ABSTRACT

The main purpose of this research work is to provide bioinformatics way of support to reduce the heart diseases globally. The four most prominent noncommunicable diseases (NCDs) are cardiovascular diseases, diabetes, cancer, and chronic obstructive pulmonary diseases. The most common cause of death in the United States is the heart disease with 28.5% based on CDC/NHS, National Vital Statistics System [Roger, 2011]. In Jayadeva Cardiological Hospital, Bangalore, India alone approximately 10000 heart patients are treated every month, whereas about every 25 seconds, an American will have a coronary event.

Echocardiography is one of the popular techniques in human heart diagnosis for the study of heart abnormalities. Generally, these images provide a wealth of clinically relevant and useful information, including the size and shape of the heart, its pumping capacity and the location and extent of any damage to its tissues. It is especially useful for assessing stenosis and regurgitation diseases of the heart. However, the current clinical practice requires manual intervention in both imaging and in interpretation. The ultrasound operator has to manually demarcate major anatomical structures like Left Ventricle (LV), Right Ventricle (RV), Left Atrium (LA), and Right Atrium (RA) and computes numerical quantities such as length, diameter, area, fractional shortening, stroke volume, ejection fraction, etc., from these images. Because of the fact that the image is analyzed manually it purely depends on the expertise of the operator in detecting the heart cavities accurately. Any error caused in the image quantification will lead to incorrect diagnosis.

The current thesis aims at developing an implementable model that can analyze the echo images of a particular patient automatically and offer clinically relevant data for making appropriate decisions. For this, a series of steps need to be accomplished starting from acquiring proper echo images of the patient and up to complex data mining and image processing tasks. Once the echo images and/or video are available, it is preprocessed and segmented to obtain the Region of Interest (ROI). For this, a number of efficient algorithms
need to be developed. One of the primary requirements of the current work is that all these algorithms should be executed within the database environment as most of the patient data is already present in the database. This is called as tightly-coupled system which is possible only through SQL and PL/SQL procedures. The best choice for the medical image segmentation is K-Means clustering, because $k$ can be set to 3 for echo images. However, the algorithm has to be modified to make it efficient and work within the database environment. Tracing the endocardial boundary through geometric active contour has proved successful in many medical image processing and analysis. The conventional methods take and to overcome these problems a novel model is proposed and evaluated in this thesis work.

As the ultrasound machine produced echo images are in different views and are in large numbers, analyzing each image qualitatively and/or quantitatively takes more time. Primarily echo image analysis is carried out using 2D and color Doppler images, as most of the heart related diseases can be investigated with these two modalities. At present, hear image study is performed based on quantitative analysis rather than qualitative analysis. The accuracy of this analysis is dependent on the skill set of the operator. Tracing the Region of Interest (ROI) such as LV, LA, etc, and developing time and space efficient algorithms is a challenging problem. This thesis work addresses some of these issues for 2D and color Doppler echo images. For analyzing color Doppler images, color histogram, texture, edge density, and statistical methods are adopted in this thesis work.

This work has lead to develop domain specific quantitative and qualitative features that form a feature vector. Two important medical areas where this could be useful are CBIR and classification. Both these applications are currently not available in the healthcare environment. Therefore, a complete CBIR framework for the echo images is built based on the universal generic feature model. This means that the user can dynamically select appropriate features so that the retrieved images are most similar with respect to the query image. Another interesting area is classifying the echo images as either normal or abnormal and within the abnormal there are further categorizations possible. In this research work, an improved naïve-Bayesian classification model, which is widely used by many researches in the medical domain, is used for this purpose. Since NBC works based on probability it is expected to give a better accuracy. Here, the training and testing of the modules are based on the database SQL and/or OLAP queries to maintain the tightly coupled design model. The implementation is carried out with
the help of the following software tools: C#.NET, AForge.NET, Oracle 10g, ODAC, ODP.NET, MATLAB, etc.

Thus, the outcome of this research is expected to reduce the gap between manual way of analysis and automated methodologies through sophisticated computer vision and data mining based algorithms. The proposed methods offer accurate, fast, content-based archiving, automatic classification, relevant clinical results, support at ICU, diagnosis at remote places, and many more.