Cardiac Diagnosis System for Heart Disease Patients Using Machine Learning Algorithms

MOE THIDA

Lecturer, CESDT Department, Computer University (Myitkyina), Kachin State, Myanmar

Abstract- ***Nowadays, machine learning techniques applied to medical image analysis have become an essential role according to many advances in digital imaging. Analyzing ECG medical images to diagnose types of arrhythmias which may cause sudden cardiac death is very demanding in healthcare sector. In this paper, J48 Decision Tree and Support Vector Machine (SVM) are utilized for cardiac diagnosis classification system. And, ECG medical images are applied to classify four different arrhythmia types and one normal condition type in this work. The proposed system is experimented using benchmark ECG images from MIT-BIH arrhythmia database. The diagnosis performance of the system is shown by using classification accuracy, sensitivity and specificity analysis. According to the experimental results, the proposed system achieves the satisfying classification accuracy over 80 percent (%).***

Indexed Terms- ***Arrhythmias, Decision Tree, ECG, Support Vector Machine, MIT-BIH.***

I. INTRODUCTION

Nowadays, the advancements in computer technologies and techniques health-care providers and specialists to research using the Internet or Telemedicine technology. According to the surveys of WHO, heart disease has become one of the primary cause of death and it kills twenty million people or patients per year. Besides, the specialists have analyzed that these patients suffer different kinds of cardiovascular failures, heart attacks and strokes severely. Therefore, there is an essential work to distinguish or classify the diagnosis of heart diseases efficiently. Arrhythmia, a form of irregularity in heart rhythms, has led to heart attack and it should be diagnosed as early as possible. There are many issues and problems in the process of identifying and classifying arrhythmias as it is necessary to analyze each heartbeat of the ECG records. Moreover, early diagnosis of arrhythmias has become essential due to different types of heart diseases which may cause the increase of mortality rate [2], [10]. An electrocardiogram (ECG) is usually applied to discover any abnormalities which are caused by damage to the heart. In addition, it sometimes can be applied as a blood test to detect abnormal levels of certain enzymes in the bloodstream. Nowadays, ECG medical images are applied in arrhythmia-based cardiac diagnosis. Arrhythmia classification using ECG medical images is the identification of normal and abnormal (arrhythmic) heartbeats which are based on heartbeat morphology of ECG signal. The P, Q, R, S, and T waves in electrocardiogram are electrical voltages generated by the heart and recorded by the electrocardiograph from the surface of the body of the humans [6], [7].

Modern computational methods, developed in the field of machine learning, offer new approaches to leveraging the growing volume of imaging data available for analyses. In addition, machine learning techniques can be used to reduce the mortality rate, improve the accuracy in disease diagnosis and mainly reduce the diagnosis time. Recently, many methods and approaches have been proposed to develop an efficient classification model for arrhythmia detection. According to the various surveys and studies, machine learning algorithms including neural networks, decision trees, random forest, gradient boosting, and SVM are demanded to apply in arrhythmia classification on ECG data. In this system, Decision Tree and Support Vector Machine (SVM) are applied to diagnose heart diseases by using ECG medical images [3], [5], [11]. The paper is organized as follows: the related works which motivating the proposed work is briefly presented in section 2. In section 3, the proposed cardiac diagnosis classification including pre-processing of ECG images and classification of cardiac diagnosis using machine learning algorithms are all discussed in detail. The

explanations about experimentation of the system is then described in section 4. The final conclusion and future works are presented in section 5.

II. RELATED WORKS

Cardiac arrhythmia is a kind of severe condition for human being and it is a primary cause of abnormal heart rhythm. Manoj Athreya A, Avani H S, Pooja, Madhu S, and K. Paramesha [11] made some effective surveys with different machine learning techniques and algorithms proposed by different authors in the detection arrhythmia using Electrocardiogram (ECG) recordings. Mohameda and Bennis Abdellatifa [1] developed a classification system using ECG images combined with image processing procedures and Artificial Neural Network (ANN). In this system, they utilized the haar-like descriptor in calculating haar features for ANN. Pocholo James M. Loresco1, Aaron Don Africa [9] showed patients' cardiovascular activity analysis that uses image processing methods as spatially oriented feature extraction to retrieve only the important features like Atrial (rate/min), Ventricular, QRS, QT, QTc and PR intervals in second. It used root-mean-square-error (RMSE) and normalized root-mean-square-error (NRMSE). Elif IZCI, Mehmet Akif OZDEMIR, Murside DEGIRMENCI, and Aydin AKAN [12] proposed the ECG arrhythmia classification method using a deep 2D convolutional neural network (CNN) to classify eight different types of the arrhythmias.

III. CARDIAC DIAGNOSIS SYSTEM

In developing cardiac diagnosis system, there are many issues and challenges to provide efficient real-life support of patients. The proposed system in this paper is intended for heart disease patients who suffer arrhythmias. In current time, arrhythmia classification using ECG medical images applied with machine learning algorithms is very demanding in health care research areas. In this system, Decision Tree and Support Vector Machine (SVM) are utilized to make some comparative studies between these two algorithms in arrhythmia classification.

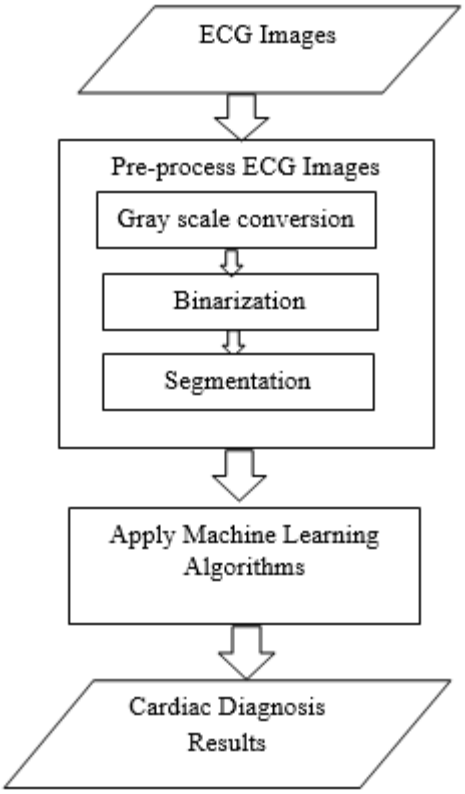

Fig. 1 The cardiac diagnosis system using ECG medical images

A. Pre-processing of ECG Images

In the preprocessing of images, ECG segmentation medical images are obtained from scanning the ECG papers of heart disease patients. The experimental data or ECG images are collected by scanning the ECG paper of patients through MIT-BIH arrhythmia database [13]. The first step of preprocessing of images is gray scale conversion. The original ECG images are transformed into gray ECG images. Binarization of images is the second preprocessing step. It removes the unnecessary information (grid lines). In the final step of preprocessing is segmentation. It segments the images including P wave, QRS wave, and T wave from scanned portions in a 60 sec interval in order to analyze cardiac diagnosis through these ECG images. Typically, the ECG images includes P, QRS, and T waves. The 'P' wave is arterial contraction rate, 'Q' wave is shifted downward just before ventricular contraction and 'R' wave is the peak of ventricular contraction. Moreover, the 'S' wave is the downward deviation immediately

after contraction of the ventricle, T is the repair of the ventricle.

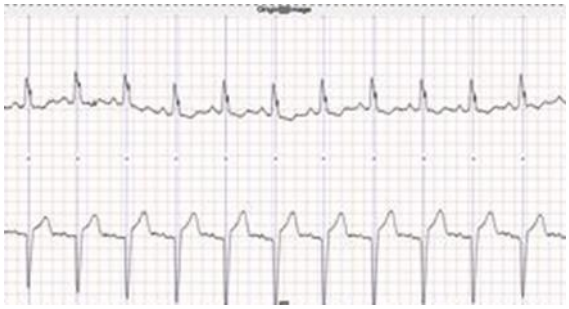

Fig. 2 The original ECG image

Fig. 3 The gray scale converted ECG image

Fig. 4 The binarization of ECG image

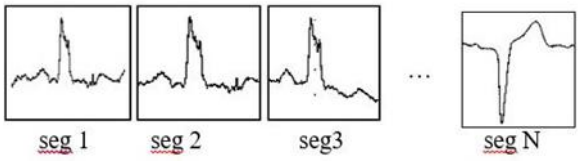

Fig. 5 The segmentation of ECG image

B. Cardiac Diagnosis Classification

After accomplishing the pre-processing procedures of ECG images, these images are passed to the next phase of classification for heart disease diagnosis. Decision Tree and Support Vector Machine (SVM) algorithms are applied to differentiate four different arrhythmia types and one normal condition including normal ECG image, four common heart diagnoses such as Atrial Fibrillation (AF), Ischemic Heart Failure (IHF), Congestive Heart Failure (CHF), and Myocardial Heart Failure (MHF) [10]. The decision tree algorithm performs two steps in the building a tree and then applying this tree to the corresponding data or dataset. There are many kinds of decision tree algorithms such as CART, ID3, C4.5, CHAID, and J48. Among them, J48 algorithm is utilized in this system. The J48 algorithm recursively classifies data until it has been categorized as perfectly as possible. To perform comparative studies between two sorts of classification algorithms, the support vector machine (SVM) is also utilized in this system. SVM classifies data by finding the best hyper plane which separates all data points of one class from those of the other class. SVM is typically based on mathematical functions intended to model in real world problems [11].

IV. EXPERIMENTATION

The main purpose of the proposed system is to develop a cardiac diagnosis system using medical images for heart disease patients who suffer arrhythmias. All experiments are performed with Intel i7 8300 CPU.

A. Applied Dataset

The experimental data are ECG medical images by scanning the ECG graph paper of patients and obtain from MIT-BIH arrhythmia database [13]. The various beat types that which are organized from 48 records of 47 volunteers in this database. Individual record includes 30 minutes duration, 360 Hz sampling rate and band pass filtered at 0.1– 100 Hz. Typically, all heartbeats of a record were annotated according to their arrhythmia types by independent domain experts.

B. Experimental Results

The diagnosis of the heart disease especially for arrhythmia patients are performed on MIT-BIH arrhythmia database. During the experimentation, it is observed that the correct and complete data collection and pre-processing procedures may provide for the choice of best classifier. In this system, 5-fold and 10-fold cross validation techniques are applied on MIT-BIH arrhythmia dataset. The comparative results of J48 Decision Tree and Support Vector Machine (SVM) algorithms using 5-fold cross validation are shown in Table 1. And, the comparative results

between them are also described in Table 2. According to the experimentation, it can be seen that SVM offers the promising results using not only 5-fold but also 10-fold cross validation than J48. It can be analyzed that the SVM can offer better classification accuracy, sensitivity, and specificity and it can be recommended as a good classifier for cardiac diagnosis of heart disease patients.

Table 1. The performance of classifiers using 5-fold cross validation

Classifier	Accuracy	Sensitivity	Specificity
SVM	85.36%	84.45%	88.27%
J48	71.23%	73.34%	70.08%

Table 2. The performance of classifiers using 10-fold cross validation

Classifier	Accuracy	Sensitivity	Specificity
SVM	85.01%	84.75%	85.62%
J48	76.58%	73.68%	74.07%

V. DISCUSSIONS AND FUTURE WORKS

In today's world, a large size of population are suffering heart diseases especially arrhythmias and early detection and accurate medical assistance can save many lives of humans. In addition, analyzing the ECG medical images using machine learning approaches to diagnose the types of arrhythmias which may cause sudden cardiac death is very demanding in healthcare sector. In this paper, machine learning based approaches such as Decision Tree and Support Vector Machine (SVM) for the cardiac arrhythmia classification are efficiently applied to classify our different arrhythmia types and one normal condition type using ECG medical images from the MIT-BIH arrhythmia database. According to the experimentation, the SVM can offer better classification accuracy, sensitivity, and specificity and it can be recommended as a good classifier for cardiac diagnosis of heart disease patients. In future works, deep learning based models will be considered for cardiac diagnosis classification for heart diseases.

REFERENCES

[1] A. Bennis, M. Atibi, and Boussaa, -ECG image classification in real time based on the haar-like features and artificial neural networks,| in Proc. Procedia Computer Science, vol. 73, pp.32-39, 2015.

[2] A. Akan, O. A. Mehmet, and D. Murside, -Cardiac Arrhythmia Detection from 2D ECG Images by Using Deep Learning Technique,| in Proc. 2019 Medical Technologies Congress (TIPTEKNO), IEEE, 2019, pp. 1-4. 2019.

[3] A. Mohammad, R. K. Hamid, -Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals,| in Proc. Biomedical Signal Processing and Control, 2014, pp. 291-296.

[4] D. A., O. O. Oludayo, and Odunaike, -Automating skin disease diagnosis using image classification,| in Proc. Proceedings of the world congress on engineering and computer science, vol. 2, pp. 850-854, 2013.

[5] Martin-Isla, Carlos, and et al., -Image-Based Cardiac Diagnosis with Machine Learning: A Review,| in Proc. Frontiers in Cardiovascular Medicine, 2020.

[6] M. U. Shaomin, X. U. Yongyu, and D. Mengping, -Image retrieval method based on CNN and dimension reduction,| in Proc. 2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), IEEE, 2018, pp. 441-445.

[7] N. Saeeda, I. Muhammad, and Z. Ahmad, -Deep learning for medical image processing: Overview, challenges and the future,| in Proc. Classification in BioApps, Springer, 2018, pp. 323-350.

[8] N. S., Tomov, and et al., -On deep neural networks for detecting heart disease, arXiv: 1808.07168, 2018.

[9] P. J. M., Loresco, and Africa (2018). ECG print-out features extraction using spatial-oriented image processing techniques. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), vol. 10, pp. 15-20, 2018.

[10] S. R. William, M. David, and et al., -ECG-based heartbeat classification for arrhythmia detection: A survey,| in Proc. Computer methods and programs in biomedicine, vol. 127, pp.144-164, 2016.

[11] K. Paramesha, S. H. Avani, and A. Manoj (2019). Detection of Cardiac Arrhythmia using Machine Learning Algorithms. International Journal of

Recent Technology and Engineering (IJRTE), vol. 8, pp.11704-11707, 2019.

[12] J. Tae, Jun, and et al., -ECG arrhythmia classification using a 2-D convolutional neural network, arXiv: 1804.06812, 2018.

[13] https://www. physionet.org

[14] https://www.dicardiology.com

[15] https://www.heart.org