Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 10, pp.40676-40678, October, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

OSMF CLASSIFICATION USING GMM

*Venkatakrishnan, S.

Assistant Professor, Mechanical Engineering, Annamalai University, Tamilnadu, India

ARTICLE INFO

Article History:
Received 28th July, 2016
Received in revised form
22nd August, 2016
Accepted 25th September, 2016
Published online 30th October, 2016

Key words:

Feature extraction, Histogram,
Image classification, GMM.

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. Gaussian Mixture Model (GMM) is a mixture of several Gaussian distributions and can therefore represent different subclasses inside one class. Using GMM, a system is developed to classify an image into normal and OSMF category. Experiments showed significantly satisfactory results with an accuracy of 91%.

Citation: Venkatakrishnan, S., 2016. "OSMF classification using GMM", *International Journal of Current Research*, 8, (10), 40676-40678.

INTRODUCTION

Oral Submucous Fibrosis is detected at a very early stage, cessation of the habit is sufficient. Most patients with oral submucous fibrosis present with moderate-to-severe disease. Moderate-to-severe oral submucous fibrosis is irreversible. Medical treatment is symptomatic and predominantly aimed at improving mouth movements. The onset and progress of the disease is believed to be multifactorial. A number of factors trigger the disease process by causing a juxtaepithelial (below/near the epithelium) inflammatory reaction in the oral mucosa. Factors include areca nut chewing, ingestion of chilies, genetic and immunologic processes, nutritional deficiencies and other factors. The flavonoid components of areca nut have been found to have some direct influence on collagen metabolism. It has been found that exposure of fibroblast cells to alkaloid results in the accumulation of collagen. Also, there is a decreased degradation of collagen in OSMF mucosa compared to normal tissue.
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)

This paper is organized as follows: Section 2 gives a brief summary of related work done in the computerized diagnosis of oral diseases. Histogram Feature Extraction is described in Section 3. The principle of Gaussian Mixture Model is described in Section 4. Experimental results are discussed in Section 5. Conclusion and future work are given in Section 6.

Related work

Application of Image Analysis methods in OSMF, other precancerous conditions and oral cancer for diagnosis, staging and classification aspects is being expertise in the recent years. In June 2013, Anuradha K and Sankaranarayanan K described feature extraction techniques to classify oral cancers using Image Processing [9]. In this work, a system is developed to segment, extract features and classify cancers. Later, a comparison is made. The proposed system consists of five steps. First, the images are enhanced and the Region of Interest (ROI) is segmented using Marker Controlled Watershed Segmentation. Feature Extraction methods like Gray Level Co-occurrence Matrix (GLCM), Intensity Histogram and Gray

Corresponding author: **Venkatakrishnan, S.,**
Assistant Professor, Mechanical Engineering, Annamalai University, Tamilnadu, India

Level Run Length Matrix (GLRLM) are used to extract features from ROI. Next, classification is made using Support Vector Machine (SVM) classifier to classify the tumor as benign or malignant mass and a comparative study is performed to identify the best feature extraction technique. In August 2012, Mitesh Amitkumar Modi, Vishal R. Dave, Viral G. Prajapati and Keyur A.Mehta described "A A Clinical Profile of Oral Submucous Fibrosis"[3]. A hospital-based study was conducted on 80 oral Submucous Fibrosis cases who visited our hospital in Jamnagar. A detailed history of each patient was recorded along with a clinical examination. Biopsy was performed for histopathological correlation. Clinical stage of the disease in terms of the ability to open one's mouth was correlated with histopathological grading.

FEATURE EXTRACTION

Feature Selection: Feature extraction plays an important role in constructing an oral pre- cancerous lesions classification system. Discriminative power of features or feature sets tells how well they can discriminate different classes. 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 study, color histogram features were extracted from both normal and OSMF microscopic images. Classification was done using GMM.

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. 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. The total area of a histogram used for probability density is always normalized to 1. If the lengths of the intervals on the x-axis are all 1, then a histogram is identical to a relative frequency plot. Color histogram is used to compare images in many applications. In this study, 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.

Classifiers

After the features are extracted, a suitable classifier must be chosen. A number of classifiers are used and each classifier is found suitable to classify a particular kind of feature vectors depending upon their characteristics. The classifier used commonly is Gaussian Mixture Model. The GMM classifier is used to compare the feature vector of the prototype with image feature vectors stored in the database. It is obtained by finding the distance between the prototype image and the database.

Gaussian Mixture Model (GMM)

Gaussian Mixture Model (GMM) is a mixture of several Gaussian distributions and can therefore represent different subclasses inside one class. 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 to10. When the number of mixtures is less, the performance is low whereas the classification performance increases, as the number of mixtures 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.

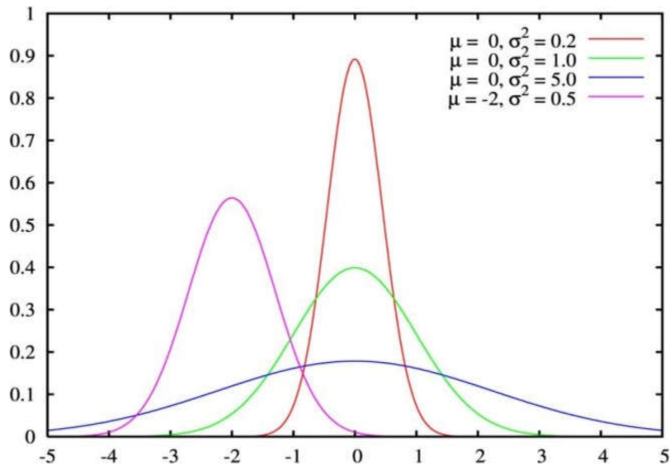

Fig. 1. Architecture of the GMM

The kernel function may be any of the symmetric functions that satisfy the Mercer's conditions (Courant and Hilbert, 1953).

Table 1.

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.

RESULTS AND DISCUSSION

The performance of Oral Submucous Fibrosis classification (OSMF) using GMM has been analyzed. Color Histogram features are extracted to characterize the OSMF images. Classifiers GMM are applied to obtain the optimal class boundary between the two classes namely normal and OSMF images by learning from training data.The classification rate using 64 histogram bins showed an accuracy of above 90% in all the models. GMM shows the highest accuracy of 91% for OSMF classification with 64 bins using Gaussian kernels. Fig. 2. 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 2.

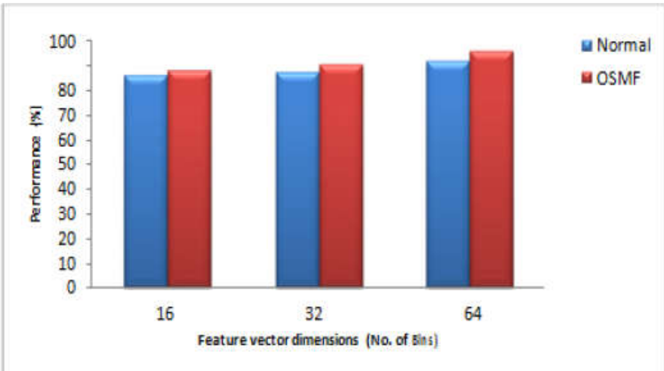

Fig. 2. Average performance of normal and OSMF classification with different bins by GMM model with 10 mixtures using Histogram features

Conclusion

In this paper, a system for classifying OSMF affected images from normal images was proposed. Color histogram features where extracted from both normal and OSMF affected images. The features where trained and tested using GMM for different bins. The system showed an accuracy of 91.0% for 64 bins. In future other pattern recognition algorithms can be analyzed and the performance can be studied for computerized diagnosis of OSMF

REFERENCES

Ahmad M. Sarhan, 2009. "Cancer Classification Based On Microarray Gene Expression Data Using DCT and ANN", *Journal of Theoretical and Applied Information Technology,* vol. 6, no. 2, pp. 208-216.

Anuradha K and Sankaranarayanan K. 2013. "Comparison of Feature Extraction Techniques to classify Oral Cancers using Image Processing", *International Journal of Application or Innovation in Engineering & Management (IJAIEM),* vol. 2, no. 6, pp. 456-462.

Artur Chodorowski and Tomas Gustavsson, Ulf Mattsson, 2002. "Support Vector Machines For Oral Lesion Classification", Biomedical Imaging, 2002 Proceedings, *IEEE International Symposium,* pp. 173-176.

Bharathi A and Natarajan A. M. 2011. "Enhanced Multicategory Classification Using Fast SVM Learning (FSVML) for Microarray Gene Expression Cancer Diagnosis", *International Journal of Computer Technology and Electronics Engineering (IJCTEE),* vol. 2, no. 2, pp.63-71.

Duda R, Hart P and Stork D. 2000. "Pattern Classification", Second Edition, John Wiley & Sons, Inc.

Ganesan N, Venkatesh K, Rama M. A and Malathi Palani A, 2010. "Application of Neural Networks in Diagnosing Cancer Disease Using Demographic Data", *International Journal of Computer Applications,* vol. 1, no. 26, pp. 76-85.

Hari Kumar R, Vasanthi N.S and Balasubramani M, 2012. "Performance Analysis of Artificial Neural Networks and Statistical Methods in Classification of Oral and Breast Cancer Stages", *International Journal of Soft Computing and Engineering (IJSCE),* vol. 2, no. 3, pp. 263-269.

Ireaneus Anna Rejani Y and Thamarai Selvi S. 2009. "Early Detection of Breast Cancer Using SVM Classifier Technique", *International Journal on Computer Science and Engineering,* vol. 1, no. 3, pp. 127-130.

Mitesh Amitkumar Modi, Vishal R. Dave, Viral G. 2012. 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.

Muthu Rama Krishnan M, Mousumi Pal, Suneel K. Bomminayuni, Chandan Chakraborty, Rajan Rashmi Paul, Jyotirmoy Chatterjee and Ajoy K. Ray, 2009. "Automated classification of cells in sub-epithelial connective tissue of oral sub-mucous fibrosis - An SVM based approach", *Computers in Biology and Medicine,* vol.39, no. 12, pp. 1096-1104.

Ramachandran P, Girija N and Bhuvaneswari T. 2013. "Cancer Spread Pattern - an Analysis using Classification and Prediction Techniques", *International Journal of Advanced Research in Computer and Communication Engineering,* vol. 2, no. 6, pp. 2363-2367.

Rusha Patra, Chandan Chakraborty and Jyotirmoy Chatterjee, 2012. "Textural Analysis of Spinous Layer for Grading Oral Submucous Fibrosis", *International Journal of Computer Applications,* vol. 47, pp. 0975-0987.

Shabana A.H.M, El-Labban N.G. and LeeK.W, 1987. "Morphometric analysis of basal cell layer in oral premalignant white lesions and squamous cell carcinoma", *Journal of Clinical Pathology,* vol. 40, pp. 454-458.

Shilpa B. Rajesh and Satheesha B. H. 2012. 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.

Tathagata Ray M. E, Shivashankar Reddy D, Anirban Mukherjee, Jyotirmoy Chatterjee, Ranjan R. 2008. 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.

****