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

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

International Journal of Current Research
Vol. 8, Issue, 09, pp.39439-39442, September, 2016

ISSN: 0975-833X

RESEARCH ARTICLE

HISTOGRAM AND BICC FEATURES FOR CLASSIFYING OLP IMAGES

*Dr. Venkatakrishnan, S.

Assistant Professor, Mechanical Engineering, Engg-Wing DDE, Annamalai University

ARTICLE INFO

Article History:
Received 20th June, 2016
Received in revised form
09th July, 2016
Accepted 18th August, 2016
Published online 30th September, 2016

Key words:
Feature extraction,
Histogram,
Image Classification, Aann,Gmm.

ABSTRACT

Oral Lichen Planus (OLP) are the most common oral cavity lesions that can progress to cancer. The only reliable method to diagnose such diseases is microscopic examination of tissue samples. The proposed method analyses the performance of pattern classification techniques such as Auto associative Neural Network (AANN) and Gaussian Mixture Model (GMM) to classify oral precancerous images, using features that have significant characteristics associated with the oral disorders. The features used in this work are than 90.0% in all the models. Histogram features are extracted from OSMF images. AANN and GMM are used to classify the images into normal and OSMF affected categories. The classification performance ranges from 88% to 97.6% for all these models. Then the proposed method extracts BICC features from OLP images and the performance of AANN and GMM in classifying OLP images are analyzed. Classification results show an accuracy ranging from 90% to 96% for all these models. Color histogram and BICC features are extracted from oral lichen planus images. The classification performance using AANN and GMM are analyzed.

Citation: Dr. Venkatakrishnan, S. 2016. "Histogram And BICC Features For Classifying OLP Images", *International Journal of Current Research*, 8, (9), 39439-39442.

INTRODUCTION

Oral Lichen Planus (OLP) is a chronic inflammatory disease of unknown etiology. Application of computer aided diagnosis in the analysis and classification of oral cancer and precancerous lesions could overcome the difficulties of subjective variations and minimize the effort and time taken for diagnosis by oncologists. Oral lichen planus presents as white striations, white papules, white plaques, erythema, erosions or blisters affecting predominantly the buccal mucosa, tongue and gingivae, although other sites are occasionally involved. Oral lichen planus affects 1-2 per cent of the general adult population and is the most common non- infectious oral mucosal disease. Oral lichen planus affects women more than men (1.4:1). Oral lichen planus occurs predominantly in adults over 40, although younger adults and children may be affected. Lesions are typically bilateral and often appear as a mixture of clinical subtypes. White or grey streaks may form a linear or reticular pattern on an erythematous background.
Symptoms of oral submucous fibrosis include:

- 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

In 2007, Kunio Doi [2007], described "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential". In this article, the motivation and philosophy for early development of CAD schemes are presented together with the current status and future potential of CAD in a PACS environment. With CAD, radiologists use the computer output as a "second opinion" and make the final decisions. CAD is a concept established by taking into account equally the roles of physicians and computers, whereas automated computer diagnosis is a concept based on computer algorithms only.

***Corresponding author:* Dr. Venkatakrishnan, S.,**
Assistant Professor, Mechanical Engineering, Engg-Wing DDE, Annamalai University.

With CAD, the performance by computers does not have to be comparable to or better than that by physicians, but needs to be complementary to that by physicians. In fact, a large number of CAD systems have been employed for assisting physicians in the early detection of breast cancers on mammograms. In February 2013, Ashraf Afifi, Zanaty E.A and Said Ghoniemy described "Improving the Classification Accuracy Using Support Vector Machines (SVMS) with New Kernel" In this paper, a new kernel function called polynomial radial basis function (PRBF) that could improve the classification accuracy of support vector machines (SVMs) has been introduced. The proposed kernel function combines both Gauss (RBF) and Polynomial (POLY) kernels and is stated in general form. It is shown that the proposed kernel converges faster than the Gauss and Polynomial kernels. The accuracy of the proposed algorithm is compared to algorithms based on both Gaussian and polynomial kernels by application to a variety of non-separable data sets with several attributes. We noted that the proposed kernel gives good classification accuracy in nearly all the data sets, especially those of high dimensions.

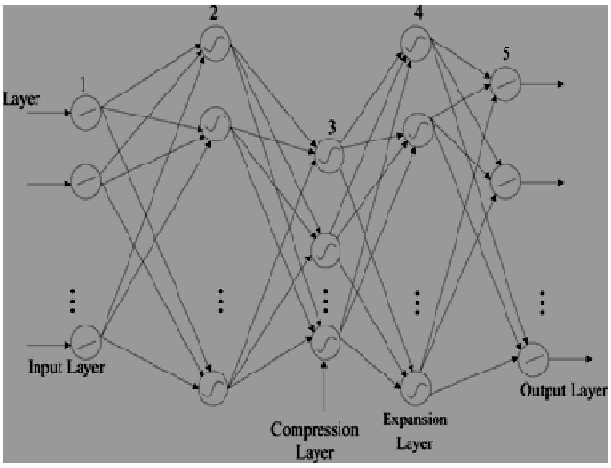
Feature Extraction

Feature Selection: Color histogram features are extracted from both OLP affected and normal images. BICC features are extracted from both OLP affected and normal images. Among the three kernels, Gaussian kernel gives a maximum performance for Histogram, BICC and combined features. Experiments were conducted for histogram features, BICC features and combined features. The distribution of 16, 32 and 64 histogram bins and 10, 45 and 105 dimensional feature vectors in the feature space for different sized blocks of BICC feature vectors is captured using an AANN and GMM model. The AANN and GMM 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 BICC feature vectors.Gaussian Mixture Models are a type of density models comprises a number of components which are combined to provide a multimodel 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. The classification performance increases, as the number of mixtures increases. The performance of OLP classification was studied using GMM for different mixtures. Expectation maximization algorithms were used to decide the parameters of the mixtures.

To construct a histogram: 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. The technique used for obtaining a uniform histogram is known as histogram equalization or histogram linearization.

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 Histogram and BICC features for classifying OLP using AANN and Gaussian Mixture Model. 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.

Autoassociative Neural Network (AANN): An Artificial Neural Network (ANN), usually called Neural Network, is a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. The feed forward neural network was the first and arguably simplest type of artificial neural network devised.

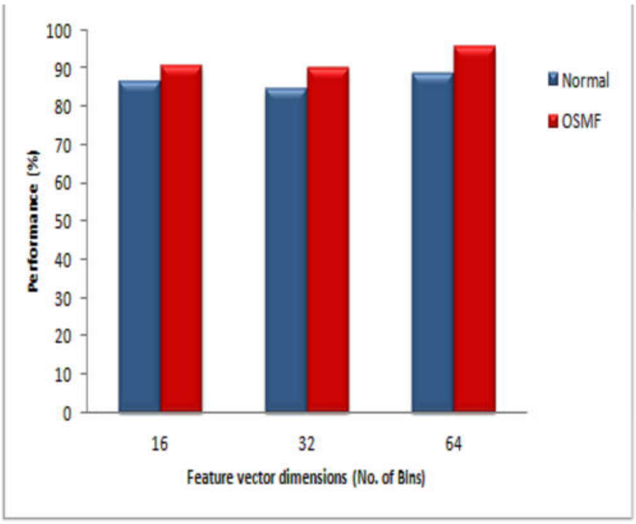


Gaussian Mixture Model (GMM): Gaussian Mixture Model (GMM) is a mixture of several Gaussian distributions and can therefore represent different subclasses inside one class. The distribution of the histogram features is captured using GMM. Gaussian Mixture Models are a type of density models comprises 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. The kernel function may be any of the symmetric functions that satisfy the Mercer's conditions (Courant and Hilbert,1953).

RESULTS AND DISCUSSION

Oral lichen planus presents as white striations, white papules, white plaques, erythema, erosions or blisters affecting predominantly the buccal mucosa, tongue and gingivae, although other sites are occasionally involved.

Architecture of the AANN



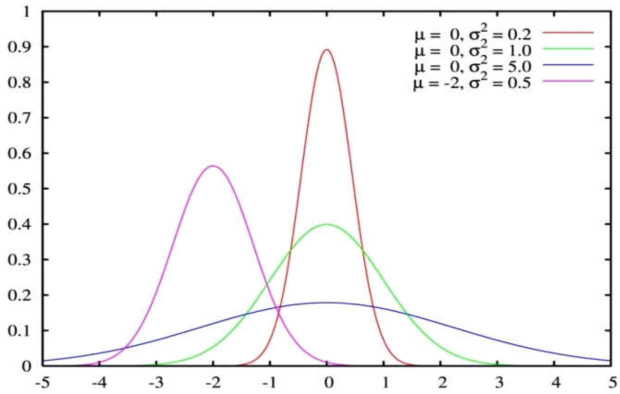


Fig. 1. Architecture of the GMM

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.

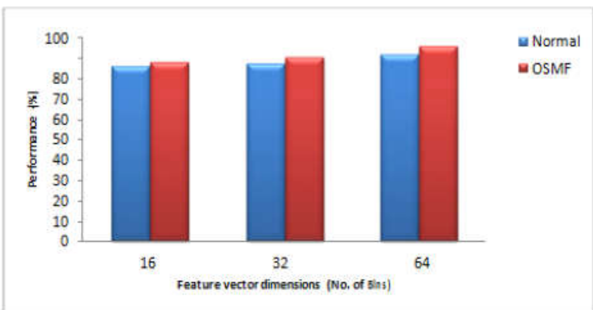


Fig. Average performance of normal and OLP classification with different bins by AANN and GMM model with 10 mixtures using Histogram features

Classifiers AANN and GMM are applied to obtain the optimal class boundary between the two classes namely normal and OLP images by learning from training data. 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.

Conclusion

In this paper, a system for classifying OLP affected images from normal images was proposed. Color histogram features where extracted from both normal and OLP affected images. The features where trained and tested using AANN and 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 OLP

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.

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.

Ashraf Afifi, Zanaty E. 2013. A and Said Ghoniemy, "Improving the Classification Accuracy Using Support Vector Machines (SVMS) with New Kernel", Journal of Global Research in Computer Science, vol. 4, no. 2, February.

Chandra Shekar and Sachin Ganesan, 2011. "Oral Lichen Planus", Journal of Dental Sciences & Research, vol. 2, no. 1, pp. 62-87, February.

Chodorowskia A, Mattsson U and Gustavsson T. 2000. "Oral lesion classification using true-color images", Doktorsavhandlingar vid Chalmers Tekniska Hogskola 1646, vol. 1, no. 12.

Fernando Augusto C, Garcia de Sousa, Paradella and Thas Cachut, 2009. "Malignant potential of oral lichen planus: A meta-analysis", Revista Odonto Ciência (Journal of Dental Science), vol. 24, no. 2, pp. 194-197, February.

Kalaiselvi Geetha M, Palanivel S and Ramalingam V. 2009. "A novel block intensity comparison code for video classification and retrieval", Expert Systems with Applications, vol. 36, no. 3, pp. 6415-6420, April.

Krishna Kanth B.B.M. 2013. "A Fuzzy-Neural Approach for Leukemia Cancer Classification", International Journal of Scientific Research (IJSR), vol. 2, no. 11, pp. 206-208, November.

Kunio Doi, 2007. "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential", Computerized Medical Imaging and Graphics, Elsevier, vol. 31, pp. 198-211.

Medhat Mohamed Ahmed Abdelaal, Muhamed Wael Farouq, Hala Abou Sena and Abdel-Badeeh Mohamed Salem,

2010. “Applied Classification Support Vector Machine for providing Second Opinion of Breast Cancer Diagnosis”, The Online Journal on Mathematics and Statistics (OJMS), vol. 1, no. 1, 2010.

Pradeep N, Girisha H, Sreepathi B and Karibasappa K. 2012. “Feature Extraction of Mammograms”, International Journal of Bioinformatics Research, vol. 4, no. 1, pp. 241-244, February.

Sanjit Mukherjee, Jay Gopal Ray and Keya Chaudhuri, 2010. “Microscopic analysis of histological and immunohistochemical sections to differentiate normal, precancer and cancerous oral squamous epithelial tissues”, Microscopy: Science Technology, Applications and Education, pp. 993-1000.

Shekhar Singh, 2012. “Cancer Cells Detection and Classification in Biopsy Image”, International Journal of Computer Applications, vol. 38, no. 3, pp. 15-21, January.

****