IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

RHEUMATIC HEART DISEASE DETECTION USING DEEP LEARNING FROM SPECTRO TEMPORAL REPRESENTATION OF UNSEGMENTED HEART SOUNDS

[1]KANCHERLA YASASWANI, [2]ANKAVARAPU SUNEEL KUMAR, [3]DARAPUREDDY HARISH,

[4]Y DINESH KUMAR, [5]KUPPILI YASODA KRISHNA

[1]B.TECH(STUDENT), [2]B.TECH(STUDENT), [3]B.TECH(STUDENT), [4]ASSOCIATE PROFESSOR,
[5] ASSOCIATE PROFESSOR
[1]Department of Computer Science And Engineering,
[1]Satya Institute Of Technology and Management , Vizianagaram , Andhra Pradesh, India

Abstract : Rheumatic Heart Disease (RHD) is an autoimmune response to a bacterial attack which deteriorates the normal functioning of the heart valves. The damage on the valves affects the normal blood flow inside the heart chambers which can be recorded and listened to via a stethoscope as a phonocardiogram. However, the manual method of auscultation is difficult, time consuming and subjective. In this study, a convolutional neural network based deep learning algorithm is used to perform an automatic auscultation and it classifies the heart sound as normal and rheumatic. The classification is done on un-segmented data where the extraction of the first, the second and systolic and diastolic heartsounds are not required. The architecture of the CNN network is formed as an array of layers. Convolutional and batch normalization layers followed by a max pooling layer to down sample the feature maps are used. At the end there is a final max pooling layer which pools the input feature map globally over time and at the end a fully connected layer is included. The network has five convolutional layers. This current work illustrates the use of deep convolutional neural network using a Mel Spectrotemporal representation. For this current study, an RHD heart sound data set is recorded from one hundred seventy subjects from whom one hundred twenty four are confirmed RHD patients. The system has an overall accuracy of 96.1% with 94.0% sensitivity and 98.1% and specificity.

Index Terms – **Rheumatic heart failure, deep learning, heart sounds, machine learning , PCG.**

1. INTRODUCTION

INTRODUCTION

Rheumatic Heart Failure (RHF), also known as congestive heart failure, is a medical condition in which the heart is unable to pump blood effectively, leading to a range of symptoms and complications. It is a long term or Rheumatic condition, as opposed to acute heart failure, which occurs suddenly. RHF can be caused by various underlying conditions, including coronary artery disease (CAD), hypertension, heart valve disorders, cardiomyopathies, and other heart-related diseases. It can also result from non-cardiac factors such as kidney disease, diabetes, and lifestyle factors like obesity and excessive alcohol consumption.

Rheumatic heart failure (RHF) is a Rheumatic, 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]. RHF has reached epidemic proportions in the population, as its incidence is increasing by 2% annually. In the developed world, RHF affects 1-2% of the total population and 10% of people older than 65 years.

1.1 AIM OF THE STUDY.

The aim for Rheumatic Heart Failure Detection Using ML/DL Model is to develop a machine learning (ML) model that can accurately and reliably detect Rheumatic heart failure (R) based on recordings of physiological signals and patient data. RHF is a serious and Rheumatic medical condition characterized by the heart's inability to pump blood effectively, and early detection and monitoring are essential for providing appropriate medical care and intervention.

1.2 SCOPE OF THE STUDY.

The scope for the study of "Rheumatic Heart Failure Detection" is both broad and significant, offering various avenues for research and potential applications. The scope can encompass data from different healthcare institutions, wearable devices, and research studies. The scope may extend to real-time monitoring and early detection systems that enable timely intervention in patients with RHF.

2. RESEARCH METHODOLOGY

The methodology section outline the plan and method that how the study is conducted. The details are as follows;

2.1 DATA DESCRIPTION

2.1.1 UKC-JSI DATASET

Seven datasets were used in our study. The first was collected by the authors of this paper, while the rest of the datasets originated from the 2016's PhysioNet challenge [1]. Our dataset (UKC-JSI, Table 1) was obtained using a professional digital stethoscope 3MTM Littmann Electronic Stethoscope Model 3200. stethoscope allows the recording of up to 12 clips of up to 30 s in length, with a sampling rate of 4 KHz.

TABLE I OVERVIEW OF OUR EXPERIMENTAL DATA RECORDED ON HEALTHY INDIVIDUALS AND ON PATIENTS IN DECOMPENSATED AND RECOMPENSATED RHF EPISODES.

	Decomp.	Recomp.	Healthy	All
# Subjects	51	22	110	183
# Recordings	52	22	159	233
Duration (min)	17	7	52	76

2.1.2 PhysioNet DATASET

The PhysioNet 2016 Cardiology Challenge database consists of six datasets (A through F, recorded by six research groups - participants of the competition) containing a total of 3,153 heart sound recordings, lasting from 5 seconds to just over 2 minutes. The recordings were obtained in either a clinical or a nonclinical environment, from both healthy subjects and pathological patients, and from different locations on the body.

2.2 METHOD

The outline of our method is presented in Fig. 2. It consists of the following two main components: a classic ML component (represented with colored squares on the right side of the figure) and an end-to-end DL component (represented with non colored squares). The input to the classic ML pipeline is the same as the input to the end-to-end DL pipeline, but the classic ML pipeline contains a feature extraction process to extract features from the raw data and to format the data into a classic ML format.

OVERVIEW OF THE SIX (A TO F) PHYSIONET CHALLENGE DATASETS.

			Proportion (%)		
Data	**#Subjects**	**#Recordings**	**Abnorm.**	**Normal**	**Unsure**
A	121	409	67.5	28.4	4.2
B	106	490	14.9	60.2	24.9
C	31	31	64.5	22.6	12.9
D	38	55	47.3	47.3	5.5
E	356	2054	7.1	86.7	6.2
F	112	114	27.2	68.4	4.4
ALL	764	3153	18.1	73.0	9.7

2.3 CLASSIC MACHINE LEARNING

The classic ML component consists of feature extraction, feature selection, and a segment-based ML model. As noises not related to heart sounds are expected to be present in the recordings, the first step is filtering. Most cardiovascular sounds are most likely to occur in the frequency range below 1 kHz [25]; thus, we applied a low-pass Butterworth filter with a threshold of 1 kHz to the raw audio files. For the segmentation of the filtered audio signal, we used a sliding window technique.

2.4 DEEP LEARNING

DL represents a class of ML algorithms that use a cascade of multiple layers of nonlinear processing units [37]. The first layer receives the input data, and each successive layer uses the output from the previous layer as input. DL architectures are able to solve complicated AI tasks (e.g., in computer vision, language, biomedicine, etc.) by learning high-level abstractions from raw data [38].

A cascade of multiple layers of nonlinear processing units, where each processing unit receives input from the previous layer, is called a Fully Connected Neural Network (FCNN). In a typical FCNN, layer i computes an output vector zi using the following equation:

$$\boldsymbol{z}_i = f(\boldsymbol{b}_i + \boldsymbol{W}_i \boldsymbol{z}_{i-1}) \qquad (1)$$

where bi (biases) and Wi (weights) are the parameters for the i th layer. $zi-1$ is the output vector of the previous layer, and $z0$ is the input data. The activation function f can be a Rectified Linear Unit (ReLU), as follows [39]:

$$f(c) = \max(0, c) \qquad (2)$$

Or some other nonlinear function, for example, sigmoid or tanh. For classification problems, the final output layer ( ) usually uses a soft max activation function, as follows:

$$z_{i_j} = \frac{e^{b_{i_j} + W_{i_j} z_{i-1}}}{\sum_j e^{b_{i_j} + W_{i_j} z_{i-1}}} \qquad (3)$$

where j represents the j th row of the weights Wi . The softmax function has a nice property, as follows:

$$\sum_j z_{I_j} = 1 \qquad (4)$$

and it is always positive; thus, it can be used as an estimator for an input data sample x to belong to the j th class for a specific problem, as follows:

$$P(y = j|\boldsymbol{x}) \qquad (5)$$

The parameters of the network (bi and) are learned using an optimization algorithm, for example, gradient descent [40]. For a binary classification problem (e.g., y is either 1 or 0), binary cross-entropy is used as a loss function that is minimized over the N pairs of data samples and labels (xn, yn).

$$J(W) = -\frac{1}{N} \sum_n [y_n log(p_n) + (1 - y_n) log(1 - p_n)] \qquad (4)$$

where, pn is the estimated probability for the Nth sample to belong to class 1. CNNs are a type of NN that are designed with three main architectural ideas to ensure some degree of shift, scale, and distortion invariance.

2.4.1 COMBINING CLASSIC MACHINE-LEARNING AND END-TO-END DEEP LEARNING

The four outputs of the components before the recording based ML model ("R x S x Fs features", "R x S predictions", "R x S x 32 DNN features" and "R x S DNN predictions") are first averaged for each recording (thus, we obtain the averaged "R x Fs features", "R predictions", "R x 32 DNN features" and "R DNN predictions") and are then used as the input to the recording-based ML model. Finally, the recording-based ML model outputs the final prediction for each recording.

Each of the three classifiers (the segment-based classifier, the Spectro-temporal ResNet and the recording-based classifier) were trained separately on the training data. Since the recording-based classifier is a meta-learner that utilizes the output of the segment-based classifier and the Spectro-temporal ResNet, a holdout set is required for its training.

2.4.2 Baseline method

The baseline method starts with heart sound segmentation using the Springer's algorithm [17] for detecting the four states, i.e., S1, systole, S2, and diastole (see Fig. 1). After that, it extracts twenty features from the position information of the four states mentioned above. The first ten features are defined as the averages and standard deviations of the beat-to-beat intervals (RR intervals), S1, S2, systolic and diastolic intervals.

The last ten features describe the averages and standard deviations of the ratios between the different intervals and the ratios between the mean absolute amplitude during systole or diastole to that during the S1 or S2 period in each heartbeat. The features are fed to a binary logistic regression classifier. To provide a better comparison, we also used an ensemble method (Random Forest) as another baseline classifier.

3. Libraries Used

Pandas: Pandas is a Python computer language library for data analysis and manipulation. It offers a specific operation and data format for handling time series and numerical tables. It differs significantly from the release3-clause of the BSD license. It is a well-liked open-source of opinion that is utilized in machine learning and data analysis.
NumPy: The NumPy Python library for multi-dimensional, big-scale matrices adds a huge number of high-level mathematical functions. It is possible to modify NumPy by utilizing a Python library. Along with line, algebra, and the Fourier transform operations, it also contains several matrices-related functions.
Matplotlib: It is a multi-platform, array-based data visualization framework built to interact with the whole SciPy stack. MATLAB is proposed as an open-source alternative. Matplotlib is a Python extension and a cross-platform toolkit for graphical plotting and visualization.
Scikit-learn: The most stable and practical machine learning library for Python is scikit-learn. Regression, dimensionality reduction, classification, and clustering are just a few of the helpful tools it provides through the Python interface for statistical modeling and machine learning. It is an essential part of the Python machine learning toolbox used by JP Morgan. It is frequently used in various machine learning applications, including classification and predictive analysis.
Keras: Google's Keras is a cutting-edge deep learning API for creating neural networks. It is created in Python and is designed to simplify the development of neural networks. Additionally, it enables the use of various neural networks for computation. Deep learning models are developed and tested using the free and open-source Python software known as Keras.
h5py: The h5py Python module offers an interface for the binary HDF5 data format. Thanks to p5py, the top can quickly halt the vast amount of numerical data and alter it using the NumPy library. It employs common syntax for Python, NumPy, and dictionary arrays.

4. RESULTS AND DISCUSSION

The metrices considered in this project is :
Accuracy, Sensitivity, Specificity

METRICS	MACHINE LEARNING ALGORITHM	DEEP LEARNING ALGORITHM
ACCURACY	86.4	95.1
SENSITIVITY	76.9	100
SPECIFICITY	90.9	92.7

It has been demonstrated that deep convolutional networks which are designed for image recognition can be successfully trained to classify heart sound spectral images. Our trained model obtained best overall accuracy ,sensitivity and 98.2% specificity in detecting RHD. This can help to develop timely, affordable and reliable access to cost-effective technologies for the detection and prevention of rheumatic heart disease.

5. References

[1] M. Gjoreski et al., "Rheumatic heart failure detection from heart sounds using a stack of machine-learning classifiers," in 2017 International Conference on Intelligent Environments (IE). IEEE, 2017, pp. 14-19.
[2] J. Voigt et al., "A reevaluation of the costs of heart failure and its implications for allocation of health resources in the United States," Clinical cardiology, vol. 37, no. 5, pp. 312-321, 2014.

[3] G. D. Clifford et al., “Classification of normal/abnormal heart sound recordings: the PhysioNet/Computing in Cardiology Challenge 2016,” in 2016 Computing in Cardiology Conference (CinC). IEEE, 2016, 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, 2016.
[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Image Net classification with deep convolutional neural networks,” in International Conference on Neural Information Processing Systems, 2012, pp. 1097-1105.
[6] C. Szegedy et al., “Inception-v4, inception-ResNet and the impact of residual connections on learning,” in Thirty-First AAAI Conference on Artificial Intelligence, 2017.
[7] T. Young et al., “Recent trends in deep learning based natural language processing,” IEEE Computational intelligence magazine, vol. 13, no. 3, pp. 55-75, 2018.
[8] Y. Bengio et al., “A neural probabilistic language model,” Journal of machine learning research, vol. 3, no. Feb, pp. 1137–1155, 2003.
[9] S. Amiriparian et al., “Recognition of echolalic autistic child vocalisations utilising convolutional recurrent neural networks,” in Inter speech, 2018, pp. 2334-2338.
[10] S. Amiriparian et al., “Bag-of-Deep-Features: Noise-robust deep feature representations for audio analysis,” in International Joint Conference on Neural Networks (IJCNN). IEEE, 2018, pp. 1-7.
[11] M. Gjoreski et al. “Deep affect recognition from R-R intervals,” in Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers. ACM, 2017, pp. 754-762.
[12] H. P. Martinez, Y. Bengio, G. N. Yannakakis, “Learning deep physiological models of affect,” IEEE Computational intelligence magazine, vol. 8, no. 2, pp. 20-33, 2013.
[13] M. Gams, “Weak intelligence: Through the principle and paradox of multiple knowledge,” Advances in computation: Theory and practice, Volume 6, Nova science publishers, inc., NY, ISBN 1-56072-898-1, pp. 245, 2001.
[14] M. Gjoreski et al., “Classical and deep learning methods for recognizing human activities and modes of transportation with smart phone sensors,” in Information Fusion. 2019, under revision.
[15] M. Gjoreski et al., “Toward early detection and monitoring of Rheumatic heart failure using heart sounds,” Intelligent Environments, pp. 336- 343, 2019.
[16] S. Schmidt et al., “Segmentation of heart sound recordings from an electric stethoscope by a duration dependent hidden-markov model,” in 2008 Computers in Cardiology. IEEE, 2008, pp. 345-348.
[17] D. B. Springer, L. Tarassenko, and G. D. Clifford, “Logistic regression-HSSM-based heart sound segmentation,” IEEE Transactions on Biomedical Engineering, vol. 63, no. 4, pp. 822-832, 2015.
[18] M. N. Homsi et al., “Automatic heart sound recording classification using a nested set of ensemble algorithms,” in 2016 Computing in Cardiology Conference (CinC). IEEE, 2016, pp. 817-820.
[16] S. Schmidt et al., “Segmentation of heart sound recordings from an electric stethoscope by a duration dependent hidden-markov model,” in 2008 Computers in Cardiology. IEEE, 2008, pp. 345-348.
[17] D. B. Springer, L. Tarassenko, and G. D. Clifford, “Logistic regression-HSSM-based heart sound segmentation,” IEEE Transactions on Biomedical Engineering, vol. 63, no. 4, pp. 822-832, 2015.
[18] M. N. Homsi et al., “Automatic heart sound recording classification using a nested set of ensemble algorithms,” in 2016 Computing in Cardiology Conference (CinC). IEEE, 2016, pp. 817-820.
[19] F. Plesinger et al., “Heart sounds analysis using probability assessment,” Physiological measurement, vol. 38, no. 8, p. 1685, 2017.
[20] C. Potes et al., “Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds,” in 2016 Computing in Cardiology Conference (CinC). IEEE, 2016, pp. 621-624.