AI-CardioCare

Artificial Intelligence Based Device for Cardiac Health Monitoring

Joshi, Rakesh Chandra; Khan, Juwairiya Siraj; Pathak, Vinay Kumar; Dutta, Malay Kishore

Published in:
IEEE Transactions on Human-Machine Systems

DOI (link to publication from Publisher):
10.1109/THMS.2022.3211460

Publication date:
2022

Document Version
Accepted author manuscript, peer reviewed version

AI-CardioCare: Artificial Intelligence Based Device for Cardiac Health Monitoring

Rakesh Chandra Joshi, Juwairiya Siraj Khan, Vinay Kumar Pathak, and Malay Kishore Dutta

Abstract—**Cardiac disorders are one of the leading causes of mortality around the globe and early diagnosis of heart diseases can be beneficial for its mitigation. In this article, an artificial intelligence (AI) based device has been proposed, which allows for an automatic and real-time diagnosis of cardiac diseases based on deep learning techniques. The heart sound (phonocardiogram) signal is acquired by a customized designed stethoscope and the signal is processed before analysis using AI methods for the classification of four major cardiac diseases (Aortic Stenosis, Mitral Regurgitation, Mitral Stenosis, and Mitral Valve Prolapse). Two deep learning-based neural networks, one-dimensional (1-D) convolutional neural network (CNN) and spectrogram based 2-D-CNN models from the analysis of these signals has been integrated with a low-cost single-board processor to make a standalone device. All data processing is done in a single hardware setup and user interface is provided allowing the user to control the data accessibility and visibility to generate the diagnostic report. As a result, the developed device has demonstrated to be a valuable low-cost diagnostic tool for both medical professionals and personal usage at home.**

Index Terms—**Body auscultation, cardiac disorders, deep neural network, health care system, phonocardiogram.**

I. Introduction

According to World Health Organization, cardiovascular diseases (CVDs) are one of the significant causes of death worldwide and around 17.9 million people losses their life every year [1]. The main cause of high morbidity and death is late identification of heart-related disease due to a lack of necessary facilities and knowledge in developing nations [2], [3]. CVD is a collective term used for a number of blood vessels or heart-related disorders, which generally includes rheumatic heart disease, cerebrovascular disease, and coronary heart disease. The severity of CVDs may be as high as in cases such as heart attacks or strokes leading to premature deaths in four out of five cases [4]. CVDs severely affect the overall health and adversely affect the lives of the general population.

Prevention, early diagnosis and treatment are considered key factors for limiting the negative impact of these deadly diseases. Vibrations in the heart produce sound and murmur throughout each cardiac cycle, which can be recorded as audio wave signals using a digital stethoscope [5]. Heart sounds of healthy person are typically composed of low-frequency signals, whereas high-frequency noises are more common in disease due to turbulent blood flow over faulty heart valves. Electrocardiogram (ECG) and phonocardiogram (PCG) signals are two widespread noninvasive methodologies for the early diagnosis of complications related to cardiac health such as detection of defects and structural abnormalities in the heart valves. The recognition of abnormality and their classification from different abnormal classes of heart sounds can be a strenuous task even for a specialist and may introduce subjectivity in the diagnostic interpretation. In this scenario, artificial intelligence-based methods can benefit in automatic interpretation of cardiac sounds, especially in underdeveloped regions of the world, where there is scarcity of physicians.

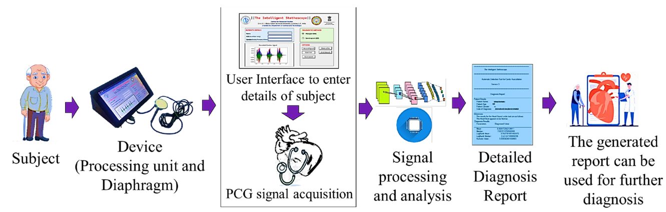

Fig. 1. Conceptual diagram of the cardiac health-monitoring device.

The applications of artificial intelligence (AI) based techniques in PCG signals are used in recent researches to evaluate whether a cardiac sound is normal or abnormal, however, this article tries to find a generalized and effective solution to numerous cardiac disorders and handle different complexities in the signals using various deep learning-based approaches. Multiclass classification of cardiac disorders becomes more challenging in real-world scenario due to the complexity of recorded PCG signals and interference due to ambient noise during heart sound auscultation. The conceptual diagram of the proposed device AI-CardioCare, is shown in Fig. 1. The goal of this article is to design a fully automatic AI-based device capable of identifying four major cardiac disorders in a noninvasive manner and tackling misclassification issue pertaining to

Manuscript received 23 February 2022; revised 19 June 2022 and 28 July 2022; accepted 28 September 2022. This work was supported by the Department of Science and Technology, India, under Grant DST/BDTD/EAG/2017. This article was recommended by Associate Editor L. L. Chen. *(Corresponding author: Malay Kishore Dutta.)*

Rakesh Chandra Joshi and Malay Kishore Dutta are with the Centre for Malay Advanced Studies, Dr. A. P. J. Abdul Kalam Technical University, Lucknow 226031, India (e-mail: rakeshchandraindia@gmail.com; malaykishoredutta@gmail.com).

Juwairiya Siraj Khan is with the Department of Neurorehabilitation Robotics and Engineering, Center for Rehabilitation Robotics, Aalborg University, 9220 Aalborg, Denmark (e-mail: jsikh@hst.aau.dk).

Vinay Kumar Pathak is with the Chhatrapati Shahu Ji Maharaj University, Formerly Kanpur University, Kanpur 208024, India (e-mail: vinay@vpathak.in).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/THMS.2022.3211460.

Digital Object Identifier 10.1109/THMS.2022.3211460

cardiac diseases. The subject needs to record the cardiac signal using an electronic stethoscope and select the requisite options in the user-friendly graphical user interface (GUI). Then, the recorded signal will be comprehensively tested with developed deep learning techniques and user can generate the diagnostic report having the information on the cardiac health.

The rest of this article is organized as follows. Section II illustrates existing related works and advancements through this article. Section III describes the system design of the proposed device. Section IV describes the details about the experimental results obtained on different scenarios of data and training models with different approaches with their discussion. Finally, Section V concludes this article.

II. Existing Related Works and Advancements Through the Current Paper

A. Related Prior Research

Advances in the field of AI with high-speed processors and efficient algorithms have made the concept of a decision support system a reality in the recent decade. Various machine learning and deep-learning methods have been developed and trained with biomedical data to have a more accurate decision-making mechanism [6]. Several studies have been reported for the development of heart disease diagnosis frameworks based on machine learning (ML) models with enhanced performance on clinical data parameters [7], [8], [9] and unsupervised learning approaches like discriminatively boosted clustering [10]. Different clinical data parameters such as age, heart rate, blood sugar, cholesterol, and blood pressure were considered to make predictive decisions and XGBoost demonstrated superior performance with an overall accuracy of 95.90% [11]. Ali et al. [12] and Shah et al. [13] used support vectors machines to increase the efficiency of the diagnosis process to select relevant features and predict cardiac disorders. In another ML-based work, 11 different features are extracted from the nonsegmented signals using instantaneous frequency and classification is done through Random Forest with 94.90% accuracy [14]. The combination of recursive feature elimination and genetic algorithm has achieved an accuracy of 86.6% after the selection of relevant feature subset [15]. Artificial neural networks were also equipped with different clinical parameters to find the cardiac abnormality [16].

Data collected from multiple wearable body sensors to measure oxygen saturation, glucose level, cholesterol, temperature, blood pressure, ECG, electromyography (EMG), and electroencephalogram (EEG) with associated medical information of the patient are used for heart disease prediction [17]. An Internet of Things (IoT) and deep learning-based patient monitoring framework for heart patients was proposed to assist in the diagnosis of cardiac disorders, and perspective medication [18]. ElSaadany et al. [19] presented a diagnostic scheme utilizing IoT with a low energy Bluetooth communication module and multiple sensors that gathered data of heart rates along with body temperature. Other such wearable sensors systems and IoT-based healthcare assistive systems for monitoring cardiac health were also designed in multiple works [20], [21], [22]. Different studies analyzed PCG signals and different AI-based techniques were used to develop some cardiac health screening systems [23], [24], [25]. A deep neural network also used to classify significantly class-imbalanced clinical data and crucial features are homogenized by using a fully connected layer. In this two-step approach, the least absolute shrinkage and selection operator and majority-voting were used where overall accuracy of 79.5% was obtained [26].

Auscultation is one of the most popular and traditional means of analyzing cardiac problems. The majority of currently available digital stethoscopes have the ability to record and transfer heart sounds. Similarly, an architecture was proposed for memory constraint mobile devices to diagnose cardiac auscultation using sequence residual and representation learning from fine-grained extracted features from the PCG and attained an accuracy of 86.57% [27]. In a recent work, deep learning with higher order spectral analysis-based approaches is utilized for multiclass classification short heart sound signals [28]. Furthermore, numerous AI algorithms have proven that auscultation data can be characterized as healthy or diseased. One of the few works in the direction to develop an end-to-end product, an AI-powered mobile application was designed that can recognize cardiac abnormalities with approximately 92% accuracy using a stethoscope and mobile but the applicability of work is limited to binary classification [29].

B. Issues With Existing Solutions and Other Challenges

IoT-based systems with multiple body sensors and clinical parameters-based diagnostic methods are not feasible for one-to-one screening in real-world scenarios. Apart from the conventional ML-based techniques, deep learning-based classification using convolutional neural network (CNN) is needed to be explored more. Instead of giving the predictive output into normal and abnormal categories, it would be more beneficial if it can identify different categories of cardiac abnormality. As only software or computer programs were designed in most of the works to analyze the data, there is a strong need to develop an AI-powered end-to-end decision-making system for real-time diagnosis of multiple cardiovascular diseases with high accuracy and robustness, which can not only help medical professionals but can also be utilized for screening of disease in the absence of a doctor in primary health care units in remote places or in rapid mass screening of the cardiac health for a large population with limited medical facilities.

C. Novel Contributions of AI-CardioCare

The main contribution of this article is to develop an end-to-end handheld, automatic, compact-size, portable, standalone, and use-friendly artificial intelligence-based solution for rapid diagnosis of different cardiac disorders. Two deep learning models have been developed for classification of low frequency cardiac sounds and is optimized in different parameters for achieving high accuracy. The deep learning model developed using the 1-D signal is complemented by another deep learning model using the spectrogram of the signal makes the process full-proof and robust. Furthermore, the models are customized in a lightweight computing framework to be integrated in a

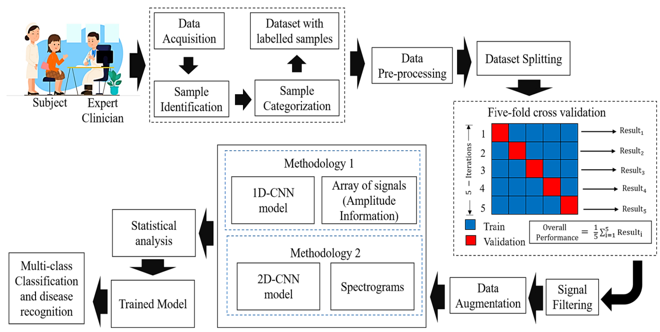

Fig. 2. Block diagram of the proposed methodology.

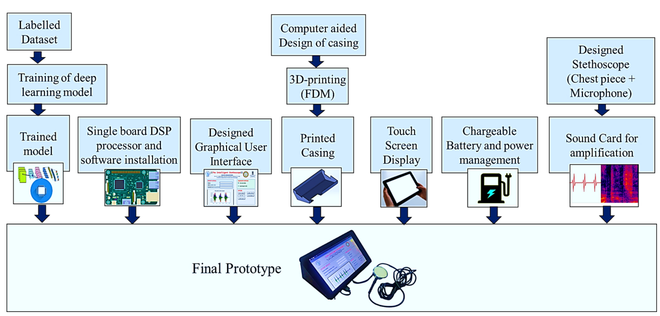

Fig. 3. Development of the final prototype.

low-cost processor to make a portable and cheap device. In the proposed modality, subjects need to record their heart signal with the designed digital stethoscope by making the selection of required options in the user-friendly GUI. The recorded sound is fed to the proposed device and that recoded signal will be processed to a complete report with diagnostic results. These results can be printed and emailed as a record.

Another key contribution is higher prediction accuracy and reduced error rate for identification of four different cardiac disorders. The proposed method has also improved recognition robustness, particularly in noisy situations with signal augmentation techniques to handle different real-world scenarios. The proposed CNN architecture optimizes the multidisease classification task with less computational complexity appropriate for real-time operations. In this method, anyone can do the requisite screening task with minimal guidance instead of a trained workforce.

III. AI-CardioCare: System Design

The methodology behind the development of the proposed AI-CardioCare device is summarized in Fig. 2, where multiple steps have been taken before porting the AI-based module in a device-based modality. The cardiac health data have been captured from multiple experts under the supervision of clinical experts. The recorded cardiac signals were preprocessed and labeled into multiple health categories and different recognized cardiac disease according to the cardiac health of the subjects and opinion of experts.

Then, the final prototype of AI-CardioCare is developed after assembling different hardware and software units in one framework to design a compact portable cardiac health-screening device, as shown in Fig. 3. Different hardware modules such

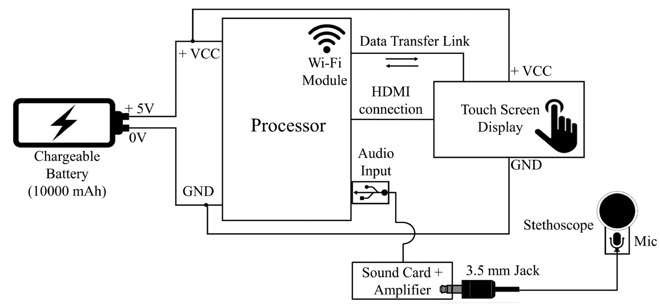

Fig. 4. Circuit diagram of the AI-CardioCare.

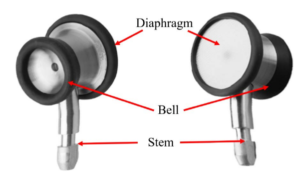

Fig. 5. Chest piece of Stethoscope.

as touch screen display, sound cards, single board computer, and power battery, are assembled together in a as 3-D printed casing for this purpose. Software installation of the necessary libraries, designing of GUI and interfacing of single board DSP processor with other hardware units is also done.

The circuit diagram of the proposed prototype is shown in Fig. 4, where different modules are connected to the processor. The Wi-Fi module can be utilized to access internet or to connect associated printers to email or print report.

A. PCG Sensing Unit

A stethoscope is designed that uses vibrations to pick up heart sounds and transmits them to a processing unit through a microphone. The stethoscope has two separate sound-receiving heads: the bell and the diaphragm, as shown in Fig. 5. Diaphragm is a flat or curved chest component, coated in a film that looks like a drum. When sound waves reach the diaphragm, it vibrates and amplifies the sound, which is then transmitted via the sealed hollow tube into the microphone. The bell has a double cup structure and is made up of stainless steel. Chrome-plated brass plate or aerospace alloy can also be used.

This stethoscope can be used to capture both pediatric and adult auscultations with diaphragm diameter around 32 mm and 44 mm, respectively. Cardiology sensitivity ranges from 3.2–26 dB in a frequency range of 50–1000 Hz. The diaphragm detects high-frequency noises with substantial pressure, whereas the bell detects low-frequency sounds. The bell carries all frequencies adequately, but any secondary low-frequency audio hides the high-frequency murmurs (e.g., aortic regurgitation) in some patients, making detection of the murmuring sound more difficult. The diaphragm does not filter out low-frequencies selectively; instead, all frequencies are attenuated uniformly, lowering the barely discernible low-frequency sounds below the human hearing threshold [30]. The bell of stethoscope should be pushed against the body wall with just enough pressure to form

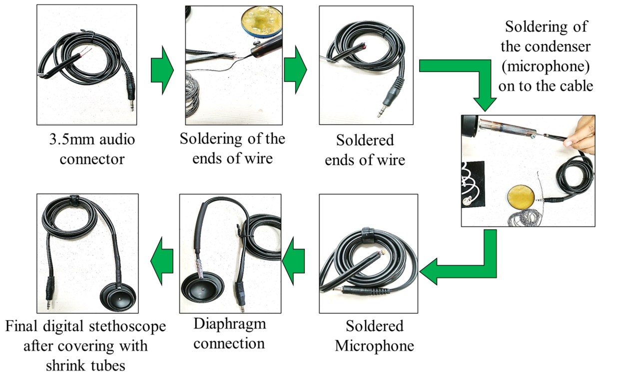

Fig. 6. Digital Stethoscope making: Connection of audio jack connector, microphone and diaphragm.

an air vacuum by excluding ambient noises in order to detect low-frequency signals.

PCG signal acquisition for AI-CardioCare can be carried out by converting the conventional analog stethoscope into a digital one. To accomplish this task, a 3.5 mm audio connector is considered with one end open so as to fix the microphone. Complete operations for making a digital stethoscope are summarized in Fig. 6.

The most common cause of poor acoustic performance is air leakage; even a little air leakage with a radius of only 0.0075 inches can reduce sound transmission by up to 20 dB, especially for frequencies below 100 Hz [31]. Thus, the entire connection is then covered using shrink tubes of varying sizes to ensure that it is not prone to external noise and disturbances.

B. Signal Conditioning Section

This section is composed of three different parts, i.e., filtering, amplification, and analog-to-digital signal conversion. Signal filtering ensures the removal of any noise components from the heart sound signal. A band-pass filter of 10–1000 Hz has been used to transmit the sounds in a given frequency range and record the cardiac sounds. The design consists of a high-pass filter followed by a low-pass filter, where the MCP604 quad operational amplifier has been used with the advantages of high-speed operation and low bias current.

The signal amplification process is carried out to magnify the input signal to yield a notably larger output signal. The amplifier circuit is made up of LM386, an integrated circuit for low voltage audio power amplifiers. It amplifies around 20 times heart sounds in the frequency range of 20–1000 Hz.

The analog heart sound signals recorded using a diaphragm need to be converted into digital format to be fed to the processor for further processing. This operation is carried out by the analog to digital converter. The PCF8591 is an 8-bit CMOS signal acquisition device having four input pins, one output, and a serial I2C-bus interface on a single chip with a single supply. Analog input multiplexing, on-chip track and hold, 8-bit analog-to-digital conversion are among the features of the device. The maximum conversion rate is determined by the I2C-maximum speed of the bus.

TABLE I
SPECIFICATIONS OF SINGLE BOARD DSP PROCESSOR

Name	Specifications
Processor	64 bit 1.5GHz ARM Quad core Cortex-A72 Broadcom BCM2711
RAM	8 GB (LPDDR) SDRAM
Memory	32 GB
Power	5V DC, 2.5A
Connectivity	Wireless IEEE 802.11b/g/n/ac 5.0 GHz and 2.4 GHz, Gigabit Ethernet, 5.0 BLE Bluetooth
Operating Temperature	0-50ºC

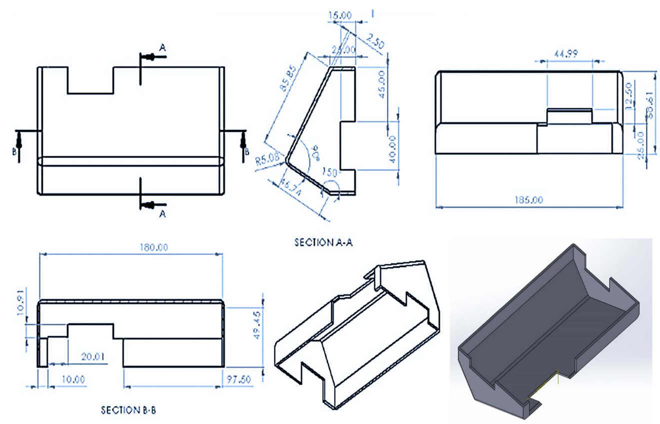

Fig. 7. Details for casing of the prototype and their dimensions (in mm).

C. Processing Unit

The processing unit of a single-board computer is comprised of Broadcom BCM2711 64-bit ARM Quad-core System on a chip having a processing speed of 1.5 GHz. Single board DSP processor supports Bluetooth and 2.4 GHz IEEE 802.11ac wireless connectivity. Interfacing can be done via different USB ports, micro-HDMI ports. The complete device specifications for AI computing module are given in Table I.

The system runs on python supporting operating system, connected with the touch screen display to perform different tasks to use the graphical user interface for visualization of the process and utilizing the predictive diagnosis facility of the proposed device. Display serial interface standard is utilized to allow high-speed communication between LCD screens having 10-point capacitive touch functionality.

D. Mechanical Design

The mechanical design of the AI-CardioCare is made according to the desired requirements of placing each component at a suitable position for making a compact and portable model. The CAD design is built in SOLIDWORKS software with appropriate dimensions after a number of iterations. The final CAD file is then saved as .SLDPRT and converted to .STL format and fed to the 3-D printer. The final casing is 3-D printed using MAKERBOT 3-D printer based on fused deposition modeling technology. The top face of the casing is extruded and left hollow for the purpose of fitting in the LCD display. The design dimensions of front, left, and right faces and slots are also given in the CAD file, as shown in Fig. 7. The rear face is given a bend at a distance of 46.74 mm about an angle of 150° for a convenient view and access to the touchscreen. The bottom face is the base of the device. All the other components are placed inside the casing below the screen.

E. AI Module

The AI-CardioCare device for the automatic diagnosis of CVDs works in two modes. One is based on 1-D classification of heart sounds i.e., raw PCG signals while the other is based on 2-D classification of spectrogram of the given/recorded heart sound signals.

The cardiac sounds in the dataset have been classified into two primary categories based on PCG signals: first having recordings normal healthy and the second containing recordings from subjects suffering from four distinct types of major CVDs i.e., aortic stenosis (AS), mitral regurgitation (MR), mitral valve prolapse (MVP), and mitral stenosis (MS), comprising five distinct classes [32]. The dataset is in .wav format and contains 1000 audio samples, 200 samples each class, and only one channel with 16 bits per sample. It features a sampling rate of 8000 Hz and a bit rate of 128 kb/s. The dataset used for training has high interclass and intraclass diversity and data samples are acquired from subjects belonging to different age group and gender with high variation in signals in terms of time, amplitude, and intensity. Noise injection is done in collected data to get efficient performance because actual cardiac signals collected by clinicians may contain noise in real-world settings. Hence, background deformations chosen at random within a frequency range of 1000 Hz and applied directly to raw PCG signals. For a signal represented by $S = [s_1, s_2, s_3, \ldots, s_n]$, having n time instances and s_i is the amplitude in those time instances for $i = [1, 2, 3, \ldots, n]$. The background deformation to add in the given signal represented as $\eta = (\eta_1, \eta_2, \eta_3, \ldots, \eta_n)$ having same length n as that of input signal. The resultant signal Υ is represented and their relation is given as

$$\Upsilon = S + \eta * \sigma \tag{1}$$

where, η ranges from 0 to 1 and σ is the control parameter is taken 1000 to keep final signal in frequency range of 1000 Hz.

The set of newly generated signals with random background noises are mixed with the raw signals to produce augmented dataset with increased size of the training set for extraction of the most discriminatory features of cardiac sounds and their authentic categorization through a deep neural network. Also, training in noisy environment increased the robustness of the model. After augmentation, final augmented dataset is composed of 2000 signals having 400 signals in each category. Thus, new version of the dataset for 5-fold cross-validation using the background deformation approach is presented in this article for validating the performance in noisy environments.

1) Diagnostic Framework for Analyzing PCG Signals Using 1-D CNN:: Automated diagnosis of the cardiac disorders is a challenging problem due to different issues such as background noises with high intensity and substantial variations in those sounds. As a result, processing these data is required in order to ready a raw PCG signal for training a CNN model [25]. This aids CNN in identifying significant and discriminating characteristics that may be used to differentiate between various cardiac issues. While acquiring PCG signals with electronic stethoscope, different noises, and other artefacts can also be recorded, which must be eliminated in order to properly identify cardiac issues. As a result, the amplitude of given signal and time length may be adjusted to various levels. Thus, all signals are subjected to 16-bit amplitude normalization, and signal duration lengths of up to 2.5 s, which have been converted to a frequency of 8 KHz, resulting in 20000 data points. Background noise, such as high-frequency noise, is typically present in recorded PCG signals. As the heart beats at a frequency of 20–150 Hz, frequencies higher than 150 Hz may be readily eliminated. Gaussian butterworth filter with high-cut at 20 Hz and low-cut at 150 Hz has been employed for the aforementioned purpose. Fig. 8 represents the normalized amplitude of PCG signals after pre-processing for different categories with respect to time.

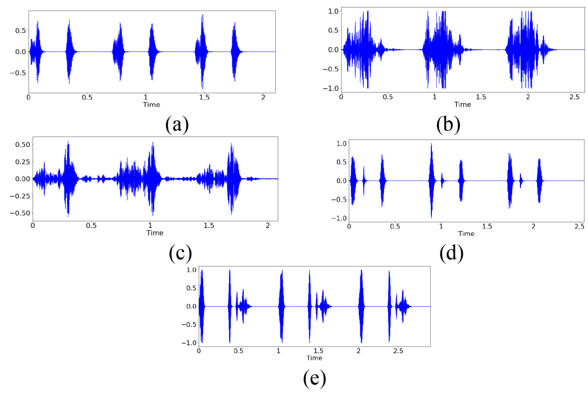

Fig. 8. Sample PCG signals after pre-processing. (a) Healthy. (b) Aortic Stenosis. (c) Mitral Regurgitation, (d) Mitral valve prolapse. (e) Mitral Stenosis.

The CNN makes use of its ability to share features or characteristics and reduce dimensionality. As a result, the number of parameters is minimized, as is the computation complexity. Here, a dataset of cardiac sounds with various amplitudes and fixed time instance is used to train the 1-D CNN for cardiac health prediction, which consists of multiple convolutional layers accompanied by few more dense layers. The data are transmitted from one layer to the next, with low-level features extracted in the first layer and more abstract information processed in the deeper layers. Fig. 9 depicts the whole architecture of the 1-D-CNN architecture. Following these convolutional layers are dense layers. Rectified linear unit (ReLU), an unsaturated nonlinear activation function, is employed to construct the proposed 1-D-CNN architecture to speed up the training process and improve accuracy because it performs better than saturated nonlinear functions like Tanh and Sigmoid. To limit the amount of network parameters, max-pooling operations are used across the region in different stages.

2) Diagnostic Framework for Analyzing PCG Signals Using 2-D CNN:: In another approach, power spectrogram of recorded cardiac signals was utilized to develop a 2-D-classification network. The power spectral density estimates the power of the signal over frequency and then, signal spectrograms have been developed where small window has been analyzed for longer time duration and plotted with respect to the time

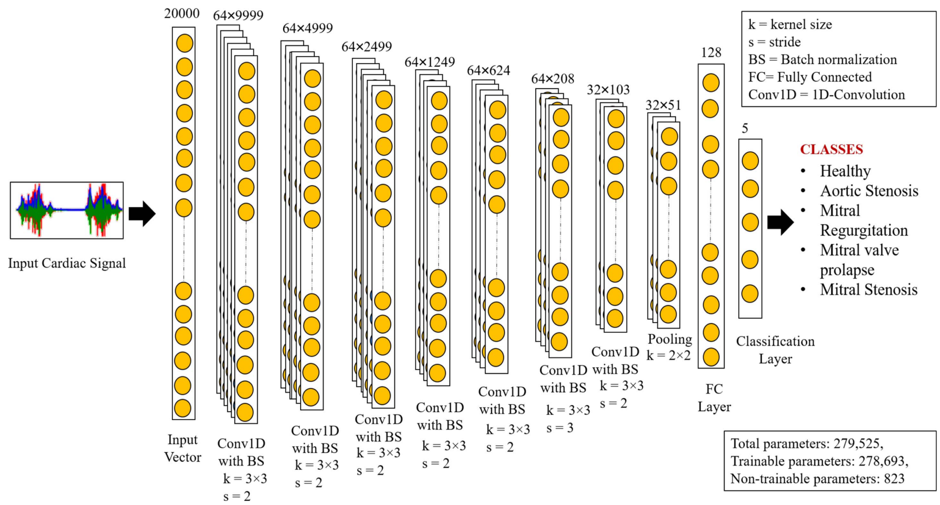

Fig. 9. CNN architecture for diagnosis of cardiac disease.

corresponding to that window [33]. Raw 1-D cardiac signals have been converted into power spectrograms in a stepwise manner with the goal of transforming the entire dataset including augmented dataset, into a dataset of spectrogram images [34]. For the time-frequency analysis of audio signals, a short-time Fourier transform (STFT) approach called Power spectrogram is utilized, with mono-audio clips feeding the algorithm as batch input. STFT allows to perform time-frequency analysis. It is used to generate representations that capture both the local time and frequency content in the signal [35]. STFT has better temporal and frequency localization properties compared with other transforms such as the Fourier transform [36]. STFT or power spectrogram can also give time-localized information about the energy content of each heartbeat. The main advantages of the STFT in the case of cardiac signal analysis is that it is possible to know the occurrence of the different frequencies in time windows and visualize nondeterministic energy in the heartbeat, which can be prominent features for deep learning methods for automatic classification [37].

The generation of STFT of a signal that changes over time requires the use of a window function to divide an elongated version of the time-varying signal into equally tiny sized portions. Each portion is then subjected to the Fourier transform. First, a 44100 Hz sampling rate raw audio file of a cardiac signal is loaded. Because a normal conversation is roughly at 60 dB, a threshold level of 60 dB is chosen to eliminate extraneous sounds. The signal is then normalized to 44100 data points by clipping or padding as needed, depending on the signal length.

The fixed normalized array was then converted into a complex-valued matrix that represents the STFT matrix, with the FFT window size set to 2048 and the hop length set to 812. The discrete STFT is represented as follows:

$$X\,(m,\,\omega)\;=\;\sum_{n\,=\,-\infty}^{\infty}\;x\,[n]\,w\,[n\;-\;mR]\,e^{-j\omega m} \qquad (2)$$

where x[n] is the input signal, w[n] is the window function of length m, X(m,ω) is the DTFT of windowed data centered about time mR, and R is the size of sample hop between consecutive DTFTs.

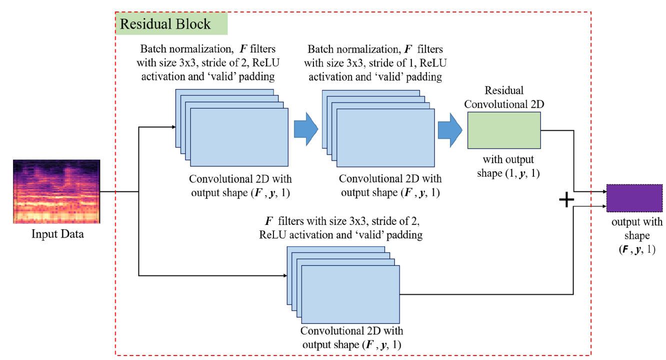

Fig. 10. Residual block used in the proposed deep learning architecture.

Then, the STFT is converted into dB-scaled STFT. Hann and Hamming window functions are widely utilized window functions in STFT. Hann window is best suited for the intended job since it has fewer side lobes and less leakage than other windowing techniques. The dimension of the output single-channel spectrogram image is $1025\times120\times1$ pixels. Various feature filters are used by the model to find local patterns in an image. The feature extraction section of the proposed 2-D-CNN architecture for processing spectrogram of cardiac signals is a combination of a convolutional layers, max-pooling layers, and batch normalization with suitable hyperparameters while the second section is a fully-connected neural network, which classifies the extracted features into different categories.

Although CNNs are efficient, but the problem of vanishing gradient emerges while updating the weights. This can be handled by employing skip connections, as seen in ResNet, and utilizing residual blocks [38]. The architecture 2-D CNN for diagnosis of cardiac signals using spectrograms in this article utilizes different skip connections. The residual block of the proposed architecture has been shown in Fig. 10, showing different skip connections and convolutional blocks that were

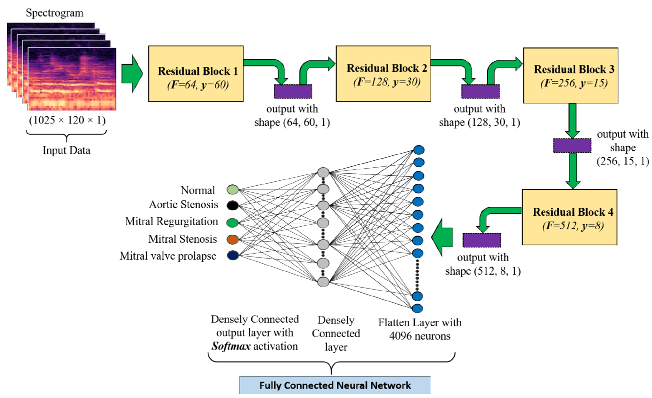

Fig. 11. Layered architecture of the proposed deep learning model containing four residual blocks and one fully connected neural network.

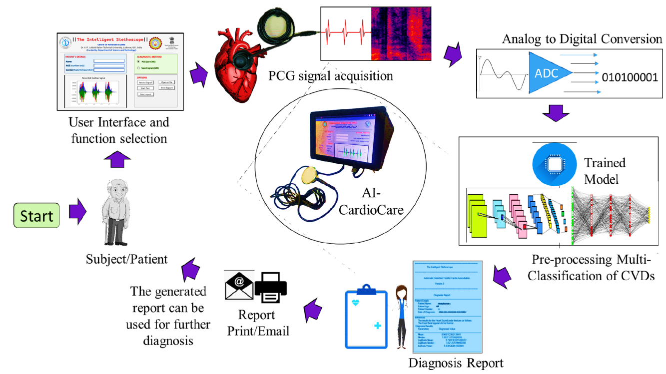

Fig. 12. Diagrammatic representation of the implementation of AI-CardioCare.

TABLE II
HYPERPARAMETERS USED IN TRAINING OF 2-D-CNN MODEL

Hyperparameter	Value
Batch size	32
Early Stopping criteria	Validation accuracy
Activation function	Softmax and ReLU
Padding	valid
Optimization Algorithm	Adadelta, Adam
Number of Epochs	500
Learning Rate	0.001
Loss	Categorical cross-entropy

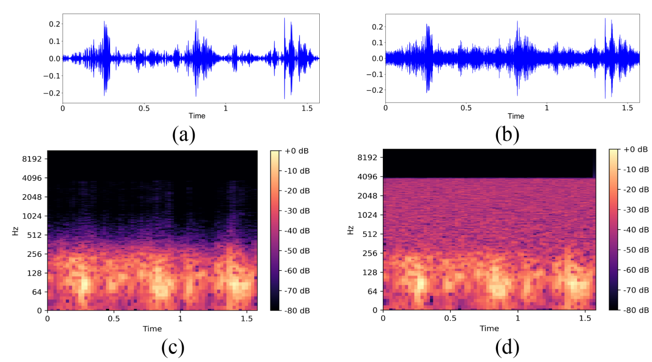

Fig. 13. Augmentation techniques. (a) Raw signal. (b) Augmented signal. (c) Raw power spectrogram. (d) Augmented power spectrogram.

used multiple times in the architecture of proposed deep learning model. It permits the gradient to flow along a second shortcut path in deep neural networks, resolving the issue of vanishing gradient.

The complete architecture has been shown in Fig. 11, where data has been processed through multiple number of layers for proper categorization. The proposed 2-D-CNN architecture for diagnosis of cardiac signals consists of 43 layers including batch normalization, activation, convolution, flattening, and dense layers.

In order to make the network efficient, responsive, and robust, the batch normalization technique is utilized for regularization to solve a major problem of internal covariate shift. Hyperparameters used for the development of the proposed CNN architecture are given in Table II. The last layer of the densely connected network exploits the softmax activation function to enable the normalization of the outputs into probabilities of the envisaged five classes.

F. AI-CardioCare Implementation

The diagrammatic representation of the implementation of AI-CardioCare has been shown in Fig. 12. It begins with switching on the AI-CardioCare device and placing the chest-piece of the stethoscope in the desired position for clear heart sounds of the patient's chest in order to acquire PCG signals.

The captured PCG sound signal is then processed within the device, in the processing unit, in a series of sequential steps. First, signal filtering is carried out to remove any noise components in the signal that is followed by signal amplification resulting in a magnification of the signal for a better yield and analog to digital conversion for further processing. The signal is then fed to the deep learning model for prediction of the normalcy of heart rate and otherwise types of CVDs by executing signal classification. The AI-CardioCare device also generates a report for further diagnosis to enable careful perusal by a medical professional.

IV. EXPERIMENTAL RESULTS

A. Experimental Configuration and Performance Metrices

The original cardiac signal dataset has 1000 audio files that are divided into five categories, i.e., Normal, AS, MR, MS, and MVP; each of which contains 200 files. To begin, the model is trained and tested using a 1000-file original dataset. The same model is trained and tested using an augmented dataset that contains a total of 2000 files, with 400 files in each category. All audio files in the .wav format are transformed into an array of signal amplitudes to train a 1-D-CNN architecture and then into power spectrogram images with dimensions of $1025 \times 120 \times 1$ pixels to process with 2-D CNN. The signals before and after augmentation are shown in Fig. 13. Then, five-fold cross-validation has been done where the data spilt into two subsets i.e., training and test set. The performance of the trained models has been accessed in each iteration on a different test set in each fold of cross-validation.

TABLE III
VARIOUS PARAMETERS, HARDWARE, AND SOFTWARE SPECIFICATIONS

Attribute Name	**Parameters Specification**
Operating System	64-bit Ubuntu
Processor	Two Intel Xenon Platinum CPUs
RAM	64GB
Graphics Server	NVIDIA DGX-II with 16 V100 GPUs of 32 GB graphics memory
Development Environment	Python. Tensorflow, Keras. librosa, Jupyter Notebook
Input Dataset	Power Spectrogram images PCG dataset
Batch Size	32
Image dimensions	$1025 \times 120 \times 1$ pixels

The accuracy, specificity, sensitivity (recall), precision, and F-1 score of the trained deep learning model are all evaluated using five-fold cross-validation. The model is tested with unseen data samples of cardiac sounds, which are labeled by human physicians. There is no overlapping between the training and testing (unseen) set during five-fold cross-validation. These performance matrices are computed from the resulted confusion matrices for both raw and augmented data using false positive (FP), true positive (TP), false negative (FN), and true negative (TN).

An exhaustive analysis has been performed using the performance matrix of the trained deep learning models using different methodologies and data configuration in order to make the model robust. The performance of the trained model is evaluated on different matrix components such as accuracy, recall (or sensitivity), specificity, and F1-score. The complete list of parameters, software, and hardware related information is mentioned in the Table III.

B. Detailed Results

All the data samples have been resized to the same dimensions, which reduces the potential ambiguity and data acquisition error. Table IV represents the compiled results for all five-folds of cross-validation for trained deep neural networks with raw and augmented data of cardiac sounds. The results are demonstrated into two parts i.e., 1-D CNN using raw PCG signals and 2-D CNN using spectrograms of those signals for both raw data as well as augmented data. Outcomes signifies high accuracy with higher values of TP and TN, whereas FP and FN got lower values.

The data augmentation and addition of noise were done for authentic categorization with low FP and FN rates. Signal acquisition in controlled conditions with some analysis of past medical history can be done to reduce false negative in clinical scenarios. Even if there are FP or FN, these rates can be reduced further after taking potential misclassified samples to retrain the deep learning model for enhancing the robustness while testing the device in real-world scenarios. Also, the proposed prototype is a screening device only and if the subject is identified as positive case with suspected disease, the subject will be recommended to medical professionals to assess the impact or grading of disease and consequent treatments.

Four confusion matrices for 1-D CNN and 2-D CNN with raw data and augmented data have been shown in the Fig. 14, where

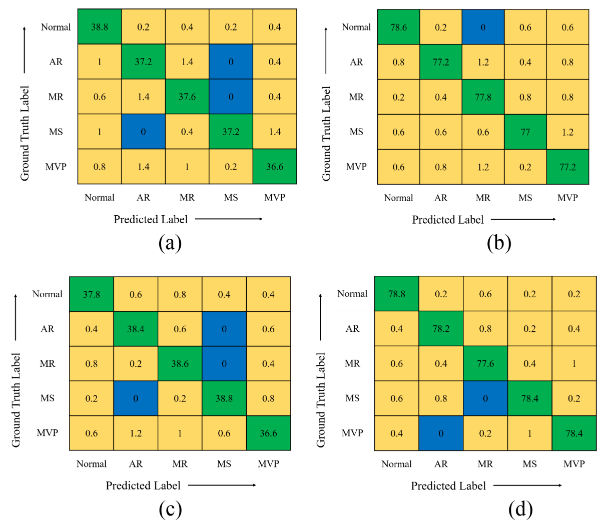

Fig. 14. Confusion matrices. (a) One-dimensional CNN with raw data. (b) 1-D CNN with augmentation. (c) Two-dimensional CNN with raw data. (d) 2-D CNN with augmentation.

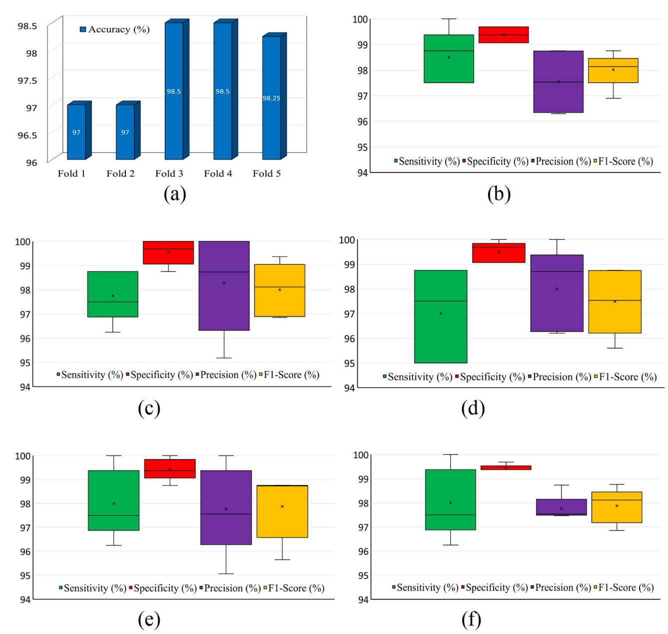

Fig. 15. (a) Accuracy in different folds of cross validation. (b) Box-plot Normal. (c) Box-plot AR. (d) Box-plot MR. (e) Box-plot MS. (f) Box-plot MVP.

the comparison of obtained multiclassification predictions and actual ground truth labels have been shown after averaging all five-folds of cross-validation.

The accuracy in different folds of cross-validation and categorized performance matrices of five different categories used in the proposed work including four abnormal categories, as shown in Fig. 15. The results are also compared with other approaches in Table V, which demonstrates superiority of the proposed method with respect to state-of-the-art methods. Overall, spectrogram-based analysis of cardiac signals results in superior results than processing the 1-D array of amplitudes of PCG signals using different deep learning architectures. More

TABLE IV
COMPILED AVERAGE FIVE-FOLD CROSS VALIDATION RESULTS USING DIFFERENT METHODOLOGIES

Method	Data	Class	TP	TN	FP	FN	Sensitivity (%)	Specificity (%)	Precision (%)	F1-Score (%)	Average Accuracy (%)
1D-CNN	Raw	Normal	38.80	156.60	3.40	1.20	97.00	97.88	91.94	94.40	93.70
		AR	37.20	157.00	3.00	2.80	93.00	98.13	92.54	92.77	
		MR	37.60	156.80	3.20	2.40	94.00	98.00	92.16	93.07	
		MS	37.20	159.60	0.40	2.80	93.00	99.75	98.94	95.88	
		MVP	36.60	157.40	2.60	3.40	91.50	98.38	93.37	92.42	
	Augmented	Normal	78.60	317.80	2.20	1.40	98.25	99.31	97.28	97.76	96.95
		AR	77.20	318.00	2.00	2.80	96.50	99.38	97.47	96.98	
		MR	77.80	317.00	3.00	2.20	97.25	99.06	96.29	96.77	
		MS	77.00	318.00	2.00	3.00	96.25	99.38	97.47	96.86	
		MVP	77.20	316.60	3.40	2.80	96.50	98.94	95.78	96.14	
2D-CNN	Raw	Normal	37.80	158.00	2.00	2.20	94.50	98.75	94.97	94.74	95.10
		AR	38.40	158.00	2.00	1.60	96.00	98.75	95.05	95.52	
		MR	38.60	157.40	2.60	1.40	96.50	98.38	93.69	95.07	
		MS	38.80	159.00	1.00	1.20	97.00	99.38	97.49	97.24	
		MVP	36.60	157.80	2.20	3.40	91.50	98.63	94.33	92.89	
	Augmented	Normal	78.80	318.00	2.00	1.20	98.50	99.38	97.52	98.01	97.85
		AR	78.20	318.60	1.40	1.80	97.75	99.56	98.24	97.99	
		MR	77.60	318.40	1.60	2.40	97.00	99.50	97.98	97.49	
		MS	78.40	318.20	1.80	1.60	98.00	99.44	97.76	97.88	
		MVP	78.40	318.20	1.80	1.60	98.00	99.44	97.76	97.88	

TABLE V
COMPARISON OF THE PROPOSED APPROACH WITH OTHER METHODS

Data	Method	Avg. Sensitivity (%)	Avg. Specificity (%)	Avg. Precision (%)	Avg. F1-Score (%)	Avg. Accuracy (%)
1-D	SVM	82.88	91.56	71.06	76.51	90.45
	Random Forest	89.78	94.75	81.04	85.19	91.80
	K-nearest Neighbour	93.20	96.77	87.83	90.44	93.95
	Proposed method	96.95	99.21	96.85	96.90	96.95
2-D	VGG-16	94.20	95.46	83.83	88.71	94.10
	Mobile-Net	95.35	97.98	92.19	93.74	95.48
	Inception-Net	93.53	95.31	83.28	88.11	93.15
	Proposed method	97.85	99.46	97.85	97.85	97.85

TABLE VI
PERFORMANCE ASSESSMENT ON TEST DATA

Category	Samples	1D-CNN		2D-CNN (Spectrogram)	
		F1-Score (%)	Accuracy (%)	F1-Score (%)	Accuracy (%)
Normal	45	95.45	93.33	96.63	95.56
AR	43	94.25	95.35	96.55	97.67
MR	37	94.74	97.30	98.67	100.00
MS	41	96.30	95.12	96.30	95.12
MVP	39	94.87	94.87	97.44	97.44
Average		95.12	95.19	97.12	97.16

training data in CNN leads to increased accuracies, as evidenced by the resulting confusion matrices, which show high accuracy with a large dataset.

Additionally, the efficacy and robustness of the developed device are tested in real-world scenarios on 205 subjects using both 1-D-CNN and 2-D-CNN approaches for acquired cardiac signals. Table VI shows the analysis of different categories of signals in terms of accuracy and F1-score. Spectrogram based 2-D-CNN approach obtained 97.16% average accuracy on these testing signals. Both 1-D and 2-D CNN models work independently that enables the user multiple analysis of cardiac signals with single handheld device without requiring additional hardware. Subjects can double-check the results with more affirmation using two completely independent methods.

C. *Computation Time*

The AI-CardioCare device acquires a 2.5 s PCG signal as input using a stethoscope. The average processing time to process a single recorded signal is 0.25 s using the 1-D-CNN technique and 0.32 s for the spectrogram based 2-D-CNN technique, including uploading time at a rate of 150 kb/s. A conversion time to convert and save a raw signal to a power spectrogram takes an average time of 0.1127 s. Thus, a total time of 2.75 s was taken with 1-D-CNN approach and 2.9327 s using 2-D-CNN approach for each cardiac signal.

D. *Developed AI Device for Cardiac Health Screening*

After training the deep neural networks for multiple iterations on given dataset, trained AI-models were ported to low-cost single-board computer for development of standalone device for having better usability in real-world situations. The proposed device consists of an electronic stethoscope (chest piece and microphone); signal preprocessing unit, DSP processor and touch screen having GUI interface for accessibility of different

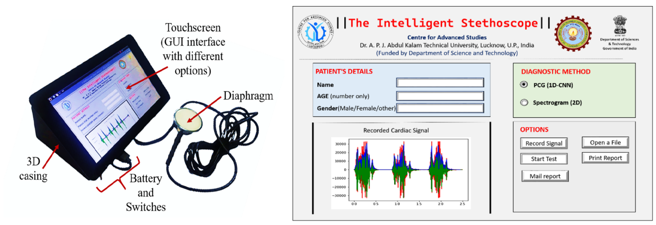

Fig. 16. (a) Developed prototype of AI-CardioCare in the lab. (b) Designed GUI for the screening device- AI-CardioCare.

functions. The developed prototype for the cardiac health assessment device is shown in Fig. 16(a). Different hardware components including battery and switches are embedded in the 3-D printed device to have a compact device and portability. A Python based user-friendly GUI application has been built for monitoring and analysis of generated reports, which can be used by any regular individual (or operator) with minimal guidance. The trained AI-model has been ported in the hardware setup with low-cost single board computer to use the designed end-to-end modality in real-world scenarios. Fig. 16(b) exhibits the GUI of the developed device for the automated cardiac health monitoring. GUI is divided into four sections i.e., patient's details, diagnostic method, options, and report visualization.

Name and age of the patient will be recorded in "patient details" section to be used in generation of a diagnostic report. Operator can choose any of the two methodologies i.e., 1-D or 2-D CNN though radio button in "diagnostic method" section. The report visualization section shows the graphs of the direct-recorded PCG signal through designed stethoscope (through record signal button) or any pastrecorded cardiac signals (through "open a file" button in "options" menu). These graphs will be visualized by pressing "start test" button after selecting or recording a PCG signal and then signal will be analyzed with deep learning models to get predictive diagnosis of cardiac health. Options such as "mail report" and "print report" can be used if the device has an active internet connection with a Wi-Fi module in common network interfacing.

V. Conclusion

This article presented an AI-based embedded device to classify and recognize the PCG of subjects as a smart healthcare system. The device was automatic and fast to recognize the potential subjects having cardiac abnormality with their possible categorization. The device was capable of identifying normal and abnormal heart conditions with four major types of diseases. The device was compact, made in 3-D printed case, and is installed the necessary processing system, software, and hardware for the proper functioning of the device. The 1-D CNN and pectrogram based 2-D CNN based approaches were analyzed with five-fold cross-validation and accuracy of approximately 96.95% and 97.85% was achieved through both methods, respectively. This article presented here has great potential of benefiting people having cardiac disorders as an efficient solution for early diagnosis. As the device holds promise for improving the initial screening mechanism at Primary Health Care centers, the research proposed will be commercialized after taking necessary approvals from respective authorities.

The future research is towards the development of AI-based single devices having the capability of identifying multiple diseases related to cardiac, pulmonary, and prenatal health and their multiclass or multigrade classification. Health-related data and values obtained from different pathological tests could be integrated together with the existing mechanism to increase the reliability of the obtained results.

References

[1] Overview: Cardiovascular diseases, Jun. 11, 2021. Accessed: Dec. 15, 2021. [Online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
[2] S. S. Virani et al., "Heart disease and stroke statistics—2020 update: A report from the American heart association," *Circulation*, vol. 141, no. 9, pp. e139–e596, Mar. 2020, doi: 10.1161/CIR.0000000000000757.
[3] J. Bostrom, G. Sweeney, J. Whiteson, and J. A. Dodson, "Mobile health and cardiac rehabilitation in older adults," *Clin. Cardiol.*, vol. 43, no. 2, pp. 118–126, Feb. 2020, doi: 10.1002/clc.23306.
[4] G. A. Roth et al., "Global burden of cardiovascular diseases and risk factors, 1990–2019," *J. Amer. College Cardiol.*, vol. 76, no. 25, pp. 2982–3021, Dec. 2020, doi: 10.1016/j.jacc.2020.11.010.
[5] M. Kupari, "Aortic valve closure and cardiac vibrations in the genesis of the second heart sound," *Amer. J. Cardiol.*, vol. 52, no. 1, pp. 152–154, Jul. 1983, doi: 10.1016/0002-9149(83)90086-3.
[6] S. B. Shuvo, S. N. Ali, S. I. Swapnil, T. Hasan, and M. I. H. Bhuiyan, "A lightweight CNN model for detecting respiratory diseases from lung auscultation sounds using EMD-CWT-based hybrid scalogram," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 7, pp. 2595–2603, Jul. 2021, doi: 10.1109/JBHI.2020.3048006.
[7] B. Keerthi Samhitha, M. Sarika Priya., C. Sanjana., S. C. Mana, and J. Jose, "Improving the accuracy in prediction of heart disease using machine learning algorithms," in *Proc. Int. Conf. Commun. Signal Process.*, Jul. 2020, pp. 1326–1330, doi: 10.1109/ICCSP48568.2020.9182303.
[8] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–11, Jul. 2021, doi: 10.1155/2021/8387680.
[9] R. Spencer, F. Thabtah, N. Abdelhamid, and M. Thompson, "Exploring feature selection and classification methods for predicting heart disease," *Digit. Health*, vol. 6, Jan. 2020, Art. no. 205520762091477, doi: 10.1177/2055207620914777.
[10] K. Balaji, K. Lavanya, and A. G. Mary, "Machine learning algorithm for clustering of heart disease and chemoinformatics datasets," *Comput. Chem. Eng.*, vol. 143, Dec. 2020, Art. no. 107068, doi: 10.1016/j.compchemeng.2020.107068.
[11] N. L. Fitriyani, M. Syafrudin, G. Alfian, and J. Rhee, "HDPM: An effective heart disease prediction model for a clinical decision support system," *IEEE Access*, vol. 8, pp. 133034–133050, 2020, doi: 10.1109/ACCESS.2020.3010511.
[12] L. Ali et al., "An optimized stacked support vector machines based expert system for the effective prediction of heart failure," *IEEE Access*, vol. 7, pp. 54007–54014, 2019, doi: 10.1109/ACCESS.2019.2909969.
[13] S. M. S. Shah, F. A. Shah, S. A. Hussain, and S. Batool, "Support vector Machines-based heart disease diagnosis using feature subset, wrapping selection and extraction methods," *Comput. Elect. Eng.*, vol. 84, Jun. 2020, Art. no. 106628, doi: 10.1016/j.compeleceng.2020.106628.
[14] A. M. Alqudah, "Towards classifying non-segmented heart sound records using instantaneous frequency based features," *J. Med. Eng. Technol.*, vol. 43, pp. 418–430, 2019, doi: 10.1080/03091902.2019.1688408.
[15] P. Rani, R. Kumar, N. M. O. S. Ahmed, and A. Jain, "A decision support system for heart disease prediction based upon machine learning," *J. Reliable Intell. Environments*, vol. 7, no. 3, pp. 263–275, Sep. 2021, doi: 10.1007/s40860-021-00133-6.
[16] J. Jeyaranjani, T. D. Rajkumar, and T. A. Kumar, "Coronary heart disease diagnosis using the efficient ANN model," *Mater. Today Proc.*, pp. 1–5, Mar. 2021, doi: 10.1016/j.matpr.2021.01.257.
[17] F. Ali et al., "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion," *Inf. Fusion*, vol. 63, pp. 208–222, Nov. 2020, doi: 10.1016/j.inffus.2020.06.008.

[18] S. S. Sarmah, "An efficient IoT-based patient monitoring and heart disease prediction system using deep learning modified neural network," *IEEE Access*, vol. 8, pp. 135784–135797, 2020, doi: 10.1109/ACCESS.2020.3007561.

[19] Y. ElSaadany, A. J. A. Majumder, and D. R. Ucci, "A wireless early prediction system of cardiac arrest through IoT," in *Proc. IEEE 41st Annu. Comput. Softw. Appl. Conf.*, Jul. 2017, pp. 690–695, doi: 10.1109/COMPSAC.2017.40.

[20] J. Lin et al., "Wearable sensors and devices for real-time cardiovascular disease monitoring," *Cell Rep. Phys. Sci.*, vol. 2, no. 8, Aug. 2021, Art. no. 100541, doi: 10.1016/j.xcrp.2021.100541.

[21] M. A. Khan and F. Algarni, "A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO-ANFIS," *IEEE Access*, vol. 8, pp. 122259–122269, 2020, doi: 10.1109/ACCESS.2020.3006424.

[22] A. S. Albahri et al., "Development of IoT-based mhealth framework for various cases of heart disease patients," *Health Technol. (Berl).*, vol. 11, no. 5, pp. 1013–1033, Sep. 2021, doi: 10.1007/s12553-021-00579-x.

[23] A. Yadav, A. Singh, M. K. Dutta, and C. M. Travieso, "Machine learning-based classification of cardiac diseases from PCG recorded heart sounds," *Neural Comput. Appl.*, vol. 32, no. 24, pp. 17843–17856, Dec. 2020, doi: 10.1007/s00521-019-04547-5.

[24] J. Wang et al., "Intelligent diagnosis of heart murmurs in children with congenital heart disease," *J. Healthcare Eng.*, vol. 2020, pp. 1–9, 2020, doi: 10.1155/2020/9640821.

[25] N. Baghel, M. K. Dutta, and R. Burget, "Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network," *Comput. Methods Programs Biomed.*, vol. 197, 2020, Art. no. 105750, doi: 10.1016/j.cmpb.2020.105750.

[26] A. Dutta, T. Batabyal, M. Basu, and S. T. Acton, "An efficient convolutional neural network for coronary heart disease prediction," *Expert Syst. Appl.*, vol. 159, 2020, Art. no. 113408, 2020, doi: 10.1016/j.eswa.2020.113408.

[27] S. B. Shuvo, S. N. Ali, S. I. Swapnil, M. S. Al-Rakhami, and A. Gumaei, "CardioXNet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings," *IEEE Access*, vol. 9, pp. 36955–36967, 2021, doi: 10.1109/ACCESS.2021.3063129.

[28] A. M. Alqudah, H. Alquran, and I. A. Qasmieh, "Classification of heart sound short records using bispectrum analysis approach images and deep learning," *Netw. Model. Anal. Health Inform. Bioinf.*, vol. 9, pp. 1–16, 2020, doi: 10.1007/s13721-020-00272-5.

[29] M. Güven, F. Hardalaç, K. Özışık, and F. Tuna, "Heart diseases diagnose via mobile application," *Appl. Sci.*, vol. 11, no. 5, Mar. 2021, Art. no. 2430, doi: 10.3390/app11052430.

[30] S. McGee, "Auscultation of the Heart," in *Evidence-Based Physical Diagnosis*. New York, NY, USA: Elsevier, 2018, pp. 327–332.

[31] J. R. Kindig, T. P. Beeson, R. W. Campbell, F. Andries, and M. E. Tavel, "Acoustical performance of the stethoscope: A comparative analysis," *American Heart J.*, vol. 104, no. 2, pp. 269–275, Aug. 1982, doi: 10.1016/0002-8703(82)90203-4.

[32] G.-Y. S. Yaseen and S. Kwon, "Classification of heart sound signal using multiple features," *Appl. Sci.*, vol. 8, no. 12, Nov. 2018, Art. no. 2344, doi: 10.3390/app8122344.

[33] G. Kłosowski, T. Rymarczyk, D. Wójcik, S. Skowron, T. Cieplak, and P. Adamkiewicz, "The use of time-frequency moments as inputs of LSTM network for ECG signal classification," *Electronics*, vol. 9, no. 9, Sep. 2020, Art. no. 1452, doi: 10.3390/electronics9091452.

[34] B. McFee et al., "Librosa: Audio and music signal analysis in python," in *Proc. 14th Python Sci. Conf.*, 2015, pp. 18–25, doi: 10.25080/Majora-7b98e3ed-003.

[35] A. M. Alqudah, S. Qazan, L. Al-Ebbini, H. Alquran, and I. A. Qasmieh, "ECG heartbeat arrhythmias classification: A comparison study between different types of spectrum representation and convolutional neural networks architectures," *J. Ambient Intell. Humanized Comput.*, vol. 13, pp. 4877–4907, 2021, doi: 10.1007/s12652-021-03247-0.

[36] B. Rajoub, "Characterization of biomedical signals: Feature engineering and extraction," in *Biomedical Signal Processing and Artificial Intelligence in Healthcare*, New York, NY, USA: Academic Press, Elsevier, 2020, pp. 29–50. [Online]. Available: https://doi.org/10.1016/B978-0-12-818946-7.00002-0

[37] V. Millette and N. Baddour, "Signal processing of heart signals for the quantification of non-deterministic events," *Biomed. Eng. Online*, vol. 10, pp. 1–23, 2011, doi: 10.1186/1475-925X-10-10.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Dec. 2015. [Online]. Available: http://arxiv.org/abs/1512.03385