Intelligent Decision Support Systems for cancer diagnosis through state of the art machine learning algorithms

Budur Ibrahim Alodhaybi
MSc in Information Technology

UNIVERSITY of STIRLING

Introduction :

Breast cancer is the most common disease in woman in developed as well as developing countries . The survival rate can be improved if early detection and adequate diagnosis are made available . One of the effective methods to detect early signs of breast cancer is mammography. Mammography is further assisted by Computer Aided Decision Support System (CADSSs) these system improve diagnosis by identifying the area in mammograms that's contain cancer .

Experiment :

1- image pre-processing : This step include breast image segmentations and filtering followed by normalization or enhancement to improve image quality and reduce noise .

2- Feature Extraction : Lesions are extracted from enhanced images using LESH techniques.

3- Feature Selection : a set of feature selected

4- classification : the selected feature set is classified to separate false signals from true one. The overall system shows below figure 1 .

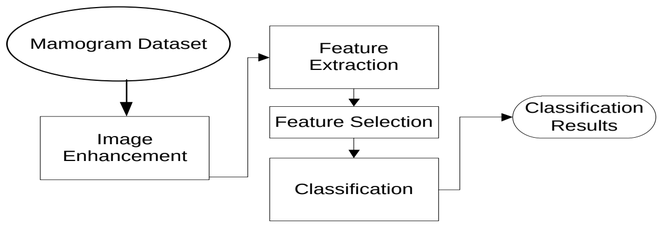

Figure 1 . Steps of experiments .

Methodology :

1- mammogram dataset : Experiments is conducted on Mammographic Image Analysis Society (MIAS). This Dataset contain images of 1024*1024 pixels size.

The Images have been labelled by radiologist and provide variety of abnormal cases such as : Circumscibed Masses(CIRC) , I11-defined Masses (MISC), Speculated Masses(SPIC), Calcification (CALC), Architec- Tual Distortion(ARCH).

2- image pre-processing : the images have low contrast which Made noise in diagnose . CLAHE (Contrast Limited Adaptive Histogram Equitation) was applied for image enhancement Showing bellow finger 2 :

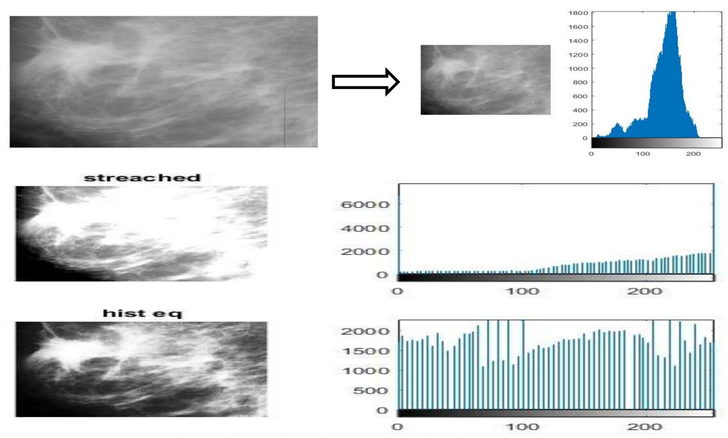

Figure 2 . Applying CLAHE and Histogram Equalization .

3- Feature Extraction and Selection : Local Energy Based Shape Histogram(LESH) feature are calculated at the point of maximum phase congruency , given as follow figure 3 .

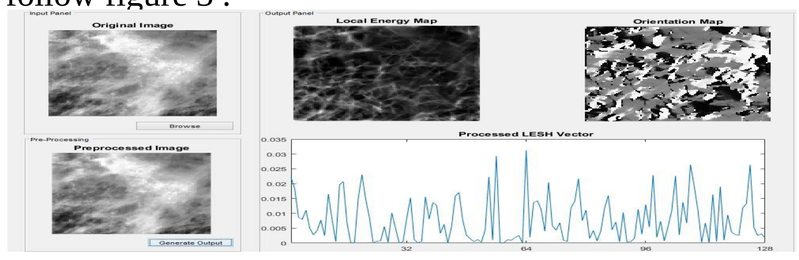

Figure 3 . Shows applying LESH feature to dataset .

4- Classification : the SVM (support Vector Machine) and ESN (Echo State Network) Classifiers has been applied to data set MIAS and the result show below table 1 and 2 :

Features selected	SVM with RBF	SVM with Linear	SVM with MLP	SVM with polynomial
50	99.09	99.09	94.44	99.17
70	99.09	**100.00**	94.54	98.18
100	99.09	99.09	93.63	96.36
All	99.09	99.09	94.62	96.51

Table 1 :classification accuracy of malignant /benign cases for SVM classifier .

Fold	Performance Measures of ESN with 100 LESH feature selected			Performance Measures of ESN with 50 LESH feature selected		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
1	95.24	100.00	83.33	90.48	100.00	83.33
2	100.00	100.00	100.00	71.43	66.67	77.78
3	100.00	100.00	100.00	100.00	100.00	100.00
4	100.00	100.00	100.00	100.00	100.00	100.00
5	95.24	100.00	88.89	85.71	100.00	70.00
Average	99.05	100.00	97.22	84.29	88.11	81.28

Table 2 :classification accuracy of malignant /benign cases for ESN classifier

5-model evaluation : The ROC plots below describe the classification performance for the MIAS datasets for different SVM kernels. The area under the curve (AUC), indicates the classification performance as one example for evaluate model . Figure 4.

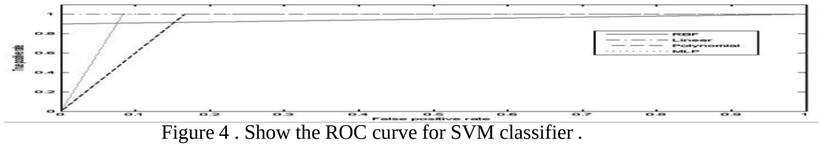

Figure 4 . Show the ROC curve for SVM classifier .

Conclusion :

ESN performs as good as SVM, when 100 feature coefficients are selected from LESH feature set, otherwise SVM performs better in case 50 feature coefficients are selected. A good combination of methods at each stage of CADSSs can enhance its performance.

Reference : Wajid,S.K,Hussain A.,luo,B.(2014) Local energy-based shape histogram feature extraction technique for breast cancer diagnosis , 2014 IEEE .