ANALYSIS OF ORAL EPITHELIAL DYSPLASIA USING MACHINE LEARNING TECHNIQUE

S. Prabavathy[1], P. Aurchana[2], K. Padmavathi[3], R. Vidhyalakshmi[4]

[1] *Assistant Professor, Department of Computer Science, A.P.C. Mahalaxmi College for Women, Thoothukudi, India.*

[2]*Assistant Professor, Department of M.C.A., Sri Manakula Vinayagar Engineering College, Puducherry, India.*

[3,4]*Assistant Professor, Department of Computer Applications, St. Joseph's College of Arts and Science, Cuddalore, India.*

Abstract: The Oral Epithelial Dysplasia (OED) lesion is referred as a pre-cancerous lesion. It is a stepwise growth to cancer within the oral mucosa. The primary occurrence of a pre-cancer lesion is consequently increases the growth of the cancer cells in its surrounding area. In the proposed work, the Data Wavelet Transformation is applied to discriminate the normal and oral epithelial dysplasia disease affected images. For this the microscopic images have collected from Raja Muthiah Dental College and Hospital. The two feature extraction techniques namely Histogram Oriented features and Local Binary Pattern are used to extract the features. The extracted features are given as the input to Back Propagation Neural Network. The histogram oriented features with Back Propagation neural network gives the satisfactory results of 85%.

Keywords—Histogram Oriented Gradient (HOG), Local Binary Pattern (LBP), Back Propagation Neural Network (BPNN), Artificial Neural Network (ANN)

1. Introduction

Oral cancer can start as a primary lesion in any of the mouth's tissues, or it can spread from a isolated origin, or by extension from a neighboring anatomic structure, such as the nasal cavity. Alternatively, the oral cancers may originate in any of the tissues of the mouth, and may be of varied histological types: teratoma, adenocar-cinoma.

Under this Oral Dysplasia is a kind of precancerous lesion and the dysplastic stages are mild Dysplasia, moderate Dysplasia and severe Dysplasia.

Grading is done using support vector machine with highest accuracy in [1][2]. Now-a-days support vector machine plays an important role in many of the research fields.

Mild dysplasia (grade I)

It shows the proliferation / hyperplasia of cells of the parabasal and basal layers that does not extend beyond the lower third of the epithelium. Architectural changes are minimal.

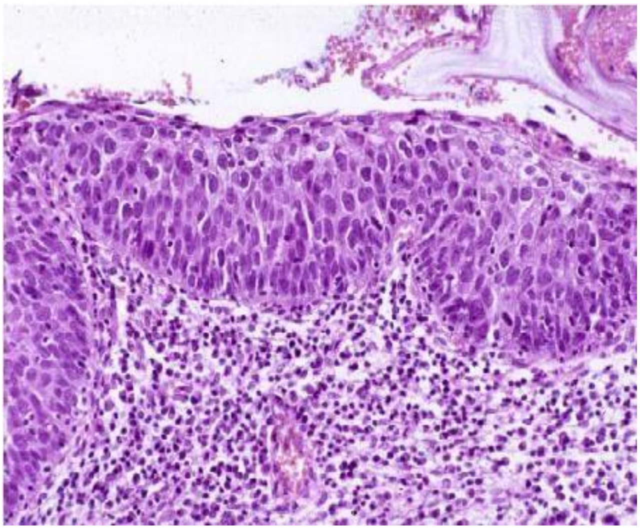

Figure 1.1. Mild Dysplasia

Moderate dysplasia (grade II)

It demonstrates a proliferation of a typical cells extending into the middle one-third of the epithelium [2][3]. Changes such as prominent cell, nuclear pleomorphism and hyperchromatism are more severe than in mild dysplasia which shows the Moderate Dysplasia in Figure 1.2.

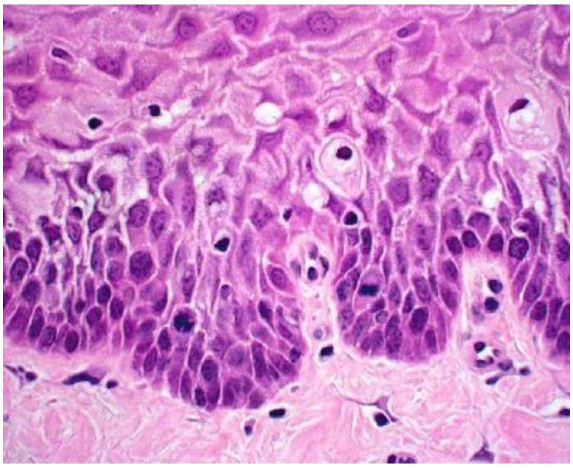

Figure 1.2. Moderate Dysplasia

In severe dysplasia (grade III)

Into the upper third of the epithelium, there is an abnormal proliferation from the basal layer. Figure 1.3. shows the Severe Dysplasia. Cytological and architectural changes can be very prominent.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

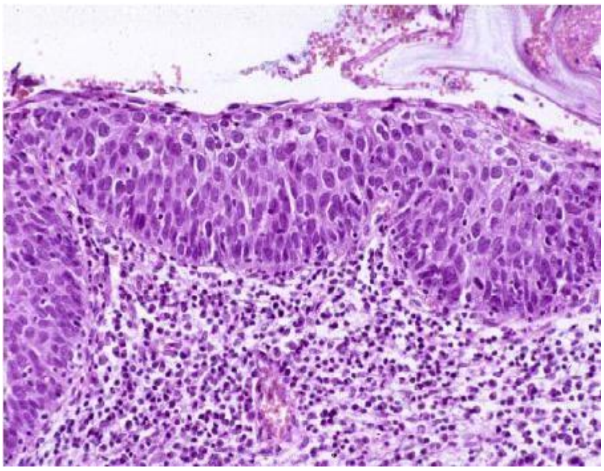

Figure 1.3. Severe Dysplasia

LITERATURE SURVEY

[4] 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 is 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 detects and classify the affected cancerous cell in the oral region by digital Image processing techniques feature extraction enables clear visualization of cancer affected areas. Here they used firefly algorithm to detect the cancer tumor in the MRI image. The cancer cells classified accurately using Expectation Maximization (EM).

[5] have characterized about oral cancer lesions using texture features. Image processing techniques with marker-controlled segmentation and feature extraction enables clear visualization of cancer affected areas with substantial resolution detecting different types of oral cancer lesions.

[6] describes reference about textural pattern classification for oral squamous cell carcinoma. They have used linear support vector machine classifier for automated diagnosis of oral cancer, which gave 100%accuracy.

[7] has surveyed about Image Processing Application. A Computer-aided Diagnosis is used to diagnosis the cancer using image processing and Artificial Neural Network and the features are extracted. H & E staining images, whose intensity and texture (Haralick) features are extracted after local preprocessing. These features are then fed to train Artificial Neural Network (ANN), for identifying whether is malignant or benign.

In [7] [8] a lung cancer is predicted based on analysis of the microscopic images of biopsy. Microscopic images of biopsy are features extracted with the Gray Level Co-Occurrence Matrix (GLCM) method and classified using back propagation neural network. This method is implemented to detect both normal and cancerous lung of biopsy samples. This system yields 81.25% accuracy.

In [9] [10] deep learning technique such as convoulutional neural network is used for the classification of musical datasets and yields satisfactory accuracy.

Outline of the Work

In this work Oral Epithelial Dysplastic feature extraction techniques and classification approaches are presented. In order to discriminate the stages of oral epithelial dysplasia the Data wavelet Ttransformations (DWT) has been applied to the microscopic images, two features namely Histogram of Oriented Gradient (HOG) and Local Binary Pattern (LBP) has been used for extract the features. The experimental result shows the classification accuracy of BPNN with HOG features can provide the better results.

Discrete Wavelet Transformation

Discrete Wavelet Transforms (DWT), which transforms a discrete time signal to a discrete wavelet representation [2]. Most of the image compression techniques use DWT based transformation. The microscopic images were taken as the input. The images were collected from 10x magnification. The images were resized into 512 x 512. And then Wavelet transformation has been applied to the image. The image was partitioned into approximation co-efficient, detailed coefficient, horizontal, vertical co-efficient and diagonal. In which, the diagonal is taken for the manipulation. The RGB images are converted into Grayscale.

FEATURE EXTRACTION

Histogram of Oriented Gradient

Histogram of Oriented Gradient (HOG) is a feature descriptor. In this method the gradient orientation is computed in the image. The image can be divided into small connected regions such as cells, and for the pixels within each cell, a histogram of the gradient directions is compiled. The descriptor is the concatenation of these histograms. In Histogram of oriented features 50 strongest points have been used for disease identification. The feature vector size is 42 X 36 dimensions.

B. Local Binary Pattern

Local Binary Pattern (LBP) was first described in 1994. It has been found to be a strong feature for the classification of texture; it has then determined that when LBP is joint with the HOG. Divide the examined window into cells (e.g. 16x16 pixels for each cell). For each pixel in a cell; compare the pixel to each of its

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

eight neighbors. Follow the pixels along a circle, i.e. clockwise or counter-clockwise. Where the middle pixel's value is larger than the value of the neighbor, write "0". Otherwise, write "1". Optionally normalize the histogram. Concatenate (normalized) histograms of all cells [9]. This generates a feature vector for the complete window.

Classifier

A. Back Propagation Neural Network (BPNN)

BPNN is used for linear as well as non-linear classification. BPNN is a supervised algorithm in which error difference between the desired output and calculated output is back propagated [10]. During learning, the process is repeated to reduce error by adjusting the weights via back propagation of error. Hidden units adjust their weights to reflect key task domain characteristics as a result of weight changes.

BPNN consists of three layers:

(1) Input Layer

(2) Hidden Layer and

(3) Output Layer.

The number of hidden layers and hidden units in each hidden layer is determined by the problem's complexity. Learning in BPNN, is a two-step process [11].

B. Step 1 (Forward Propagation):

In the first step, depend upon the inputs and weights, the outputs are calculated. For the calculation, each hidden and output unit calculates net excitation which depends on the values of previous layer that are linked to the unit in consideration.

- Weights between the previous layer unit and unit in consideration.
- Threshold value.
- This net excitation is used by activation function which returns calculated output value for that unit. This activation function must be continuous and differentiable. There are various activation functions which can be used in BPNN. Sigmoid is widely used activation function.[13] [14]

C. Step 2 (Backward Propagation of Error):

During this step, error is calculated by the difference between the targeted output and actual output of each output unit. These errors are back propagated to the previous hidden layer. For each unit in the hidden layer (N), error at that node is calculated. In similar way, error at each node of previous hidden layer, N-1 is calculated. These calculated errors are used to correct the weights so that the error at each output unit is minimized [15][16]. Until the error is minimized up to the expected level the forward and backward steps are repeated.

V. Experimental Results

HOG and LBP features are extracted from the microscopic images of Dysplasia. The image datasets were collected from Raja Muttiah Dental College and Hospital. 155 microscopic images were collected in which 55 images were used for testing and 100 images were used for training.

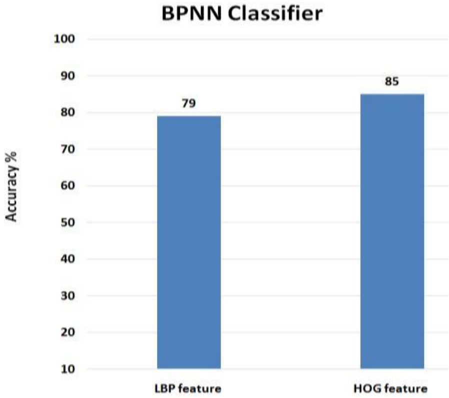

Figure 2. Accuracy of BPNN w.r.t. LBP and HOG features

The dataset were collected from 100 men and 55 women. The Hematoxylin and eosin stained images were collected in 10x magnification under the electron microscopic.

Figure 2 shows the accuracy level of BPNN when compared with LBP and HOG, in which HOG gives the satisfactory results.

VI. Conclusion

In this paper the proposed system is used to find Oral Epithelial Dysplasia and normal stages were proposed. Using HOG and LBP features were extracted and which is computed into BPNN. In which HOG using BPNN shows the highest accuracy of 85 %. In future, extracting different features and developing better classification algorithms and integration of classifiers to reduce the classification errors will be implemented.

VII. References

[1] P. Aurchana and P. Dhanalakshmi, ”SVM Based Classification of Epithelial Dysplasia Using SURF and SIFT Features”, International Journal of Pure and Applied Mathematics, Volume 117, No15, pp 11631175, 2017.

[2] P. Aurchana and P. Dhanalakshmi, “Grading of Oral Epithelial Dysplasia using Multi-SVM” Journal of Advanced research in Dynamical and Control Systems, Volume 10, pp-1479-1486, 2018.

[3] Anuradha K and Sankaranaranyanan K, “Comparison of Feature Extraction Techniques to Classify Oral Cancers using Image Processing”, International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, no 2, pp 456-462, June 2013.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

[4] Ahmed, C. M., Narayanan, M., Kalaivanan, S., Narayanan, K. S., and Reshmy, A. 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, Vol. 116 (21), pp. 149-154.

[5] M. M. R. Krishnan, .R. Acharya, C. Chakraborty, A. K. Ray, "Automated Diagnosis of Oral Cancer Using Higher Order Spectra Features and Local Binary Pattern: A

Comparative Study", TCRT, 10(5):391-504, October 2011.

[6] Rusha Patra, Chandan Chakraborty, Jyotirmoy

Chatterjee, "Textural Analysis of Spinous Layer for Grading Oral Submucous Fibrosis", International Journal of Computer Applications (0975–8887) Volume 47– No., June 2012.

[7] 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, Vol. 11(2), pp. 8-12.

[8] S. Prabavathy, V. Rathikarani, P. Dhanalakshmi, An Enhanced Musical Instrument Classification using Deep Convolutional Neural Network, International Journal of

Recent Technology and Engineering (IJRTE), Volume-8

Issue-4,

November 2019.

[9] Muthu Rama Krishnana M Pal M, Bomminayuni SK, Chakraborty C, Paul RR, Chatterjee J, Ray AK, "Automated Classification of Cells in Subepithelial Connective Tissue Of Oral Sub-Mucous Fibrosis - An SVM based Approach", ComputBio Med 2009 Dec. 39 (12), pp 1096-104.

[10] M. Muthu Rama Krishnan, Vikram Venkatraghavan, U. Rajenfran Acharya, Mousumi Pal, Ranjan Rashmi Paul, Lim Choo Min, Ajoy Lumar Ray, Jyotrimoy Chatterjee,

Chandan Chakraborty, "Automated Oral Cancer

Identification using Histopathological Images: A Hybrid Feature Extraction Paradigm, Elsevier, Volume 43, pp 352-364, 2012.

[11] M. Muthu Rama Krishnan, Anirudh Choudhary, Chandan Chakraborty, Ajoy K Ray, Ranjan R Paul,

"Texture based Segmentation of Epithelial Layer from Oral Histological Images", Elsevier, Volume 42, Issue 6, pp 632-641, 2011.

[12] S. Venkatakrishnan, V. Ramalingam, S. Palanivel,

"Classification of Oral Submucous Fibrosis using SVM,

"International Journal of Computer Applications" (0975 – 8887) Volume 78 – No.3, September 2013.

[13] Rojas, Raúl. Neural Networks: a Systematic Introduction. Chapter 7, Springer, 1996.

[14] Mitra, R., and Menaka, R., (2016), Characterization of Oral Cancer Lesions Using Texture Features, National Conference on Science, Engineering and Technology, Vol. 4(6), pp.33-35.

[15] Rahman, T. Y., Mahanta, L. B., Chakraborty, C., Das, A., and Sarma, J. D., (2018), Textural Pattern Classification for Oral Squamous Cell Carcinoma, Journal of

Microscopy, Vol. 269(1), pp. 85-93.

[16] Kulhalli, K. V., Patil, V. T., and Udupi, V. R., (2016), Image Processing Application for Computer Aided Diagnosis of Cancer, Image Processing Application for Computer Aided Diagnosis of Cancer, Vol. 1, pp. 297308.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731