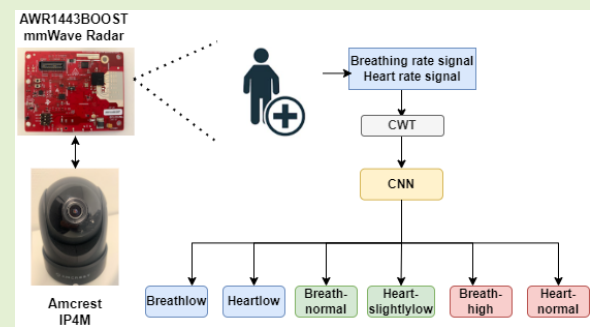


# Automatic Contact-Less Monitoring of Breathing Rate and Heart Rate Utilizing the Fusion of mmWave Radar and Camera Steering System

Khushi Gupta, Srinivas M. B. <sup>ID</sup>, Member, IEEE, Soumya J. <sup>ID</sup>, Om Jee Pandey <sup>ID</sup>, Senior Member, IEEE, and Linga Reddy Cenkeramaddi <sup>ID</sup>, Senior Member, IEEE

**Abstract**—The demand for noncontact breathing and heart rate measurement is increasing. In addition, because of the high demand for medical services and the scarcity of on-site personnel, the measurement process must be automated in unsupervised conditions with high reliability and accuracy. In this article, we propose a novel automated process for measuring breathing rate and heart rate with mmWave radar and classifying these two vital signs with machine learning. A frequency-modulated continuous-wave (FMCW) mmWave radar is integrated with a pan, tilt, and zoom (PTZ) camera to automate camera steering and direct the radar toward the person facing the camera. The obtained signals are then fed into a deep convolutional neural network to classify them into breathing and heart signals that are individually low, normal, and high in combination, yielding six classes. This classification can be used in medical diagnostics by medical personnel. The average classification accuracy obtained is 87% with precision, recall, and an F1 score of 0.93.

**Index Terms**—Breathing rate classification, breathing rate sensor, continuous wavelet transform, deep convolutional neural network (CNN), heart rate classification, heart rate sensor, machine learning, mmWave sensor.



## I. INTRODUCTION

BREATHING and heart rate, blood pressure, and temperature are the four main vital signs that must be accurately measured and classified in order to diagnose a variety of diseases [1]. A respiration rate is the number of breaths taken

Manuscript received 6 September 2022; accepted 25 September 2022. Date of publication 4 October 2022; date of current version 14 November 2022. This work was supported by the Indo-Norwegian Collaboration in Autonomous Cyber-Physical Systems (INCAPS) Project: 287918 of the International Partnerships for Excellent Education, Research and Innovation (INTPART) Program from the Research Council of Norway. The associate editor coordinating the review of this article and approving it for publication was Dr. Ioannis Raptis. (Corresponding author: Linga Reddy Cenkeramaddi.)

Khushi Gupta was with the Autonomous and Cyber-Physical Systems (ACPS) Research Group, Department of Information and Communication Technology, University of Agder, 4879 Grimstad, Norway. She is now with the Columbia University, New York, NY 10027 USA (e-mail: kg3023@columbia.edu).

Linga Reddy Cenkeramaddi is with the Autonomous and Cyber-Physical Systems (ACPS) Research Group, Department of Information and Communication Technology, University of Agder, 4879 Grimstad, Norway (e-mail: khushiag@student.uia.no; linga.cenkeramaddi@uia.no).

Srinivas M. B. and Soumya J. are with the Birla Institute of Technology and Science, Pilani, Hyderabad 500078, India (e-mail: mbs@hyderabad.bits-pilani.ac.in; soumyaj@hyderabad.bits-pilani.ac.in).

Om Jee Pandey is with the Department of Electronics Engineering, Indian Institute of Technology (BHU) Varanasi, Varanasi 221005, India (e-mail: omjee.ece@iitbhu.ac.in).

Digital Object Identifier 10.1109/JSEN.2022.3210256

TABLE I

CATEGORIZATION OF BREATHING RATE [2]

S.No.	Type	BPM(Breaths per minute)
1.	Low	Under 12
2.	Normal	12 - 25
3.	High	Over 25

TABLE II

CATEGORIZATION OF HEART RATE [3]

S.No.	Type	BPM(Beats per minute at rest)
1.	Low	< 90
2.	Slightly Low	90 - 109
3.	Normal	110 - 159
4.	High	160 - 179

in 1 min by a person. Table I shows the respiration rates as normal, low, and high. An abnormal breathing rate is an early sign of an underlying disease that necessitates prompt medical attention [1]. A pulse or heart rate, on the other hand, refers to the number of heartbeats per minute and can be classified as low, normal, slightly low, and high, as shown in Table II.

Monitoring these vital signals is critical in medical diagnosis and treatment. As a result, real-time monitoring of breathing and heart rate is essential. There are numerous methods for measuring them in the literature. However, some lack accuracy or necessitate the presence of trained personnel,

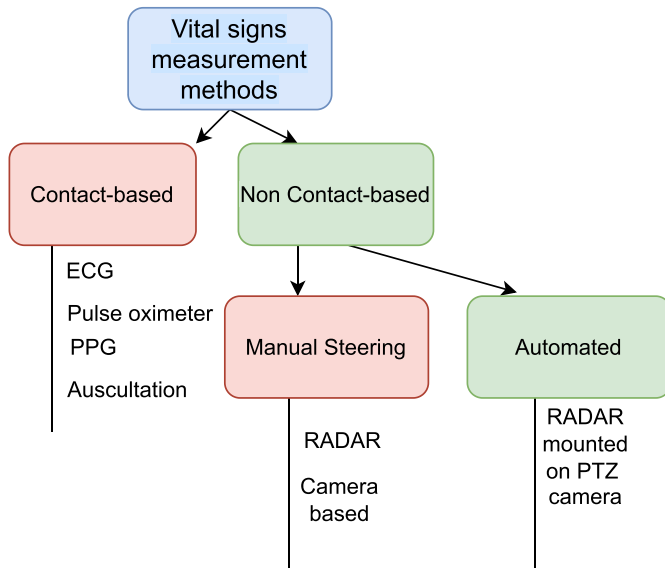


Fig. 1. Existing measurement techniques in the literature and market.

while others become prohibitively expensive [1]. Furthermore, there is a high demand for measuring these vital signs during COVID-19. Active detection and diagnosis of abnormalities become critical, especially if the underlying issue points to a chronic disease, as the situation may worsen over time [1]. As a result, a solution that detects accurately and on time becomes essential. COVID-19 has significantly altered the medical landscape. Assessing these vital signs and detecting abnormalities have long been practiced. Due to the high demand for the medical system during COVID, the presence of trained personnel cannot be guaranteed, emphasizing the importance of such automated and simple-to-use systems. Expanding the reach of medical care by incorporating modern sensors and signal processing methods is a timely solution [1].

The breathing rate is one of the earliest indicators of an underlying disease and is a better indicator than temperature, blood pressure, and pulse [4] for cardiac abnormalities, severe burn, elderly healthcare, sleep monitoring [5], psychological wellness, or any infectious diseases [6] whose presence can be confirmed with abnormalities in vital signs. As a result, its detection has many medical tracking applications where the setup can be used. Radar technology not only ensures non-contact monitoring but raw radar signals can also be processed and more centralized information extracted, which is useful for medical applications, as well as detecting changes in breathing and heartbeat signals to predict disease at an early stage [5].

### A. Existing Solutions

1) **Contact-Based**[7]: Contact Based Monitoring is the traditional method of measuring vital signs which usually involves skin contact of patients to the sensors.

- 1) The majority of contact-based models, as illustrated in Fig. 1, are based on optical or electronic sensors that require skin contact [7].
- 2) Depending on the contact type, the contact methods may be less accurate. For example, a darker shade of skin due to the presence of melamine may cause the oximeter to give a higher value of oxygen. As a result, there could

be a variety of biological factors influencing the contact outcomes [7].

- 3) Contact-based methods frequently necessitate the presence of medical personnel, for example, to measure electrocardiogram (ECG) signals.
- 4) Because of the numerous contacts with the body, patients are frequently uncomfortable.
- 5) Doctors use an auscultation technique to measure breathing rates with a stethoscope. This process is doctor-dependent, which reduces accuracy [7].

As a result of their subjective nature, some contact-based methods may misdiagnose.

2) **Noncontact-Based**: Noncontact vital sign monitoring became popular, owing to the inconvenience caused to patients by contact-based methods, such as skin irritation and a general sense of discomfort.

- 1) Because camera-based systems require adequate light for capture and have high processing costs, radar-based systems are an excellent choice due to their robustness [7].
- 2) Existing solutions, as shown in Fig. 1, use radar systems but require the person to be in the field of view of the radar, which is known as manual steering [8].

In light of these challenges, this article proposes a system that improves the accuracy of detecting vital signs within the required time, employs noncontact architecture, and automates the entire process, thereby eliminating the need for medical personnel. This work's major contributions are given as follows.

- 1) A novel automatic noncontact method of measuring breath and heart rate is proposed, which makes use of mmWave frequency-modulated continuous-wave (FMCW) radar and a pan, tilt, and zoom (PTZ) camera.
- 2) Using the proposed novel setup, a new dataset is created.
- 3) The dataset is preprocessed using the continuous wavelet transform time–frequency analysis technique.
- 4) The measured vital signals are classified into six distinct classes using a lightweight convolutional neural network (CNN): low, normal, and slightly low for breathing and low, normal, and slightly low for heart rate.

The rest of this article is organized as follows. Section II discusses the related work. Section III provides the system overview. Section IV presents the hardware design and description of the components involved in the setup. Section V highlights the algorithms used in the working of the system. Section VI describes how the setup was assembled. Section VII gives the details of the dataset and how continuous wavelet transformation (CWT) was used to preprocess it. Section VIII shows the proposed CNN model used to classify the obtained dataset in six classes and how it was trained. Section IX presents the accuracy, loss, and precision recall, and F1 scores of the model. Finally, the conclusion and possible future works are discussed in Section X.

## II. RELATED WORK

This section provides a review of the related works toward monitoring of breathing and heart rate in order to assist scholars in making future contributions in this field. It also describes how different parameters in medical diagnosis have

been measured and various challenges have been overcome by several researchers.

A BCG monitoring system is used to analyze a person's sleep pattern by monitoring vital signs, such as breathing and heart rate, as well as muscle movement in [9]. This system was later compared to existing market solutions, such as ECG, respiratory inductance plethysmography (RIP), and pulse oximetry. Similarly, a Wi-Fi-based system is used to collect vital signs in order to track sleeping patterns. A novel attempt is made in [10] to use Wi-Fi devices in conjunction with antennas to obtain breathing and heartbeat signals in order to monitor different sleep postures. A noncontact monitoring method to estimate the breathing and heart rate, and detect some related diseases has been proposed in [11]. Different sensors, such as image, infrared, and depth sensors, record video streams in order to analyze the features of face and thorax movements in real time. Later, the video streams from the face and thorax were compared to determine the breathing rate, and the similarity between them was confirmed.

A radar-based monitoring system is used by implementing the algorithm on an ultrawideband (UWB) radar in [12]. In [12], a novel method is proposed to extract heart and breathing information together using peak energy functions and empirical decomposition methods. Vital radio is a wireless technology in which the device will be attached to the walls of houses and can accurately measure breathing and heart rate [13]. To measure vital signs, such as BR and HR, a polymer optical fiber sensor is being developed. The sensor will be worn by the subject, and the signals will be extracted and filtered using various signal processing techniques [14]. Because of the onset of the COVID-19 pandemic, noncontact monitoring of breathing and heart signals has become critical. A contact-free monitoring system is developed in [15], which uses processing techniques, such as peak detection and fast Fourier transforms (FFTs), to classify signals as fast, slow, or normal.

A proximity sensor is used in [16] to measure respiration rate even when the person is unconscious, emphasizing the importance of noncontact methods. A noncontact method for measuring vital signs that capture the person's chest region using a face detection algorithm is proposed in [17]. The proposed algorithm filters the signals using peak-to-peak measurement with an accuracy of 84.66%. A review of how radar can be used for noncontact continuous monitoring of vital signs, as well as the potential consequences of mmWave radar exposure, is given in [18]. A low-cost and portable PPG system is presented to detect cardiorespiratory activity in [19]. Subsequently, a comparison between sensor and ECG readings is calculated, which is less than 3%. Another work focusing on PPG technology, in which the sensor is incorporated in a wristband to monitor heart waveforms during physical activities, is provided in [20]. This work reviews the growing demand for PPG sensings for vital signs monitoring. The comparison of noncontact and contact-based monitoring, highlighting the benefits and drawbacks of both by stating that noncontact methods are more comfortable but contact-based methods are less complex, is given in [21]. A review of contact-less methods while keeping respiration frequency in

mind and explaining where the use of contact-less methods is superior to contact methods is provided in [22].

A seismocardiogram-based sensor that measures breathing state and categorizes it as normal, long and labored, or breathless is highlighted in [23]. The reflected SCG signals are used to obtain chest wall vibrations, and the SCG signal is projected using ECG signals. The performance is then assessed using classifiers for 1147 SCG beats, yielding an accuracy of around 91% [23]. The study demonstrates how the use of contact-based methods can make subjects feel uneasy, so a noncontact method is proposed in which chest images are taken, and it is discovered that chest inhalation and deflation have different visual appearances [24]. A salient chest region is obtained using image processing techniques, and then, an optical signal is used to obtain a breath signal. The obtained dataset is compared with an existing dataset on a classifier to compare and evaluate the dataset. A sleep monitoring system based on a piezoresistive sensor and ML with an accuracy of about 97% is designed with the importance of regular monitoring of respiratory rate in mind [25].

### III. SYSTEM OVERVIEW

The proposed idea uses a camera mount as shown in Fig. 2 mounted over an FMCW radar as shown in Fig. 3 to capture a person's heart rate and breathing rate by automatic means. In this case, as shown in Fig. 4, the person must face the camera in any orientation. This will be accomplished by mounting the radar on a PTZ 360° rotation camera. The camera will rotate in the direction of the person and focus on the chest in order to fix the radar on it. Radar will then use chest displacements to accurately measure breathing and heart rate. The automatic steering will be enabled by integrating an Amcrest API that connects to the camera's PTZ property. After obtaining the heart and breathing rate signals and preprocessing using time–frequency technique, the signals will be passed to a CNN model in the form of images to be classified as normal, high, and low (breath) and low, slightly low, and normal (heart) and vital signs breathing and heartbeat.

#### A. Working of Setup

The above described setup works sequentially as follows:

- 1) The PTZ camera detects a person in the room and steers toward the person using face detection and steering code input via OpenCV and API.
- 2) When the camera detects a person, it stops moving and focuses on the person, remaining stationary until the radar begins and stops measuring chest displacements.
- 3) The radar senses vibrations or chest displacements by measuring the phase difference over a specific range. The graphical user interface (GUI) is used to measure the heart rate and breathing rate.
- 4) The data points for the breathing and heart signals are obtained and stored in a tabular format before being passed to the time–frequency technique, CWT for data preprocessing.
- 5) The preprocessed data are fed into a deep CNN to classify the breathing and heart rates.

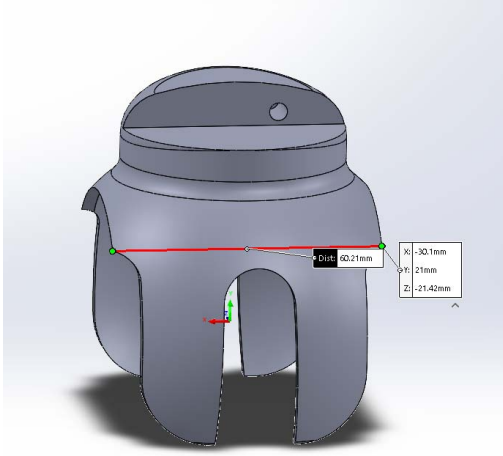


Fig. 2. AutoCAD design of camera mount.

TABLE III  
TECHNICAL DESCRIPTION OF AWR1443BOOST

S.No.	Characteristic	Values
1.	Receivers	4
2.	Transmitters	3
3.	CPU	ARM-Cortex R4F 200MHZ
4.	Maximum ADC Sampling Rate	12.5 MSPS
5.	RAM	576 MB
6.	Operating Temperature Range	-40 <sup>0</sup> C to 125 <sup>0</sup> C

#### IV. HARDWARE DESIGN

The details of radar mounting over the PTZ camera, technical description of AWR1443BOOST and PTZ camera, and FMCW radar along with radar integration are provided in this section.

##### A. Camera Mount

The radar was mounted on the curved surface of the PTZ camera in the novel setup proposed in this article. This is possible by creating a camera mount that sits on the camera surface and provides a flat surface for the radar to sit on. As a result, a camera mount design is created in AutoCAD, as shown Fig. 2, and then 3-D printed.

Later, it was designed on AutoCAD with various iterations because the inner surface of the camera mount should fit exactly over the camera, allowing for the least amount of sliding along the surface even when the radar sits on top of it, as shown in Fig. 3, adding weight to it.

##### B. Technical Description

Table III shows the technicalities of AWR1443BOOST [26]. Table IV shows the technicalities of Amcrest IP4M PTZ Camera [27].

##### C. FMCW Radar

FMCW radar [28] is a radar sensor that works by modulating the transmitted signal to change its operational frequency over time [29], [30], [31]. FMCW radars are widely used due to their ability to measure small ranges with high accuracy,



Fig. 3. Radar and mount setup.

TABLE IV  
TECHNICAL DESCRIPTION OF PTZ CAMERA

S.No.	Characteristics	Values
1.	Pan	360 degrees
2.	Tilt	90 degrees
3.	FoV	120 diagonal
4.	Resolution	4MP(2688x1520)
5.	Storage	1
6.	Alpha filter value for breath value	MicroSD Card-128 GB

simultaneous measurement of range and velocity, and low processing because of low intermediate frequencies [30], [32]. It detects through the reflection of the transmitted signal, which, as mentioned above, is frequency-modulated [30], [33]. This frequency-modulated transmitted signal [34], together with the reflected signals, produces the beat frequency (frequency difference between received and transmitted signals), as shown in Fig. 5. The distance between the radar antenna and the reflecting object can be calculated using beat frequency by determining the change in phase of the beat signal, as shown in the following [35].

The transmitted signal can be written as [35]

$$p(t) = e^{j(2\pi f_s t + \pi \frac{B}{T} t^2)} \quad (1)$$

where  $p(t)$  is the transmitted signal,  $B$  is the bandwidth, and  $T$  is the time period.

The received signal is given as follows [35]:

$$r(t) = e^{j(2\pi f_s(t-t_w) + \pi \frac{B}{T}(t-t_w)^2)} \quad (2)$$

The beat signal is obtained by transforming the above signals and filtering them and can be written as [35]

$$b(t) = p'(t) r(t) \approx e^{j(4\pi \frac{BR}{cT} t + \frac{4\pi}{\lambda} R)} = e^{j(2\pi f_b t + \phi_b)} \quad (3)$$



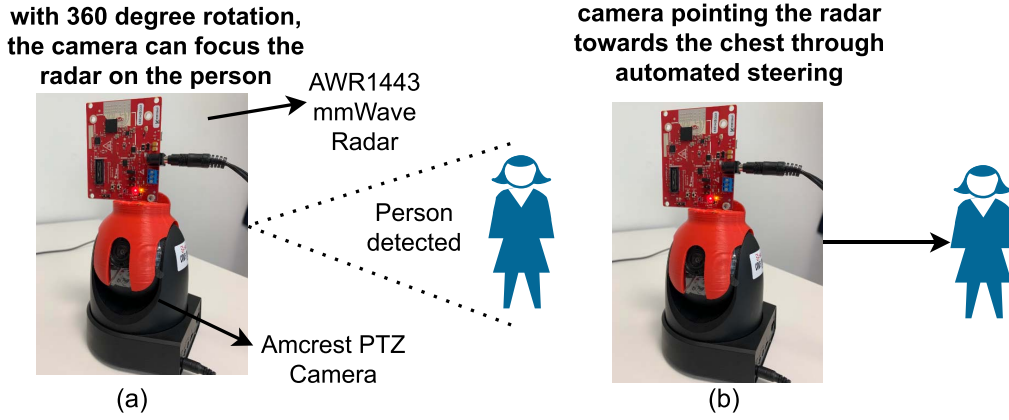


Fig. 4. Illustration of (a) camera capturing the person using its 360° feature and (b) camera pointing the mounted radar over the person's chest.

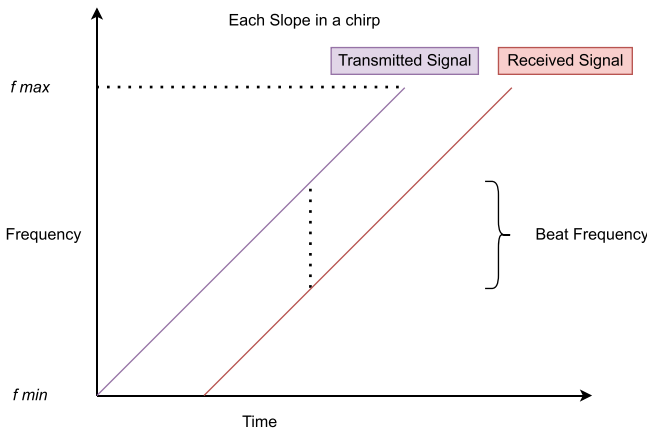


Fig. 5. Beat frequency obtained from FMCW single chirp.

where  $t_w = (2R/c)$ ,  $f_b$  is the beat frequency, and  $f_s$  is the starting frequency.

The following equation of the beat signal gives us the frequency  $f_b$  and phase  $\phi_b$ . Phase is derived by FFT of the beat signal  $b(t)$ , and the change in this phase can be used to find the change in range using the following equation [36]:

$$\Delta\phi_b = \frac{4\pi}{\lambda} \Delta R \quad (4)$$

where  $\Delta\phi_b$  is the phase change,  $\lambda$  is the wavelength, and  $\Delta R$  is the change in range.

The radar signal is measured in chirps (a signal in which frequency changes (increases or decreases) with time) [37] due to the high responsiveness of transducers to chirp signals [8]. Here, the frequency moves or changes while being in a frequency band range [37].

FMCW radar calculates the distance  $R$  to the object that reflects the transmitted signal, as given in [30]

$$R = \frac{c_0 |\Delta t|}{2} = \frac{c_0 |\Delta f|}{2 \left( \frac{df}{dt} \right)} \quad (5)$$

where  $c_0$  is the speed of light, which is equal to  $3 \times 10^8$  m/s,  $\Delta t$  is the time delay (s),  $\Delta f$  is the change in frequency (Hz),  $R$  is the distance (antenna and the object that reflects the signal), and  $df/dt$  is the change in the frequency with respect to change in time [30].

The AWR1443 FMCW radar is an excellent choice for the radar module because it can operate at low power levels and provides better detection accuracy for smaller movements [26]. As a result, it is ideal for a self-monitored automated system. The radar detects the vibrations of chest displacement in this work, capturing the activity of the heart and lungs, and providing the breathing and heart rates. Radar detects reflected electromagnetic waves transmitted from the human chest wall and measures chest displacement or vibrations from the heartbeat or breathing [8]. Using the radar measurements and unwrapping of the phase, a chest displacement-dependent signal can be obtained using (4). This signal is useful as it contains the signal corresponding to the breathing and heartbeat of the body.

#### D. Radar Integration

TI mmWave sensor measures the chest displacement or vibrations of the person in front of the radar. The AWR1443 cores filter out the noise and give a clean breathing and heartbeat graph from the vibrations of the chest and give an estimated value of heart and breathing rate. Fig. 6 mentions the step-by-step process of obtaining the BR and HR signals from the person's chest using AWR1443BOOST [35]. The detailed description is mentioned in the following [35].

- 1) The radar provides ADC data, which is FFT transformed to produce a range profile.
- 2) The maximum value of the range bin is calculated, which aids in determining the actual subject for which readings must be taken.
- 3) The phase value is extracted from the range data and then measured in relation to different time values.
- 4) This phase is modified by unwrapping it with the addition or subtraction of  $2\pi$  if the difference between consecutive phase values is greater than  $\pi$  or less than  $-\pi$ , and later displacement is calculated with respect to this unwrapped phase.
- 5) The phase difference between unwrapped phases is computed. To remove impulsive noises from these unwrapped phase estimations, a backward and forward phase difference is computed. Assuming the current phase to be  $b(m)$ , the forward phase difference can be computed as  $b(m) - b(m+1)$  and the backward phase

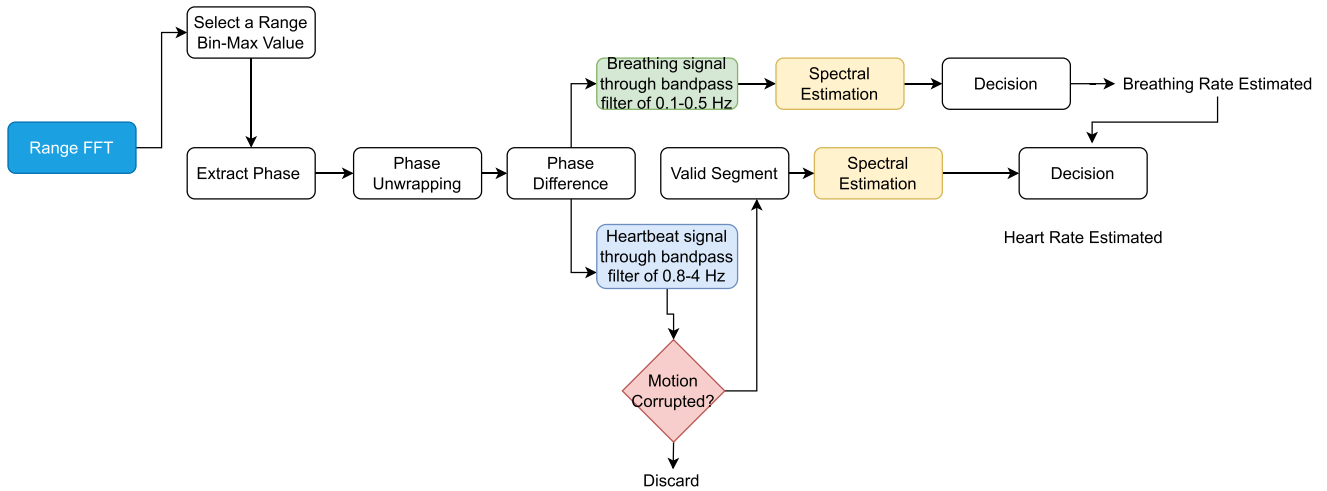


Fig. 6. Radar processing for vital signs' extraction.



Fig. 7. GUI snapshot and data extraction for further processing [35].

difference as  $b(m) - b(m - 1)$ . If the computed phase difference still exceeds the threshold value, then it is substituted with an approximate value [35].

To obtain breathing rate waveforms, a bandpass filter with a frequency range of 0.1–0.5 Hz is applied continuously to real-time input data. Then comes spectral estimation, which includes FFT, autocorrelation, and peak interval determination. Finally, the breathing rate is calculated using the most likely confidence interval from all three spectral estimation filtering techniques.

To obtain heart rate waveforms, the following steps are performed in addition to the steps above [35].

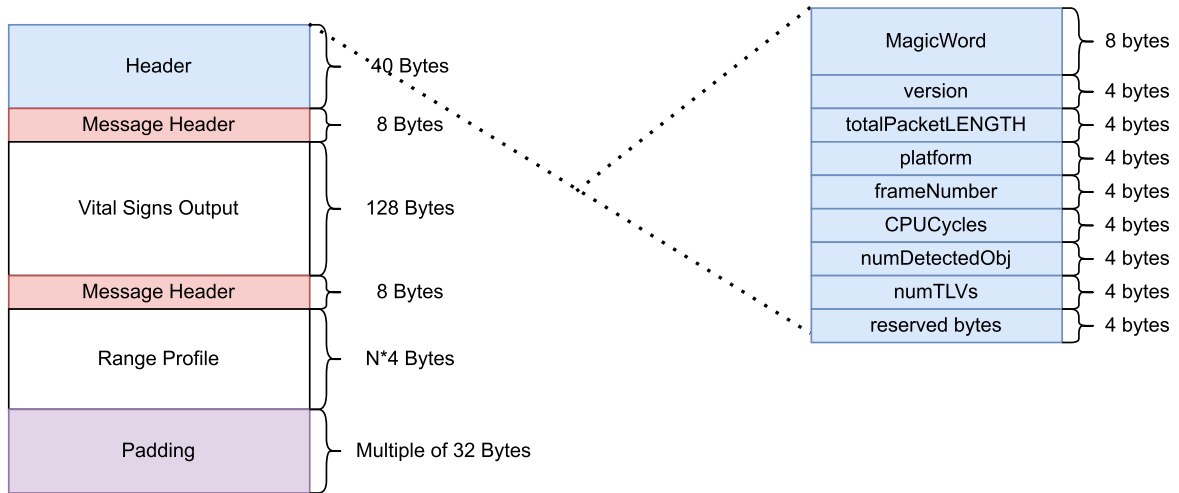
- 1) A bandpass filter of the range of 0.8–4 Hz is placed continuously on the real-time input data.
- 2) To improve heart rate estimation, it is passed through a motion corrupted segment removal block in order to reduce large amplitude impulses in the signal. The signal is divided into 20 segments, and if these segments have an energy exceeding the defined energy threshold

$(E > E_{Th})$ , then either its scaled by  $(E_{Th}/E)^{1/2}$  or the segments are neglected.

- 3) Spectral estimation is performed which consists of FFT, autocorrelation, and determining the peak interval on the autocorrelated data.
- 4) The heart rate decision is based on the most likely confidence interval considering all three different spectral estimation filtering techniques. The heart rate decision is also based on the obtained breathing rate decision.

A GUI, as shown in Fig. 7, displays the BR and HR along with chest displacements, breathing, and heartbeat patterns.

Vital signs' waveform is obtained by extracting the phase from the beat signal and unwrapping the phase and sampling it along the slow time axis [35] using the one out of two chirps for processing with parameters mentioned in Table V. Radar is then integrated by sending the configurations shown in Tables VI and VII through universal asynchronous receiver transmitter (UART) packet as shown in Fig. 8. These data have been extracted for further processing and classification using machine learning.



Where N is number of processed range-bins

Fig. 8. UART packet sent to the AWR1443BOOST.

TABLE V  
CHIRP CONFIGURATION PARAMETERS

Parameters	Values
Bandwidth	4 GHz
Start Frequency	77 Hz
Time Axis sampling rate(Frame-rate)	20 Hz
ADC Sampling rate	2 MHz
ADC Samples in a chirp	100
Chirp duration	50 micro-s

TABLE VI  
VITAL SIGNS' CONFIGURATION

S.No.	Parameters	Values
1.	Start Range	0.3 m
2.	End Range	1 m
3.	Waveform length of vital signs	128
4.	Receiver number	4
5.	Alpha filter value for heart-beat value	0.5
6.	Alpha filter value for breath value	0.5
7.	Alpha filter value for range-bin value	0.1

TABLE VII  
MOTION DETECTION CONFIGURATION

S.No.	Parameters	Values
1.	Enable	1
2.	Segment Length	32 Bytes
3.	Energy threshold on breathing waveform	3
4.	Energy threshold on heart waveform	0.05
5.	Threshold on the heartrate CM	0.4
6.	Threshold on the difference between heartRate and peakbased estimate	15
7.	Gain Control	0

## V. SOFTWARE ALGORITHM

This section discusses the steps of face detection and HAAR classifier along with PTZ camera integration, where the method of connecting the camera and camera steering is described.

### A. Face Detection and HAAR Classifier

This algorithm [38] has two significant steps:

- 1) training;
- 2) classification.

Training is done with negative and positive labeled images (no person and person). Identification is done using “HAAR features” [39] that are rectangles of features with dark and light pixels in them. Each feature is assigned a value, which is the difference between the sum of dark and the sum of light pixels [38]

$$\text{Pixelvalue} = (\text{SODP}/\text{NODP}) - (\text{SOLP}/\text{NOLP}) \quad (6)$$

where SODP is the sum of dark pixels, NODP is the number of dark pixels, SOLP is the sum of light pixels, and NOLP is the total number of light pixels. The trained classifier on prelabeled images extracts relevant features from the image to reduce error. The final classifier is a summation of other weighted weak classifiers (these classifiers cannot classify an image on its own but rather work in summation to classify with accuracy) [38]. Instead of applying all the obtained features to the image, a cascade of classifiers is used, which groups features according to predefined criteria to different classifiers and then is applied to an image to reduce computations [38]. A relevant feature is called a weak feature, which, when added up, constitutes the classifier [38]. Later, many such classifiers are cascaded to improve accuracy and detection speed, and remove negative features.

Face detection is implemented using OpenCV [40]. To detect the object, the HAAR classifier extracts features from the image, and every pixel is calculated using (6). The HAAR classifier is widely used because it is computationally less expensive and performs better than other classifiers, such as the LBP classifier [40].

### B. PTZ Camera Integration

A PTZ camera by Amcrest offers 360° rotation, 90° pan, and 16 times zoom [27], which, when integrated with the radar, increases its FoV and enables it to capture the data from a person in any orientation (stationary). It has other features, such as night vision, camera recording, and motion detection [27]. The camera was connected to the code via RSTP protocol.

The PTZ camera is the key component in the novel design because it automates the capture by steering toward the person, eliminating the need for the person to be in the radar's frame for accurate detection. It enables the radar to automatically steer toward the person and capture vital signs, such as breathing and heart rate. The PTZ camera used in the system is actually an RGB camera. Under low-light conditions, the RGB camera has some limitations. In the future, a thermal camera and/or night vision camera can be integrated into the system to improve system reliability in low-light conditions.

**1) Connecting the Camera:** The PTZ IP camera can connect to the internet via Ethernet cable, Wi-Fi, or cellular network, and it supports web app integration. In the proposed design, the IP camera has been connected via Ethernet cable to the router on a 2.4-GHz frequency. After connecting the camera, the live stream can be accessed via the mobile application on the desired channel.

Once the camera was connected to the internet, it was linked to the OpenCV code on Jupyter via the RSTP protocol, and its IP address was obtained (obtained via the mobile application). The live video frames can be accessed on the connected PC after a successful connection.

**2) Passing Video Frames and Threading:** The video frame is passed through `cv2.VideoCapture()` function and is connected to the camera via random spanning tree protocol (RSTP) and the URL. The video frame is then passed to the `rescaleframes` function because the obtained video frame was a zoomed-in frame. The function rescales the frame shape by the scaling factor, which is passed as 25, which means 25% of the frame shown. This number was chosen because it provides the entire view of the video streaming. The `cv2.VideoCapture()` collects every frame and then blocks the remaining Python code, causing significant latency. To reduce latency while increasing frame rate, threading was employed. Threading was used to create two threads that are independent of one another. The frames were captured by the first thread, and the remaining steering code was executed by the second thread.

**3) Automatic Camera Steering:** Automatic steering has two major steps.

- 1) First, it detects a person's face and continues to scan until it finds one; once a person is detected, it stops at the person's face. The face detection algorithm computes the center of the rectangle to find the center of the face and returns the coordinates of the face using the bounding rectangle.
- 2) Thereafter, it calls the position API using `moveCamera()`, passes the coordinates of the center of the face, and adjusts the camera to the chest region according to the  $x$ - and  $y$ -coordinates passed from the faces variable to the position calling function. The camera stops at the person using the `StopTour()` and `StartTour()` API called in a `movecamera()` user defined function which changes the API return values based on the parameters passed to it. The radar can then be directed toward the chest region. The person must remain stationary, while the readings are taken; this is

a requirement of the entire system. The individual can be in any orientation but must be stationary.

The `moveCamera()` function uses the URL syntax from Amcrest documentation [41] using GET, requests, and hyper-text transfer protocol (HTTP) digest authentication to connect to the server using the username and password assigned to the camera by the user through Amcrest Application.

## VI. ASSEMBLY OF THE SETUP

Creating an appropriate camera mount to place the radar over the camera was the first step in the final assembled setup as shown in Fig. 9. The camera and radar are then individually integrated for their respective functionality and role in the measurement setup. To minimize the load on the components over the radar, a flat heightened surface is provided for the radar to be placed above the camera. The radar is screwed at one end with the other end out of the flat surface. The stationary person is in the same room as the setup but within 0.3–1 m of the setup for accurate measurements by the radar, as shown in Fig. 10. The setup points to the chest area, and the readings are taken by launching the GUI. The GUI's functionality can be verified by observing a maximum peak in the range profile, which should correspond to the subject whose data are being collected (other smaller peaks are from object interferences).

## VII. DATASET DETAILS AND PREPROCESSING

The data were obtained from 12 people (in which one of them had chest surgery and the rest were healthy). Each subject was asked to breathe in the low, high, and normal tempos for nearly a minute for each reading in ten different orientations, resulting in 30 different readings per subject and a dataset of 720 measurements. These data are unbalanced, as shown in Table VIII. The ten different orientations are sitting near to the radar (approximately 0.3 m), sitting away from the radar (approximately 1 m), sitting left, sitting right, and any random orientation (for example, a person using his/her mobile phone while reading and being unaware of the collection process), and then, the same orientations are repeated with the person standing. The data obtained were monitored on the GUI, and the raw waveforms' displacement (a.u) versus time was plotted. This dataset is then preprocessed using CWT to obtain frequency versus time plots; this also reduces the image size substantially and makes the dataset lightweight.

### A. Continuous Wavelet Transformation

Wavelet transformations are accurate for transient signals, such as ECG signals, breathing waveforms, heart waveforms, or any other real-time vital sign waveform because they provide transformed information about the signal at each time step with accurate information over frequencies [42]. In this type of transformation, a small part of the signal, namely, a wavelet, is varied in amplitude and frequency by changing its scaling factor, and it is then slid across the main signal by multiplying the scaling factor by it [43]. As the name implies, CWT is a transformation in which the wavelet transforms through a scale factor (having infinite values and possibilities,



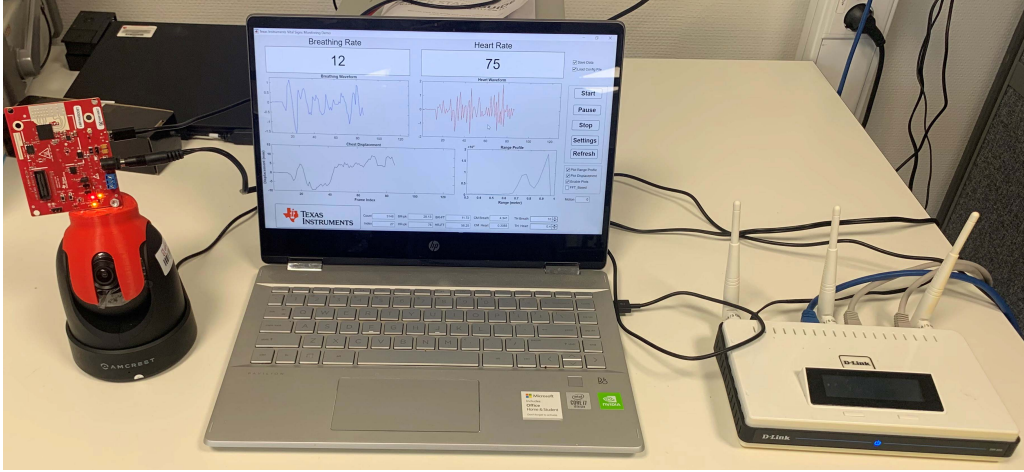


Fig. 9. Final measurement setup.

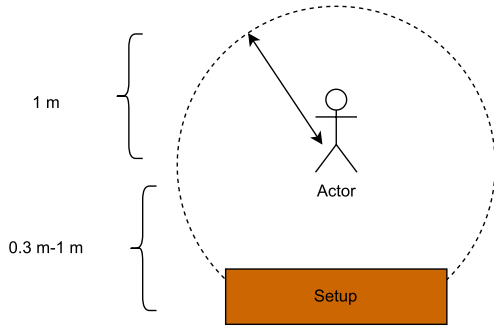


Fig. 10. Distance requirements for the setup.

TABLE VIII  
DATASET DISTRIBUTION

S.No.	Classes	Images per class
1.	low-heart	111
2.	normal-heart	116
3.	slilow-heart	132
4.	normal-breath	125
5.	low-breath	106
6.	high-breath	126

making it continuous rather than discrete) and compares the resemblances between the sliding function and the signal [44]. By varying  $c$  and  $d$ , as shown in the following equation, the coefficients of the CWT signal  $T(c, d)$  are obtained [42]:

$$T(c, d) = \frac{1}{\sqrt{c}} \int_{-\infty}^{\infty} x(t) \bar{\psi}^* \left( \frac{t-d}{c} \right) dt \quad (7)$$

where  $x(t)$  represents the signal,  $\psi$  represents the mother wavelet, and the operation involves the complex conjugate of it,  $c$  is the scale factor, and  $d$  is the shifting factor. Scaling is the compression and dilation of a signal by multiplying it with a scale factor and obtaining a number of new signals (in this case, infinite as there are infinite possibilities of scaling), and this proves to be an advantage of wavelet transformations as there are many options to choose from, depending on the application [42], [43]. A small-scale factor produces a compressed signal with rapidly changing values and a high frequency, whereas a large-scale factor produces a dilated signal with

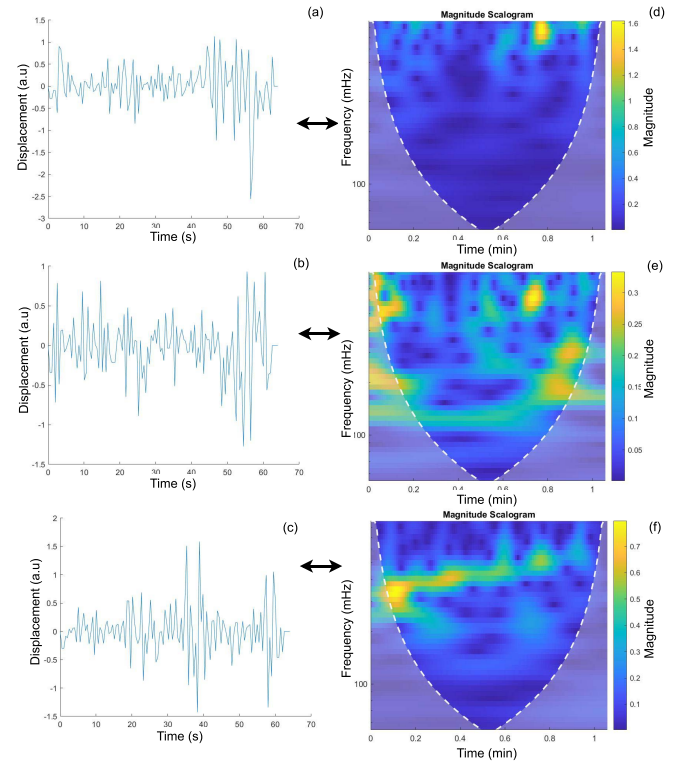


Fig. 11. Collected dataset waveforms from the GUI in (a) low (b) high and (c) normal. The transformed waveforms are shown in (d) low (e) high and (f) normal in setting orientation.

slow changing values and a low frequency [42]. A large- or small-scale factor should be chosen depending on the application. CWT is an excellent choice for image compression because the type of compression required can be obtained by adjusting the scale factor [44]. The measured heart rate data and its corresponding preprocessed CWT images are shown in Fig. 11. Each image is constructed with a fixed number of data points, i.e., 128 over a span of 64 s giving the sampling frequency (number of samples in 1 s) of 2 Hz ( $128/64 = 2$ ). The dataset collected is in the form of data points that are then converted to plots and passed through CWT.

TABLE IX

SUMMARY OF THE PROPOSED MODEL IN TERMS OF OUTPUT SHAPE AND THE NUMBER OF PARAMETERS IN EACH LAYER

S.No.	Layer (type)	Output Shape	Param #
1.	input_1 (Input Layer)	(None, 128, 128, 3)	0
2.	conv2d (Conv2D)	(None, 124, 124, 32)	896
3.	batch_normalization (BatchNormalization)	(None, 124, 124, 32)	128
4.	max_pooling2d (MaxPooling2D)	(None, 62, 62, 32)	0
5.	dropout (Dropout)	(None, 62, 63, 32)	0
6.	re_lu (ReLU)	(None, 63, 63, 32)	0
7.	conv2d_1 (Conv2D)	(None, 58, 58, 64)	18496
8.	batch_normalization_1 (BatchNormalization)	(None, 58, 58, 64)	256
9.	max_pooling2d_1 (MaxPooling2D)	(None, 29, 29, 64)	0
10.	dropout_1 (Dropout)	(None, 29, 29, 64)	0
11.	re_lu_1 (ReLU)	(None, 29, 29, 64)	0
12.	flatten (Flatten)	(None, 53824)	0
13.	dense (Dense)	(None, 128)	6889600
14.	dropout_3 (Dropout)	(None, 128)	0
15.	dense_1 (Dense)	(None, 128)	16512
16.	dropout_4 (Dropout)	(None, 128)	0
17.	dense_2 (Dense)	(None, 64)	8256
18.	dropout_5 (Dropout)	(None, 64)	0
19.	dense_3 (Dense)	(None, 64)	4160
20.	dropout_6 (Dropout)	(None, 64)	0
21.	dense_4 (Dense)	(None, 6)	390
<b>Total params: 6,938,694</b>			
<b>Trainable params: 6,938,502</b>			
<b>Non-trainable params: 192</b>			

### VIII. PROPOSED CNN MODEL AND TRAINING

A deep CNN is highly used for image classification, which makes use of convolutional layers [45], in which the features are convolved over input data. In CNN, the features are learned during training unlike manual training, which makes them an ideal choice for CV applications [46]. After the convolutional layer, a feature map is created and is normalized using batch normalization layer [47] to stabilize the output. On this normalized output, the activation layer [48] is applied to increase the nonlinearity of the model, making it more fit for the real-world data [46]. The max-pooling [49] layer is the next layer in which the maximum value of the pixel intensity is calculated and returned to the kernel, reducing the number of computations, overfitting, and increasing the variance of the model to predict in all conditions [46].

Next is the flatten layer, in which the feature map is converted into a 1-D vector to further put it as input to the final layers [50]. Final layers usually consist of a combination of dropout and dense layers. The dropout layer is used to slow down the training process by introducing a fraction of the available dataset, which will be put into training or not [51]. The final layer is the fully connected layer or the dense layer [48], in which every neuron is connected to every other neuron in the network. It uses the output from the previous image to input in the next one [46]; it uses softmax activation and later classifies it into the final number of classes.

The proposed 2-D CNN model is the lightweight model, which has the layer architecture in the order of the convolutional layer and the batch normalization layer followed by max-pooling, dropout, and ReLU activation layers. Two such sets of this layer organization are used in the reference model, as shown in Fig. 12. The convolutional layer has a filter size of 3, and the input images will be set to the dimension of  $128 \times 128$ , giving the input shape to the model

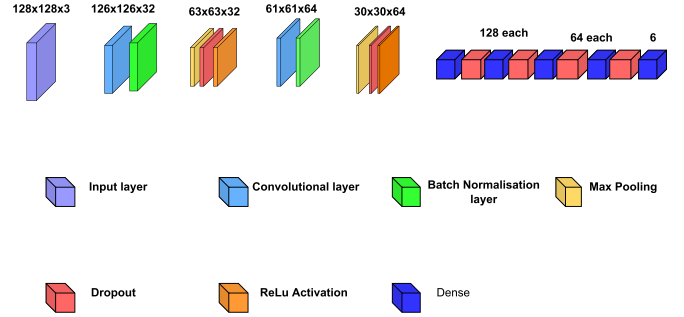


Fig. 12. Reference model CNN architecture.

as  $128 \times 128 \times 3$ . The batch normalization layer is used after the convolutional layer to slow down the training and make the learning stable by normalizing the input layer. The max-pooling layer has a size of  $2 \times 2$ , and the dropout layers are set to 0.2 and 0.3 and some layers with 0.4. The layers with the highest number of parameters use L2 regularization with a rate of 0.005 to decrease the overfit. The final layers are a sequence of dropout and dense layers leading to the output classification into six classes, namely, breathlow, breathnormal, breathhigh, heartlow, hearthigh, and heartnormal. Finally, the output layer is used, which consists of the dense layer with softmax activation as it helps in proper identification in multiclass scenarios (more than 2) in our case 6. The layerwise architecture and output shape of each layer of the reference model are tabulated in Table IX.

#### A. Training

The model shown in Table IX is trained on the Google Compute Engine using the Adam Optimizer and softmax activation in the output layer at a learning rate of 0.00001, a batch size of 16, and 100 epochs. The dataset is split

TABLE X  
ACCURACIES OVER DIFFERENT TRAINING FOLDS

S.No.	Model	Training Accuracy	Validation Accuracy	Testing Accuracy
1.	Model fold 1	91.34	82.24	78.12
2.	Model fold 2	80.602	87.53	87.5
3.	Model fold 3	91.92	79.78	78.12
4.	Model fold 4	92.29	83.06	87.5
5.	Model fold 5	89.58	81.69	79.69
6.	Model fold 6	90.10	83.33	73.34
7.	Model fold 7	90.62	80.07	81.25
8.	Model fold 8	90.97	77.05	79.69
9.	Model fold 9	90.45	80.60	75.00
10.	Model fold 10	90.10	79.23	69.00

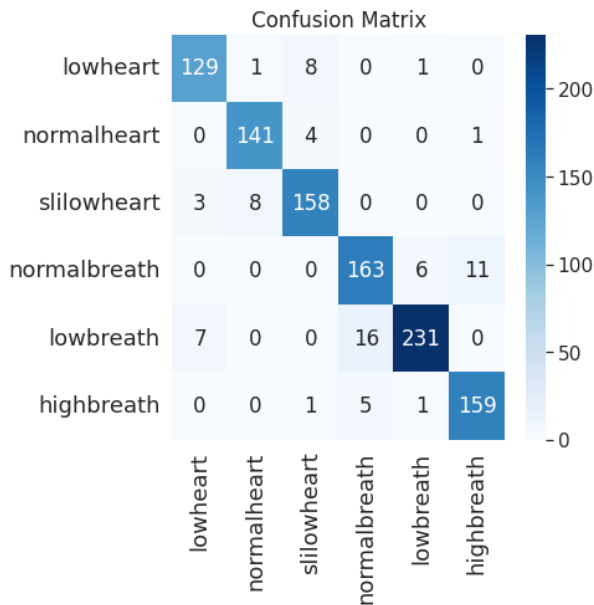


Fig. 13. Confusion matrix of the proposed deep CNN model.

into test and training automatically using the  $k$ -fold cross-validation method [52]. The results obtained from the training are discussed in the results section. The “ $k$ ” value selected was 10, giving rise to ten different data splits between test and train resulting in ten different models. The model with the best accuracy and less overfit is selected as the model fit for classification.

## IX. RESULTS AND DISCUSSION

The proposed deep CNN is trained on six different classes of our own obtained radar dataset. Fig. 13 shows the confusion matrix of the six classes—low breath, normal breath, high breath, low heart, normal heart, and slightly low heart. The matrix shows that low heart is classified 93% correctly, the normal heart is classified 96.5% correctly, the slightly low heart is classified 93% correctly, the normal breath is classified 91% correctly, the low breath is classified 91% correctly and the class high breath is classified with accuracy 96%. The low and normal breath classes are quite close to each other in sample type and, due to further imbalance low breath, get wrongly classified as normal breath in some instances.

Fig. 14 highlights the train and validation accuracies versus epochs for the model with the training fold of 9. The graph

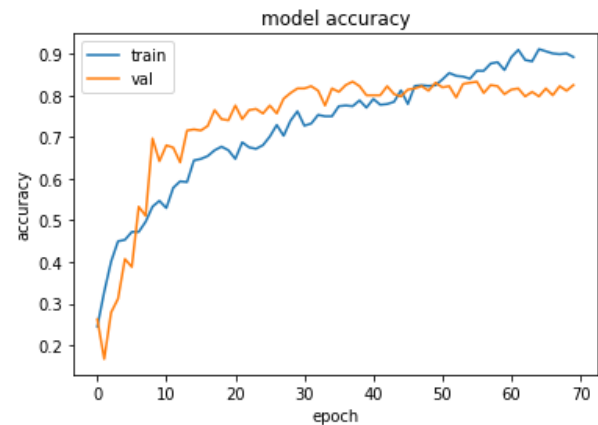


Fig. 14. Training and validation accuracies of the proposed deep CNN model.

TABLE XI  
COMPARING PRECISION (P), RECALL (R), AND F1 SCORES  
FOR ALL MODELS OVER THE DATASET

S.No.	Model	Precision(P)	Recall(R)	F1
1.	Class 0	0.92	0.94	0.93
2.	Class 1	0.93	0.96	0.95
3.	Class 2	0.92	0.92	0.92
4.	Class 3	0.90	0.89	0.89
5.	Class 4	0.95	0.92	0.93
6.	Class 5	0.92	0.96	0.94
Average Precision:0.93				
Average Recall:0.93				
Average F1:0.93				

shows that the increase in validation accuracy is quite steady resulting in training accuracy to overtake the accuracy after certain epochs, showing that there is minor overfit, which can be solved with a more varied dataset. The validation loss becomes constant after a while, unlike the training loss that decreases exponentially.

Table X shows the validation, test, and training accuracies of the proposed CNN model over all the ten training splits. In some folds, validation accuracy is greater than testing; in some folds, testing is better, but training accuracy is mostly greater than both, whereas some folds have less difference in the three accuracies, showing a stable model fold, such as fold 2. Table XI shows the precision, recall, and F1 score of all the models considered in this article. Due to unbalanced data, precision, recall, and F1 also become important indicators

of the model's performance. It can be observed that class 4 that is low breath has the highest precision, which is 0.95, indicating that this class was very precise to its true value. Class 1 and Class 5 (normal heart and high breath) have the highest recall of around 0.96, showing that these classes were correctly found. F1 score of 0.95 was found the highest in class 1 (normal heart), showing that this class was predicted the closest with respect to even unseen data. The average model classification went down to a value of 0.93 for all precision, recall, and F1, which is a good value showing the model's success in predicting the obtained data.

## X. CONCLUSION AND FUTURE WORK

In this article, a novel noncontact automatic method for measuring and classifying heart rate and breathing rate has been proposed. A dataset was created using the proposed novel setup. The dataset contains 720 measurements that were classified into six classes using the CNN model with an accuracy of 87%. The accuracy of the system can be improved in the future by expanding the dataset.

## REFERENCES

- [1] K. McClure, B. Erdreich, J. H. Bates, R. S. McGinnis, A. Masquelin, and S. Wshah, "Classification and detection of breathing patterns with wearable sensors and deep learning," *Sensors*, vol. 20, no. 22, p. 6481, 2020.
- [2] *Vital Signs*. Accessed: May 15, 2022. [Online]. Available: <https://my.clevelandclinic.org/health/articles/10881-vital-signs>
- [3] J. Skinner, J. Garibaldi, and E. Ifeakor, "A fuzzy system for fetal heart rate assessment," in *Proc. Int. Conf. Comput. Intell.*, vol. 1625, Jul. 2007, pp. 20–29.
- [4] H. Liu, J. Allen, D. Zheng, and F. Chen, "Recent development of respiratory rate measurement technologies," *Physiol. Meas.*, vol. 40, no. 7, 2019, Art. no. 07TR01.
- [5] G. Paterniani et al., "Radarbased monitoring of vital signs: A tutorial overview," *TechRxiv*, 2022, doi: [10.36227/techrxiv.19212918.v1](https://doi.org/10.36227/techrxiv.19212918.v1).
- [6] A. Nicolò, C. Massaroni, E. Schena, and M. Sacchetti, "The importance of respiratory rate monitoring: From healthcare to sport and exercise," *Sensors*, vol. 20, no. 21, p. 6396, Nov. 2020.
- [7] A. T. Purnomo, D.-B. Lin, T. Adiprabowo, and W. F. Hendria, "Non-contact monitoring and classification of breathing pattern for the supervision of people infected by COVID-19," *Sensors*, vol. 21, no. 9, p. 3172, 2021.
- [8] E. Turppa, J. M. Kortelainen, O. Antropov, and T. Kiuru, "Vital sign monitoring using FMCW radar in various sleeping scenarios," *Sensors*, vol. 20, no. 22, p. 6505, Nov. 2020.
- [9] D. C. Mack, J. T. Patrie, P. M. Suratt, R. A. Felder, and M. Alwan, "Development and preliminary validation of heart rate and breathing rate detection using a passive, ballistocardiography-based sleep monitoring system," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 1, pp. 111–120, Jan. 2009.
- [10] Y. Gu, X. Zhang, Z. Liu, and F. Ren, "WiFi-based real-time breathing and heart rate monitoring during sleep," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.
- [11] A. Procházka, M. Schätz, O. Vyšata, and M. Vališ, "Microsoft Kinect visual and depth sensors for breathing and heart rate analysis," *Sensors*, vol. 16, no. 12, p. 996, Jun. 2016.
- [12] K.-K. Shyu, L.-J. Chiu, P.-L. Lee, T.-H. Tung, and S.-H. Yang, "Detection of breathing and heart rates in UWB radar sensor data using FVPIEF-based two-layer EEMD," *IEEE Sensors J.*, vol. 19, no. 2, pp. 774–784, Jan. 2019.
- [13] S. Gauni, M. Megalari, K. Kalimuthu, and M. Anjanasree, "Detection of breathing and heartrate using a simple vital radio system," *Int. J. Eng. Technol.*, vol. 7, p. 311, Mar. 2018.
- [14] A. G. Leal-Junior, C. R. Díaz, C. Leitão, M. J. Pontes, C. Marques, and A. Frizera, "Polymer optical fiber-based sensor for simultaneous measurement of breath and heart rate under dynamic movements," *Opt. Laser Technol.*, vol. 109, pp. 429–436, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0030399218312064>
- [15] M. Rehman et al., "Contactless small-scale movement monitoring system using software defined radio for early diagnosis of COVID-19," *IEEE Sensors J.*, vol. 21, no. 15, pp. 17180–17188, May 2021.
- [16] S. D. Min, J. K. Kim, H. S. Shin, Y. H. Yun, C. K. Lee, and M. Lee, "Noncontact respiration rate measurement system using an ultrasonic proximity sensor," *IEEE Sensors J.*, vol. 10, no. 11, pp. 1732–1739, Nov. 2010.
- [17] G. O. Ganfure, "Using video stream for continuous monitoring of breathing rate for general setting," *Signal, Image Video Process.*, vol. 13, no. 7, pp. 1395–1403, 2019.
- [18] A. Singh, S. U. Rehman, S. Yongchareon, and P. H. J. Chong, "Multi-resident non-contact vital sign monitoring using radar: A review," *IEEE Sensors J.*, vol. 21, no. 4, pp. 4061–4084, Nov. 2021.
- [19] E. Piuze, S. Pisa, E. Pittella, L. Podestà, and S. Sangiovanni, "Low-cost and portable impedance plethysmography system for the simultaneous detection of respiratory and heart activities," *IEEE Sensors J.*, vol. 19, no. 7, pp. 2735–2746, Dec. 2019.
- [20] D. Biswas, N. Simões-Capela, C. Van Hoof, and N. Van Helleputte, "Heart rate estimation from wrist-worn photoplethysmography: A review," *IEEE Sensors J.*, vol. 19, no. 16, pp. 6560–6570, May 2019.
- [21] M. Ali, A. Elsayed, A. Mendez, Y. Savaria, M. Sawan, and M. Sawan, "Contact and remote breathing rate monitoring techniques: A review," *IEEE Sensors J.*, vol. 21, no. 13, pp. 14569–14586, Jul. 2021.
- [22] C. Massaroni, A. Nicolò, M. Sacchetti, and E. Schena, "Contactless methods for measuring respiratory rate: A review," *IEEE Sensors J.*, vol. 21, no. 11, pp. 12821–12839, Jun. 2021.
- [23] T. Choudhary, L. N. Sharma, M. K. Bhuyan, and K. Bora, "Identification of human breathing-states using cardiac-vibrational signal for m-Health applications," *IEEE Sensors J.*, vol. 21, no. 3, pp. 3463–3470, Feb. 2021.
- [24] K. Y. Lin, D. Y. Chen, and W. J. Tsai, "Image-based motion-tolerant remote respiratory rate evaluation," *IEEE Sensors J.*, vol. 16, no. 9, pp. 3263–3271, May 2016.
- [25] Z. Wang et al., "A piezoresistive array based force sensing technique for sleeping posture and respiratory rate detection for SAS patients," *IEEE Sensors J.*, early access, Dec. 13, 2022, doi: [10.1109/JSEN.2021.3134823](https://doi.org/10.1109/JSEN.2021.3134823).
- [26] *Awrl443boost*. Accessed: May 15, 2022. [Online]. Available: <https://www.ti.com/tool/AWR1443BOOST>
- [27] *Amcrest 4MP IP Camera WiFi UltraHD Wireless Outdoor Security Camera Bullet—IP67 Weatherproof, 98FT Night Vision, 4-Megapixel (2688 TVL), Rep-IP4M-1026 (Black) (Certified Refurbished)*. Accessed: May 15, 2022. [Online]. Available: <https://amcrest.com/4-megapixel-wifi-wireless-ip-security-bullet-camera-rep-ip4m-1026b-certified-refurbished.html>
- [28] S. Gupta, P. K. Rai, A. Kumar, P. K. Yalavarthy, and L. R. Cenkeramaddi, "Target classification by mmWave FMCW radars using machine learning on range-angle images," *IEEE Sensors J.*, vol. 21, no. 18, pp. 19993–20001, Sep. 2021.
- [29] L. R. Cenkeramaddi, J. Bhatia, A. Jha, S. K. Vishkarma, and J. Soumya, "A survey on sensors for autonomous systems," in *Proc. 15th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Nov. 2020, pp. 1182–1187.
- [30] D.-I. F. C. Wolff, *Radar Basics*. Accessed: May 15, 2022. [Online]. Available: <https://www.radar-tutorial.eu/02.basics/Frequency/20Modulated/20Continuous/20Wave/20Radar.en.html>
- [31] L. R. Cenkeramaddi et al., "A novel angle estimation for mmWave FMCW radars using machine learning," *IEEE Sensors J.*, vol. 21, no. 8, pp. 9833–9843, Apr. 2021.
- [32] P. K. Rai et al., "Localization and activity classification of unmanned aerial vehicle using mmWave FMCW radars," *IEEE Sensors J.*, vol. 21, no. 14, pp. 16043–16053, Jul. 2021.
- [33] S. S. S. Panda, T. Panigrahi, S. R. Parne, S. L. Sabat, and L. R. Cenkeramaddi, "Recent advances and future directions of microwave photonic radars: A review," *IEEE Sensors J.*, vol. 21, no. 19, pp. 21144–21158, Oct. 2021.
- [34] B. Pardhasaradhi and L. R. Cenkeramaddi, "GPS spoofing detection and mitigation for drones using distributed radar tracking and fusion," *IEEE Sensors J.*, vol. 22, no. 11, pp. 11122–11134, Jun. 2022.
- [35] T. Instruments. (May 2017). *Mmwave Vital Signs Lab*. [Online]. Available: <https://training.ti.com/jp/mmwave-vital-signs-lab>
- [36] W. Lv, W. He, X. Lin, and J. Miao, "Non-contact monitoring of human vital signs using FMCW millimeter wave radar in the 120 GHz band," *Sensors*, vol. 21, no. 8, p. 2732, Apr. 2021.
- [37] S. Jurdak et al., "Detection and localization of multiple short range targets using FMCW radar signal," in *Proc. Global Symp. Millim. Waves (GSMW) ESA Workshop Millimetre-Wave Technol. Appl.*, Jun. 2016, pp. 1–4.



- [38] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Feb. 2001, p. 1.
- [39] M. Rezaei, H. Z. Nafchi, and S. Morales, "Global Haar-like features: A new extension of classic Haar features for efficient face detection in noisy images," in *Proc. Pacific-Rim Symp. Image Video Technol.* Cham, Switzerland: Springer, 2013, pp. 302–313.
- [40] A. B. Shetty, Bhoomika, Deeksha, J. Rebeiro, and Ramyashree, "Facial recognition using Haar cascade and LBP classifiers," *Global Transitions Proc.*, vol. 2, no. 2, pp. 330–335, Nov. 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666285X21000728>
- [41] *Amcrest+http+api+9.2020*, Amcrest, New York, NY, USA, 2020.
- [42] Accessed: Apr. 8, 2022. [Online]. Available: <https://in.mathworks.com/help/wavelet/gsc/continuous-wavelet-transform-and-scale-based-analysis.html>
- [43] L. Debnath and F. A. Shah, *Wavelet Transforms and Their Applications*. Cham, Switzerland: Springer, 2002.
- [44] J.-P. Antoine, "The continuous wavelet transform in image processing," *CWI Q*, vol. 11, no. 4, pp. 323–345, 1998.
- [45] T. Bezdán and N. Bacanin, "Convolutional neural network layers and architectures," in *Proc. Int. Sci. Conf. Inf. Technol. Data Related Res.*, Jan. 2019, pp. 445–451.
- [46] A. Bonner. (Jun. 2019). *The Complete Beginner's Guide to Deep Learning: Convolutional Neural Networks*. [Online]. Available: <https://towardsdatascience.com/wtf-is-image-classification-8e78a8235acb>
- [47] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*.
- [48] A. M. Javid, S. Das, M. Skoglund, and S. Chatterjee, "A ReLU dense layer to improve the performance of neural networks," 2020, *arXiv:2010.13572*.
- [49] J. Nagi et al., "Max-pooling convolutional neural networks for vision-based hand gesture recognition," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Nov. 2011, pp. 342–347.
- [50] J. Jin, A. Dunder, and E. Culurciello, "Flattened convolutional neural networks for feedforward acceleration," 2014, *arXiv:1412.5474*.
- [51] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014. [Online]. Available: <http://jmlr.org/papers/v15/srivastava14a.html>
- [52] Y. Wen, M. Kalander, C. Su, and L. Pan, "An ensemble noise-robust K-fold cross-validation selection method for noisy labels," 2021, *arXiv:2107.02347*.



**Khushi Gupta** received the bachelor's degree in electrical and electronics engineering (EEE) from the Birla Institute of Technology and Science (BITS), Pilani, Hyderabad, India, in 2022. She is currently pursuing the master's degree in computer engineering with Columbia University, New York, NY, USA.

She was a Visiting Student Researcher with the Autonomous and Cyber-Physical Systems (ACPS) Research Group, University of Agder (UiA), Grimstad, Norway. Her research interests

are deep learning, computer networks, the Internet of Things (IoT), computer architecture, computer vision, and embedded systems.



**Srinivas M. B.** (Member, IEEE) received the Ph.D. degree from the Indian Institute of Science, Bengaluru, India, in 1991.

He is currently a Professor with the Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani, Dubai Campus, Dubai, UAE. His research interests include nanoscale circuit design, data converters, and Information and Communication Technology (ICT) for healthcare.

Prof. M. B. was a recipient of the Microsoft Research Digital Inclusion Award in 2006 and the Stanford Medicine MedTech Innovation Award in 2016. He has served as the Chairperson of the IEEE Hyderabad Section from 2007 to 2008 and the Founding Chairman of the CAS/EDS Joint Chapter, Hyderabad Section, from 2012 to 2013.



**Soumya J.** received the bachelor's degree in electronics and communication engineering from Jawaharlal Nehru Technological University, Hyderabad, India, in 2007, and the master's and Ph.D. degrees in electronics and electrical communication engineering from the Indian Institute of Technology Kharagpur, Kharagpur, India, in 2010 and 2015, respectively.

From 2011 to 2012, she was a Scientist "SC" with the Indian Space Research Organisation (ISRO), Bengaluru, India. From 2014 to 2015, she was a Faculty Member with the National Institute of Technology (NIT) Goa, Ponda, India. Since 2015, she has been an Assistant Professor with the Department of Electrical and Electronics Engineering (EEE), Birla Institute of Technology and Science (BITS), Pilani (BITS-Pilani), Hyderabad, India. As a Principal Investigator, she has been implementing several funded projects from Department of Science and Technology (DST), Government of India, and has been collaborating with various research groups in India and abroad. Her research interests led to a credit of more than 45 publications in peer-reviewed journals and reputed international conferences. Her research interests include Network-on-Chip (NoC) design, fault-tolerant system design, NoC and processor verification, Reduced Instruction set Computer- V (RISC-V)-based interconnect design, embedded systems, and real-time systems.



**Om Jee Pandey** (Senior Member, IEEE) received the Ph.D. degree from the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, India, in January 2019.

He worked as a Postdoctoral Fellow at the Communications Theories Research Group, Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada. He is currently an Assistant Professor with the Department of Electronics Engineering, Indian Institute of Technology (BHU) Varanasi, Varanasi, India. His research interests include wireless sensor networks, low-power wide-area networks, unmanned aerial vehicle (UAV) networks, mobile and pervasive computing, cyber-physical systems, the Internet of Things, cloud and fog computing, UAV-assisted optical communications, and social networks.



**Linga Reddy Cenkeramaddi** (Senior Member, IEEE) received the master's degree in electrical engineering from the Indian Institute of Technology Delhi (IIT Delhi), New Delhi, India, in 2004, and the Ph.D. degree in electrical engineering from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2011.

He worked at Texas Instruments, Bangalore, India, on mixed-signal circuit design before joining the Ph.D. Program at NTNU. After finishing his Ph.D. degree, he worked on radiation imaging for an atmosphere-space interaction monitor (ASIM mission to the International Space Station) at the University of Bergen, Bergen, Norway, from 2010 to 2012. He is currently the Leader of the Autonomous and Cyber-Physical Systems (ACPS) Research Group and a Professor with the University of Agder, Grimstad, Norway. He has coauthored over 100 research publications that have been published in prestigious international journals and standard conferences. His main scientific interests are in cyber-physical systems, autonomous systems, and wireless embedded systems.

Dr. Cenkeramaddi is also a member of the editorial boards of various international journals and the technical program committees of several IEEE conferences. He is the Principal Investigator and a Co-Principal Investigator of many research grants from the Norwegian Research Council.