N_9-1}\sum_{j=0}^{N_{9-1}} Pd,\theta(i,j)\frac{(i-\mu_x)(j-\mu_y)}{\sigma_x\sigma_y} \qquad (11)$$

Homogeneity,

$$\sum_{i,j=0}^{N-1}\frac{P_{ij}}{1+(i-j)^2} \qquad (12)$$

D. Probabilistic neural network (PNN)

For applications like categorization and pattern recognition probabilistic neural network(PNN) gives an optimum solution. Tn this algorithm Parzen window and a non-parametric function is used to approximated the parent probability distribution function(PPDF) of each class. PPDF of the previous class is assisted to estimate the PPDF of new data set. The maximum posterior probability of new data input is obtained by Bayes rule. By this way, wrong classification is avoided. PNN works with multilayered feed forward network such as Input layer,Hidden layer, Pattern layer/Summation layer and Output layer.

III. EXPERIMENTAL RESULTS

The system's efficiency is assessed using a variety of metrics. Classification Accuracy (AC) and Mathews Correlation Coefficient are the metrics used (MCC). The Confusion Matrix is used to measure these numbers. A uncertainty matrix provides details about a classification system's current and expected classifications. To assess the behaviour of the given structures this matrix data is used. The confusion matrix for a two-class classifier is shown below.

Actual	**Predicted**	
	Negative	**Positive**
Negative	TN	FN
Positive	FP	TP

Table.1: Two-class classifier

TN, FN, FP and TP corresponds to True negative, False negative, False positive and True positive respectively. The classification accuracy of the system is given by,

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \qquad (13)$$

Matthews correlation coefficient (MCC) plays the role of appraising the consistency of binary classifications. It pay back a numeric between 1 and 0. This gives a correlation coefficient value which gives the mutuality between observed and expected quantity. The coefficients +1, 0 illustrates a ideal forecast, random prediction.

$$MCC = \frac{\text{TP x TN} - \text{FP x FN}}{(\text{TP}+\text{FP})\,(\text{TP}+\text{FN})\,(\text{TN}+\text{FP})\,(\text{TN}+\text{FN})}$$

MCC = TP x TN – FP x FN (13)

The competence to accurately classify oral cancer cells from the data set of images is termed as sensitivity.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP}+\text{FN}} \qquad (14)$$

The competence to classify the parts which do not match the condition is termed to be specificity of the system.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN}+\text{FP}} \qquad (15)$$

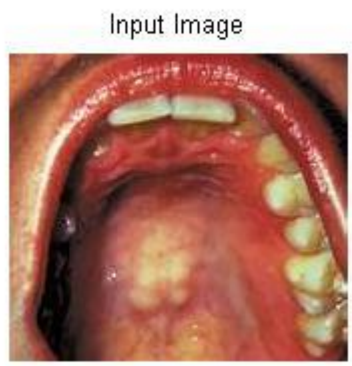

Fig.3 Input Image

Fig 4. After preprocessing and converted gray scale image of pixel resolution 255 x 255

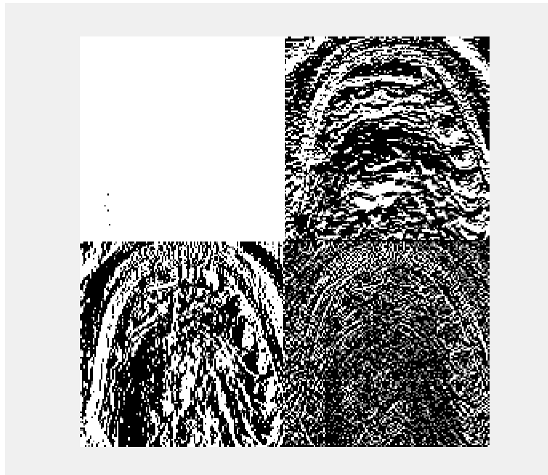

Fig 5. Region of Interest and segmentation

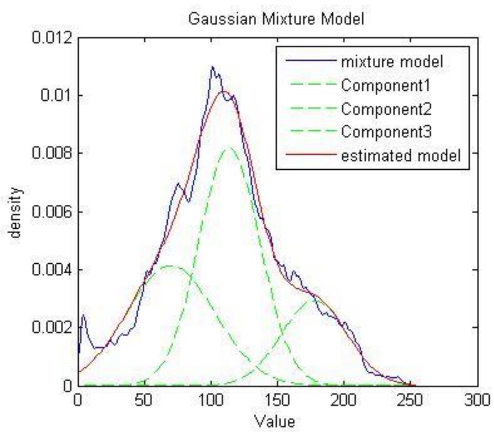

Fig 6. Gaussian Mixture Model

In fig.6, this component 1 is the mean (μ) , component 2 is the Standard Deviation(σ) obtained from the Co-variance matrix, component 3 is the data points in the image set.

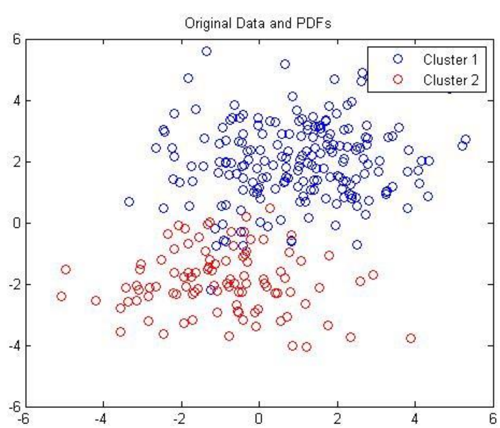

Fig 7. Data points from original data set

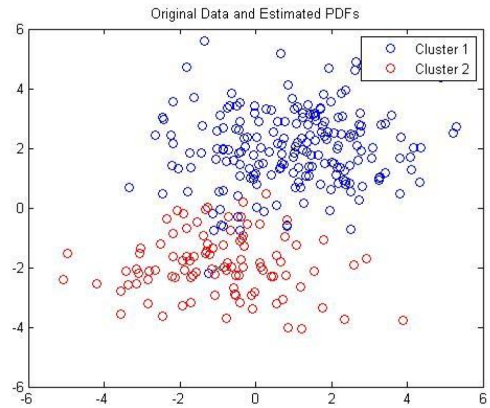

Fig 8. Data points from Estimated model

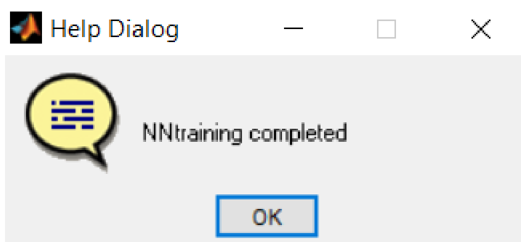

Fig 9. Completion of PNN training

Parameters	**Value**
TP value	1
TN Value	1
FP Value	1
FP Value	0
Sensitivity	100
Specificity	50
Accuracy	93.3333

Table.2 : PNN parameters

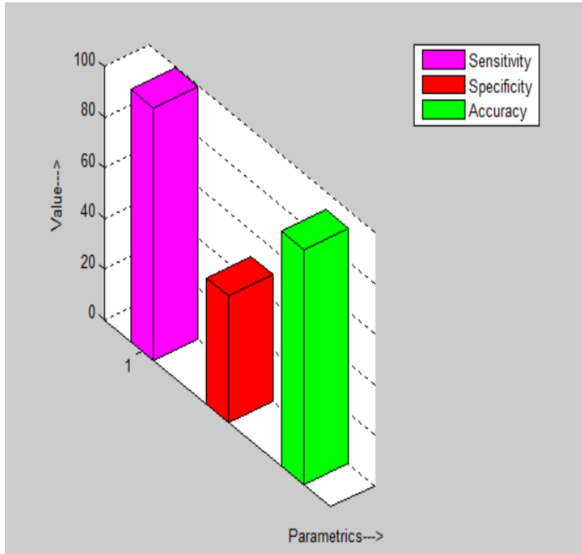

Fig 10. Comparison of PNN parameters

IV. CONCLUSION

The method we proposed gives us an accuracy of 93.3% for the spotting of oral cancer which proves that the suggested model is faster and more accurate than the clinical tests or other machine learning models in diagnosing oral cancer in the premature stages. The collected image dataset was pre-processed using the Expectation Maximation algorithm and Gaussian mixture model which aided in edge detection and noise removal. The image features were extracted with the assist of DWT(Discrete Wavelet Transform) and GLCM(Grey Level Co-Occurrence Matrix). We employed the PNN(Probabilistic Neural Network)classifier for benign or malign oral cancer classification and the entire system was successfully implemented and the aspirated output was achieved.

V. REFERENCES

[1] Prabhakaran, R. and Mohana, Dr. J., Detection of Oral Cancer Using Machine Learning Classification Methods (June 30, 2020). International Journal of Electrical Engineering and Technology, 11(3), 2020, pp. 384-393.

[2] A. T. Azar and S. A. El-Said, “Performance analysis of support vector machines classifiers in breast cancer mammography recognition,” Neural Comput. Appl., vol. 24, no. 5, pp. 1163–1177, 2014.

[3] Panigrahi, Santisudha & Swarnkar, Tripti. (2020). Machine Learning Techniques used for the Histopathological Image Analysis of Oral Cancer-A Review. The Open Bioinformatics Journal. 13.

[4] Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, Dimitrios I. Fotiadis, Machine learning applications in cancer prognosis and prediction, Computational and Structural Biotechnology Journal, Volume 13, 2015,Pages 8-17,

[5] N.R. Pal and S. K. Pal, “A Review on Image Segmentation Techniques,” Pattern Recognition”, Vol.26, No. 9, 1993, pp.1277-1294.

[6]Cheng, H., Jiang, X., et al, “Color image segmentation: advances and prospects, Pattern Recognition”, Vol. 34, 2001,pp. 2259-2281.

[7]Tu, Z. and Zhu, S.C.: Image Segmentation by Data-Driven Markov Chain Monte Carlo, IEEE Trans. on PAMI, Vol.24, No.5, 2002, pp. 657-673.

[8]J.-M. Jolion, P. Meer, S. Bataouche, “Robust clustering with applications in computer vision,” IEEE Trans. P AMI., Vol. 13, 1991, pp.791-802.

[9] Alfonso Rojas Dominguez, Asoke K. Nandi,“Toward Brest Cancer Diagnosis Based on Automated Segmentation of Masses inMammograms” Elsevier, Pattern Recognition, 42,1138-1148, 2009

[10] M. Hhadhoud, M. Amin, W. Dabbour, “Detection of Breast Cancer Tumor Algorithm Using Mathematical Morphology and Wavelet Analysis” ICGST International Journal on Graphics, Vision and Image Processing(GVIP). pp. 75-80. 2006

[11] D. Cascio, F. Fauci, R. Magro, G. Raso, “Mammogram Segmentation by Contour Searching and Mass Lesions Classification With Neural Network” IEEE TRANSACTIONS ON NUCLEAR SCIENCE, VOL. 53, NO. 5, OCTOBER 2006

[12] Brijesh Verma, Peter McLeod, Alan Klevansky, “A novel soft cluster neural network for the classification of suspicious areas in digital mammograms” Pattern Recognition 42 (2009) 1845 – 1852

[13] Arabi P M, Joshi G, and Vamsha Deepa N, 2016, Performance evaluation of GLCM and pixel intensity matrix for skin texture analysis, Perspect. Sci., 8, 203–206.

[14] Abd Latif M H, Md Yusof H, Sidek S N, and Rusli N, 2015, Implementation of GLCM Features in Thermal Imaging for Human Affective State Detection, Procedia Comput. Sci., 76, 308–315.

[15]. R. A. Welikala et al., "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," in IEEE Access, vol. 8, pp. 132677-132693, 2020,

[16]. A. Copeland, "Texture synthesis using gray-level co-occurrence models: algorithms, experimental analysis, and psychophysical support", Optical Engineering, vol. 40, no. 11, pp. 2655, 2001

[17]K. Anuradha, K. Sankaranarayanan “Comparison of Feature Extraction Techniques to classify Oral Cancers using Image Processing”, International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 6, June 2013.

[18] K. H. Ghazali, M. F. Mansor, M. M. Mustafa and A. Hussain, "Feature Extraction Technique using Discrete Wavelet Transform for Image Classification," 2007 5th Student Conference on Research and Development, 2007, pp. 1-4.

[19]. J. Zhang, X. Hong, S. Guan, X. Zhao, H. Xin and N. Xue, "Maximum Gaussian Mixture Model for Classification," 2016 8th International Conference on Information Technology in Medicine and Education (ITME), 2016, pp. 587-591.