International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958, Volume-8 Issue-5, June 2019

Design of A Decision Support System for Detection of Oral Cancer using Matlab

N. Kripa, R. Vasuki, Prasath Alias Surendhar

Abstract: Oral squamous cell carcinoma (OSCC) represents the predominant neoplasm of the neck and head region, particularly featuring aggressive nature with unfavorable prognosis which is associated along with it. In this research the support system detects and classifies the oral squamous cell carcinoma [1]. The algorithms GLCM and Feed Forward Neural Network (FNN) are utilized to predict the occurrence of oral cancer through the analysis using MATLAB. We try to systematically study and analyze the basis of oral cancer evolvement by the image processing technique using neural network toolbox in Matlab.Images of normal and abnormal images were collected and the feature extraction for all these images was carried out using GLCM. These features extracted were compared and selected accordingly for the classification. The images were initially classified based on thresholding according to variations shown in the images. Later a Feed Forward Network (FNN) was developed and trained to predict the occurrence of oral squamous cell carcinoma [2]. The values of the features selected were given as input to train based on this training the classifier gives the output as cancer or normal image. The Results were obtained as our initial goal and it was observed that the accuracy in the results is better when we use FNN classifier than thresholding.

Keywords – Oral Squamous Cell Carcinoma (OSSC), Feed forward Neural Network, MATLAB, Gray-Level Co-Occurrence Matrix (GLCM), Oral Cancer, Feature Extraction.

I. INTRODUCTION

According to the world wide cancer ranking the oral cancer stands in 8th place and its one of the major reasons in causing the oral cancer is the malignant neoplasm. Oral squamous cell carcinoma is found to be more predominant in individuals who do more smoking along with consumption of alcohol. As with this factor when the sex of the patients is been taken into account wherein the men is only been more affected than the women population because of this factor. Another factor which is adjoin to it is the exposure to sun which aggravates the OSCC development especially in the lips and the occurrence rate of locoregional relapses even leads to inability and death. The prediction and detection of the cancer is a significant and risk involved process , hence we are required to build a decision support system that is efficient to detect the cancer. To enhance the task of detection of oral cancer, we used the Artificial Neural Network (ANN) technique under the Neural Fuzzy logic Network [3].

The most usual method is taking the histopathological evaluation by taking biopsy from the tissue to diagnose the oral cancer. This sometimes leads to show the wrong diagnostic symptom like asymptomatic lesion, so that the patients will be guided to see the dentist for the treatment because, these oral cancers will normally does not show early signs of cancer [4]. So most oral screening programs include the visual examination and the usage of chemiluminesce, toluidine blue, brush biopsy and fluorescence imaging. Several classifiers are been developed in the neural network to detect the cancer images by image processing of the cell images of the patients.

II. METHODOLOGY

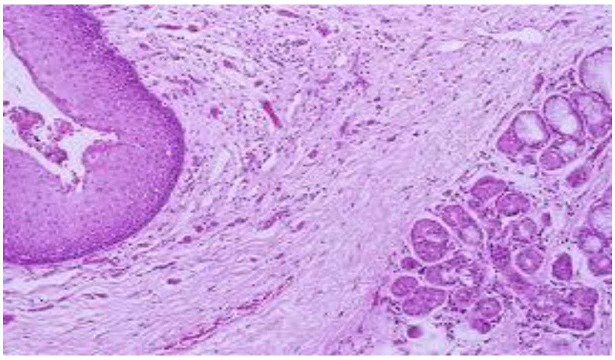


Fig 1: Histology of Normal image

Fig 2: Histopathology Cancer image

It is evident that we can differentiate by just looking, but its not accurate and we can go wrong, hence we have to follow the methods [5].

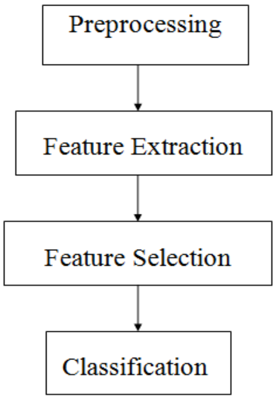


Fig 3: Methodology of Image Processing

II.A. Pre-processing

Pre-processing of the image is the first step to be carried out. Filtering the gray scaled image is the significant process to get a image with a lesser signal to noise ratio.

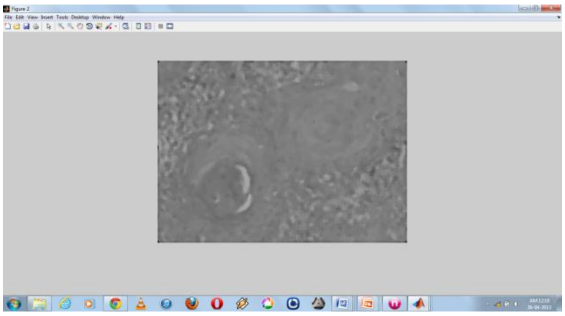
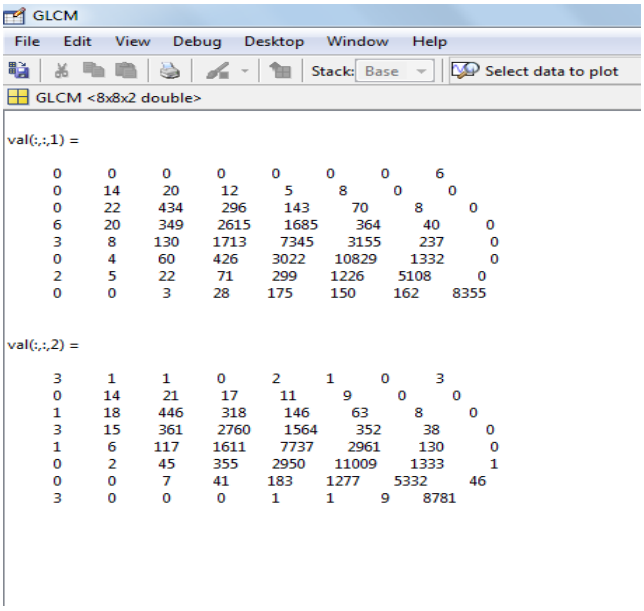


Figure 4: Cancer image after pre-processing

II.B. Feature Extraction

We have extracted various features of the image to perform the classification. Intensity-t, Variance –y, Skewness-X, Area-l, Mean-M, Standard deviation-s, Entropy –E, Histogram-D. Results of the image being converted into a GLCM matrix is displayed below. The code was developed from the glcm toolbox in the matlab[6]. The stats function gives the four parameters Contrast, Energy, Homogeneity and Correlation under the GLCM.

Table 1: GLCM matrix



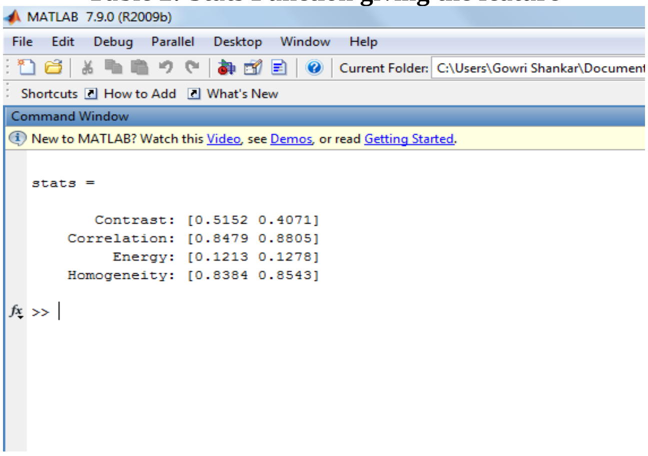
val(:,:,1) =

0	0	0	0	0	0	0	6
0	14	20	12	5	8	0	0
0	22	434	296	143	70	8	0
6	20	349	2615	1685	364	40	0
3	8	130	1713	7345	3155	237	0
0	4	60	426	3022	10829	1332	0
2	5	22	71	299	1226	5108	0
0	0	3	28	175	150	162	8355

val(:,:,2) =

3	1	1	0	2	1	0	3
0	14	21	17	11	9	0	0
1	18	446	318	146	63	8	0
3	15	361	2760	1564	352	38	0
1	6	117	1611	7737	2961	130	0
0	2	45	355	2950	11009	1333	1
0	0	7	41	183	1277	5332	46
3	0	0	0	1	1	9	8781

Table 2: Stats Function giving the feature



II.C. Feature Selection

We concluded that Entropy, Energy and Homogeneity are the features that can be selected for accurate classification.

Table 3: Results for feature selection- Normal image

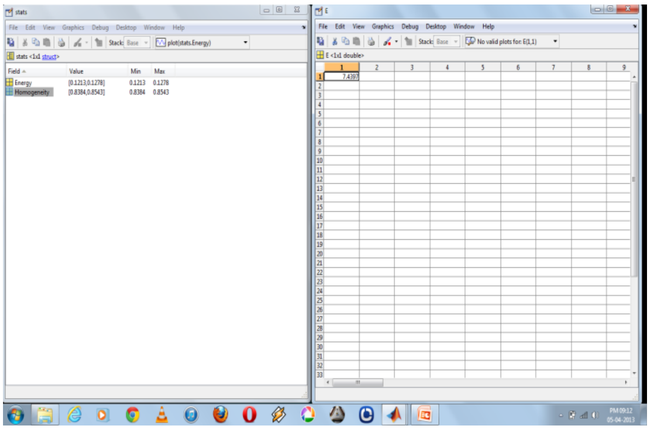
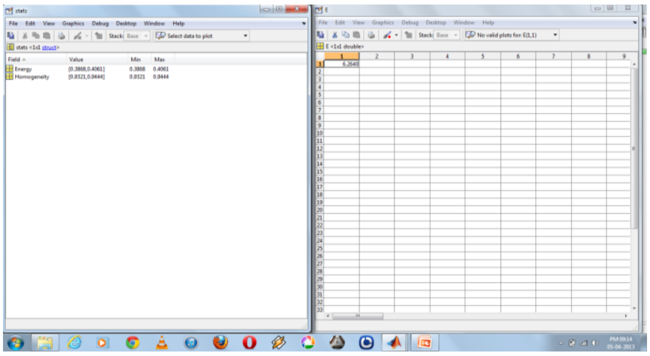


Table 4: Results for feature selection- Cancer image



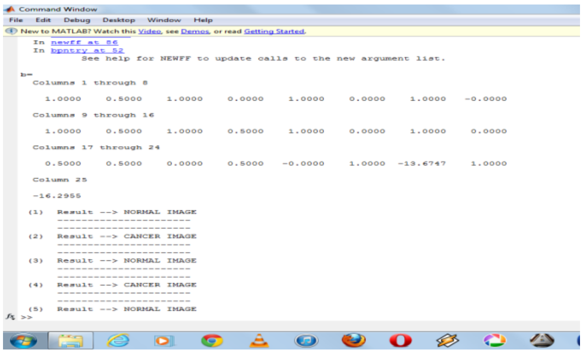
II.D. Classification

Feed Forward Network (Fnn)

In this classifier we feed the feature values of the images in an array and train the classifier using those values . Energy and Homogeneity features were taken. Array of input vectors (features) were taken and a target was set for first 20 images and 5 images for testing. The target was set as 1 for cancer images and 0 for normal images. Net=newff (P,T..) function was used to perform this classification[7].Once the training is done the trained values are stored in a variable b, which is used for testing images. The simout() function was used to store the values in b.

The testing is done by grouping the values of b as cancer images for 1 and normal images for 0[8].

Table 5: FNN Classifier Results



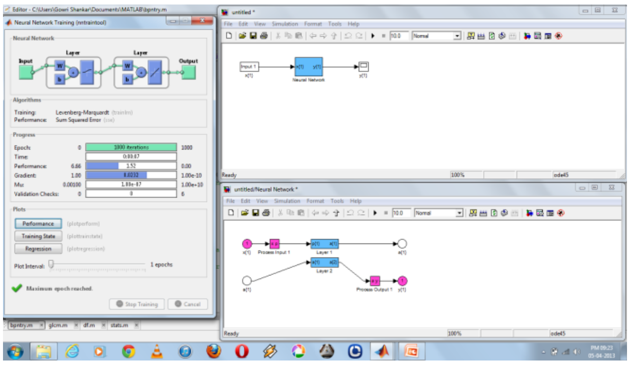


Figure 5: Feed forward Network Developed

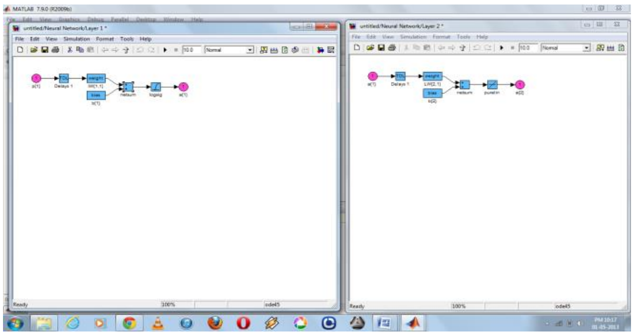


Figure 6: Network with weights and biases

Finally the classifier displays the images as cancer or normal images and gives the feed forward network that was trained and developed [9]. The code shown below is used to train the feed forward network into groups as 0 and 1 so that the images are trained and when we test it classifies correctly.

III. RESULTS

The Network performance is shown in the following figure with the Sum squared error (sse) against the number of Epochs . It is seen that the error decreases as the number of epoch increases.

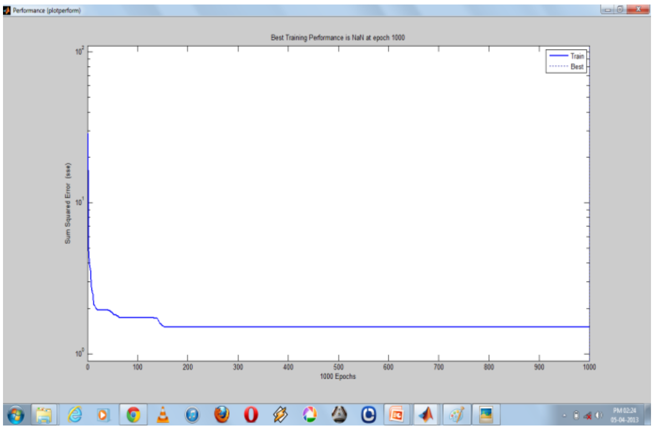


Figure 7: Network Performance

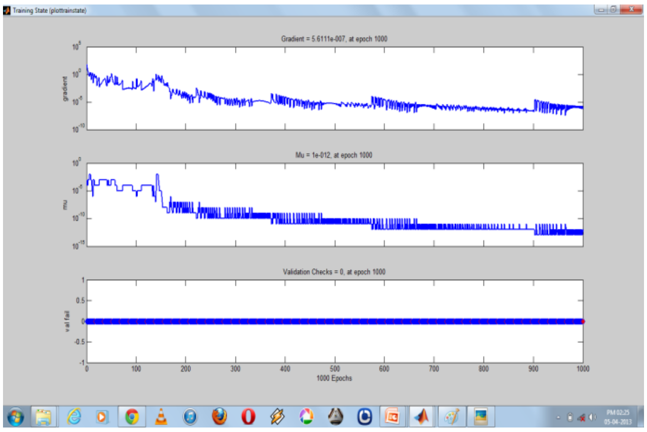


Figure 8: Training state performance

The capability to generalize is one of the major advantages of neural networks. So for this the trained net should divide the data from the same class. It is defined as the learning data which it has not seen before. In application wise the program developers have particular parts only for all possible patterns for neural network generalization. In order to have the maximum generalization, the data sets are considered into three parts. The main thing is that the error is been driven to minimization during training. This set of training is used to train to a neural network. The neural network patterns performances which are not done, training determines the validation set. When the validation set error reaches the minimum the learning and training is stopped. But if in the case the training is not stopped in the appropriate time, overtraining occurs and so the whole data net work decreases, while even the error still decreases on the training data. After finishing the learning phase, the net should be finally checked with the third data set, the test set and classification is done.

IV. CONCLUSION

Usage of neural network is not appropriate to be used for all the cases because some may be having many complex equations which could not be solved with this method. Even though the ANN method is the good and powerful tool, but when compared to the classical methods like PCA, cluster analysis, MLRA, pattern recognition, etc. gives even better results in minimized time. The application based on these empirical models involving the neural networks chooses only minimal set of data.

Even of the tolerance of error is met the trained network data should be subjected to examination for sign of bias [10]. We were able to successfully design a decision support system for detecting the occurrence of oral squamous cell carcinoma. We had formulated a GLCM matrix and extracted features for classification.Thresholding technique was initially implemented for classification for classification of normal/abnormal (cancer) images. To verify the results and to improve the accuracy, a Feed Forward Neural Network was developed. The Accuracy of the Feed forward Network is above 90% and requires less computation time.

REFERENCES

1. K. Anuradha, K.Sankaranarayanan, "Detection of Oral Tumor based on Marker – Controlled Watershed Algorithm", *International Journal of Computer Applications,* Volume 52– No.2, August 2012, 15-18.
2. Cameron Goertzen, Hayder Mahdi, Catherine Laliberte, Tomer Meirson, Denise Eymael, Hava Gil-Henn, and Marco Magalhaes, "Oral inflammation promotes oral squamous cell carcinoma invasion", Oncotarget. 2018 Jun 26; 9(49): 29047–29063.
3. Karla Maria Carvalho, Poonam Ramnath Sawant, Anita Dhupar, and Anita Spadigam, "A Case of Oral Squamous Cell Carcinoma in a Nontobacco Habitué*", Int J Appl Basic Med Res*. 2017 Oct-Dec; 7(4): 278–280.
4. Guang Li,1, Xian Li,2, Meng Yang, Lvzi Xu, Shixiong Deng, and Longke Ranb, "Prediction of biomarkers of oral squamous cell carcinoma using microarray technology*", Sci Rep*. 2017; 7: 42105.
5. Patricia Soo-Ping Thong ; Malini Olivo ; Stephanus Surijadarma Tandjung ; Muhammad Mobeen Movania, "Review of Confocal Fluorescence Endomicroscopy for Cancer Detection", *IEEE Journal of Selected Topics in Quantum Electronics*, Volume: 18 , Issue: 4 , July-Aug. 2012, 1355 – 1366.
6. K.A. Shahul Hameed ; A. Banumathi ; G. Ulaganathan, "Segmentation of immunohistochemical staining of β-catenin expression of oral cancer using Gabor filter technique", *IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM - 2012)*
7. J. C. Melchor, H. Navas M, Marcos A. Iza, M. De Diego, D. Rando, I. Melchor J. Burgos, "Predictive performance of PAM vs fFN test for risk of spontaneous preterm birth in symptomatic women attending an emergency obstetric unit: retrospective cohort study", *Ultrasound in Obstetrics and Gynecology*, 29 August 2017.
8. Thomas E. Yankeelov, Richard G. Abramson, and C. Chad Quarles, "Quantitative multimodality imaging in cancer research and therapy", *Nat Rev Clin Oncol.* 2014 Nov; 11(11): 670–680.
9. Supreeta Arya, Devendra Chaukar,1 and Prathamesh Pai, "Imaging in oral cancers", *Indian J Radiol Imaging*. 2012 Jul-Sep; 22(3): 195–208.
10. Brian M. Trotta , Clinton S. Pease, Jk John Rasamny, Prashant Raghavan, Sugoto Mukherjee, "Oral Cavity and Oropharyngeal Squamous Cell Cancer: Key Imaging Findings for Staging and Treatment Planning*", RadioGraphics*, Vol. 31, No. 2, 599-610.