Chapter 8

CONCLUSION

This chapter gives the conclusions and main findings of the thesis. In the beginning, main highlights of thesis and the contributions are explained. This is followed by the limitations and future scope of the work is presented.
8.1 Conclusion

The outcome of this thesis is a set of novel algorithms implemented in various stages of the Computer Aided Diagnosis systems for early detection and classification of abnormalities in mammogram images. This thesis mainly focused on the different aspects of the detection and classification of abnormalities in mammogram images using different machine learning algorithms. The different classification methods such as Distance measure, MLP, ELM and Lazy classifiers are used in our proposed algorithms. A novel feature extraction method is used for constructing feature vector, which acts as the input for the classifiers, from the Wavelet transformation coefficients and Gray Level Co-occurrence matrix. In addition to the above features, dimensionality reduced DWT coefficients using PCA are also used as a feature vector for the classification. Comparing the performance of the above classification algorithms, excellent classification results are obtained for ELM, MLP and Lazy classifiers (K*/IBL) for Wavelet Transform coefficients as well as GLCM feature vector. Two clustering methods such as K-means and Fuzzy C means are also implemented and used for validating the images obtained from local hospitals. The results obtained by the clustering methods are in agreement with the expert opinion. Novel mammogram enhancement algorithms (WSTA) and automatic mammogram segmentation algorithms are also introduced in this thesis as part of the preprocessing and segmentation of abnormalities in mammogram images.

8.2 Major contributions

- A novel feature extraction method is introduced in this thesis. The texture pattern of an image can be isolated by viewing the abnormal area of the mammogram images in different orientations. Feature
vectors of the ROIs of the mammogram images are constructed using this concept. Two different texture feature extraction methods such as Wavelet Transformation and GLCM are used. The fractional part of the biggest approximation coefficients of DWT decomposition are used for creating the feature vector. In GLCM features, four different GLCM matrices are constructed in four different orientations of the ROIs. Four important features from each GLCM are derived and combined to form a single feature vector comprising of 16 features. This is a novel as well as the most significant feature vector that we constructed and used for the classification of mammogram images into different categories.

- A novel mammogram enhancement method WSTA is introduced in this thesis for enhancing the mammogram images. Wavelet Semi Thresholding Algorithm (WSTA) based on morphological top hat filtering and Hit-or-miss transformation is used for the enhancement of the mammogram images. This algorithm enhances the low intensity borders and fine details of the image in a better way.

- Extracting the Region of Interest where the abnormality lies is one of the most difficult tasks in medical image analysis. In this thesis, fully automated and efficient algorithm for extracting the ROIs of the mammogram images is implemented.

- We identified and established that the classification of mammogram images using Euclidean distance measure is better than any other distance measures.

- Choosing the fractional part of the biggest wavelet transformation coefficient influences the classification accuracy.
Wavelet filters such as $db8$, $db16$ and Biorthogonal wavelet filters are the most suitable wavelet filters for classification.

The classification performance is increased as we applied dimensionality reduction technique such as PCA on wavelet transform coefficients.

Introduced a recently developed Neural Network classifier named Extreme Learning Machine (ELM) to the world of mammogram classification.

The classification results obtained by ELM are also promising. Even though the classification accuracy obtained by ELM is slightly lesser than MLP, the time and cost incurred by the ELM is minimum compared to MLP.

Introduced three different fast and instance based Lazy classifiers for the classification of mammogram images.

The performance of the classification of mammogram images using GLCM and Lazy classifier also stood in top for K* and IBL classifiers.

8.3 Limitations

In the case of mammogram enhancement, we limited our study by focusing only on the wavelet based thresholding algorithms. Soft computing based enhancement algorithms can be experimented for getting better enhancement results.

In mammogram enhancement, the thresholding method are applied only on the detailed wavelet decomposition coefficients in the assumption that noise as well the unwanted components of
mammogram image always reside in the high frequency components. But there may be a possibility for applying thresholding on approximation coefficients of the WT for obtaining better enhancement.

- In the proposed automatic segmentation algorithm for extracting the ROIs from mammogram images, multiple seed points are being identified. But we did not utilize all the seed points identified by our algorithm for extracting multiple region of interest in the image.

- Using the fractional part of the biggest wavelet transform coefficients and the machine learning algorithms, we classified most of the mammogram images into different categories depending upon the abnormality in the image. But our classification algorithms do not have the provision to identify which approximation coefficients are significant for the classification of mammogram images.

- Using the GLCM feature extraction method we constructed GLCMs based on four different orientations such as $0^\circ$, $45^\circ$, $90^\circ$, $135^\circ$ and the feature vector is constructed using only these four orientations. Better classification results may be achieved, if we construct four GLCM matrices in other orientations too.

- In this thesis, we considered only four features from GLCM matrices, constructed based on four different orientations. So the experiments can be conducted and evaluated for better classification performance by including additional features from the GLCM matrices.

- A detailed study need to be conducted for increasing the performance of the classification using ELM, leading to better classification results.
with minimum time and effort compared to conventional MLP and other classifiers.

- The ultimate aim of dimensionality reduction is to reduce the number of features from the dataset. Using PCA, the dimensionality of the feature set is reduced by representing them into different Eigen values and Eigen vectors. These Eigen values and Eigen vectors are not direct wavelet approximation coefficients that we obtained from the decomposition.

### 8.4 Future scope of the work

Although the present work yields good results, certain proposals for future works are listed below.

- Enhancement of mammogram images may be performed using some soft computing based techniques for obtaining better improvement.

- The automatic segmentation algorithms implemented in this thesis can be extended to deal with the extraction of multiple ROIs from the different seed points identified by our algorithm.

- A fast machine learning algorithm such as Extreme Learning Machine (ELM) introduced in this thesis is to be studied in detail for improving the organization and performance of the ELM for mammogram classification with better accuracy in minimum time and effort.

- A possibility may be analyzed for identifying and reducing the attribute which have little significance in the classification process.