Cardio Vascular classification using Machine Long Short-Term Memory

D.Abinaya[1], Y.Abinaya[2,] M.Arunmozhi[3,] S.Jagadeeswari[4]
[1, 2, 3] *IV CSE, Department of Computer Science and Engineering,M.A.M. College of Engineering and Technology, Trichy*
[4]*AP, Department of Computer Science and Engineering, M.A.M. College of Engineering and Technology, Trichy-621 105*

Abstract

Over the past decades a considerable amount of time and effort has been expended researching and developing to recognize heart valves diseases. The purpose of this project is to set out an investigation into Echocardiography for the development of a heart valves diagnosis system. Artificial neural networks are finding many uses in the medical diagnosis application. Classification is an important tool in medical diagnosis decision support. Feed-forward back propagation neural network is used as a classifier to detect type of valve diseases.In this study, the data where obtained from Echocardiography file to diagnosed diseases. The data is separated into inputs and targets the diagnosis of heart diseases; the percent correctly classified in the simulation sample by the feed-forward back propagation network is 95percent.Machine learning algorithm and deep learning opens new door opportunities for precise predication of heart attack. Paper provides lot information about state of art methods in Machine learning and deep learning. An analytical comparison has been provided to help new researches working in this field.

Keywords - *Artificial Neural Networks, Medical Diagnosis, Feed-Forward Back Propagation Network, Artificial Intelligence, and Decision Support Systems, ECO Echocardiography.*

I. INTRODUCTION

Cardiovascular diseases have become the top cause of death in industrialized countries. They are responsible for up to 48% of the total number of deaths, and are expected to reach about 25 million a year by 2030 [1]. It is therefore of critical importance to improve techniques of cardiac function assessment, thus facilitating diagnosis and treatment of these diseases. There are a variety of methods used to evaluate the health of the heart. Among the noninvasive techniques, medical imageing is used to assess its mechanical action bymeans of various modalities such as magnetic resonance imaging (MRI) and ultrasound imaging (UI). However, because of its relatively high temporal resolution, UI is more adapted to the rapid motion of the heart. In addition, it presents advantages such as low budget requirements and reduced discomfort for the patient. This makes UI, particularly echocardiography, the most widely used modality in cardiology. Furthermore, the acquired ultrasound (US) images provide information that is essential for cardiac function evaluation. US images can be exploited either through direct visualization or using post-processing methods that extract valuable qualitative and measurable features. In this context, 2D automatic cardiac motion estimation as well as the associated strain measurements has been proved to be efficient tools for the diagnosis of cardiovascular diseases.

II. LITERATURE SURVEY

A. Motion estimation in echocardiography

Motion estimation is a difficult task because of its ill-posed nature, *i.e.*, it generally does not have a unique solution. In order to tackle this issue, additional constraints are often added to regularize this problem. These constraints represent *a priori* knowledge about the behavior of motion and force the estimation to narrow down to the type of displacements expressed by this information. In multiple motion estimation schemes, the ill-posedness is overcome by introducing a prior in the form of spatial or temporal smoothness. Regularization can also be achieved by using a parametric motion model. Depending on the problem at hand, these models limit the motion to be, *e.g.*, rigid or affine. In cardiac motion estimation, a common regularization is the B-spline parameterization. In this approach, the transformation is limited by the presence of control points and smoothness is added due to the B-spline basis function.

B. Tissue doppler imaging

The TDI method depicts myocardial motion (measured astissue velocity) at specific locations in the heart. Tissuevelocity indicates the rate at which a particular point in themyocardium moves toward or away from the transducer. Integration of velocity over time yields displacement or the absolute distance moved by that point. Tissue Doppler– derived velocity can be obtained via pulsed Doppler (by placing a sample volume at a particular location), M-mode Doppler, or 2-dimensional color Doppler. Color Doppler acquires tissue velocity information from the entire sector, and thus, multiple sites can be interrogated simultaneously. Individual segments are analyzed *ex post facto*. Although all of these methods yield the same mechanical information, differences in the peak values exist. Frame rates are highest with the M mode, lower withpulsed Doppler, and lowest with color Doppler TDI.

C. Strain rate estimation

RF data were reconstructed from the IQ data set. Afteraxial and lateral interpolation of the RF data set (by afactor 2 and 8, respectively, resulting in spatial samplingof, approximately, 20 and 150 μm in the axial and lateraldirections at the maximal image depth), the motion of theRF patterns between two successive RF frameswas estimated using a newly developed algorithm basedon the methodology described by previous authors.However, to track the motion of a RF pattern, the sum ofabsolute differences (SAD) function rather than thecross-correlation function was used. Moreover, temporalstretching (range, 0.97–1.03 in steps of 0.01) of the first signal window, as presented, was indispensable.

D. Potential clinical application ofultrasonic _ and sr imaging

Ischemic Heart Disease

Despite several attempts to implement new cardiacultrasound methods to quantify ischemia, the routineclinical evaluation of regional function in ischemicheart disease has remained firmly on the basisof visual assessment of wall motion and wall thickening.However, the eye has been shown to havelimitations in assessing the timing of the complexchanges in regional myocardial deformation thatoccur in differing ischemic substrates. Thus, post systolic thickening (PST), which isan important parameter to measure when attemptingto quantify regional ischemia (and which usuallylasts only 50-60 milliseconds), may neither be displayedby ultrasound systems with low frame ratesnor, if displayed, be appreciated visually. Even inultrasound systems with sufficiently high framerates, it would be necessary to time aortic valveclosure to determine the relative amounts of regionalthickening/shortening that occur during orafter ejection.

III. EXISTING SYSTEM

A. Learning Strategies

The dictionary can be either fixed inadvance, *i.e.*, learned offline from a set of ground-truth motiondata, or learned in an adaptive way from the current estimation,*i.e.*, using an online scheme. More details about these twostrategies are provided below.

1. Offline dictionary learning

Given a set of ground-truthdata, the motion dictionary is learned and fixed beforethe estimation process. This strategy was used for image denoising, face recognition and texture segmentation. The offline learning supposes thatthe training set is adapted to the problem and spans thedifferent types of motions that might occur during theestimation. Since the learning process is done only once, the offline strategy has the advantage of being less timeconsuming.

2. Adaptive dictionary learning

This strategy, also knownas online learning, is commonly used in the area ofnatural image denoising. Online learningconsists in extracting training patches from the noisyimage itself. Note that the initial dictionary can eitherbe an offline learned dictionary, a predefined dictionary,or simply a set of random patches extracted froman initial motion estimate $\boldsymbol{u0}$, *i.e.*, the result of a firstrough motion estimation. Since the dictionary is jointlyestimated with the motion field, the adaptive learningstrategy is more time consuming. However, it remainsthe most appropriate approach in the absence of asuitable or sufficient training set.

IV. PROPOSED WORK

A. Deep Learning for Predication inEchocardiography

Deep learning can be defined as subfield of machine learning which is based on learning at multiple level of representation and abstraction ,each level contains multiple processing unit for multiple processing between the input and output layer. Deep learning works on the principle of feature hierarchy where higher level hierarchy is formed by composition lower level features. Deep learning bring renaissance to the neural network model major work is going in the field of in its implementation through stacked restricted Boltzmann machine and auto encoder-decoder technique. This method impress researches with their performance in field of image processing and layer wise pre training techniques other areas of its application include Natural language processing,acoustic processing., Recurrent Neural Network (RNN) is consider to be best suited for sequential feature and sequential data their exist various method working on these two version Long Short-Term Memory **(**LSTM) was proposed by Hochreiter and Schmidhuber, the performance is quite impressive in the field related to sequence based task . Other contemporary method to LSTM is gated recurrent unit (GRU), it is simpler than LSTM but the result is quite impressive. A temporal based heart disease prediction has been done where author used GRE to achieve the high accuracy. Researchers have begun to use deep learning technique for medical dataset. is used encoder-decoder type pattern form serum of uric acid .similar kind of works have been discussed in great detail by the author. In generalised approach of deep learning has been illustrated. In the flow chart there are five modules it has their own specific operation, the goal is to present the above flow chart in most general way. Data collection is phase in which dataset from standard repository is get collected followed stage of pre-processing which include functionality of noise reduction and feature selection. Next step is core for deep learning because it implements the basic algorithmic approach adapted for manipulation of data set, the algorithms may vey from deep belief network [15] torecurrent neural network. performance analysis of above data mining technique has been the major module because it illustrated about basic comparison of above adapted method , in the last discovery of knowledge module will get our desire goal which include percentage or probability of happening the instances. In our case it is the probability of heart attack in the patient.

B. Data pre-processing

One of the main steps that affect the performance and quality of prediction of machine learning models is data quality and data pre-processing. Data pre-processing includes handling missing values, smooth noisy data, identify or remove outliers, normalization, transformation, etc. Therefore, several steps have been applied to handle some issues on the dataset.

All frameworks that foresee heart maladies use clinical dataset having parameters what's more, inputs from complex tests directed in labs. None of the framework predicts heart ailments in view of danger variables, for example, age, family history, diabetes, hypertension, elevated cholesterol, tobacco smoking, liquor admission, heftiness or physical idleness, and so on. Heart ailment patients have part of these obvious danger elements in like manner which can be utilized adequately for determination. Framework in light of such hazard variables would help medicinal experts as well as it would give patients a notice about the plausible vicinity of coronary illness even before he visits a healing centre or goes for excessive restorative checkups.

• Outliers: a value of an attribute is considered as an outlier if it deviates from the expected value for this attribute. Outliers has been handled using inter-quartile range (IQR). The IQR is a good choice for handling the outliers since the dataset used in this study is nearly symmetric means that its median equals its midrange

•Missing values: The only attribute that has missing values is Peak Diastolic blood pressure and the number of individuals with missing Peak Diastolic blood pressure is 72. All the values for this attribute are replaced by the mean value of this attribute.

• Discretization: Aims to reduce the number of values for continuous attributes. This is done by splitting the range of the continuous attribute into intervals. Discretization reduces the time needed to build the prediction model and improve the prediction results. The following attributes have been discretized: Age, METS, Resting Systolic blood pressure, Resting Diastolic blood pressure, The Percentage of Heart Rate Achieved, Peak Heart Rate and Peak Diastolic blood pressure.

• Sampling: The dataset used in this study consists of 23,095 with 8,090 patients with experienced hypertension and the rest did not. The most common metric used to evaluate machine learning techniques is accuracy. This measure does not work properly when the data is imbalanced (the variance between patients who experienced hypertension and those who did not experience hypertension is considerably high). However, the nature of our prediction problem requires a high rate of correct detection of patients who are at high risk of developing hypertension. In general, there are two different methods to address the imbalanced dataset and obtain a balanced dataset (the number of patients who experienced hypertension is close to the number of patients who did not experienced hypertension). The first method is over-sampling the minority class (patients who experienced hypertension) and the second method is under-sampling the majority class. In this study, we used both under-sampling and over-sampling to handle the imbalanced data problem and compare the performance of both techniques. We used the Synthetic Minority Over-sampling (SMOTE) Technique. It is an over-sampling technique in which the minority class is over-sampled by creating “synthetic” examples rather than by over-sampling with replacement. SMOTE selects the minority class samples (patients experienced hypertension) and creates “synthetic” samples along the same line segment joining some or all k nearest neighbours belonging to the minority class.

V. ANALYSIS OF AVAILABLE LEARNING ALGORITHM

When it comes to comparing two or more machine learning algorithm, it is most difficult because two algorithms is differ in many ways. Reason for difficulty in comparison because algorithm are highly depended on dataset , it is not wise to decide properly which algorithm is perform for the particular dataset , there is only one way to know about the efficiency of algorithm for the particular dataset is implement them. Analytical comparison is require to properly decide the difference between different machine learning algorithm this type ofwork could be useful for researchers who want to work in this fieldComparison will highlight the key difference on different background this paper has tried to reflect majority of comparison between different algorithms so that beginner and new

VI. CONCLUSION

Heart attack is crucial health problem in human society. This paper has summarised state of art techniques and available methods for predication of this disease. Deep learning an emerging area of artificial intelligence showed some promising result in other field of medical diagnose with high accuracy. It is still an open domain waiting to get implemented in heart diseasepredication.Some methods of deep learning has been discussed which can be implemented for heart disease predication, along with pioneer machine learning algorithms. An analytical comparison has been done for finding out bestavailable algorithm for medical dataset. In future our aim is to carry forward the work of temporal medical dataset, where dataset varies with time and retraining of dataset is required.

REFERENCES

[1] W.H. Organization. (2012). World Health Statistics. [Online]. Available:

[2] T. P. Abraham, V. L. Dimaano, and H.-Y. Liang, "Role of tissue Dopplerand strain echocardiography in current clinical practice," Circulation,vol. 116, no. 22, pp. 2597–2609, 2007.

[3] C. Cottrell and J. N. Kirkpatrick, "Echocardiographic strain imagingand its use in the clinical setting," Expert Rev. Cardiovascular Therapy,vol. 8, no. 1, pp. 93–102, 2010.

[4] J. D'hooge et al., "Two-dimensional ultrasonic strain rate measurementof the human heart in vivo," IEEE Trans. Ultrason., Ferroelect., Freq.Control, vol. 49, no. 2, pp. 281–286, Feb. 2002.

[5] A.M. Shah and S. D. Solomon, "Myocardial deformation imaging:Current status and future directions," Circulation, vol. 125, no. 2,pp. e244–e248, 2012.

[6] G. R. Sutherland, G. Di Salvo, P. Claus, J. D'Hooge, and B. Bijnens,"Strain and strain rate imaging: A new clinical approach to quantifyingregional myocardial function," J. Amer. Soc. Echocardiography, vol. 17,no. 7, pp. 788–802, 2004.

[7] C. B. Composet al., "Radial basis functions for combining shape andspeckle tracking in 4D echocardiography," IEEE Trans. Med. Image.,vol. 33, no. 6, pp. 1275–1289, Jun. 2014.

[8] C. B. Compas et al., "A combined shape tracking and speckle trackingapproach for 4D deformation analysis in echocardiography," in Proc.9th IEEE Int. Symp. Biomed. Imag. (ISBI), Barcelona, Spain, May 2012,pp. 458–461.

[9] E. Y. Wong et al., "4-D echocardiography assessment of local myocardialstrain using 3-D speckle tracking combined with shape tracking,"in Proc. IEEE Int. Ultrason. Symp. (IUS), Jul. 2013, pp. 100–103.

[10] O. Mirea, J. Duchenne, and J. Voigt, "Recent advances in echocardiography:Strain and strain rate imaging," Research, vol. 5, no. 787, p. F1000,2016.

[11] M. Alessandrini et al., "A pipeline for the generation of realistic 3Dsynthetic echocardiographic sequences: Methodology and open-accessdatabase," IEEE Trans. Med. Imag., vol. 34, no. 8, pp. 1436–1451, Jul. 2015.

[12] Y. Nagata et al., "Intervendor variability of two-dimensional strainusing vendor-specific and vendor-independent software," J. Amer. Soc.Echocardiograph., vol. 28, no. 6, pp. 630–641, 2015.

[13] J.-U. Voigt et al., "Definitions for a common standard for 2Dspeckle tracking echocardiography: Consensus document of theEACVI/ASE/industry task force to standardize deformation imaging,"Eur. Heart J.-Cardiovascular Imag., vol. 16, no. 1, p. 1, Dec. 2014.

[14] J. D'hooge, et al., "Two-dimensional speckle tracking echocardiography:Standardization efforts based on synthetic ultrasound data," Eur. HeartJ.-Cardiovascular Image ., vol. 17, no. 6, pp. 693–701, Aug. 2015.

[15] K. Farsalinos, A. Daraban, S. Ünlü, J. Thomas, L. Badano, and J. Voigt,"Head-to-head comparison of global longitudinal strain measurementsamong nine different vendors: The EACVI/ASE inter-vendor comparisonstudy," J. Amer. Soc. Echocardiograph., vol. 28, no. 10, pp. 1171–1181, 2015.