Chapter VI

CONCLUSIONS AND SCOPE OF FUTURE OF WORK

There are many more issues that could have been investigated further and included in this thesis but the work had to be drawn to a conclusion at some point. This present work builds a framework which can be built on. It includes a theoretical foundation and development of neuro-fuzzy networks with genetic optimization and weight initialization techniques. This chapter summarizes the distinguishing features of this research work on CANFIS with genetic optimization and opens a window to the future research that could stem from this work.

6.1 CONCLUSIONS

The outcome of this research encompasses the following

In this research, a new application of CANFIS model (multiple inputs and multiple outputs) with genetic optimization of parameters in layer 2 and layer 4 for the diagnosis of various diseases was designed.

To reduce the complexity of CANFIS while testing with multiple input classification problem, dimensionality reduction was used to reduce the number of input attributes and results are statistically tested using ANOVA to prove that there is no significant difference between them (Table 4.6).

The performances of various ANN classification models were also investigated for the breast cancer diagnosis problem. The performance level of SVM was not as high as those of the SOM and PCA. This may be attributed to several factors including the training algorithms, estimation of the network parameters, and the scattered and mixed nature of the features. This work indicates that CANFIS gives better classification accuracy than all other neural classifiers analyzed and can be effectively used for breast cancer diagnosis to help oncologists.
For the detection of erythematous-squamous diseases which is a typical application of CANFIS with multiple input-output problem, the obtained total classification accuracy of 96.65% is better than all accuracy reported in literature (Latha Parthiban and Subramanian 2009). These results indicate that the proposed CANFIS model has great potential in detecting the erythematous-squamous diseases. So, CANFIS will be a second opinion guide to the doctors in evaluating their own diagnosis by re-examining it with the CANFIS network.

Dataset of heart disease from UCI machine learning repository has been studied and preprocessed and cleaned out to prepare it for classification process. CANFIS was proposed as a dependable and robust method developed to identify a non-linear relationship and mapping between the different attributes of heart disease dataset. It has been shown that of GA is a very useful technique for auto-tuning of the CANFIS parameters and selection of optimal feature set. The fact is that computers cannot replace humans and by comparing the computer-aided detection results with the pathologic findings, doctors can learn more about the best way to evaluate areas that computer-aided detection highlights.

For e-coli bacteria classification, CANFIS gives the best result for both training and testing data and gives MSE of 0.0021, but consumes more time compared to ANNs like PCA, SOM and SVM. For this reason, if time factor is considered, according to the simulation results, the most suitable model for classifying e-coli bacteria data is PCA. The training dataset accuracy of 94.11% and test data accuracy for e-coli of 87.23% is obtained.

Dataset of mammographic mass data from UCI machine learning repository has been studied and preprocessed and cleaned out to prepare it for classification process. The classification results with CANFIS are very encouraging and could help doctors to use computer-aided detection results effectively. Initializing the weight of ANNs using Stern series speeds up training, which is very useful for classifying huge dataset.

To improve the training speed of neural networks, weight initialization has been done with stern sequence (Figure 4.18), which shows better performance for heart disease database (Table 4.30). CANFIS with preprocessing shows better performance.
accuracy on SPECTF Heart and SPECT heart dataset (Table 4.31) than available literature (Roya Asadi 2009).

Many algorithms for preprocessing medical images have been studied and NSCT is found to be very efficient in reducing noise in medical images. Table 5.8 and Table 5.9 shows the comparative performance of wavelets, curvelets, contourlets and NSCT for different FNA images. An algorithm for enhancing edges of medical image is analyzed experimentally (Table 5.10) and hypothesis testing using ANOVA proves that it enhances only the DV and not the BV.

So an expert system for medical diagnosis that uses a knowledge base and inference engine to make decisions has been successfully developed. Input for the knowledge base is gathered through real time data from hospitals and the inference engine uses a series of If-then statements to derive intelligence from the knowledge base and the designed system shows better accuracy.

6.2 SCOPE FOR FUTURE WORK

In the course of pursuing this work, it was noted that certain aspects need deeper analysis. Hence some of the aspects are mentioned below so that further work can be done on these lines.

i. Further work on e-coli bacterium is likely to yield much more comprehensive and accurate models of the protein localization sites.

ii. For a huge database with many attributes which has to be trained using ANNs, dimensionality reduction can be done through GA of CANFIS, which has not yet been attempted.

iii. Telemedicine research platform have been explored by many medical personals, software engineers, and researchers in an attempt to reach the patients in remote areas, to create a paperless and film less technology for hospitals; and in another perspective to create a platform for medical specialists to communicate or discuss regarding a patient via teleconferencing,
which will improve management. The novel technique that has been developed can be implemented for telemedicine applications also.

iv. Some process to improve the speed of inference system using CANFIS network can be designed.

v. Initialization of tuning parameters using Stern series can be tested on other standard ANN architectures and their results can be analyzed.

vi. Through multi-agent systems, knowledge can be extracted from specific disease database and can be used to give a coordinated master report for multiple diseases. It will be very efficient when multiple standard databases from different hospitals are interconnected to give coordinated result.

vii. In the clinical setting, CAD methods might be used as a tool invoked by the radiologist on viewing a particular medical case or as a routine screening procedure performed on all examinations. The workstation could be configured to allow the radiologist to control the sensitivity and specificity of the computer output. A choice of fewer false-positive lesions would be achieved at the cost of a lower number of true-positive lesions and vice versa. This trade-off could be adjusted by the radiologist, depending on the nature of the case material and personal preference. For example, a radiologist might choose an output with high sensitivity for examining high-risk patients being screened for cancer, whereas a lower sensitivity and lower false-positive rate might be desired for patients at low risk for cancer.