Chapter Three

Literature Review

The art of image processing, specifically image enhancement and pattern recognition and classification have progressed remarkably over the past few decades, improving the efficacy of mammography vision and interpretation. This chapter details the different techniques and methodologies in previous literature which evolved applications of image enhancement and pattern recognition and classification.

As a preprocessing stage of a CAD system, Papadopoulos et al. [35] in 2008 employed five enhancement methods; contrast-limited adaptive histogram equalization (CLAHE), the local range modification (LRM) and redundant dyadic wavelet transform (RDWT) utilizing linear stretching and wavelet shrinkage (WSRK) techniques for image improvement.

Another preliminary stage to a CAD system that has been proposed by Rahmati et al. [36] is a preprocessing filter named fuzzy contrast-limited adaptive histogram equalization (FCLAHE). It is based on refining aspects of the enhancement algorithm, CLAHE found in the works of Pisano et al. [30] followed by employing a nonlinear fuzzy function to overcome inhomogeneities that compromise image quality.

Fuzzy logic has been used once again in the works of Basha et al. [37] in 2005 – 2009 as part of an automatic detection of breast cancer mass in mammograms. Initially, the differentiation of masses and microcalcifications is achieved by segmentation using morphological operators. Fuzzy c which implies clustering for intensity based segmentation was utilized to highlight features of cancer masses.

Discrete wavelet transform (DWT) was used to derive statistical wavelet features, co-occurrence wavelet features and a hybrid of the two into three different feature databases for texture classification, as proposed by Arivazhagan et al. [28] in 2003. The first two feature databases achieved high mean success rates but the hybrid approach superseded with the highest mean success rate for texture classification.

Mavroforakis et al. [29] in 2006 created a standard feature set that was used as an annotation list for the description and classification of clinical properties of mammographic tumors in quantitative terms. In addition, the feature set was proposed to be possibly used as a benchmark for the input of any CAD system in automatic tumor evaluation. Classifiers such as Support Vector Machine (SVM), linear classifiers and neural networks investigated the efficiency of each feature to create rankings according to their discriminating power. Accuracy rates up to almost 95% were achieved.

Giger et al. [38] in 1994 presented a computerized classification scheme of masses that demonstrated similar performance to that of a radiologist's. Extraction of the margin of masses which determined the degree of spiculation and hence, the likelihood of malignancy, was achieved using a non-linear bilateral subtraction method that enhanced asymmetries. Two features were used for the input of the artificial neural network. Results proved the method to be effective in distinguishing between benign and malignant masses.

A much common approach up taken as seen so far, is the computer aided mass classification method into benign and malignant tissue characterized by textural features as proposed for instance, by Islam et al. [14] in 2010. The benign-malignant classification on the region of

interest (ROI) that contains mass using statistical textural features; mean standard deviation, entropy, skewness, kurtosis and uniformity was achieved as the inputs to the artificial neural network classifier. A 90.91% sensitivity and 83.87% specificity was accomplished.

The extraction of textural features of the segmented region of interest (ROI) was also demonstrated by Abdallah et al. [30] in 2011 using gray level co-occurrence matrices (GLCM) for tumor classification. Features were extracted from four spatial orientations; (0º, 45º, 90º and 135º) and two pixel distance for three different block size windows (8x8, 16x16 and 32x32). The method produced results at an accuracy of 91.67% sensitivity and 84.17% specificity from the output of the Artificial Neural Network based classifier (ANN). The outcome was reported comparable to using the state-of-the-art Computer-Aided Detection system.

Qutaishat Munib et al [39] in 2005 presented two techniques for building a computer-aided diagnosing system for classification of abnormality in digital mammograms they have investigated and analyzed wavelet transform for image enhancement and features extraction, and the ANFIS algorithm for classification process. Their results have shown that this method is very effective for the automatic detection and classification of abnormalities in digital mammogram.

The proposed system by Elif Derya Übeyli [40] in 2009 used ANFIS classifier for classification of breast masses. The ANFIS classifier was used to detect the breast cancer when nine features defining breast cancer were used as inputs. The presented ANFIS model combined the neural network adaptive capabilities and the fuzzy logic qualitative approach. The classification results and statistical measures were used for evaluating the ANFIS. The total classification accuracy of the ANFIS model was 99.08%.

The study of Mei-Ling Huang et al [41] in 2012 compared the particle swarm optimizer (PSO) based artificial neural network (ANN), the adaptive neuro-fuzzy inference system (ANFIS), and a case-based reasoning (CBR) classifier with a logistic regression model and decision tree model. It also applied three classification techniques to the Mammographic Mass Data Set, and measured its improvements in accuracy and classification errors. The experimental results showed that, the best CBR-based classification accuracy is 83.60%, and the classification accuracies of the PSO-based ANN classifier and ANFIS are 91.10% and 92.80%, respectively.

In the study of Ali Keles et al [42] in 2011, they have developed an expert system that they called as an Ex-DBC (Expert system for Diagnosis of Breast Cancer), because differentiating between benign and malignant mammographic findings, however, is quite difficult. Only 15–30% of biopsies performed on nonpalpable but mammographically suspicious lesions prove malignant. The golden standard for diagnosis of breast cancer is biopsy. But, biopsy can be a source of patient discomfort, bleeding and infection, and can burden the health care system with extra costs. Thus, to reduce unnecessary biopsy rate have acquired big importance. The fuzzy rules which they used in inference engine of Ex-DBC system were found by using neurofuzzy method. Ex-DBC can be used as a strong diagnostic tool with 97% specificity, 76% sensitivity, 96%positive and 81% negative predictive values for diagnosing of breast cancer. That the developed system's positive predictive is high is very important. By means of this system can be prevented unnecessary biopsy. Beside it can be benefited from this system for training of students in medicine.

A computer-aided diagnosis method was proposed in the work of Weidong Xu et al [43] in 2007. DWT was used to extract the high-frequency signal of the images firstly, and thresholding with hysteresis was applied to locate the suspicious MCs. Then, filling dilation was applied to segment those desired regions. During the detection, ANFIS was used to adjust the parameters, making the CAD algorithm more adaptiv, and precise. At last, the suspicious MCs were classified with MLP, and the experiments showed the advantages of the proposed method over the conventional ones.

The hybrid system that was created by Manisha Arora and Dinesh Tagra [44] in 2012 used Neuro-Fuzzy (ANFIS-MATLAB) which is a combination of Neural Network and Fuzzy Logic. As an extension of this research and curiosity to evaluate the hybrid approach they implemented a Fuzzy Inference System (FIS) in MATLAB using fuzzy toolbox. The hybrid system trained on equally distributed dataset outperforms all other approaches discussed in literature. Specifically the sensitivity obtained in our Neuro-Fuzzy system is 100% which outperforms sensitivity of 99.37% in the SVM (Support Vector Machine) model used by E. D. Ubeyli [45] in 2007.