Available online at www.journalijmrr.com

INTERNATIONAL JOURNAL OF MODERN RESEARCH AND REVIEWS

Int. J. Modn. Res. Revs.

Volume 6, Issue 1, pp 7-11, January, 2018

ISSN: 2347-8314

COMPARISON OF PATTERN BASED CLASSIFICATION SYSTEM IN ORAL SUBMUCOUS FIBROSIS (OSMF)

S.Venkatakrishnan and Dr.V.Ramalingam

*[1]Assistant Professor, Mechanical Engineering, Annamalai University, Chidambaram
[2]Professor, Dept of CSE (FEAT), Annamalai University, Chidambaram

Article History: *Received 7th january,2018, Accepted 30th January,2018, Published 31st January,2018*

ABSTRACT

Oral Submucous Fibrosis (OSMF) is an insidious chronic progressive precancerous condition of the oral cavity and oropharynx with a high degree of malignant potential. It is characterized by a generalized fibrosis or stiffening of the mucosal covering of the mouth tissue. Using GMM and AANN, a system is developed to classify an image into normal and OSMF category. Experiments showed significantly satisfactory results with an accuracy ranging from 90% to 95%.

Keywords: Feature extraction, Histogram, Image classification, GMM,AANN.

1.INTRODUCTION

Oral Submucous Fibrosis (OSMF) is a precancerous condition associated with the use of areca nut in various forms. Worldwide, estimates of OSMF shows a confinement to Indians and Southeast Asians, with overall prevalence rate in India to be about 0.2% to 0.5%. Ingestion of chillies, genetic susceptibility, nutritional deficiencies, altered salivary constituents and autoimmunity and collagen disorders may be involved in the pathogenesis. The condition is well recognized for its malignant potential rate of 7.6% and is particularly associated with use of areca nut in various forms with significant duration and frequency of chewing habits. In OSMF, there is a progressive inability to open the mouth and tongue movement gets restricted to varying degrees depending up on the severity of the disease process
Symptoms of oral submucous fibrosis include:

- Deficiency of iron (anemia), Vitamin B complex, minerals, and malnutrition are promoting factors that disturbs the repair process of the inflamed oral mucosa, thus leads to deranged healing and resultant scarring and fibrosis.
- Oral pain and a burning sensation upon consumption of spicy foodstuffs
- Increased salivation
- Change of gustatory sensation
- Hearing loss due to stenosis of the eustachian tubes.
- Dryness of the mouth
- Nasal tonality to the voice
- Dysphagia to solids (if the oesophagus is involved)
- Impaired mouth movements (eg, eating, whistling, blowing, sucking)

Fig1.: Intraoral photograph of the buccal mucosa showing blanched oral mucosa with erosions in the initial stages of oral submucous fibrosis

Corresponding author:* ***Prof. S.Venkatakrishnan Asst.**Professor, Mechanical Engineering, Annamalai University, Chidambaram*

Fig.1 shows a sample intraoral photograph of a patient in the initial stages of oral submucous fibrosis. Histopathologically OSMF is characterized by the concomitant presence of less vascularized collagenous connective tissue with overlying atrophic epithelium in oral mucosa. Traditionally, the pathologists use histopathological images of biopsy tissue and examined them under light microscope to detect OSMF which is a highly qualitative process . Early detection of premalignant and cancerous mucosal lesions improves the survival and morbidity of patient's sufferings . Currently one of the greatest challenges to oral oncobiologists is to determine and identify the degree of tissue damage or stages of various precancerous states of oral tissue and to detect the exact transition of a normal tissue to precancerous state .

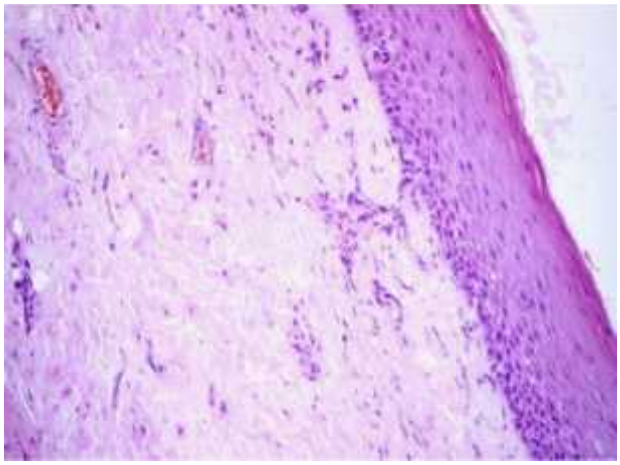


Fig.2: Microscopic image of OSMF showing an atrophic epithelium and fibrous connective tissue; E-Epithelium; CT-Connective Tissue; K-Keratin

Fig.2 shows the microscopic image of OSMF. Atrophic epithelium with loss of retepegs and dense fibrous connective tissue are seen. Presently, no specific diagnostic test is available for OSMF except for histopathological studies. The main histopathological characteristic of OSMF is the deposition of collagen in the subepithelial connective tissue leading to epithelial atrophy. It has been found that exposure of buccal mucosal fibroblasts to alkaloid may result in the accumulation of collagen. Collagenase activity has been found to be lower in OSMF than in normal oral mucosa. These findings suggest that OSMF should be considered as a collagen metabolic disorder resulting from alkaloid exposure and individual variations in collagen metabolism.

In the more advanced stage of the disease, the essential feature is a fibrous band restricting mouth opening and causing difficulty in mastication, speech, swallowing and maintaining oral hygiene. Development of fibrous bands in the lip makes the lip thick, rubbery and difficult to retract or avert; a band around the lips gives the mouth opening an elliptical shape. Fibrosis makes cheeks thick and rigid.

Microscopic changes

The initial pathology of OSMF is characterized by juxta-epithelial inflammation including edema, large fibroblasts and an inflammatory infiltrate, consisting primarily of neutrophils and eosinophils. Later, collagen bundles with early hyalinization are seen and the acute inflammatory infiltrate contains more chronic cell types, such as lymphocytes and plasma cells, occasionally resembling lichenoid mucositis. In more advanced stages, OSMF is characterized by formation of thick bands of collagen and hyalinization extending into the submucosal tissues and decreased vascularity. Outcomes of OSMF are characterized by two features: the persistence of the disease and its potential to become malignant. OSMF is strongly associated with a risk of oral cancer

OSMF has been graded by many pathologists based on clinical features or histopathology or both. It has been graded into Grades I, II and III based on the microscopic features. The simple pathological evaluation procedure currently used does not provide a quantitative analysis of the vital changes in the tissues; that is, epithelial dysplastic changes, subepithelial fibrosis .Thus a computer based diagnostic approach will enhance the accuracy of diagnosis as well as may assist in grading OSMF.

FEATURES FOR OSMF CLASSIFICATION

Color histogram features

The histogram of a digital image with gray levels in the range [0, L − 1] is a discrete function, where L is the number of discrete gray levels.

$$P(r_k) = \frac{n_k}{n}, 0 \le k \le L-1$$

where r_k is the k^{th} gray level, n_k is the number of pixels in the image with that gray level, n is the total number of pixels in the image with k = 0, 1, 2, .., L − 1. In short, $P(r_k)$ gives an estimate of the probability of occurrence of gray level r_k.

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 (rk) in the input image.

(b) Use the transformation function sk = T(rk) to obtain the histogram equalized image. Histogram equalization significantly improves the visual appearance of the image. Similar enhancement results could be achieved using the contrast stretching approach. However the advantage of histogram equalization is that it is fully automatic.

The gray level image uses 8-bits per pixel and the intensity value at a pixel varies from 0 (black) to 255 (white). Color image uses 24 bits per pixel i.e, 8 bits for each of the three components red, green and blue in the RGB color space. Hence the total number of colors is 224=16,777,216. R=255, G=0 and B=0 corresponds to pure red color. R=0, G=0, B=0 is dark black and R=255, G=255, B=255 is pure white.

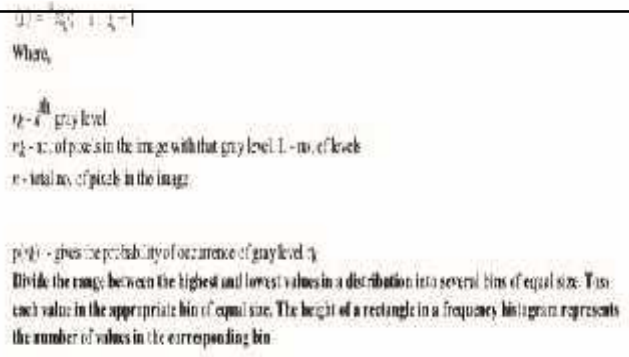
The feature is defined as a function of one or more measurements, each of which specifies some quantifiable property of an object, and is computed that it is quantifying some of the important characteristics of the object. Feature selection helps to reduce the feature space which improves the prediction accuracy and minimizes the computation time. Quantitative evaluation of histopathological features is not only vital for precise characterization of any precancerous condition but also crucial in developing automated computer aided diagnostic system. Sub- epithelial hyalinization and fibrosis are characteristic histological features of OSMF. In this research work, color histogram features were extracted from both normal and OSMF microscopic images and classification was done using SVM, RBFNN, AANN and GMM models.

Constructing 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 are presentation 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. 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 dens function of the underlying variable. The total area of histogram used for probability density is always normaliz to 1. If the lengths of the intervals on the x-axis are all then a histogram is identical to a relative frequency plot.

Color histogram is used to compare images in many applications. In this work, RGB color space is quantized into 64-dimensional feature vectors which are used as features. The image histogram is a simple bar graph of pixel intensities. The pixels are plotted along the x-axis and the number of occurrences for each intensity is represented along the y-axis:



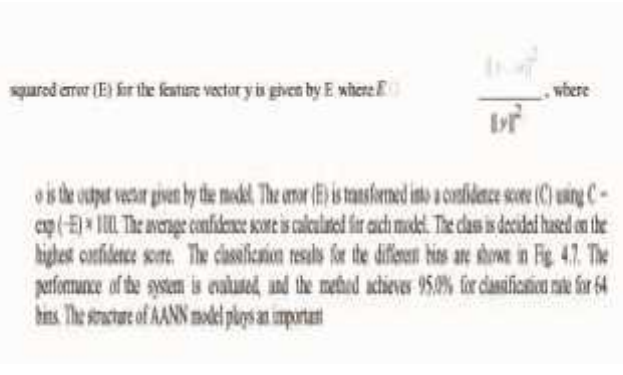
EXPERIMENTAL RESULTS

A total of 200 microscopic images which consists of 100 OSMF images and 100 normal images are considered. For four fold cross validation training data gfi (i=1,2,3,4) consisting of 150 microscopic images [50 images (25 Normal + 25
OSMF) + 50 images (25 Normal + 25 OSMF) + 50 images (25 Normal + 25

OSMF)] are used. For testing, 50 microscopic images (25 normal and 25 OSMF) are used.

Evaluation using AANN

AANN models perform an identical mapping of the input space. The distribution of 16, 32 and 64 dimensional feature vectors in the feature space for different bins is captured using an AANN model. Separate AANN models are used to capture the distribution of feature vectors of each class and the network is trained for 500 epochs. One epoch of training is a single presentation of all the training vectors to the network. For evaluating the performance of the system, the feature vector is given as input to each of the models. The output of the model is compared with the input to compute the normalized squared error. The normalized



role in capturing the distribution of the feature vectors. After some trial and error, the network structure 64L - 128N - 8N -128N - 64L is obtained. This structure seems to give good performance in terms of classification accuracy.

The number of units in the third layer (compression layer) determines the number of components captured by the network. The AANN model projects the input vectors onto the subspace spanned by the number of units (nc) in the compression layer. If there are nc units in the compression layer, then the histogram feature vectors are projected onto the subspace spanned by nc components to realize them at the output layer.

The effect of changing the value of nc on the performance of OSMF classification is studied. There is no major change in the performance if nc is between 4 and 6 and the performance of the system decreases if it is less than 2 or more than 6. The decrease in the performance for nc< 2 indicates that there may not be a boundary between the components representing the microscopic image information. The decrease in the performance for nc> 6 indicates that the training data may not be sufficient for capturing the distribution of feature vectors. The results are shown in Table 4.4 and maximum performance is obtained with the structure 64L - 128N - 8N - 128N - 64L as shown in Fig. 4.7.

Similarly, the performance is obtained by varying the number of units in the expansion layer keeping the number of units in the compression layer to 4. When the number of units in the expansion layer is increased from 128 to 146, there is no considerable increase in the performance.

Evaluation uing GMM

The distribution of the histogram features is captured using GMM. Gaussian Mixture Models are a type of density models comprises a number of components which are combined to provide a multi-model density. The performance of the system is studied for a mixture of Gaussians varying from 2 to 10. When the number of mixtures is less, the performance is low whereas the classification performance increases, as the number of mixtures increases.

Fig. shows the performance of GMM for different mixtures. When the number of mixtures is 2 the classification performance is very low. When the mixtures are increased from 2 to 4, the classification performance slightly increases. When the number of mixtures varies from 5 to 10 there is considerable increase in the performance and the maximum performance is achieved. When the number of mixtures is above 10 there is no considerable increase in the performance. With GMM the best performance is achieved with 10 Gaussian mixtures as shown in Fig..

A comparison of the performance of OSMF classification using AANN and GMM is done .the maximum performance of AANN in OSMF classification was obtained for the structure 64L-128N-6N-128N-64L with GMM classification the maximum performance was 91% with 10 mixtures.

2.DISCUSSION

The performance of Oral Submucous Fibrosis classification (OSMF) using AANN and GMM has been analyzed. Color Histogram features are extracted to characterize the OSMF images. Classifiers such as, AANN and GMM are applied to obtain the optimal class boundary between the two classes namely normal and OSMF images by learning from training data. The performance is compared among AANN and GMM from the research it was observed that AANN classifiers could give a better accuracy for OSMF classification.

3.CONCLUSION

Few other pattern classifiers such as SVM, RBFNN and others could also be applied in the same manner to classify precancer and cancer images. This could bring out a better technique for OSMF classification in the future.

4.REFERENCES

[1] Paissat DK, “Oral submucous fibrosis”; Int J Oral Surg. 1981 Oct;10(5):307-12.

[2] Yusaku Nishi and Keiichi Horio, Kentaro Saito, Manabu Habu and Kazuhiro Tominaga, “Discrimination of Oral Mucosal Disease Inspired by Diagnostic Process of Specialist”, Journal of Medical and Bioengineering, vol. 2, no. 1, pp.57-61, March 2013.

[3] Byakodi R, Byakodi S, Kiremath S, Bya Kodi J, Adaki S, Marathe K and Mahiral P, “Oral cancer in India: an epidemiologic and clinical review”, Journal of Community Health, vol. 37, no. 2, pp. 316-319, April 2012.

[4] Rusha Patra, Chandan Chakraborty and Jyotirmoy Chatterjee, “Textural Analysis of Spinous Layer for Grading Oral Submucous Fibrosis”, International Journal of Computer Applications, vol. 47, pp. 0975-0987, June 2012.

[5] Rajendran R. and Sivapathasundaram B, “Shafer’s Textbook of Oral Pathology”, 6th Edition, Elsevier India, 2009.

[6] Shilpa B. Rajesh and Satheesha B. H. Reddy, “Cytomorphometric Analysis of Obtained Squames Obtained from Normal Oral Mucosa and Lesions of Oral Submucous Fibrosis”, Journal of Indian Academy of Oral Medicine and Radiology, vol. 24, no. 3, pp. 200-205, July-September 2012.

[7] Smitha T, Sharada P and Hc Girish, “Morphometry of the basal cell layer of oral leukoplakia and oral squamous cell carcinoma using computer-aided image analysis”, Journal of Oral Maxillofacial Pathology, vol. 15, no. 1, pp. 26-33, January-April 2011.

[8] Kramer I. R. H, El-Labban N. G and Sonkodi S, “Further Studies On Lesions Of The Oral Mucosa using Computer-Aided Analyses Of Histological Features”, British Journal of Cancer, vol. 29, p. 223, 1974.

[9] Tathagata Ray M. E, Shivashankar Reddy D, Anirban Mukherjee, Jyotirmoy Chatterjee, Ranjan R. Paula and Pranab K. Dutta, “Detection of constituent layers of histological oral sub-mucous fibrosis: Images using the hybrid segmentation algorithm”, Oral Oncology, vol. 44, no. 12, pp.1167-1171, December 2008.

[10] Krishnan M. M. R, Shah P, Ghosh M and Pal M, “Automated characterization of sub-epithelial connective tissue cells of normal oral mucosa: Bayesian approach”, Proceedings of IEEE Student Technology Symposium, pp. 44-48, April 2010.

[11] Gao S, "Cell morphometric analysis in oral submucous fibrosis, leukoplakia and squamous cell carcinoma", Zhonghua Kou Qiang Yi Xue Za Zhi, vol. 27, no. 3, pp. 145-147, May 1992.
[12] Muthu Rama Krishnan M, Acharya U. R, Chakraborty C and Ray A. K, "Automated Diagnosis of Oral Cancer Using Higher Order Spectra Features and Local Binary Pattern: A Comparative Study", Twin Cities Radiation Therapists, vol. 10, no. 5, pp. 391-504, October 2011.
[13] Mitesh Amitkumar Modi, Vishal R. Dave, Viral G. Prajapati and Keyur A. Mehta, "A Clinical Profile of Oral Submucous Fibrosis", National Journal of Integrated Research in Medicine, vol. 3, no. 3, pp. 152-155, July-August 2012.
[14] Crispian Scully, Jose V. Bapan, Colin Hopper and Joel B. Epstein, "Oral cancer: Current and future diagnostic techniques", American Journal of Dentistry, vol. 21, no. 4, pp. 199-209, August 2008.
[15] Shabana A.H.M, El-Labban N.G. and LeeK.W, "Morphometric analysis of basal cell layer in oral premalignant white lesions and squamous cell carcinoma", Journal of Clinical Pathology, vol. 40, pp. 454-458, April 1987.

Structure of AANN	Feature vector dimensions (No. of bins)					
	16		32		64	
	Normal	OSMF	Normal	OSMF	Normal	OSMF
16L 76N 2N 76N 16L	87.0	88.0	87.0	89.0	72.0	78.0
32L 76N 4N 76N 32L	88.5	89.5	91.5	95.0	82.5	85.0
64L 128N 6N 128N 64L	88.5	90.7	92.0	94.5	93 .0	95.0

Table: Average performance of normal and OSMF classification in terms of number of mixtures by GMM model using Histogram features

No. of mixtures	Accuracy (%)					
	Feature vector dimensions (No. of bins)					
	16		32		64	
	Normal	OSMF	Normal	OSMF	Normal	OSMF
2	56.0	61.0	70.0	73.0	77.0	80.0
5	70.0	72.0	76.5	78.0	84.0	86.0
10	85.0	88.0	78.1	80.0	89.0	91.0

Table shows the performance of normal and OSMF classification in terms of number of mixture in GMM using Histogram features.