



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ECG-iCOVIDNet: Interpretable AI model to identify changes in the ECG signals of post-COVID subjects

Amulya Agrawal^a, Aniket Chauhan^a, Manu Kumar Shetty^b, Girish M. P^c, Mohit D. Gupta^c, Anubha Gupta^{a,*}

^a SBILab, Department of ECE, IIT-Delhi, Delhi, India

^b Department of Pharmacology, MAMC, Delhi, India

^c Department of Cardiology, GIPMER, Delhi, India

ARTICLE INFO

Keywords:

COVID
Electrocardiogram (ECG)
Post-COVID
CNN
AI in ECG
Interpretability
Shapley additive exPlanations (ShAP)

ABSTRACT

Objective: Studies showed that many COVID-19 survivors develop sub-clinical to clinical heart damage, even if subjects did not have underlying heart disease before COVID. Since Electrocardiogram (ECG) is a reliable technique for cardiovascular disease diagnosis, this study analyzes the 12-lead ECG recordings of healthy and post-COVID (COVID-recovered) subjects to ascertain ECG changes after suffering from COVID-19.

Method: We propose a shallow 1-D convolutional neural network (CNN) deep learning architecture, namely ECG-iCOVIDNet, to distinguish ECG data of post-COVID subjects and healthy subjects. Further, we employed ShAP technique to interpret ECG segments that are highlighted by the CNN model for the classification of ECG recordings into healthy and post-COVID subjects.

Results: ECG data of 427 healthy and 105 post-COVID subjects were analyzed. Results show that the proposed ECG-iCOVIDNet model could classify the ECG recordings of healthy and post-COVID subjects better than the state-of-the-art deep learning models. The proposed model yields an F_1 -score of 100%.

Conclusion: So far, we have not come across any other study with an in-depth ECG signal analysis of the COVID-recovered subjects. In this study, it is shown that the shallow ECG-iCOVIDNet CNN model performed good for distinguishing ECG signals of COVID-recovered subjects from those of healthy subjects. In line with the literature, this study confirms changes in the ECG signals of COVID-recovered patients that could be captured by the proposed CNN model. Successful deployment of such systems can help the doctors identify the changes in the ECG of the post-COVID subjects on time that can save many lives.

1. Introduction

The first Case of COVID-19 was registered in the Wuhan City of China. COVID-19 is a type of viral disease that is contagious and rapidly spreads through spilled respirational material (cough, sneeze) present in the exhaled air of the infected people. Reverse Transcription-quantitative PCR (RT-qPCR) is a gold standard test for diagnosis of COVID-19 [1]. This disease crossed the geographical boundaries with devastating effects in a short period, and today the entire world is fighting against this pandemic and hence, COVID-19 has caused immense social and economic losses throughout the world [2]. Based on the global statistics, until August 12, 2021, more than 205 million people suffered from this infection, and 4.3 million people lost their battle to COVID-19 (<https://www.worldometers.info/coronavirus>).

Researchers have been trying to develop time-series models to predict the statistics in order to support the agencies with appropriate policy decisions [2,3].

The infection impacts the respiratory tract and causes lung pneumonia, fever, cough, and loss of taste and smell [4]. Although medical science is focused on developing effective medication and preventive therapy like vaccines, there is no effective therapy available for COVID-19. Early diagnosis, patient isolation, supportive therapy are the primary modes of management of COVID-19. Recently, some studies have indicated cardiac problems in patients recovered from COVID [5, 6]. It has been observed that even after recovering from COVID-19, survivors develop sub-clinical to clinical heart damages, even though subjects did not have underlying heart disease before COVID-19 [7]. Type 1 heart attack, which is caused due to blockage in heart arteries

* Corresponding author.

E-mail address: anubha@iiitd.ac.in (A. Gupta).

<https://doi.org/10.1016/j.combiomed.2022.105540>

Received 28 January 2022; Received in revised form 26 March 2022; Accepted 15 April 2022

Available online 30 April 2022

0010-4825/© 2022 Elsevier Ltd. All rights reserved.

because of a blood clot, is rarely reported during or after COVID-19 infection. Type 2 heart attacks, caused by stress or low oxygen levels, are most commonly reported in subjects with COVID-19 [8]. It has been discovered that during the COVID-19 blood report, some people have elevated levels of a substance called “troponin” in their blood, along with ECG changes and chest pain. Elevated “troponin” levels are a sign of damaged heart tissues, and this can cause a heart attack [9].

Electrocardiogram (ECG) is used to identify cardiac abnormalities. A 12-lead ECG is generated using six unipolar chest leads (V_1 to V_6), three bipolar limb leads (I, II, and III), and three unipolar limb leads (AVR, AVL, and AVF) placed on the specific locations of the body surface. Each ECG wave consists of P, Q, R, S, T and U waves. To diagnose or detect heart abnormalities, cardiologists analyze the Electrocardiogram (ECG) recordings of the subjects that is time-consuming. Thus, it is required to develop methods to analyze and interpret the variability in ECG signals of post-COVID subjects.

Several deep learning models have been developed to diagnose or predict a disease in early stage using the signals generated from human body such EEG, ECG, and non-invasive images [10–14]. Antczak [15] trained an Inception network, generated synthetic ECG data from time-domain Wasserstein GAN, and trained a denoising encoder to perform ECG denoising. Ullah et al. [16] first pre-processed the ECG data for denoising to remove the drift noise and then transformed them into two-dimensional spectrogram images. These spectrograms were fed to the 2D convolutional neural network (CNN), which extracts and represents prominent features and classifies the ECG recordings into eight major cardiovascular diseases. Li et al. [17] suggest transforming the ECG recordings into two-dimensional spectrogram images. These transformed images carry the patient’s heartbeat morphology and the temporal relation between two adjacent heartbeats. These images are input to a 2D CNN that performs classification using the information fusion techniques.

Jun et al. [18] suggest that there is no need to pre-process the ECG signals manually because they can be directly converted into two-dimensional gray-scale images by plotting. These images are input to a two-dimensional neural network with an architecture similar to VGGNet. Zhang et al. [19] proposed a novel deep learning technique for multi-class arrhythmia classification using a spatio-temporal attention-based convolutional recurrent neural network. The feed-forward CNN extracts only the local features of ECG. Spatial attention-based pooling extracts the more significant channels. All the local features then combine to form the global features learned using a bi-directional gated recurrent unit (GRU). Avanzato & Beritelli [20] used a convolutional neural network (CNN) to diagnose cardiovascular diseases using the subjects’ ECG data.

Zhang et al. [21] proposed a deep learning architecture that includes stacking of residual blocks and convolutional layers. The model performs better when considering 12 leads of the ECG simultaneously. CNNs extract the temporal features of the ECG. The use of recurrent neural networks has also brought significant results, but that deals with the time-series aspect of ECG. Xu et al. [22] proposed combining CNN and RNN to analyze the ECG beat patterns and diagnose heart diseases. The first two layers of the convolutional network extract the ECG morphology patterns and feed them to the RNN. Transfer learning technique in training the RNN gives outstanding accuracy and an optimal global solution for abnormal ECG classification into different cardiovascular diseases.

Borra et al. [23] performed several experiments on decoding ECG signals using deep learning techniques on the standard dataset of PTBXL. They applied Inception Time, ResNet, and XResNet models to classify the ECG abnormalities into 27 different categories and reported the Inception Time model to perform the best with the 12-lead ECG data. Jo et al. [24] presented an explainable artificial intelligence mechanism to detect irregularities in the heartbeat pattern, atrial fibrillation, and the absence of P-waves using 8-seconds ECG of subjects of multiple hospitals. The usage of explainability of multi-labelled data was observed as

helpful in validating the deep learning models. One module detects irregular heartbeats, and the other detects the absence of P waves. Another recent study has built a CNN based interpretable AI model for cardiac disorders using ECG wave analysis on PTBXL dataset [25]. Similarly, interesting machine learning and deep learning studies have been conducted recently to detect stress in COVID healthcare workers using ECG signal analysis [26,27]. Thus, we observe that CNN based DL models are increasingly being used for ECG analysis.

Heart Rate Variability (HRV) indicates variation in the consecutive heartbeats of ECG signals. The maximum upwards deflection of a normal QRS complex is called the R wave peak in the ECG and the duration between two adjacent R wave peaks is termed the R-R interval. The time period between the adjacent QRS complexes is termed the N–N (normal-normal) interval. HRV is the measurement of the variability of these N–N intervals. Some recent studies have indicated change in heart rate variability (HRV) in COVID-recovered subjects [28–30]. This indicates that tracking of the heart status of post-COVID recovered subjects can help in providing timely assistance to these subjects for better survival.

In general, the analysis of HRV involves preprocessing of ECG data including noise removal [31–33], feature extraction, normalization. Traditionally, ECG signals are analyzed using the time domain and the frequency domain features, say HRV features, extracted from the one-dimensional waveforms of different leads. The manual examination of ECG signals requires expertise in the field and is a time-consuming process. Recent advances in AI can help to analyze and interpret ECG data accurately.

Motivated with the above discussion, we employed DL model to classify post-COVID subjects from healthy using the ECG data. Overall, the salient points of this work are as follows:

- 1 ECG data of COVID-recovered and healthy subjects were collected at two hospitals in Delhi, India.
- 2 Several traditional and deep learning models are trained and evaluated on the ECG data of healthy and post-COVID subjects.
- 3 Two shallow convolutional neural network architectures are proposed for the classification task. The first model, ECG-iCOVIDNet, works only on the raw ECG data, while the second model, ECG-HiCOVIDNet carries out the late fusion of the HRV features with the latent space embedding of the CNN features extracted from the raw ECG data. In general, traditional ML methods are used on HRV features, while DL methods utilize only raw ECG waveforms. In this paper, we have designed a DL architecture, ECG-HiCOVIDNet, that works on the raw ECG signals and on HRV features. To the best of our knowledge, this is one of the first studies that carries out the late fusion of HRV features in the DL model for ECG data analysis. Both the proposed models are shown to outperform the standard state-of-the-art CNN models on the ECG data.
- 4 SHAP technique is used to evaluate interpretability at the patient and the population level. At the patient level, the segments of ECG wave contributing to the classification are highlighted. The lead-wise contribution to the classification is identified.
- 5 To the best of our knowledge, this is one of the first studies to analyze the raw ECG signals of COVID-recovered patients for detecting cardiac abnormalities.

2. Materials

Data were collected by the Department of Cardiology, G.B. Pant Hospital, Delhi, India and Lok Nayak Hospital, Delhi, India. COVID-19 patients who had recovered (30–60 days after the date of infection) were initially screened for the eligibility criteria. Patients with pre-existing cardiac conditions and pathological conditions before COVID-19 infections were excluded from the study. After screening, 117 subjects were eligible for the study. A 12 lead, 500 Hz, 60 s ECG data was collected. These data were recorded during supine paced breathing using VESTA 301i (500 Hz). Similarly, ECG data of 430 healthy subjects

recorded in the study [34] at the same hospitals using the same machines were used as the control group data. We removed 12 post-COVID-19 and 3 healthy samples because their ECG data were very noisy. Finally, the ECG data of 105 post-COVID subjects (labelled as class ‘1’) and 427 healthy subjects (labelled as class ‘0’) were included for analysis in the study.

The dataset is divided into five folds corresponding to which five classifiers are trained. Each time while training a new classifier, one fold is used as the test set and the rest of the 4 folds are used for training the model. For each classifier, the training data of 4 folds is further divided into 80% as the training data and 20% as the validation data. The distribution of samples into training set, validation set and test set is shown in Table-1.

2.1. Feature extraction

Heart rate features were extracted using the HRV-analysis module Tarvainen et al. [35]. This tool removes outliers and ectopic beats from a signal using Malik’s rule Acar et al. [36]. The following time-domain HRV features were extracted: mean heart rate (Mean-HR), standard deviation of heart rate (STD-HR), mean of NN intervals (Mean-NNI), where R peak of ECG is also called the N point, median of the successive difference between NN intervals (Median-NNI), range NNI (Range-NNI), PNNI-20 (percentage of successive NN interval greater than 20 ms), PNNI-50 (percentage of successive NN interval greater than 50 ms) and standard deviation of the NN intervals (STD-NNI). HRV features derived from the NN intervals included RMSSD (root mean square NN intervals), CVNNI (Co-efficient of variation equal to the ratio of standard deviation of the NN intervals divided by mean NN interval) and CVSD (Coefficient of variation of successive difference equal to the root mean square NN intervals divided by mean NN interval). Frequency domain HRV features included High frequency (HF), Low Frequency (LF), Very Low Frequency (VLF), HFNU (normalized high frequency power), LFNU (normalized low frequency power value), and LF/HF (ratio of low frequency and high frequency power).

3. Methods

3.1. Models

In this subsection, we present the existing state-of-the-art DL models, traditional ML classifiers, and the proposed architectures that were trained and tested using five-fold cross-validation, on the above described dataset. In general, traditional ML methods are used on HRV features, while DL methods are used on only raw ECG waveforms. Hence, we used the existing state-of-the-art DL models on the raw ECG data, the traditional ML classifiers on the HRV features, and our proposed DL models with and without HRV features.

3.1.1. Existing standard state-of-the-art DL models

Spatio-Temporal CNN Model (ST-CNN-8): Attia et al. [37] proposed a spatio-temporal CNN model for ECG data analysis that considers the temporal aspect of ECG signal of the leads along time using eight temporal layers and spatial aspect across the leads using one spatial

layer. These temporal and spatial layers are followed by two fully connected dense layer with sigmoid activation function. We implemented this architecture and also named it as ST-CNN-8.

ResNet50: ResNet architecture was introduced by He et al. [38] for image recognition problem. ResNet50 is a 50-layer deep architecture. These networks have a general architecture of convolution, pooling, activation and fully-connected layers stacked one over the other. Although this stacking allows better feature to be learned, a deeper architecture can still show degradation owing to multiple reasons including the problem of vanishing or exploding gradients. Thus, each layer of ResNet learns a residual function instead of fitting a desired underlying function via the use of skip connections. These skip connections solve the problem of vanishing gradients and enable the model to learn an identity function. This ensures good performance in the deeper layers as well. Thus, ResNet provides better performance, in general. Here, ResNet50 model is adapted for one-dimensional inputs. We have used the publicly available implementation by Kotikalapudi [39].

SENet: Squeeze-and-Excitation Network or SENet [40] explicitly model the independencies between channels and adaptively gives importance to them according to the relevance. SENet applies global average pooling to generate channel-wise statistics to squeeze global spatial information into a channel vector of size equal to the number of convolutional channels. This squeezed vector is passed through a two-layer neural network and the output can be used as weight on original feature maps and, thus, one can adaptively recalibrate channel-wise feature responses.

Attention-56: This architecture was proposed by Wang et al. [41]; wherein attention modules are stacked between the residual units for modeling an attention-aware network that can learn attention-aware features. We have used the publicly available implementation by Sourajit2110 [42].

3.1.2. Traditional machine learning models

Logistic Regression: Logistic regression is a supervised learning method that is used to implement binary classification. It predicts the probability of the input sample belonging to each class. It is computed by fitting an “S” shaped logistic function to the data. The output probability indicates the likelihood of a subject belonging to the “post-COVID class”.

SVM: Support vector machine (SVM) is a very popular traditional supervised machine learning classifier. In SVM, a data sample is plotted as a point in an n -dimensional space (where n is the number of features). These data points are divided into different classes via finding the hyperplane that maximizes the distance of the nearest data point of each class (on opposite sides of this plane) from the hyperplane. Thus, it is also called the maximum-margin classifier. We utilized SVM classifier with RBF kernel.

Decision Tree: Decision tree is a supervised tree-structured classifier that takes decision by asking binary questions. Based on the answers (yes/No), it splits the branches of the tree. Features are present at the nodes and the branches represent the decision rules. The outcome is represented by the leaf nodes. This is also one of the most efficient traditional machine learning method. We utilized GINI criterion in the decision tree.

Table 1

In each split, one of the folds is treated as the test set and the remaining 4 folds are divided into 80% as training data and 20% as validation data.

Split	Training Data		Validation Data		Test Data		Total Data	
	healthy	post-COVID	healthy	post-COVID	healthy	post-COVID	healthy	post-COVID
1	273	67	68	17	86	21	427	105
2	273	67	68	17	86	21	427	105
3	273	67	69	17	85	21	427	105
4	273	67	69	17	85	21	427	105
5	273	67	69	17	85	21	427	105
				Total	427	105		

3.1.3. Proposed DL models

Convolutional Neural Network without HRV: ECG-iCOVIDNet (Fig. 1) The architecture of the proposed ECG-iCOVIDNet model comprises of three convolutional blocks stacked sequentially after each other. Each convolutional block comprises of a 1D-convolutional layer with ReLu activation followed by a batch-normalization (BN) layer. BN layer is used to deal with the internal covariance-shift problem. After the BN layer of the first convolutional block, a dropout layer is also used. Dropout layer discards some nodes randomly from a layer by removing all their connections, and helps in preventing overfitting of the model. The third convolutional block is followed by a global average-pool layer that produces the final feature set of the raw ECG data. These features are also called as the latent space embedding and are fed as input to the fully-connected (FC) layer of 50 nodes. The FC layer also uses Relu activation and is followed by a hidden layer with a single output node with sigmoid activation. This layer outputs the probability value that is used to determine the class of an input data sample as healthy or post-COVID. For the classification, the raw ECG data of a subject is fed as input to the proposed network. A block diagram of ECG-iCOVIDNet is shown in Fig. 1.

Convolutional Neural Network with HRV: ECG-HiCOVIDNet (Fig. 2) Since HRV features are also important from the medical point of view, late fusion of these features is carried out to the latent space embeddings of CNN blocks of the ECG-iCOVIDNet. In other words, features extracted by the flatten layer of the ECG-iCOVIDNet are concatenated with the HRV features. These concatenated features are passed to the fully connected dense layer as shown in Fig. 2. This modified model is named as ECG-HiCOVIDNet.

3.2. Evaluation metrics

For evaluating the proposed model, we have used six evaluation metrics: accuracy, precision, recall, AUC, F_1 -score, and Matthews correlation coefficient (MCC). These evaluation metrics are derived from true positive (TP), false positive (FP), true negative (TN) and false negative (FN). Here, a sample is defined as TP if it is class '1' (post-COVID) and also predicted by the model as class '1' label; a sample is defined as FP if it is class '0' (healthy) and predicted as class '1' label; a sample is defined as TN if it is class '0' (healthy) and predicted as class '0' label; and a sample is defined as FN if it is class '1' (post-COVID) and predicted as class '0' label. A brief description of these evaluation metrics is given as below:

- Accuracy is the ratio of correctly classified samples (both positive and negative class) to the total number of samples. It informs about the percentage of correct predictions of the model.

$$Accuracy = \frac{Correctly\ classified\ samples}{Total\ number\ of\ samples} = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision is the ratio of correctly classified positive samples to all the positive samples claimed by the model. It is used to decide how precise a model is on the positive predicted samples.

$$Precision = \frac{Correctly\ positive\ classified\ samples}{Total\ number\ of\ positive\ claimed\ samples} = \frac{TP}{TP + FP}$$

- Recall is the ratio of correctly classified positive samples to the total number of positive samples in the data. It tells how many of actual positives a model is able to capture of the total number of ground truth positives.

$$Recall = \frac{Correctly\ positive\ classified\ samples}{Total\ number\ of\ positive\ samples\ in\ data} = \frac{TP}{TP + FN}$$

- F_1 -score is used to find a balance between precision and recall, particularly, when the dataset contains uneven class distribution.

$$F_1 - score = 2 \frac{Precision \times Recall}{Precision + Recall}$$

- Area Under Curve (AUC) represents the measure of separability and how much the model is capable of distinguishing between the classes. It does this by plotting the ROC (receiver operating characteristics) curve with True Positive Rate (TPR) on the y-axis and False Positive Rate (FPR) as the x-axis. AUC of 1 indicates complete separability, 0.5 indicates no model separability and AUC of 0 indicates wrong predictions.

- Matthews Correlation Coefficient (MCC) is used to measure the quality of binary classifications. It is similar to the correlation coefficient between the observed and predicted binary classifications and considers TP, TN, FP and FN as a balanced measure. MCC of +1 indicates perfect prediction and -1 indicates completely wrong prediction.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

4. Results and analysis

All the models were trained on the above explained dataset using five-fold cross validation. Google Colab, a cloud-based Jupyter notebook environment, was utilized. The data split is provided in Table 1 describing the number of subjects in the training, validation and test phase for each of the fold's classifier. It was made sure that no test set sample was shown during the training phase. GPU was used as the hardware accelerator. Keras API, which runs on top of the Tensor flow framework, was used to implement the models.

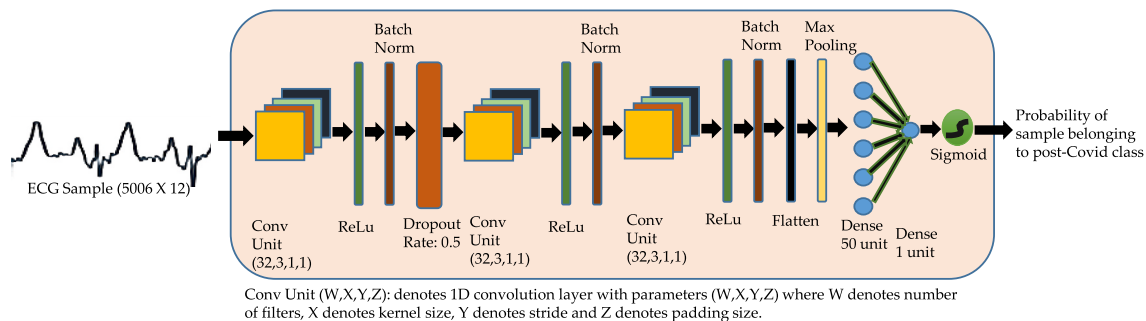


Fig. 1. Block Diagram of ECG-iCOVIDNet Architecture: The model comprises of three convolutional blocks followed by a global average pool layer. The output of global average pool layer is flattened and passed through a fully connected layer, which is connected to a single neuron with sigmoid activation function to obtain the probability of the sample belonging to each class.

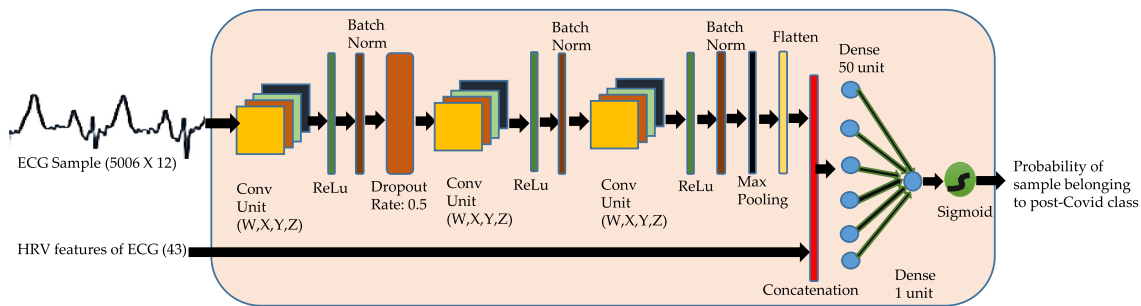


Fig. 2. Block Diagram of ECG-HiCOVIDNet Architecture: The model comprises of three convolutional blocks followed by a global average pool layer. The output is flattened and concatenated with the HRV features and then passed through fully connected dense layers. The output is passed through the sigmoid activation function to obtain the probability. For each split of the data, one classifier of ECG-HiCOVIDNet is trained. The details of the CNN blocks in these classifiers is as follows. Classifier-1: (W = 32,X = 3,Y = 1,Z = 1); Classifier-2: (W = 32,X = 5,Y = 1,Z = 1); Classifier-3: (W = 32,X = 5,Y = 1,Z = 1); Classifier-4: (W = 64,X = 5,Y = 1,Z = 1); Classifier-5: (W = 96,X = 3,Y = 1,Z = 1).

Since traditional ML methods are used on HRV features, these are trained and tested on HRV features, existing state-of-the-art DL models and the proposed ECG-iCOVIDNet are tested on raw ECG data and HRV features. In Both ECG-iCOVIDNet and ECG-HiCOVIDNet, a dropout rate of 0.5 was used and the models were trained using 100 epochs for each of the five splits of the data. The models use Adam optimizer as the optimization algorithm that combines the best of the AdaGrad and RMSProp algorithms and performs much better than other optimizers. Binary cross-entropy is used as the loss metric.

The parameter settings for different DL models are described next. Resnet-50 was trained with the learning rate of 0.0001 for 100 epochs using ADAM optimizer with binary cross entropy loss function. ST-CNN-8 was trained with the learning rate of 0.0005 for 100 epochs using ADAM optimizer with binary cross entropy loss function. A batch size of 64 was chosen. We also used a dropout rate of 0.05. Both SENet and Attention-56 were trained with the learning rate of 0.0005 for 100 epochs using ADAM optimizer with binary cross entropy loss function. A batch size of 64 and a dropout rate of 0.2 was chosen.

4.1. Performance

Results of all models that are described above are shown in Table-2. The table contains all the evaluation metrics described above, namely accuracy, precision, recall, F_1 -score, AUC and MCC, calculated on the test fold for each of the fold’s classifier and compiled for all five folds. Results show that our proposed architecture yields the best performance with 100% accuracy, F_1 -score and AUC as 1, on the test data. To visually demonstrate the ability to distinguish between the healthy and post-COVID classes, t-SNE plots are shown in Figs. 3 and 4 that demonstrate that initially the input data is not distinguishable in two classes. The samples from different classes start forming clusters as we moves from the first convolution block to the last convolution block. Eventually, the data gets segregated into two different classes as seen from the tsne plots made of data after the flatten layer. Thus, we can infer from

this that both the proposed models have the ability to separate healthy and post-COVID samples.

The deep learning models show improvement over the traditional models. The SVM approach applied using only on the HRV features of the ECG dataset shows best performance among these traditional ML models with an accuracy of 80.26%. Logistic Regression tries to classify the linearly separable data, but performs poorly with only 69.81% accuracy on the test data. Attention-based model, namely Attention-56 scored an accuracy of 98.07% and, hence, performed better than ST-CNN-8 and SENet models. The ST-CNN-8 model yielded an accuracy of 97.93% and demonstrated an improvement over the SENet model. The reason behind the improvement could be the use of spatial and temporal layers that could exploit the information of all the channels as well as the information present across channels.

The proposed approach of using CNN architecture features concatenated with the HRV features in the ECG-HiCOVIDNet model demonstrated better performance than the attention-based model. ECG-HiCOVIDNet model scored an accuracy of 99.28%. Resnet model with 50 layers gained higher accuracy with 99.81% on the test data. The proposed ECG-iCOVIDNet model yielded the best results with an accuracy of 100% on the test data. It scored an AUC of 100% and F_1 -score as 1. It outperformed the traditional ML models and the state-of-the-art DL models. Global Average Pooling (GAP) layer after the third convolutional block outputs the average of each feature map and reduced the vector size to 32 before the dense layers. This also reduced the total trainable parameters to nearly 9, 500 in all the five classifiers for five folds. In the ECG-HiCOVIDNet model, the GAP layer outputs the average of each feature map with vector size to 32. These 32 features when concatenated with 43 HRV features, result in a total of 75 features after the concatenation layer. The total trainable parameters increased to nearly 66, 500 in all the five classifiers for five folds. The proposed ECG-HiCOVIDNet model demonstrated an improvement of 19.02% over the traditional models, and 1.14% improvement over Attention-56 model, while the ECG-iCOVIDNet model displayed an improvement of 19.74% over the traditional model, and 0.19% over the ResNet-50 model.

Table 2
Comparative performance of the proposed models with the existing models.

Models	Accuracy	Precision	Recall	F_1 -Score	AUC	MCC
ECG-iCOVIDNet	100%	1.0	1.0	1.0	1.0	1.0
Resnet 50	99.81%	0.995	0.999	0.997	0.9988	0.9942
ECG-HiCOVIDNet	99.28%	0.994	0.985	0.989	0.9809	0.9783
Attention 56	98.07%	0.966	0.977	0.971	0.9483	0.9418
ST-CNN-8	97.93%	0.977	0.958	0.965	0.9583	0.934
SENet	95.38%	0.873	0.887	0.878	0.8845	0.7795
SVM using HRV	80.26%	0.8026	1.0	0.8905	0.5	0.0
Decision Trees using HRV	74.62%	0.8438	0.8407	0.8414	0.6013	0.1975
Logistic Regression using HRV	69.81%	0.8442	0.7647	0.8024	0.5512	0.5966

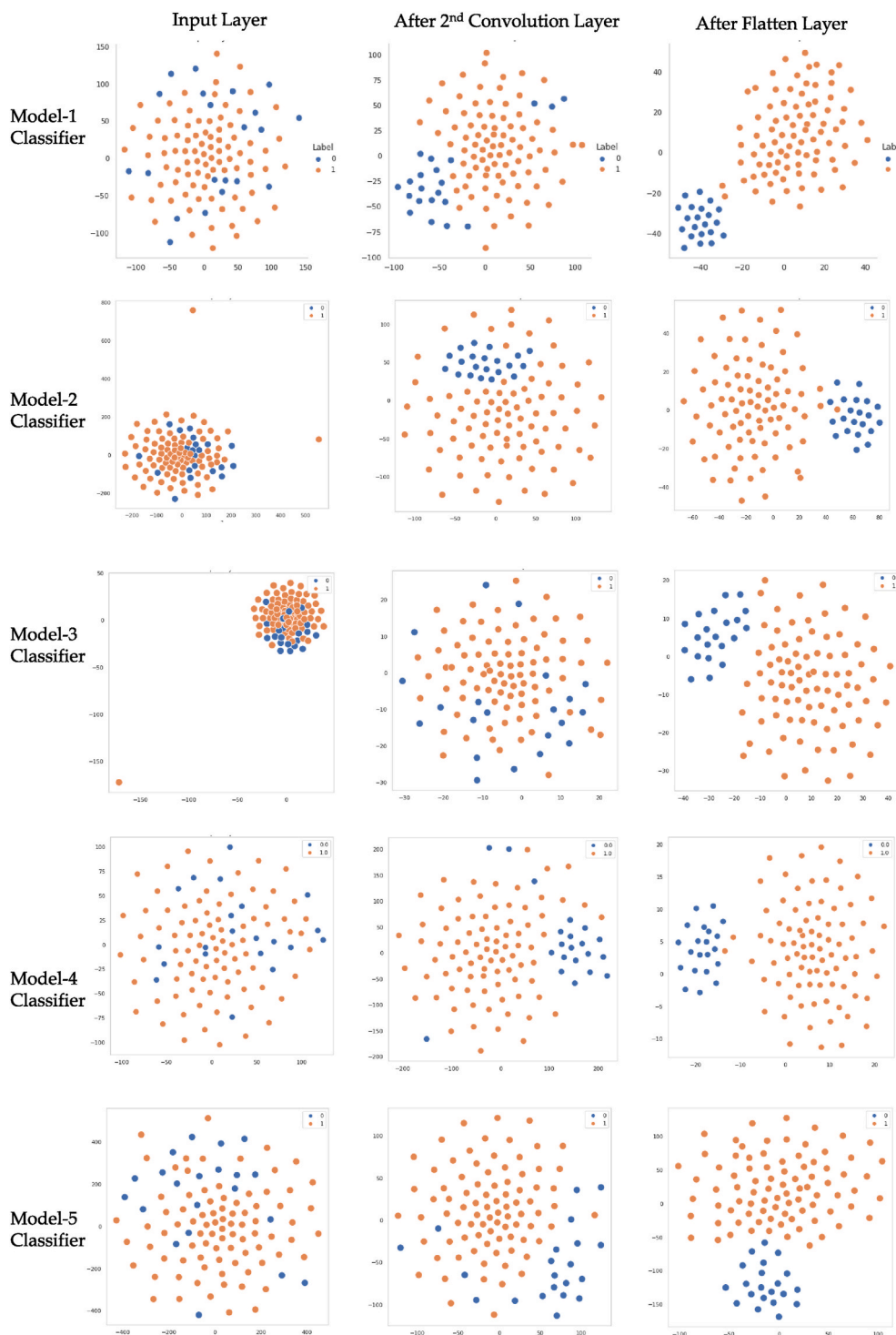


Fig. 3. t-SNE plots of ECG data being separated layer after layer when features are learnt gradually by the ECG-iCOVIDNet model.

4.2. Interpretability

Although we have seen that CNN based architecture, ECG-iCOVIDNet, performed best for ECG data classification, it is difficult for humans to understand the features learned by DL models due to their complex architecture and non-linear behaviour. CNNs are considered as “black boxes” due to the lack of interpretability. ShAP (Shapley Additive Explanations) developed by [43], is an excellent way to interpret the features learnt by the deep neural network. It provides visual explanation of the classification done by our CNN models. We employed ShAP

Gradient Explainer for interpreting the (relevant) distinguishing features. It is based on integrated gradients method, which is a feature attribution method for deep neural networks. From these ShAP values, we used the top 500 ShAP values as the important features for the diagnosis of the particular class. The regions in the ECGs corresponding to these important features are highlighted in red, while the features with lesser importance are seen in the blue color. The analysis is done at two levels as described below.

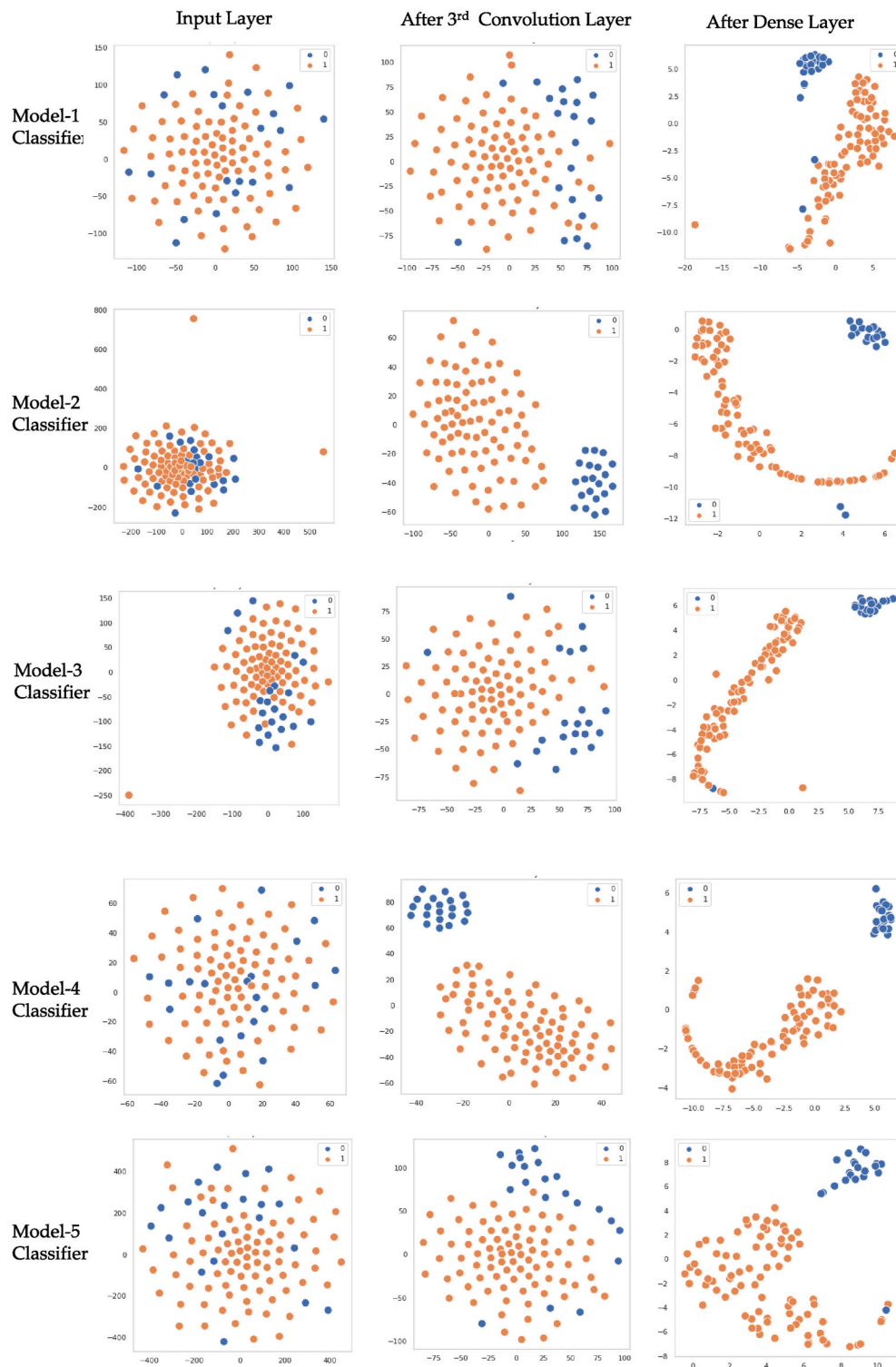


Fig. 4. t-SNE plots of data being separated layer after layer when features are learnt gradually by the ECG-HiCOVIDNet model. Here, HRV features are concatenated with the latent space embedding of the convolutional blocks.

4.2.1. Patient level

At the level of single patient, it is important to identify the features from the data of a particular patient that help in classifying a patient into a particular class.

ShAP applied on the CNN model accepts raw ECG data and generates an output of the same size as the input, with ShAP value for each position of input ECG data. A ShAP value of $S > 0$ indicates positive contribution of the corresponding input position towards the

classification of that patient into its predicted class. Top 500 ShAP values of each lead are used to highlight the most contributing features with red colour as shown in Figs. 5 and 6.

4.2.2. Lead-wise importance

To compute the lead-wise importance for each class, we added the ShAP values of one lead for all the patients of each class separately. This is the total contribution of each lead in each class. Next, we averaged this

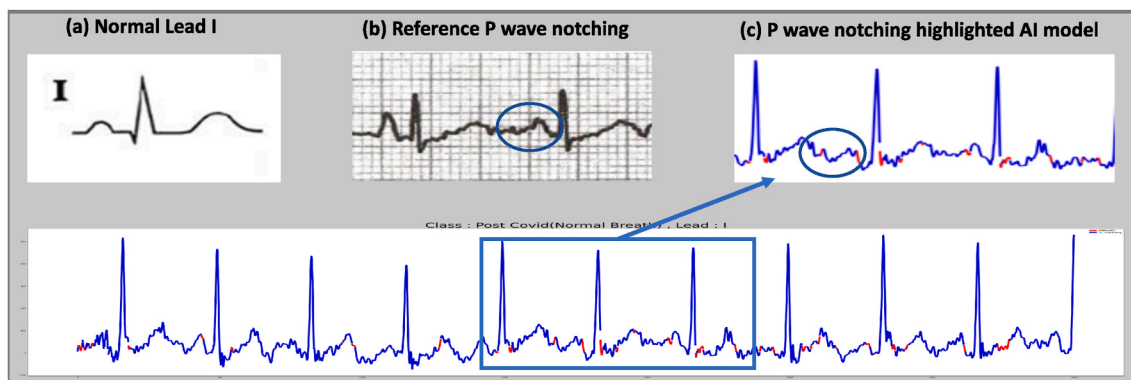


Fig. 5. Wider P wave and/or notching of P wave highlighted in lead I of ECG, in subjects having ejection fraction less than 45%: (a) Showing healthy lead I ECG, (b) wide/notching P wave as reported in the literature, and (c) segments of ECG highlighted in red color by our ECG-iCOVIDNet model correspond to wide/notching P wave regions.).

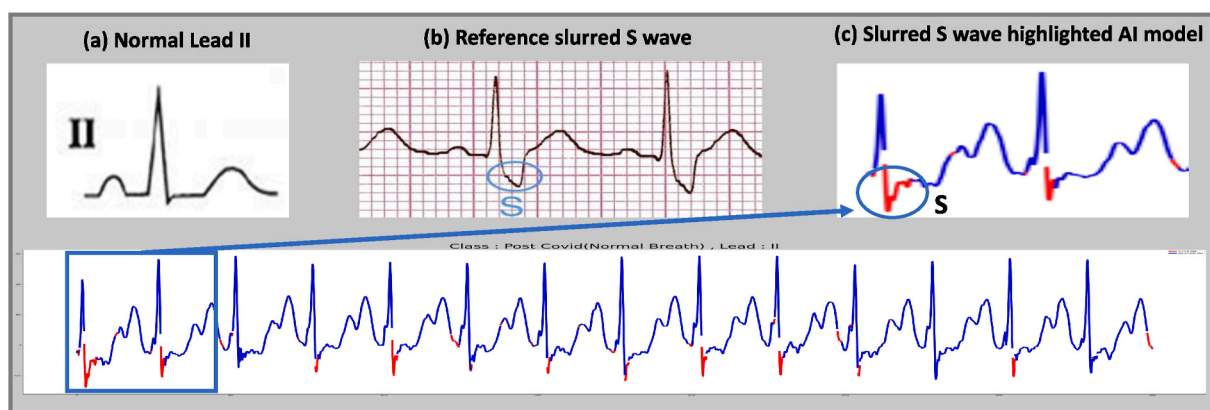


Fig. 6. Slurred S wave highlighted in lead II of ECG, in subjects having left ventricular dysfunction: (a) Showing healthy lead II ECG, (b) slurred S wave reported in the literature, and (c) segments of ECG highlighted in red color by our ECG-iCOVIDNet model correspond to slurred S wave.

lead-wise contribution for each class. The average contribution of each lead is also calculated. This process is documented in Algorithm-1. Fig. 7 shows the impact of all the 12 leads on the two classes and also the average contribution of each lead. It is observed that in the classification/prediction of “post-COVID” class, lead-aVL has the highest contribution, whereas in the prediction of “healthy” class, lead-aVR has the highest contribution. The average contribution of lead-aVL is highest amongst all the leads followed by lead-aVR. It shows that lead aVL contributes maximum to the classification of healthy versus post-COVID. Indeed, this lead is known to have good predictive capability.

We also worked towards the explainability/interpretability of the ECG-iCOVIDNet classifier. In our study, subjects had Left Ventricular

Diastolic Dysfunction (LVDD) that is measured in terms of Global Longitudinal Strain (GLS) during ECHO. A GLS value of less than 16% indicates LVDD and correspondingly, change should have been observed in ECG as slurred S wave. These slurred S wave changes were indeed highlighted by our ECG-iCOVIDNet AI model via ShAP analysis in such patients as shown in Fig. 6. Similarly, the ECG data of subjects having ejection fraction less than 45% has notching or wider P wave. These changes were also detected by the ShAP interpretable AI model and were highlighted in the red color as shown in Fig. 5).

Algorithm 1. Algorithm for Lead Wise Analysis

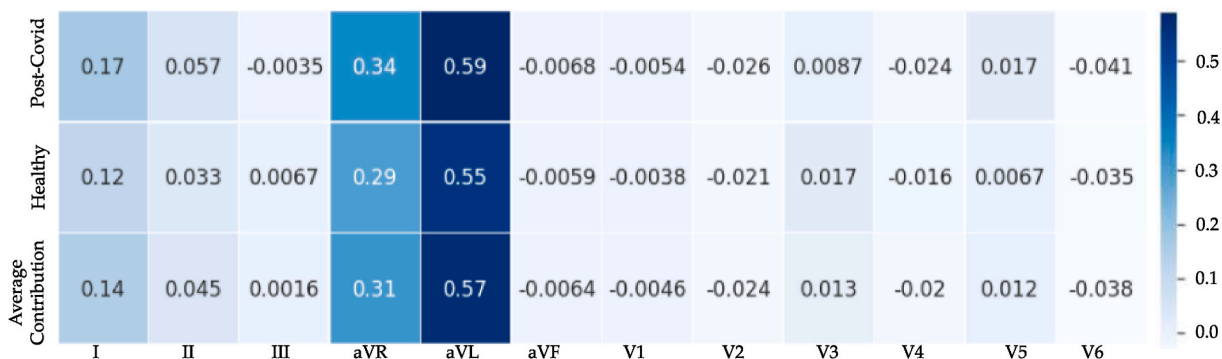


Fig. 7. Lead wise Importance: showing the impact of all the 12 leads on the two classes. The average contribution of lead-aVL is the highest followed by lead aVR.

Algorithm 1 Algorithm for Lead Wise Analysis**Input:** ShAP Values of Test data: $ShAP$, Predicted labels (0 or 1) for test data: $Pred$ **Output:** Total and average Contribution of each lead in classification

```

 $c[i][k]$                                 ▷ Contribution of  $k^{th}$  lead for  $i^{th}$  class
 $sum[i]$                                 ▷ Sum of contribution of all leads for  $i^{th}$  class
 $r[i][k]$                                 ▷ Normalized Contribution of  $k^{th}$  lead on  $i^{th}$  class
 $rdash[k]$                                 ▷ Average Contribution of  $k^{th}$ 
healthyindex  $\leftarrow$  Indices where  $Pred = 1$ 
COVIDindex  $\leftarrow$  Indices where  $Pred = 0$ 

for  $i \leftarrow 0$  to 1 do
    ▷ Get ShAP Values of a particular class
    if  $i = 0$  then
         $S \leftarrow ShAP[healthyindex]$ 
    else if  $i = 1$  then
         $S \leftarrow ShAP[PostCOVIDindex]$ 
    ▷ Getting total contribution of  $k^{th}$  lead on  $i^{th}$  class
    for each lead  $k$  do
        for each patient  $d$  do
            for each timestamp  $j$  do
                 $c[i][k] \leftarrow c[i][k] + S[d][j][k]$ 
    ▷ To normalize the contribution
    ▷ Calculate the sum
    for each class  $i$  do
        for each lead  $k$  do
             $sum[i] \leftarrow sum[i] + c[i][k]$ 
    ▷ Normalize the contribution
    for each class  $i$  do
        for each lead  $k$  do
             $r[i][k] \leftarrow c[i][k]/sum[i]$ 
    ▷ Each lead average contribution
    for each lead  $k$  do
         $rdash[k] \leftarrow (r[0][k] + r[1][k])/2$ 

```

5. Discussion and conclusion

The 12-lead ECG is the most common screening test to check heart diseases. However, most of the time underlying heart disease can not be seen with ECG and require higher diagnostic methods such as ECHO and CT scan. Our ECG-iCOVIDNet model is able to predict the underlying heart problem in post-COVID subjects using only ECG data with 100% accuracy. In this paper, we presented results of various models trained using the ECG data of healthy and post-COVID subjects. It is evident that deep learning models perform better with ECG dataset. We observed that Spatio-temporal CNN (ST-CNN-8), Attention-56, and ResNet, were good architectures for the classification of the samples. The proposed ECG-iCOVIDNet and ECG-HiCOVIDNet models use convolution blocks with global average pool layer to learn features from the 12-lead raw ECG data and demonstrate outstanding performance. This observation is aligned with the visual inferences drawn from the t-SNE plots also. ECG-HiCOVIDNet model that uses HRV features derived from ECG waves along with the raw ECG data yields 99.28% accuracy, while the ECG-iCOVIDNet model without HRV features score 100% accuracy. This shows that ECG-iCOVIDNet could extract relevant features from the raw ECG waveforms and hence, addition of HRV features derived from the raw ECG waveforms did not yield any advantage, again affirming the good performance of the trained model.

Secondly, we also worked towards the explainability/interpretability of the best classifier developed. This can help the medical teams to trust

the decisions made by the ECG-iCOVIDNet model. Our model highlighted the important abnormal segments of ECG that help to distinguish between classes using ShAP at patient-level and population-level. Recently, non-invasive ECG finding P-wave dispersion (Pd) and Wide/notching P wave have been shown as the sign of various pathological conditions in the literature, where Pd is calculated by the shortest and the longest P-wave duration recorded from ECG waves [44]. Identification of Pd for human eye is really impossible. Here, our model highlighted wide/notching P wave in lead I of ECG of Post-COVID in ShAP (Fig. 5). Left ventricular dysfunction is seen in some COVID recovered subjects without previously diagnosed heart disease. Slurred S wave is a sign of left ventricular dysfunction [45,46]. This is highlighted in lead II of ECG of post-COVID subjects (Fig. 6). This explainable AI model with interpretable ShAP figures showing the abnormal segments in ECG waves yields relevant medical results that would help doctors in primary and secondary healthcare centers to trust this AI model that can help to diagnose post-COVID heart abnormalities using the ECG.

Our AI model with visualization of the abnormal ECG segments helps the cardiologists in finding any underlying heart irregularities with less human error, especially, in overloaded healthcare setup in low/middle-income countries such as India. Therefore, this proposed architecture of deep neural networks can be easily deployed at clinical setups, where the entire nation is struggling with a high burden of heart issues after suffering from COVID-19. Moreover, explainable AI models developed on 12-lead ECG data and its HRV features can help the non-cardiologist

diagnose the issues faster and timely, which can improve the efficiency of primary and secondary healthcare services to early diagnose the heart pathology accurately.

6. Benefits and limitations

This work can be helpful for doctors to screen the post-COVID patients who come for follow-up care for addressing the heart issues using ECG without costlier investigation as ECHO, especially, for the doctors at primary and secondary healthcare centers, where no cardiologist is available. Furthermore, the model can be used to analyze the ongoing changes in post-COVID patients and treat them before this changes into major heart problems. Last but not the least, such a study can also be used to identify and develop a broader understanding of the cardiac abnormalities due to coronavirus.

This study has certain limitations. First of all, although an elaborate effort was made to collect the ECG data of post-COVID subjects, the dataset is small as of now. Further, there are many interpretability methods that can be employed to draw inferences on the decisions made by the AI model. We employed SHAP analysis, which is one of the most widely used methods. Here, models are able to learn the abnormalities that exist in the patient's data, but it is possible that inference for all such abnormalities is not easy to generate. In the future, it would be interesting to conduct a benchmarking study on the available interpretability methods on ECG application to figure out which model(s) work best on ECG datasets. A good question to answer is: Is there an interpretability method that works best on ECG data, in general? Researchers can do such a benchmarking study using multiple ECG dataset and multiple interpretability methods. Further, a webapp can be made and installed for use by COVID healthcare workers/doctors, where they can upload the ECG in a tabular format as an input and obtain the results on whether the person is normal or has suffered from COVID earlier. Further, the app can also be used to inform the ECG wave regions where the changes are observed. This work can also be extended to study the nature of COVID and its action on the human heart.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgement

We would like to thank the Centre of Excellence in Healthcare, IIIT-Delhi, India for providing the financial support to carry out this research work.

References

- [1] S. Yamayoshi, Y. Sakai-Tagawa, M. Koga, O. Akasaka, I. Nakachi, H. Koh, K. Maeda, E. Adachi, M. Saito, H. Nagai, et al., Comparison of rapid antigen tests for COVID-19, *Viruses* 12 (2020) 1420.
- [2] P. Singh, A. Gupta, Generalized SIR (GSIR) epidemic model: an improved framework for the predictive monitoring of COVID-19 pandemic, *ISA Trans.* (2021), <https://doi.org/10.1016/j.isatra.2021.02.016>. In press.
- [3] P. Singh, A. Singhal, B. Fatimah, A. Gupta, An improved data driven dynamic SIRD model for predictive monitoring of COVID-19, in: *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2021, pp. 8158–8162.
- [4] J. Majumder, T. Minko, Recent developments on therapeutic and diagnostic approaches for COVID-19, *AAPS J.* 23 (2021) 1–22.
- [5] A. Gasecka, M. Pruc, K. Kukulka, N. Gilis-Malinowska, K.J. Filipiak, M. J. Jaguszewski, L. Szarpak, Post-COVID-19 heart syndrome, *Cardiol. J.* 28 (2021) 353–354.
- [6] J. Hall, K. Myall, J.L. Lam, T. Mason, B. Mukherjee, A. West, A. Dewar, Identifying patients at risk of post-discharge complications related to COVID-19 infection, *Thorax* 76 (2021) 408–411.
- [7] J.P. Lang, X. Wang, F.A. Moura, H.K. Siddiqi, D.A. Morrow, E.A. Bohula, A current review of COVID-19 for the cardiovascular specialist, *Am. Heart J.* 226 (2020) 29–44.
- [8] M. Cameli, M.C. Pastore, G.E. Mandoli, F. D'ascenzi, M. Focardi, G. Biagioni, P. Cameli, G. Patti, F. Franchi, S. Mondillo, et al., COVID-19 and acute coronary syndromes: current data and future implications, *Front. Cardiovasc. Med.* 7 (2021) 369.
- [9] V.S. Mahajan, P. Jarolim, How to interpret elevated cardiac troponin levels? *Circulation* 124 (2011) 2350–2354.
- [10] A. Dogan, M. Akay, P.D. Barua, M. Baygin, S. Dogan, T. Tuncer, A.H. Dogru, U. R. Acharya, PrimePatNet87: prime pattern and tunable q-factor wavelet transform techniques for automated accurate EEG emotion recognition, *Comput. Biol. Med.* 138 (2021), 104867.
- [11] E. Kaplan, S. Dogan, T. Tuncer, M. Baygin, E. Altunisik, Feed-forward LPQNet based automatic alzheimer's disease detection model, *Comput. Biol. Med.* 137 (2021), 104828.
- [12] M. Baygin, S. Dogan, T. Tuncer, P.D. Barua, O. Faust, N. Arunkumar, E. W. Abdulhay, E.E. Palmer, U.R. Acharya, Automated ASD detection using hybrid deep lightweight features extracted from EEG signals, *Comput. Biol. Med.* 134 (2021), 104548.
- [13] O. Attallah, ECG-BiCoNet: an ECG-based pipeline for COVID-19 diagnosis using bi-layers of deep features integration, *Comput. Biol. Med.* 142 (2022), 105210.
- [14] P. Aggarwal, N.K. Mishra, B. Fatimah, P. Singh, A. Gupta, S.D. Joshi, COVID-19 image classification using deep learning: advances, challenges and opportunities, *Comput. Biol. Med.* (2022), 105350.
- [15] K. Antczak, A Generative Adversarial Approach to ECG Synthesis and Denoising, 2020 *arXiv preprint arXiv:2009.02700*.
- [16] A. Ullah, S.M. Anwar, M. Bilal, R.M. Mehmood, Classification of Arrhythmia by using deep learning with 2-D ECG spectral image representation, *Rem. Sens.* 12 (2020) 1685, <https://doi.org/10.3390/rs12101685>. URL.
- [17] J. Li, Y. Si, T. Xu, S. Jiang, Deep convolutional neural network based ECG classification system using information fusion and one-hot encoding techniques, *Math. Probl Eng.* (2018), 7354081.
- [18] T.J. Jun, H.M. Nguyen, D. Kang, D. Kim, D. Kim, Y.-H. Kim, ECG Arrhythmia Classification Using a 2-D Convolutional Neural Network, 2018 *arXiv preprint arXiv:1804.06812*.
- [19] J. Zhang, A. Liu, M. Gao, X. Chen, X. Zhang, X. Chen, ECG-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network, *Artif. Intell. Med.* 106 (2020), 101856.
- [20] R. Avanzato, F. Beritelli, Automatic ECG diagnosis using convolutional neural network, *Electronics* 9 (2020) 951.
- [21] D. Zhang, S. Yang, X. Yuan, P. Zhang, Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram, *iScience* 24 (2021), 102373.
- [22] X. Xu, S. Jeong, J. Li, Interpretation of electrocardiogram (ECG) rhythm by combined CNN and BiLSTM, *IEEE Access* 8 (2020) 125380–125388.
- [23] D. Borra, A. Andaló, S. Severi, C. Corsi, On the application of convolutional neural networks for 12-lead ECG multi-label classification using datasets from multiple centers, in: 2020 *Computing in Cardiology*, 2020, pp. 1–4, <https://doi.org/10.22489/CinC.2020.349>.
- [24] Y.-Y. Jo, Y. Cho, S.Y. Lee, J.-m. Kwon, K.-H. Kim, K.-H. Jeon, S. Cho, J. Park, B.-H. Oh, Explainable artificial intelligence to detect atrial fibrillation using electrocardiogram, *Int. J. Cardiol.* 328 (2021) 104–110.
- [25] A. Anand, T. Kadian, M.K. Shetty, A. Gupta, Explainable AI decision model for ECG data of cardiac disorders, *Biomed. Signal Process Control* 75 (2022), 103584.
- [26] M.D. Gupta, M.K. Jha, A. Bansal, R. Yadav, S. Ramakrishnan, M. Girish, P. G. Sarkar, A. Qamar, S. Kumar, S. Kumar, et al., COVID 19-related burnout among healthcare workers in India and ECG based predictive machine learning model: insights from the BRUCEE-Li study, *Indian Heart J.* 73 (2021) 674–681.
- [27] A. Gupta, D. Kansal, V. Gupta, M.K. Shetty, M. Girish, M.D. Gupta, X-ECGNet: an interpretable DL model for stress detection using ECG in COVID-19 healthcare workers, in: 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART), IEEE, 2021, pp. 1–5.
- [28] T.E. Adler, L. Norcliffe-Kaufmann, R. Condos, G. Fishman, D. Kwak, N. Talmor, H. Reynolds, Heart rate variability is reduced 3-and 6-months after hospitalization for COVID-19 infection, *J. Am. Coll. Cardiol.* 77 (2021), 3062–3062.
- [29] S. Kunal, M.K. Shetty, B. Shah, M. Girish, A. Bansal, V. Batra, S. Mukhopadhyay, J. Yusuf, A. Gupta, M. Gupta, Heart rate variability in post-COVID-19 recovered subjects using machine learning, *Circulation* 144 (2021). A14096–A14096.
- [30] B. Shah, S. Kunal, A. Bansal, J. Jain, S. Poundrik, M.K. Shetty, V. Batra, V. Chaturvedi, J. Yusuf, S. Mukhopadhyay, et al., Heart rate variability as a marker of cardiovascular dysautonomia in post-COVID-19 syndrome using artificial intelligence, *Indian Pacing Electrophysiol J.* 22 (2022) 70–76.
- [31] S. Agrawal, A. Gupta, Fractal and EMD based removal of baseline wander and powerline interference from ECG signals, *Comput. Biol. Med.* 43 (2013) 1889–1899.
- [32] S. Agrawal, A. Gupta, Removal of baseline wander in ECG using the statistical properties of fractional Brownian motion, in: 2013 IEEE International Conference on Electronics, Computing and Communication Technologies, IEEE, 2013, pp. 1–6.
- [33] P. Singh, I. Srivastava, A. Singhal, A. Gupta, Baseline wander and power-line interference removal from ECG signals using Fourier decomposition method, in: *Machine Intelligence and Signal Analysis*, Springer, 2019, pp. 25–36.
- [34] M.D. Gupta, A. Bansal, P.G. Sarkar, M. Girish, M. Jha, J. Yusuf, S. Kumar, S. Kumar, A. Jain, S. Kathuria, et al., Design and rationale of an intelligent algorithm to detect Burnout in Healthcare workers in COVID era using ECG and artificial intelligence: the BRUCEE-LI study, *Indian Heart J.* 73 (2021) 109–113.
- [35] M.P. Tarvainen, J.-P. Niskanen, J.A. Lipponen, P.O. Ranta-Aho, P.A. Karjalainen, Kubios HRV—heart rate variability analysis software, *Comput. Methods Progr. Biomed.* 113 (2014) 210–220.
- [36] B. Acar, I. Savelieva, H. Hemingway, M. Malik, Automatic ectopic beat elimination in short-term heart rate variability measurement, *Comput. Methods Progr. Biomed.* 63 (2000) 123–131.

- [37] Z.I. Attia, P.A. Friedman, P.A. Noseworthy, F. Lopez-Jimenez, D.J. Ladewig, G. Satam, P.A. Pellikka, T.M. Munger, S.J. Asirvatham, C.G. Scott, et al., Age and sex estimation using artificial intelligence from standard 12-lead ECGs, *Circulation: Arrhythmia Electrophysiol.* 12 (2019), e007284.
- [38] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [39] R. Kotikalapudi, keras-resnet. <https://github.com/raghakot/keras-resnet>, 2017. (Accessed 9 October 2020).
- [40] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.
- [41] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, X. Tang, Residual attention network for image classification, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3156–3164.
- [42] Sourajit2110, Residual-attention-convolutional-neural-network. <https://github.com/Sourajit2110/Residual-Attention-Convolutional-Neural-Network>, 2018. (Accessed 7 February 2021).
- [43] Lundberg, M. Scott, and Su-In Lee. A unified approach to interpreting model predictions, *Adv. Neural Inf. Process. Syst.* 30 (2017).
- [44] S. Okutucu, K. Aytemir, A. Oto, P-wave dispersion: what we know till now? *JRSM Cardiovasc. Dis.* 5 (2016), 2048004016639443.
- [45] M. Tudoran, C. Tudoran, V.E. Lazureanu, A.R. Marinescu, G.N. Pop, A.S. Pescariu, A. Enache, T.G. Cut, Alterations of left ventricular function persisting during post-acute COVID-19 in subjects without previously diagnosed cardiovascular pathology, *J. Personalized Med.* 11 (2021) 225.
- [46] H. Takamatsu, H. Tada, H. Okaniwa, H. Toide, H. Maruyama, R. Higuchi, K. Kaseno, S. Naito, M. Kurabayashi, S. Oshima, et al., Right bundle branch block and impaired left ventricular function as evidence of a left ventricular conduction delay, *Circ. J.* 72 (2008) 120–126.