JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING AND END-TO-END DEEP LEARNINGFOR THE DETECTION OF CHRONIC HEART FAILURE FROM HEART SOUNDS

CHELLURI SASI
Department of Master of Computer Science
Miracle Educational Society Group of Institutions
Vizianagram– 535216 (AP) India

SARAGADAM SRIDHAR
Department of Master of Computer Science
Miracle Educational Society Group of Institutions
Vizianagram– 535216 (AP) India

ABSTRACT

Chronic heart failure (CHF) affects over 26 million of people worldwide, and its incidence is increasing by 2% annually. Despite the significant burden that CHF poses and despite the ubiquity of sensors in our lives, methods for automatically detecting CHF are surprisingly scarce, even in the research community. We present a method for CHF detection based on heart sounds. The method combines classic Machine-Learning (ML) and end-to-end Deep Learning (DL). The classic ML learns from expert features, and the DL learns from a spectro-temporal representation of the signal. The method was evaluated on recordings from 947 subjects from six publicly available datasets and one CHF dataset that was collected for this study. Using the same evaluation method as a recent PhysoNet challenge, the proposed method achieved a score of 89.3, which is 9.1 higher than the challenge's baseline method. The method's aggregated accuracy is 92.9% (error of 7.1%); while the experimental results are not directly comparable, this error rate is relatively close to the percentage of recordings labeled as ''unknown'' by experts (9.7%). Finally, we identified 15 expert features that are useful for building ML models to differentiate between CHF phases (i.e., in the decompensated phase during hospitalization and in the recompensated phase) with an accuracy of 93.2%. The proposed method shows promising results both for the distinction of recordings between healthy subjects and patients and for the detection of different CHF phases. This may lead to the easier identification of new CHF patients and the development of home-based CHF monitors for avoiding hospitalizations.

Keywords:ML(MachineLearning),DL(DeepLearning), CHF(Chronic Heart Failure)

INTRODUCTION

Chronic heart failure (CHF) is a chronic, progressive condition underscored by the heart's inability to supply enough perfusion to target tissues and organs at the physiological filling pressures to meet their metabolic demands [1]. CHF has reached epidemic proportions in the population, as its incidence is increasing by 2% annually. In the developed world, CHF affects 1-2% of the total population and 10% of people older than 65 years. Currently, the diagnosis and treatment of CHF uses approximately 2% of the annual healthcare budget. In absolute terms, the USA spent approximately 35 billionUSD to treat CHF in 2018 alone, and the costs are expected to double in the next 10 years [2]. Despite the progress in medical- and device-based treatment approaches in the last decades, the overall prognosis of CHF is still dismal, as 5-year survival rate of this population is only approximately 50%. In the typical clinical course of CHF, we observe alternating episodes of compensated phases, when the patient feels well and does not display symptoms and signs of fluid overload, and decompensated phases, when symptoms and signs of systemic fluid overload (such as breathlessness, orthopnea, peripheral edema, liver congestion, pulmonary edema) can easily be observed. During the latter episodes, patients often require hospital admission to receive treatment with intravenous medications (diuretics, inotropes) to achieve asuccessful negative fluid balance and return to the compensation state. Early detection of HF worsening would allow a treating physician to adjust the patient's medical management on an outpatient basis in a timely manner and thus avoid the need for a hospital admission. Currently, an experienced physician can detect the worsening of HF by examining the patient and by characteristic changes in the patient's heart failure biomarkers, which are determined from the patient's blood. Unfortunately, clinical worsening of a CHF patient likely means that we are already dealing with a

fully developed CHF episode that will most likely require a hospital admission. Additionally, in some patients, characteristic changes in heart sounds can accompany heart failure worsening and can be heard using phonocardiography. An example of a phonocardiogram (PCG) recording of a healthy subject is presented in Fig. 1. In healthy subjects, 2 heart sounds are typically heard (called S1 and S2). S1 is caused by the closure of the mitral valve and ventricular wall in the early systole, S2 is caused by the closure of the aortic and pulmonary valves at the beginning of the diastole. Here, the interval between S1 and S2 is called systole, i.e., the contraction phase of the cardiac cycle, and the interval between S2 and S1 is called diastole, i.e., the relaxation phase of the cardiac cycle. Additional heart sounds (such as S3 and S4) can be heard in certain cardiac conditions and are never regarded as normal. In the case of CHF (in the course of decompensation), we can often hear a third sound (S3) that typically appears 0.1-0.2 s after the second sound, i.e., S2. Recently, it has been demonstrated that some physiological parameters, such as the occurrence of additional heart sounds or increased blood pressure in the pulmonary circulation, already start to appear several weeks before the CHF patient develops a clinically evident decompensation episode. This is also an important therapeutic window where outpatientbased treatment interventions can reverse CHF deterioration and return the patient to the compensated state without the need for a hospital admission. In recent years, many studies have proposed MachineLearning (ML) approaches for the automatic detection of different heart conditions using PCG signals recorded with a digital stethoscope [1]. Nevertheless, methods that explicitly focus on CHF detection are quite scarce. The typical ML pipeline for the detection of different heart conditions is as follows: segmentation of the signals by detecting the ''typical'' heart sounds (i.e., S1 and S2), denoising of the signals, extracting individual frequency-domain and timedomain features, and learning a feature-based ML model (e.g., using ML algorithms, such as Random Forest or Support Vector Machine - SVM) that is capable of classifying healthy vs. unhealthy sounds. Most of the features currently used are based on medical and audio/signal analysis knowledge. However, a PCG recording that sounds unhealthy to one expert may sound healthy to another one; therefore, doctors never diagnose a CHF patient using only heart sounds, but rather use a holistic view of the patient instead (i.e., extensive medical history, blood pressure, laboratory tests, etc.). This uncertainty is one reason why 9.7% of the recordings in the recent PhysioNet cardiology challenge [3] were actually labeled as ''unknown'' by experts, while the rest of the recordings were labeled as healthy or unhealthy. The recent advancements in Deep Learning (DL) suggest that end-to-end learning (i.e., ML models that learn directly from the raw data and no features are needed) can outperform the classic, feature-based ML. For example, DL has achieved breakthrough performance in tasks such as pattern recognition problems [4], image processing [5], [6], natural language processing [7], [8], speech and audio processing [9], [10], and sensor data processing [11], [12]. For CHF detection, a successful combination of classic ML and end-to-end DL can outperform each single approach [13]. The classic ML approach learns from a large body of expert-defined features, and the DL approach learns both from a time-domain (the raw PCG signal) representation of the signal and a temporaldomain representation (the spectrogram) of the signal. This approach was successful in our previous study of human activity recognition from smartphone sensor data [14]. In addition to distinguishing the CHF patients and healthy individuals, we focus on detecting the CHF state (compensated vs. decompensated) based on the analysis of heart sound recordings. Our work builds upon the initial studies, where we demonstrated that it is possible to distinguish between healthy individuals and patients in a decompensated CHF episode using a stack of machine-learning classifiers and expert features, showing promising results on a limited dataset [15]. We expand upon this approach using a considerably larger patient dataset, including six additional PhysioNet datasets, and an improved ML method that uses end-to-end DL. Furthermore, we investigate the differences in the heart sounds during the transition between the decompensated and recompensated states of CHF, with the aim of developing personalized monitoring models. Early detection of the worsening of CHF has the potential to reduce hospitalizations due to the worsening of the condition, which both improves the quality of life of patients and decreases the financial and logistic burden on the patient and the health system.

PROPSED SYSTEM:

In medical imaging the interest in deep learning is mostly triggered by convolutional neural networks (CNNs) [56],14apowerful way to learn useful representations of images and other structured data. Before it became possible to use CNNs efficiently, these features typically had to be engineered by hand, or created by less powerful machine learning models. Once it became possible to use features learned directly from the data, many of the handcrafted image features were typically left by the wayside as they turned out to be almost worthless compared to feature detectors found by CNNs.

ADVANTAGES:

- Better prediction
- Better accuracy

LITERATURE SURVEY

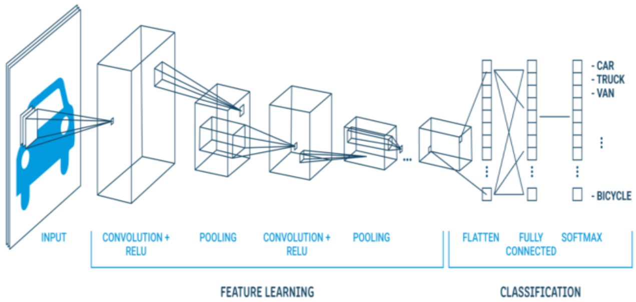
Several different machine learning based diagnostic approaches have been proposed in the literature for used on the Cleveland dataset. Using multi-layer Perceptron and support vector machine (SVM) algorithms, Gudadhe et al. built a diagnosis system for HD categorization and attained an accuracy of 80.41 percent. The HD categorization method was created by Humar et al. using a neural network combined with Fuzzy logic. The overall accuracy of the classification system was 87.4 percent. Using the statistical measurement system enterprise miner (5.2), Resul et al. created an ANN ensemble-based HD diagnosis system with an accuracy of 89.01%, sensitivity of 80.09%, and specificity of 95.91%. This HD diagnostic system was developed by Akil et al. using machine learning. The ANN-DBP algorithm was combined with the FS method, and the results were promising. An expert medical diagnosis method for HD detection was proposed by Palaniappan et al. [7]. The machine learning predictive models NB, DT, and ANN were employed in the system's creation. Accuracy for NB was 86.12%, accuracy for ANN was 88.12%, and accuracy for DT was 80.4%. By adapting the artificial neural network method, Olaniyi et al. created a three-stage method for HD prediction in angina, which showed an accuracy of 88.89%. For HD diagnosis, Samuel et al. created a unified medical decision support system using artificial neural networks and Fuzzy AHP. The proposed approach achieved an accuracy of 91.10 percent. Using relief and rough set methods, Liu et al. suggested an HD classification system. Overall, the proposed technique was 92.32 percent accurate in its classifications. In this paper, we offer a method for identifying HDs based on the use of feature selection and classification tools. When it comes to choosing which features to use, the Sequential Backward Selection (SBS FS) algorithm is the way to go. K-Nearest Neighbor (K-NN) performance as a classifier has been evaluated using both the entire feature set and a subset of features. Accuracy was successfully achieved with the proposed strategy. Another study by MOHAN et al. uses mixed machine learning approaches to create an HD prediction tool.

CNN is a powerful image processing algorithm. Right now, these are the best algorithms available for automatically processing photos. Businesses widely use these algorithms to do tasks like object identification in images.

HD. To illustrate why the proposed work is so crucial, this article presents variou

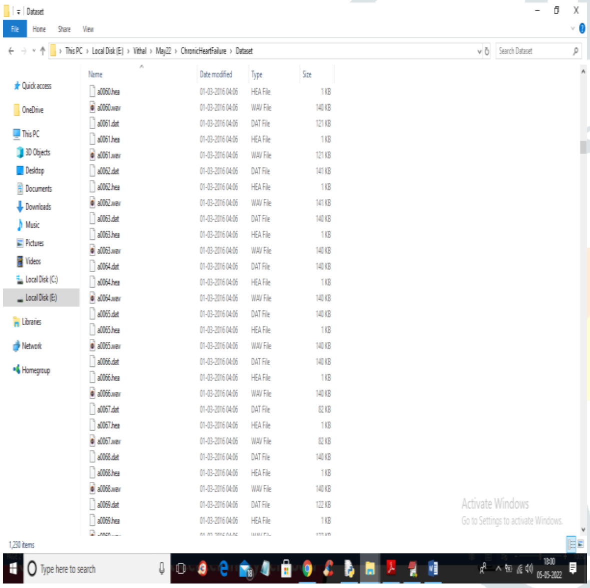
RGB combination data is present in images. An image from a file can be loaded into memory using Matplotlib. The computer only sees a series of numbers; it cannot perceive an image. 3-dimensional arrays are used to store colored images. The first two dimensions match the image's height and width (the number of pixels). The final dimension represents each pixel's red, green, and blue hues.

Convolutional Neural Networks with three layers, specifically designed for use in image and video recognition applications. CNN is primarily utilized for image analysis applications such as instance segmentation, object detection, and picture recognition.



Results

Below is the dataset screen used in this project

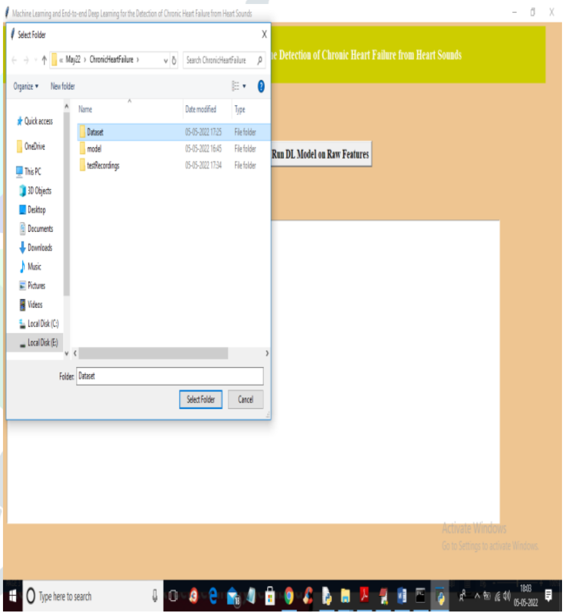


In above screen we have 3 files where .hea file contains class label as Normal or Abnormal and .dat file contains PCG signals and .wav file contains heart sound recording and by using all files we will train all algorithms

SCREEN SHOTS

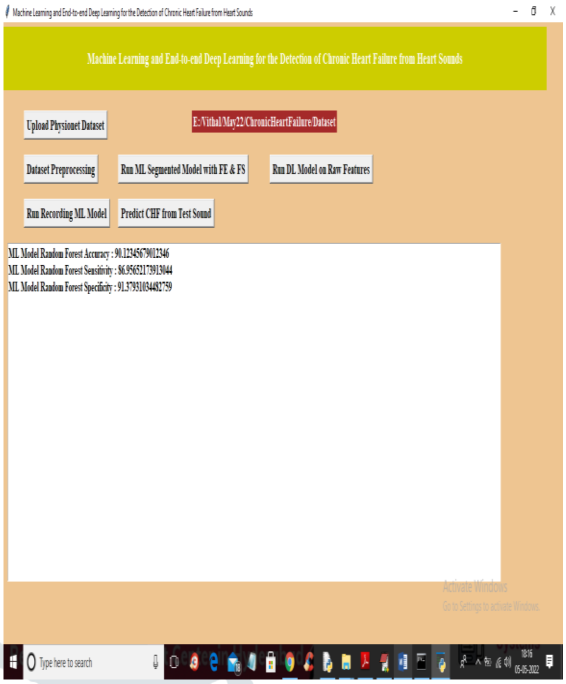
To run project double click on 'run.bat' file to get below screen

In above screen click on 'Upload Physionet Dataset' button to upload dataset



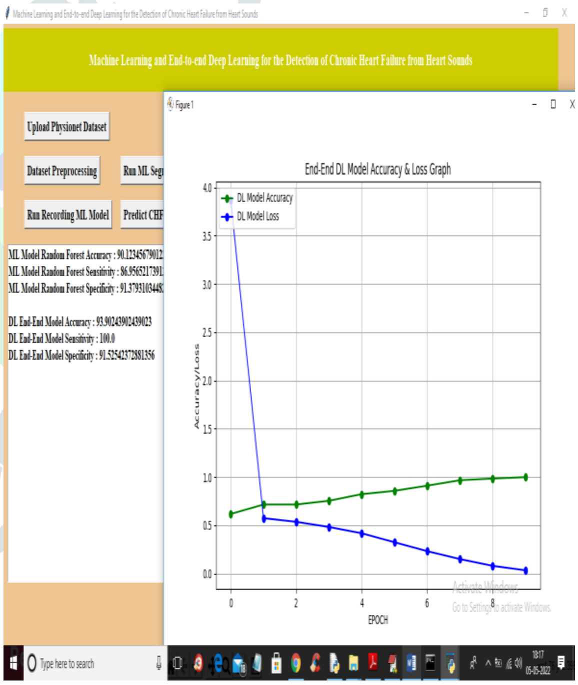
In above screen selecting and uploading 'Dataset' folder and then click on 'Select Folder' button to load dataset and to get below output

In above screen dataset loaded and now click on 'Dataset Preprocessing' button to read all dataset file and then extract features from it

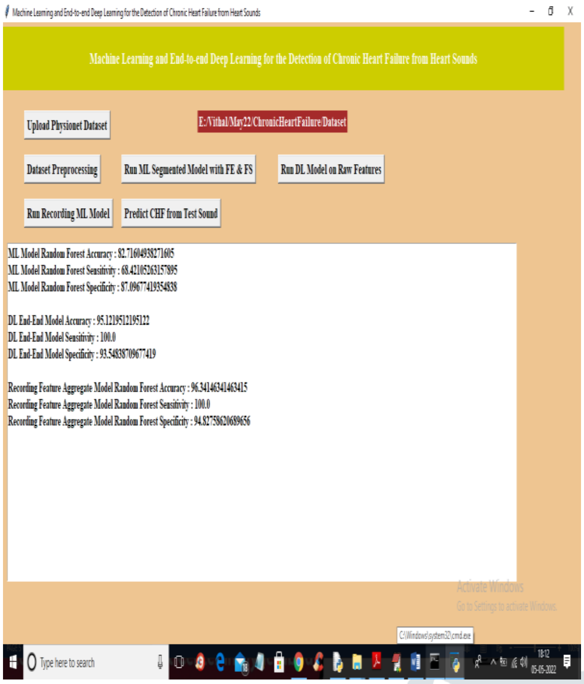


In above screen with Classic ML we got 90% accuracy and now click on 'Run DL Model on Raw Features' to get below output

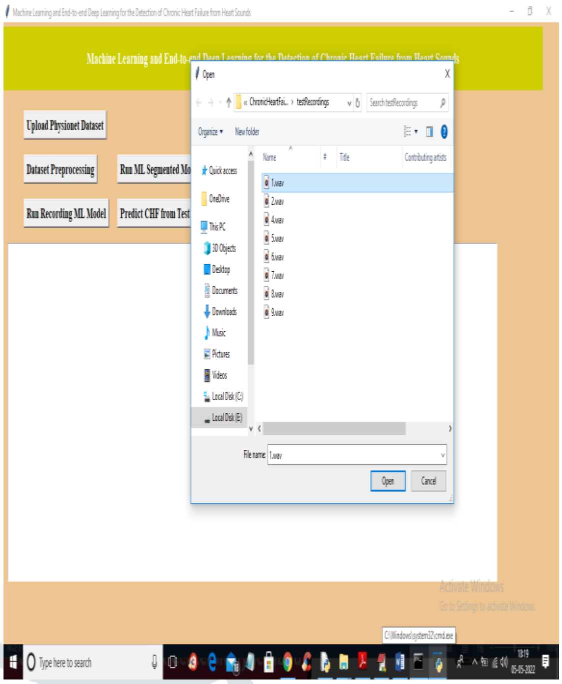
In above screen we can see dataset contains 405 heart sound files from 405 different person and 117 are the Normal sound and 288 are abnormal and in graph x-axis represents normal or abnormal and y-axis represents number of persons for normal or abnormal. Now close above graph and then click on 'Run ML Segmented Model with FE & FS' button to train Classic ML segmented model on above dataset and get below output



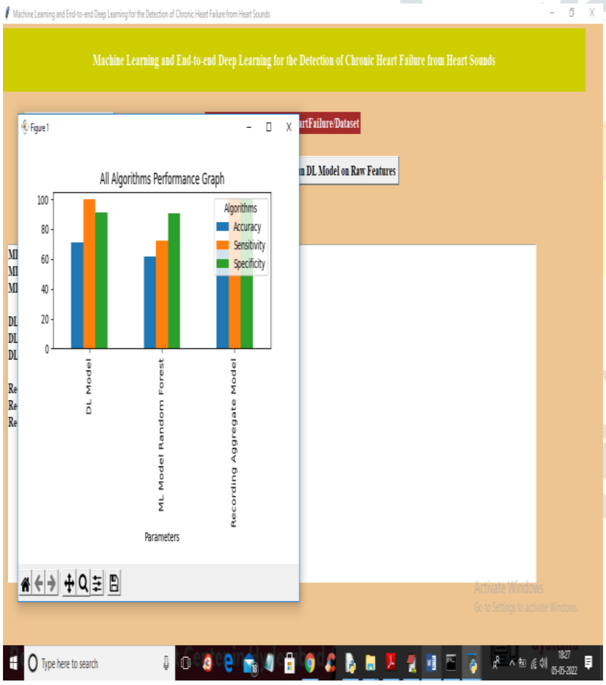
In above screen with DL model we got 93% accuracy and in graph x-axis represents epoch or iterations and y-axis represents accuracy or loss values and green line represents accuracy and blue line represents LOSS and we can see with each increasing epoch accuracy got increase and loss got decrease and now close above graph and then click on 'Run Recording Model' button to get below output



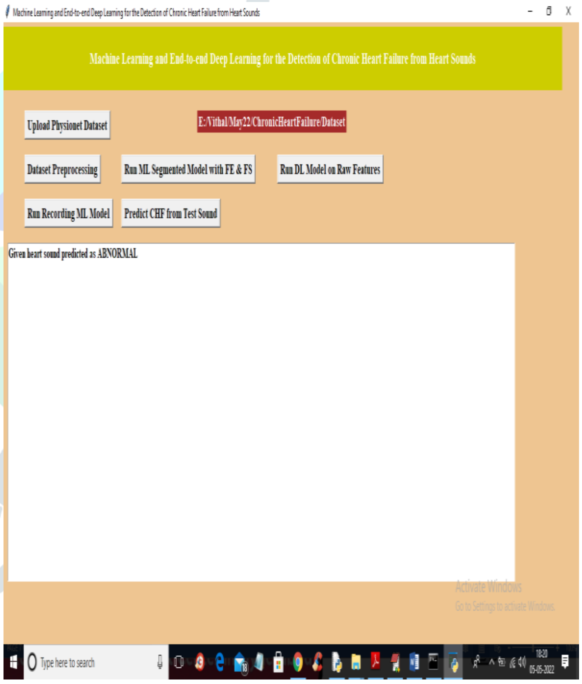
In above screen with recording model we got 96% accuracy and we can see all algorithms performance graph in below screen



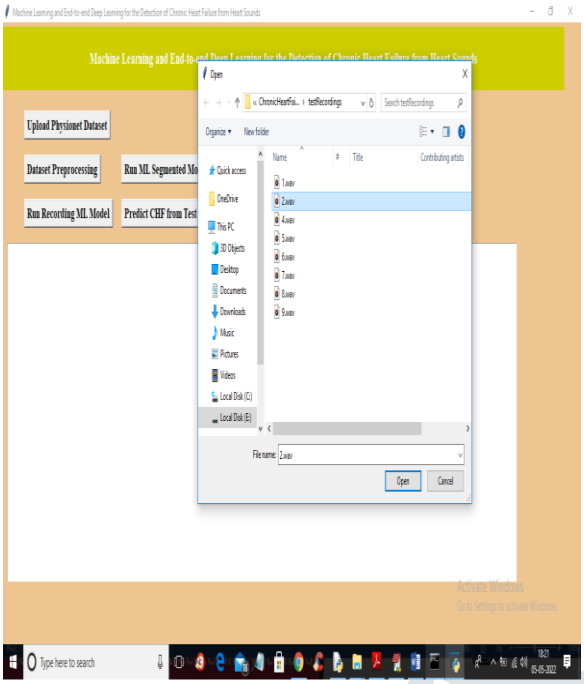
In above screen selecting and uploading '1.wav' file and then click on 'Open' button to get below output



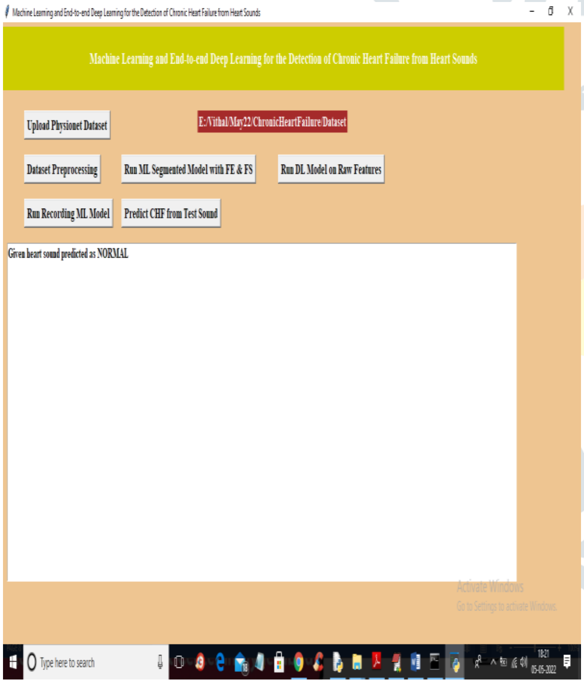
In above graph x-axis represents algorithm names and y-axis represents accuracy, sensitivity and specificity and in all algorithms Recording model has got high accuracy. Now close above graph and then click on 'Predict CHF from Test Sound' button to upload test sound file and get predicted output as Normal or Abnormal



In above screen uploaded heart sound file predicted as ABNORMAL and similarly you can upload other files and test



For 2.wav' file below is the output



CONCLUSION

In this study, we introduced a novel approach to identifying CHF in PCG recordings. The technique integrates traditional ML with full-stack DL. The time-domain (i.e., the raw PCG signal) and spectral representations of the signal are used by DL for learning, while the classical ML relies on a vast set of features defined by experts. Both our personal dataset for CHF detection and six public PhysioNet datasets used in the recent PhysioNet Cardiology Challenge were utilised to assess the method's efficacy. We were able to thoroughly assess the method's efficacy on analogous domains thanks to the challenge datasets. All dataset evaluation findings demonstrated that our method outperforms the challenge baseline methods (see the PhysioNet experiments section). Given that the PCG audio is recorded from a different body position in most of the datasets (e.g., aortic area, pulmonic area, tricuspid area, and mitral area), and that the datasets are labelled for different types of heart-related conditions, the proposed method is quite robust and useful for detecting differenttypes of heart-sound classification problems and not just for CHF detection, provided that domain-specific labelled data is available. Finally, we went above and beyond the traditional healthy vs. sick dichotomy and investigated individualised models for recognising the two stages (recompensated and decompensated) of CHF (i.e., when the patient needs medical attention). We found 15 characteristics whose distributions vary with the time of day. With just these two features, we were able to construct a straightforward decision tree classifier (see Fig. 3) that, according to a LOSO test, achieves a 93.2% accuracy in differentiating between the recompensated and the decompensated phases. Although we recognise the possibility of overfitting in these last tests, given that the dataset contains only 44 samples, we find the results highly encouraging and feel they provide a good foundation upon which to build additional tailored models. As far as we are aware, this is the first study to tackle this issue head-on.

REFERENCES

[1] M. Gjoreski, M. Simjanoska, A. Gradisek, A. Peterlin, M. Gams, and G. Poglajen, ''Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers,'' in Proc. Int. Conf. Intell. Environ. (IE), Aug. 2017, pp. 14–19.

[2] J. Voigt, M. S. John, A. Taylor, M. Krucoff, M. R. Reynolds, and C. M. Gibson, ''A reevaluation of the costs of heart failure and its implications for allocation of health resources in the United States,'' Clin. Cardiol., vol. 37, no. 5, pp. 312–321, May 2014.

[3] G. D. Clifford, C. Liu, B. Moody, D. Springer, I. Silva, Q. Li, and R. G. Mark, ''Classification of normal/abnormal heart sound recordings: The PhysioNet/computing in cardiology challenge 2016,'' in Proc. Comput. Cardiol. Conf. (CinC), May 2017, pp. 609–612.

[4] X. Jiang, Y. Pang, X. Li, and J. Pan, ''Speed up deep neural network based pedestrian detection by sharing features across multi-scale models,'' Neurocomputing, vol. 185, pp. 163–170, Apr. 2016.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ Commun. ACM, vol. 60, no. 6, pp. 84–90, May 2017.

[6] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, ‘‘Inception-v4, inception-resnet and the impact of residual connections on learning,’’ in Proc. 31st AAAI Conf. Artif. Intell., 2017.

[7] T. Young, D. Hazarika, S. Poria, and E. Cambria, ‘‘Recent trends in deep learning based natural language processing,’’ IEEE Comput. Intell. Mag., vol. 13, no. 3, pp. 55–75, Aug. 2018.

[8] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, ‘‘A neural probabilistic language model,’’ J. Mach. Learn. Res., vol. 3, pp. 1137–1155, Feb. 2003.

[9] S. Amiriparian, A. Baird, S. Julka, A. Alcorn, S. Ottl, S. Petrovic, E. Ainger, N. Cummins, and B. W. Schuller, ‘‘Recognition of echolalic autistic child vocalisations utilising convolutional recurrent neural networks,’’ in Proc. INTERSPEECH, Aug. 2018, pp. 2334–2338.

[10] S. Amiriparian, M. Gerczuk, S. Ottl, N. Cummins, S. Pugachevskiy, and B. Schuller, ‘‘Bag-of-deep-features: Noise-robust deep feature representations for audio analysis,’’ in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2018, pp. 1–7.

[11] M. Gjoreski, H. Gjoreski, M. Luštrek, and M. Gams, ‘‘Deep affect recognition from R-R intervals,’’ in Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput. ACM Int. Symp. Wearable Comput. (UbiComp), 2017, pp. 754–762.

[12] H. P. Martinez, Y. Bengio, and G. N. Yannakakis, ‘‘Learning deep physiological models of affect,’’ IEEE Comput. Intell. Mag., vol. 8, no. 2, pp. 20–33, May 2013.