Classification of Oral Sub mucous Fibrosis using SVM

P.Aurchana[1], P. Dhanalakshmi[2], C. Kalyani (MDS)[3]
[1, 2, 3] Department of CSE
[1, 2, 3] Annamalai University, Annamalainagar, Tamilnadu, India.

Abstract- *Oral sub mucous fibrosis (OSMF), is Pre-dominantly found among the people of Asian descent and the disorder is chronic, progressive and the stage of the disease depends on the clinical presentation. The paper takes the OSMF microscopic images in which the histogram equalization is used for image enhancement and the feature extraction in OSMF has been achieved using Gray Level Co-occurrence Matrix (GLCM). After the feature extraction has been done the SVM classifiers the OSMF affected images from the normal. The performance has been analyzed.*

Keywords- Oral sub mucous fibrosis (OSMF), Gray-level Co-occurrence Matrix (GLCM), Histogram, Image classification, Feature extraction, Support Vector Machine (SVM).

I. INTRODUCTION

Oral sub mucous fibrosis (OSMF) is a chronic, complex, premalignant lesion of the oral cavity, characterized by juxta epithelial inflammatory reaction and progressive fibrosis of the submucosal tissues [1]. As the disease progresses, the jaws become rigid to the point that the person is unable to open the mouth [1]. The condition is remotely linked to oral cancers and is associated with areca nut or betel quid chewing, a habit similar to tobacco chewing, is practiced predominantly in Southeast Asia and India, dating back thousands of years. The pathogenesis of the disease is not well established. A number of factors trigger the disease process by causing inflammatory reaction in the oral mucosa. Areca nut chewing, ingestion of chilies, genetic and immunologic processes, nutritional deficiencies, etc, are a few among them.

Symptoms of oral sub mucous fibrosis included Progressive inability to open the mouth (trismus) due to oral fibrosis and scarring, Oral pain and a burning sensation upon consumption of spicy foodstuffs, Increased salivation, Change of gustatory sensation, Hearing loss due to steno sis of the Eustachian tubes Dryness of the mouth, Nasal tonality to the voice, Dysphasia to solids Impaired mouth movements (e.g., eating, whistling, blowing, sucking).Figure1shows the overall architecture of the proposed classification system

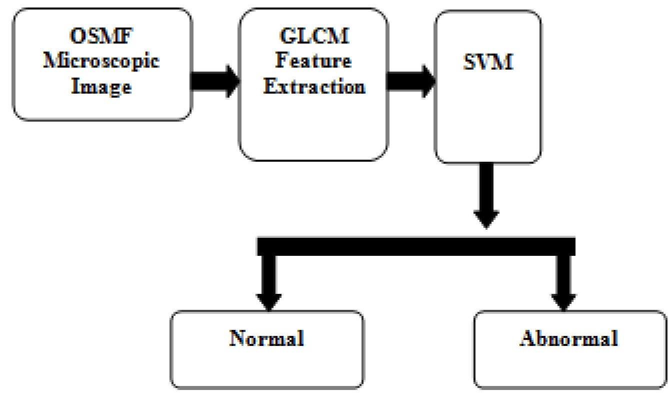


Figure 1: Block Diagram of the OSMF Classification System.

II. FEATURE EXTRACTION

A. To construct a histogram

In statistics, a histogram is a graphical representation of the distribution of data. It is an estimate of the probability distribution of a continuous variable. A histogram is a representation of tabulated frequencies, shown as adjacent rectangles, erected over discrete intervals (bins), with an area equal to the frequency of the observations in the interval. The height of a rectangle is also equal to the frequency density of the interval, i.e., the frequency divided by the width of the interval. The total area of the histogram is equal to the number of data.

A histogram may also be normalized displaying relative frequencies. It then shows the proportion of cases that fall into each of several categories, with the total area equaling 1. The categories are usually specified as consecutive, non-overlapping intervals of a variable. The categories (intervals) must be adjacent, and often are chosen to be of the same size [4]. The rectangles of a histogram are drawn so that they touch each other to indicate that the original variable is continuous. Histograms are used to estimate the probability density function of the underlying variable.

$$p(r_k) = {}^{n_k}/_{n,} \quad 0 \le k \le L-1 \qquad (1)$$

Where r_k is the k^{th} gray level, n^k no. of pixels in the image with the gray level, L no. of level n total no. of pixels in the image p(r_k) gives the probability of occurrence of gray level r_k.

Divide the range between the highest and lowest values in a distribution into several bins of equal size. Toss each value in the appropriate bin of equal size. The height of a rectangle in a frequency histogram represents the number of values in the corresponding bin.

B. Histogram Equalization

The technique used for obtaining a uniform histogram is known as histogram equalization or Histogram linearization.

The steps to perform histogram equalization are:

(a) Find the probability of occurrence of each gray level (r_k) in the input image using equation (1).

(b) Use the transformation function Sk= T(rk) = to obtain the histogram equalized image Histogram equalization significantly improves the visual appearance of the image.

C. Gray level Co-occurrence Matrix Features (GLCM)

Gray Level Co-occurrence Matrix exploits the higher order distribution of gray values of pixel that are defined with a specific distance or neighborhood criterion. In [6] the simplest form, the GLCM P(i,j) is the distribution of the number of occurrence of a pair of gray values i and j separated by a distance vector d = [dx;dy]. The GLCM normalizes each value in the matrix by dividing the total number of occurrence, providing the probability of occurrence of a pair of gray values separated by a distance vector.

Statistical texture features are computed from the normalized GLCM as the second-order histogram H(yq,yr,d) representing the probability of occurrence of a pair of gray values yq and yr separated by a distance vector d. Texture features can also be described by a difference histogram, Hd(ys,d), where ys = |yq −yr|. Hd(ys,d) indicates the probability that a difference in gray levels exists between two distinct pixels [7]. We used 22 textural features in our study. The following equations define these features. Let (a, b) be the (a, bj)th entry in a normalized GLCM [8], [9].

Table 1: Haralick texture features (characteristics 1 to 5)

S. No	Characteristics
1.	Contrast
2.	Correlation
3.	Energy/angular second moment
4.	Homogeneity
5.	Autocorrelation

D. Support Vector Machines (SVM)

SVM is a statistic machine learning technique that has been successfully applied in the pattern recognition area and, is based on the principle of structural risk minimization. SVM constructs a linear model to estimate the decision function using non-linear class boundaries based on support vectors. SVM learns an optimal separating hyper plane from a given set of positive and negative

Figure 1: shows the architecture of the SVM. It maps the input patterns into a higher dimensional feature space through some nonlinear mapping chosen a priori. A linear decision surface is then constructed in this high dimensional feature space. Thus, SVM is a linear classifier in the parameter space, but it becomes a non-linear classifier as a result of the nonlinear mapping of the space of the input patterns into the high dimensional feature space.

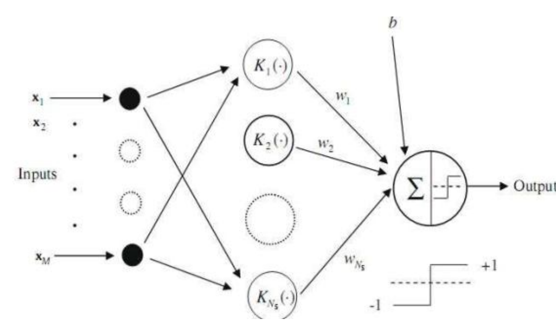


Figure 2: Architecture of the SVM (Ns is the number of support vectors).

The kernel function may be any of the symmetric functions that satisfy the Mercer’s conditions. There are several SVM kernel functions as given in table 1.

Table 2 : Types of SVM inner product kernel

Types of kernels	**Inner Product Kernel K(x^T , x_i)**	**Details**
Polynomial	$(x^T x_i + 1)^p$	Where x is input patterns, x_i is support vectors, σ^2 is variance, $1 \le i \le N_s$, N_s is number of support vectors, β_0, β_1 are constant values. p is degree of the polynomial
Gaussian	$\exp\left[-\frac{\lvert x^T - x_i\rvert^2}{2\sigma^2}\right]$	
Sigmoidal	$\tan h(\beta_0(x^T x_i) + \beta_1)$	

III. PERFORMANCE MEASURES

Sensitivity and Specificity are statistical measures used for studying the performance of classification. Sensitivity measures the proportion of actual positives which are correctly identified

$$sensitivity = \frac{\text{True positive}}{(\text{True positive} + \text{False negative})} \quad (2)$$

IV.EXPERIMENTAL RESULTS

A. Dataset

Normal and OSMF affected tissue images were collected from patients of Raja Muthiah Dental College and Hospital (RMDC & H).The dataset of 200 Microscopic images was collected 100 were of normal microscopic images and 100 were of affected images which was classified by SVM classifier.

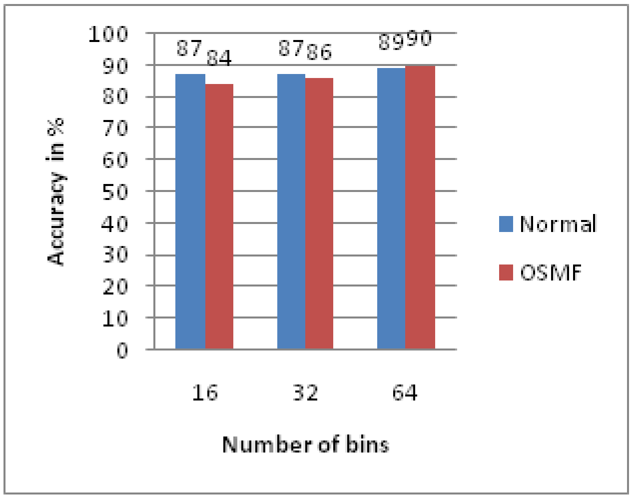
A total of 200 microscopic images which consists of 100 OSMF images and 100 normal images are considered.

Table2: Performance comparison of OSMF classification in terms of different kernel functions using SVM.

Types of Kernel	**Accuracy (%)**
Polynomial	79
Gaussian	89
Sigmoidal	71

Table 3: Shows the sensitivity and specificity for OSMF Microscopic images.

Performance Measures	**Accuracy (%)**
Sensitivity	74 %
Specificity	88 %



The above figure shows that maximum performance is achieved for 64 bins. Repeated experiments were carried Out for 16, 32 and 64 bins.

Table 2: shows the performance of OSMF classification for different kernel function using SVM.

V. CONCLUSION

In this paper, a system for classifying OSMF affected images from normal images was proposed. Using histogram texture features where extracted from both normal and OSMF affected images which is computed using GLCM. The features where trained and tested using SVM for different bins. The system showed an accuracy of 97.0% for 64 bins.

REFERENCES

[1] 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.

[2] M. Muthu Rama Krishnan A, Mousumi Pal B, Suneel K Bomminayuni C, Chandan ChakrabortyA,RanjanRashmi Paul B, JyotirmoyChatterjee A, Ajoy K. Ray A,C, “Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis An SVM

based approach”,Elsevier,www.elsevier.com/locate/cbm, September 2009.

[3] P.Mohanaiah,P.Sathyanarayana,L.GuruKuma, “Image Texture Feature Extraction Using GLCM Approach”, International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013 1,ISSN 2250-3153.

[4] Van der Waal.I, “Potentially malignant disorders of the oral and oropharyngeal mucosa; present concepts of management”, Oral Oncology,46(6):423-5,June 2010.

[5] 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.

[6] 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.

[7] Shah.P. Sch. of Med. Sci. & Technol., IIT,- Kharagpur, Kharagpur, India ; Krishnan, M.M.R. Chakraborty, C,Ray, A.K.,Morphometric”, Pattern Analysis of Basal Cell Nuclei for Oral Cancer Screening”, Bioinformatics and Biomedical Engineering [iCBBE] 2010 National conference, 1-4, June 2010.

[8] Ray A Krishnan, M M. Shah.P, Choudhary A A. Chakraborty. C. Paul R R. “ Textural characterization of histopathological images for oral sub-mucous fibrosis detection”, Tissue cell, 43(5):318–30, October 2011.

[9] AnirbanMukherjeeJyotirmoyChatterjeeRanjan.R.PaulPranab.K.Dutta,M.E.Tathagata Ray, D.Shivashanker Reddy,“Detection of constituent layers of histologicaloral sub-mucous fibrosis images using the hybrid segmentation algorithm”,Oral Oncology, 44(12):1167–1171,December 2008.

[10] M.Muthuramakrishnan,Pal,Suneel.K.Bomminayuni,ChandanChakraborthy,”Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis-An SVM based approach”, Computers in Biology and medicine,39(12):1096-1104. December 2009.

[11] Chandan. K. Ra M. Muthu Rama Krishnan, Pratik Shah,“Statistical analysis of textural features for improved classification of oral histopathological images”, Journal of Medical Systems, 36(2):865–881, April 2012.

[12] Chandan Chakraborty Ajoy. K. Ray Ranjan. R. Paul M. Muthu Rama Krishnan, Anirud Choudhry.”Texture based segmentation of epithelial layer from oral histological images. Micron, 42(6):632–641,August 2011.

[13] U. Rajendra Acharya, Mousumi Pal, Ranjan Rashmi Paul, Lim Choo Min Ajay kumar Ray Jyotirmoy Chatterjee Chandan Chakraborty M. Muthuramakrishnan, VikramVenkatraghavan. “Automated oral cancer identification using histopathological images: A hybrid feature extractionparadigm”, 43(2-3):352–364, February 2012.

[14] Pratik Shah M.Tech. ChandanChakraborty- MuthuRama Krishnan Mookiah,sAjoy K. Ray.Brownian motion curvebased textural “classification and its application in cancer diagnosis”, Analytical and Quantitative Cytology and Histology, May 2013.

[15] MuthuramaKrishnan.M.Shah. P, “Structural markers for normal oral mucosa and oral sub-mucous fibrosis Micron”,41(4):312–20, June 2010.

[16] Shweta Jain,” Brain Cancer Classification Using GLCM Based Feature Extraction in Artificial Neural Network”, International Journal of Computer Science & Engineering Technology (IJCSET), ISSN : 2229-3345 ,Vol. 4 No. 07 Jul 2013

[17] JsyotirmoyChatterjeeRushaPatra,ChandanChakraborty.Textural analysis of spinous layer for grading oral submucousfibrosis”, International Journal of Computer Applications, 47:975-8887,June 2012.

[18] R R Paul, A Mukherjee, P K Dutta, S. Banerjee, M Pal, J Chatterjee, K. Chaudhuri, K. Mukkerjee,”A novel wavelet neural network based pathological stage detection technique for an oral precancerous condition”, IJCATM ,58(2):932-938,February2005.