

Research

Wearable bracelet and machine learning for remote diagnosis and pandemic infection detection

Ayah Abdel-Ghani¹  · Amira Abdalla¹  · Zaineh Abughazzah¹ · Mahnoor Akhund¹ · Khalid Abualsaud¹  · Elias Yaacoub¹ 

Received: 4 June 2024 / Accepted: 13 August 2024

Published online: 19 August 2024

© The Author(s) 2024 [OPEN](#)

Abstract

The COVID-19 pandemic has highlighted that effective early infection detection methods are essential, as they play a critical role in controlling the epidemic spread. In this work, we investigate the use of wearable sensors in conjunction with machine learning (ML) techniques for pandemic infection detection. We work on designing a wristband that measures various vital parameters such as temperature, heart rate, and SPO2, and transmits them to a mobile application using Bluetooth Low Energy. The accuracy of the wristband measurements is shown to be within 10% of the readings of existing commercial products. The measured data can be used and analyzed for various purposes. To benefit from the existing online datasets related to COVID-19, we use this pandemic as an example in our work. Hence, we also develop ML-based models that use the measured vital parameters along with cough sounds in order to determine whether a case is COVID-19 positive or not. The proposed models are shown to achieve remarkable results, exceeding 90% accuracy. One of our proposed models exceeds 96% performance in terms of accuracy, precision, recall, and F1-Score. The system lends itself reasonably for amendment to deal with future pandemics by considering their specific features and designing the ML models accordingly. Furthermore, we design and develop a mobile application that shows the data collected from the wristband, records cough sounds, runs the ML model, and provides feedback to the user about their health status in a user-friendly, intuitive manner. A successful deployment of such an approach would decrease the load on hospitals and prevent infection from overcrowded spaces inside the hospital.

Keywords Convolutional neural network (CNN) · COVID-19 · Deep neural network (DNN) · Machine learning (ML) · MHealth · Pandemic infection detection · SpO2 sensor · Wearable devices · XGboost

1 Introduction

Despite the fact that pandemics seem to be rare events, reports have shown that the pace at which pandemics emerge is accelerating. Ever since 1980, more than 35 infectious diseases have appeared, at a rate of one disease every 8 months [1]. The world's recent experience with COVID-19 has demonstrated the impact pandemics have on the human health and the global economy. Moreover, they lead to a huge load on healthcare institutions and affect the services they provide. Figs. 1 and 2 from [2] show the burden healthcare professionals were facing by the number of tests that needed to be performed (Fig. 1), and the number of patients that were being admitted to the hospitals (Fig. 2). In fact, the COVID-19

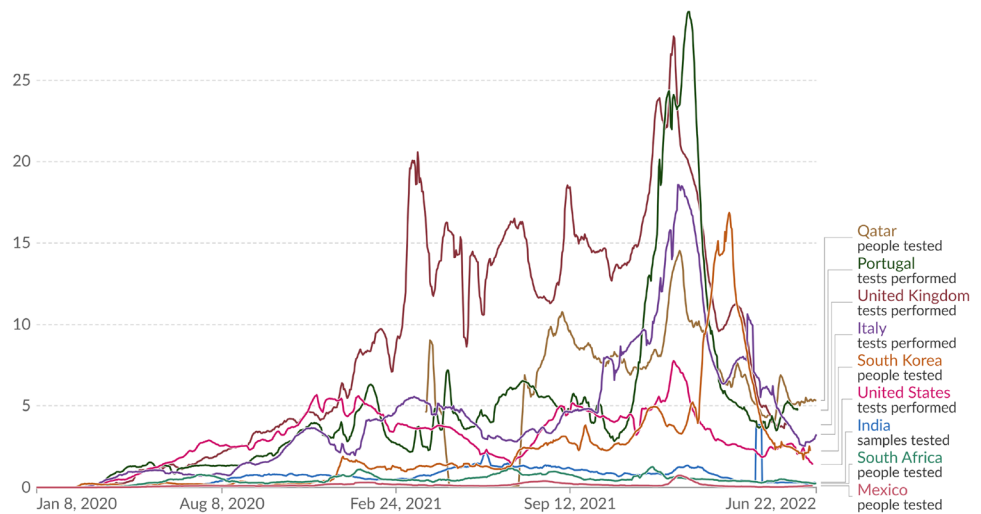
✉ Elias Yaacoub, eliasy@ieee.org; Ayah Abdel-Ghani, aa1901410@qu.edu.qa; Amira Abdalla, aa1802220@qu.edu.qa; Zaineh Abughazzah, za1908270@qu.edu.qa; Mahnoor Akhund, ma1802426@qu.edu.qa; Khalid Abualsaud, k.abualsaud@qu.edu.qa | ¹Computer Science and Engineering Department, Qatar University, Doha, Qatar.



Fig. 1 Daily COVID-19 tests per a thousand people [2]

Daily COVID-19 tests per 1,000 people

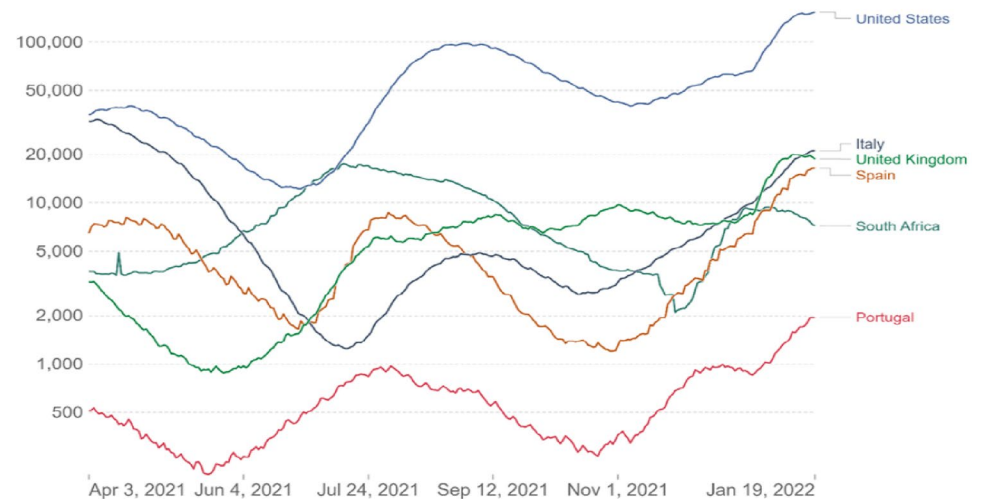
7-day rolling average. Comparisons across countries are affected by differences in testing policies and reporting methods.



Data source: Official data collated by Our World in Data - Last updated 10 April 2024
 Note: Our data on COVID-19 tests and positive rate is no longer updated since 23 June 2022.

Fig. 2 Number of COVID-19 patients in hospitals [2]

Number of COVID-19 patients in hospital



Source: Official data collated by Our World in Data - Last updated 18 September 2022

pandemic has highlighted the need for effective early infection detection methods, as they play a critical role in controlling the epidemic spread. Existing infection detection methods, like Polymerase Chain Reaction (PCR), and Rapid Antigen (RAT) tests, and diagnostic procedures involving blood samples, X-rays, and CT scans, require individuals to visit the hospital. This leads to overcrowded waiting areas in healthcare facilities, which increases the infection rate and poses a health risk on patients. Therefore, a more effective infection detection method is required to deal with the influx of patients that arrive in hospitals, and simultaneously reduce the number of extra infections caused as a result of the packed hospitals. A remote preliminary diagnostic test using common symptoms like Cough Sounds, Body Temperature, Heart Rate, and Saturation of Peripheral Oxygen (SpO₂) levels can be created to determine whether a person needs to visit the hospital. This helps individuals to determine whether they are suspected of an infection, and only visit hospitals in this case, consequently reducing the load on healthcare facilities.

Studies have shown the importance of using Artificial Intelligence in the medical field, as it can efficiently analyze medical data and improve digital technology [3]. In our previous work [4], we present a solution using Machine

Learning (ML) algorithms to help in the early detection of coronavirus (COVID-19) using patient cough sounds. Patient symptomatic data is also collected through a wristband, including Body Temperature, Heart Rate, and SpO₂ levels. This data is continuously transmitted to the Mobile Application via BLE technology for analysis. In this paper, we present a generalized approach that uses ML algorithms by jointly including the vital parameters data with the cough sounds in the classification process. The ML algorithms are used to identify any abnormalities in the readings to alert the user to take a cough test or visit a doctor for further evaluation.

Thus, the contributions of this paper can be summarized as follows: (i) Implementing a wristband that uses wearable sensors to measure vital signs parameters; (ii) Deriving ML models that use cough sounds and vital signs measurements to detect infection, and showing that they perform well compared to the existing literature; and (iii) Designing a mobile application that collects the wristband's readings, allows to user to perform a cough test, and implements the ML models to detect infection and inform the user accordingly.

The remainder of the paper is organized as follows: Sect. 2 reviews the existing literature, Sects. 3 presents our proposed solution and explains our work, and 4 showcases the obtained results. Finally, the conclusion is drawn in Sect. 5, and future research advancements are explored.

2 Literature review

This section presents an overview of the most relevant works in the literature, and outlines the novelty of the proposed approach.

2.1 Mobile health technologies

The integration of wearable sensors and intelligent applications, exemplified by research such as that conducted by [5], marks a significant advancement in healthcare. This amalgamation enables the creation of innovative detection methods capable of identifying various diseases based on common symptoms. The early detection facilitated by these methods is vital for prompt medical intervention, crucial in curbing the spread of infectious diseases and potentially saving lives. Recent studies underscore the significant impact of these technologies on enhancing healthcare practices.

Currently, there are a lot of wearable sensors on the market that are used to track important health metrics. Fitness trackers, exemplified by Fitbit and Garmin Vivosmart as examined in the study by [6], wearable Electrocardiography (ECG) monitors like the Apple Watch (with ECG feature) enable continuous monitoring, detecting irregular heart rhythms as mentioned by [7], and smartwatches, like those covered in the study by [8], are a few instances of mHealth technology available in the market. Moreover, several studies have explored wearable medical-grade sensors, including remote patient monitoring systems as discussed in [9]. Due to developments in wearable sensor technology, people may now continuously monitor health metrics outside of conventional clinical settings, giving them more autonomy to take charge of their health and well-being.

An important symptom indicating a COVID-19 infection is the low percentage of oxygen in the blood. SpO₂ levels can be measured using pulse oximeters found in wearable sensors, and readings that are equivalent to or below 92% indicate that the person might be infected by the disease and needs to urgently visit the hospital [10]. Efforts were made to create a pulse oximeter in contact-tracing devices like Samsung, Fitbit, and Apple watches, however, the readings are unreliable due to poor contact between the wristband and the watch, and continuous wrist movements [11]. Therefore, more focus is put into the reading of heartrate in these devices than the SpO₂ levels.

Authors in a recent work [12] implement a wristband prototype that gives accurate measurements for blood oxygen levels. The MPU9250 Inertial Measurement Unit (IMU) motion sensor is used to track the wrist rotation and the acceleration of the wrist. The values measured are then used to decrease the error rate of the SpO₂ measurements, creating a reliable blood oxygen level monitoring device. We refer to the authors' work in our paper, and we build on it by including other vital parameters and creating a connection with the mobile application via BLE, as will be explained later in Sect. 3.1.

2.2 Sensor connectivity

A Near Field Communication (NFC) ring and an Android application are used in the Rose-Hulman study Institute of Technology to show the viability of a Wireless Body Area Networks (WBAN) system [13]. A system prototype was created and tested as proof of concept. One of the sensors used in the system's implementation, the TCRT1000, transforms the heart rate data it receives into an electrical output. A microcontroller (MSP430G2553) receives the signal after it has been filtered and amplified. The microcontroller calculates the timing difference between two successive pulses. The NFC tag is written with the computed value. An Android phone's application launches and reads the data from the NFC tag when the phone is in the tag's transmission range. After that, the received data is put into a local SQLite database and processed further to create trends.

In our work, we benefited from the insights provided by the research paper [13] while diverging from its approach by adopting BLE technology instead of NFC. By incorporating BLE, we capitalized on its numerous advantages, including extended range and compatibility with our wristband design. This decision allowed us to overcome challenges encountered with NFC components.

2.3 Cough and symptom data classification

Various types of machine learning techniques are utilized for cough classification and detection of COVID-19. The authors in [14] display the techniques and findings of many researchers in the case of Cough classification and respiratory disease detection. The results of 10 papers that perform COVID-19 diagnosis with cough are shown in a table, and different models, such as Support Vector Machine (SVM), Deep Neural Network (DNN), and Convolutional Neural Network (CNN), and DTL-BC (Deep Transfer Learning-based Binary Classifier) are used. The results shown range from 68% to as high as 98.5%. In [15], a paper included in the comparison, the authors use a DNN model to classify Cough Sounds and other Symptom data. Good results are achieved, being [Precision: 95.04%, Recall: 90.1%, Accuracy: 96.83%]. The authors dataset consists of around 30,000 audio segments collect from 150 patients. The features extracted from the audio segments include: Mel Frequency Cepstral Coefficients (MFCCs), Log Energy, Zero Crossing Rate, Skewness, Entropy, Format Frequencies, Kurtosis, and Fundamental Frequency. The other symptom data used include the common symptoms like Fever [1, 0], Difficulty Breathing [1, 0], etc. In our work, one of the experiments will consist of extracting the MFCC features from the audio files, as will be explained later on in Sect. 3.2.5.

Other researchers used different types of models, including CNNs and SVMs [16–19]. In [16] and [17], CNNs are used to train on Cough sounds only. In [16], the CNN is made of two 2D Convolutional Layers with Max Pooling 2D and are trained on the image representation of 1,811 Cough samples. In [17], pre-processing is performed on the 5,320 audio samples and passed to a biomarker to be later passed to the CNN model. The CNN model created is made of three parallel Residual Neural Networks-50, and the output of the three networks is pooled using a 2D Global Average Pooling layer.

On the other hand, [18] and [19] used SVMs to train on the Cough samples. The authors in [18], an SVM is used to train on 4 different classes, including Normal, COVID-19, Pertussis, and Pneumonia. Log-based Mel-Spectrograms of the Cough sounds are used to compute a vector (m) which is passed to the SVM. In [19], both Cough sounds and symptoms are used for the classification. From the Cough sounds, 16-framelevel descriptors are extracted, including: Zero-Crossing-Rate, Root Mean Square frame energy, Pitch Frequency, Harmonics-to-Noise Ratio, and 12 Mel-Frequency Cepstral Coefficients. The features are combined with the most common COVID-19 symptoms, specifically: chills, dizziness, dry cough, fever, headache, muscleache, short of breath, loss of smell/taste, sore throat, tightness in chest, and wet cough. A vector is used to indicate the presence or absence of these features using [1, 0]. Data augmentation using Synthetic Minority Oversampling Technique (SMOTE) is performed to balance the dataset, and the SVM was trained on the 828 resulting samples.

In [20], the authors test with multiple model structures, including, CNN, Long Short-Term Memory (LSTM), Hybrid CNN-LSTM, and Attend-based Hybrid CNN-LSTM. The dataset used is the COUGHVID Crowdsourcing Dataset [21], also used in our work. The researchers filter out the dataset to keep only the label cough segments, remove silent sections in the audio segments, and keep only the cough segments with a higher probability of a cough. Since the dataset classes are unbalanced, the researchers perform Frequency and Time masking to increase the number of samples in the smaller "COVID-19" class. 10-fold Cross Validation is used to train the models, and hyper-parameter

tuning is performed, and the best results are achieved by the Hybrid CNN-LSTM model structure. The authors work included only cough sounds, and no other symptom data were used. In our work, a similar model structure is used in the Initial Approach, and in the first experiment of the Second Approach, as will be explained later in detail.

2.4 Novelty of the proposed approach

Most of the previous works either consider cough sounds only, or vital parameters only to detect infection. The few works that consider both, e.g., [19], achieve relatively low accuracy, as shown in the comparisons of the results section. In this work, we provide a more complete diagnosis by considering both cough sounds and vital parameters, while designing ML models that achieve significantly better results than the literature. Moreover, we provide a complete solution based on a prototype that uses sensor measurements from a wearable wristband, in conjunction with a mobile application that displays these measurements and allows the user to perform a cough test. Indeed, existing commercial wearable devices provide some vital parameter readings, without providing the capability, in their accompanying mobile apps, to use cough sound testing with ML to detect infection.

As mentioned above, we had previously developed a machine learning model that can diagnose a person's health using their cough sounds [4], and achieved high results of around 92%. Moreover, we developed a basic prototype of a wristband with a MAX30102 sensor to collect the vital parameters of a user and process them to be sent to the mobile application. Building on that, the main contributions of the work mentioned in this paper can be summarized as follows:

- Develop a Machine Learning model that can detect pandemic infections from the user's vital parameters, including Body Temperature, Heart Rate, and SpO2 Levels.
- Develop a single Machine Learning model that can diagnose a case as either COVID-19 or Healthy by utilizing the two different infection symptom types; the user's Cough Sounds (audio data), and the three main Vital parameters: Body Temperature, Heart Rate, and SpO2 Levels (numbers).
- Develop a complete efficient diagnosis system, including a mobile application that allows the user to record their cough and is dedicated to the wristband to check a person's health from the data collected. The application provides feedback to the user using colors that symbolize their health state (Green: Healthy, Red: Immediately visit the Hospital).
- Develop a wristband prototype that has a sensing unit to collect the user's vital parameters and can process the data (microcontroller) and transmit it continuously to the application using BLE.
- Provide a diagnosing method that individuals can use to self-monitor their symptoms during a pandemic, and a solution for medical facilities to reduce the burden of in-person infection testing and patient visits.

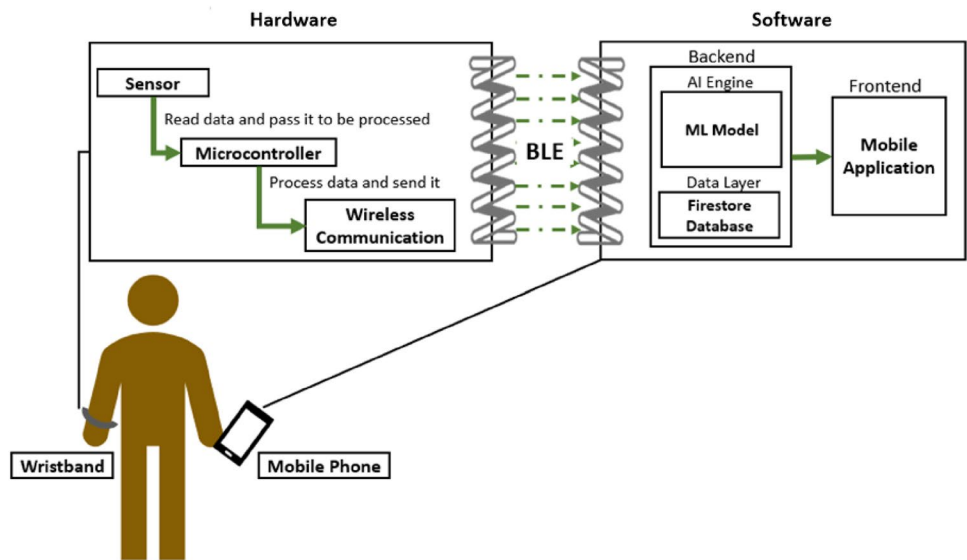
3 Proposed solution

In this work, we aim to create a system that can be used to conduct remote diagnostic tests to detect infections during a pandemic. We apply our concept on the latest pandemic COVID-19 due to the availability of a large open-source crowdsourcing dataset COUGHVID [21] and a vital parameter dataset [22]. The system we create is a complementary tool to avoid pressure on healthcare facilities as the symptoms are tested remotely without the need for a COVID-19 RAT or PCR test.

The system is divided into two main sections, a Hardware section which consists of the Wristband, and a Software section which includes the Mobile Application as the Frontend, the Machine Learning model, the Database, and the Server as the Backend. To perform the diagnostic test, 4 symptoms are collected from the user, the Cough Sound, Body Temperature, Heart Rate, and SPO2 levels. The first symptom is collected through the Mobile Application through an interface that allows the user to record their cough while testing. The remaining three symptoms are collected through the wristband, processed, and sent to the backend server. When all the data is collected, the classification process begins, and the result is displayed to the user. A high-level architecture of the system can be seen in Fig. 3.

In our work, we create the Machine Learning model using two approaches. Initially, we create two separate machine learning models for each data type (Cough Sounds—Vital Parameters), however, we work on a more reliable diagnostic method in the second approach. Our final system uses the model created in the second approach. More details about the two approaches are explained in Sect. 3.2.

Fig. 3 The structure of the proposed system



3.1 Hardware

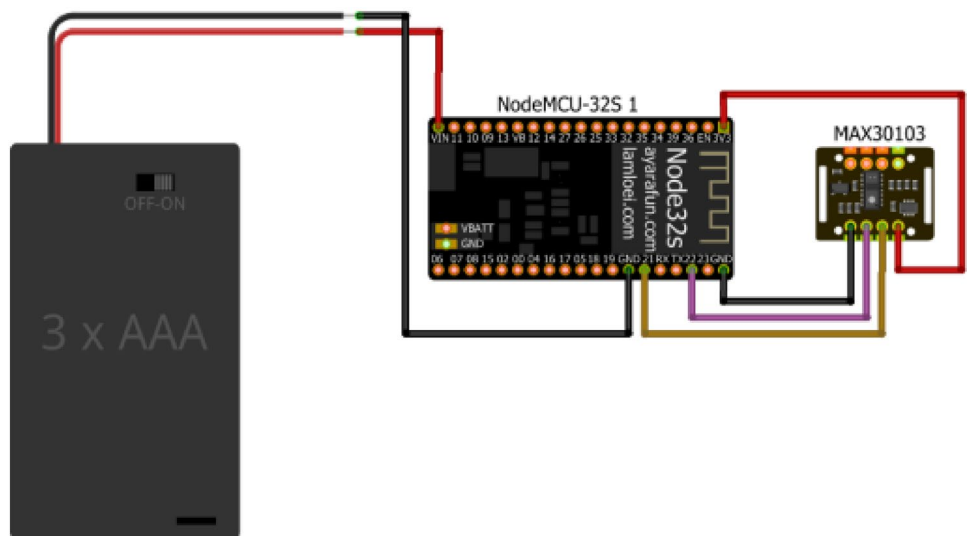
3.1.1 Components

The system’s hardware comprises three primary parts: the MAX30102 sensor, the Arduino Nano microcontroller (NodeMCU ESP32S), and the BLE protocol-based wireless communication subsystem. The modularity of the system has been thoughtfully considered throughout its design. The following circuit diagram, Fig. 4, illustrates the connections between these parts. The block MAX30102, which is in charge of gathering physiological data, is highlighted in the diagram. This module measures body temperature, heart rate, and SPO2 levels by integrating photo-detectors and LED emitters.

The data obtained from the MAX30102 sensor is then passed to the NodeMCU ESP32S block which serves as the system’s central control unit. The NodeMCU ESP32 handles the processing tasks related to the received sensor data. It processes the data and performs necessary algorithms or calculations to extract meaningful information.

A BLE is responsible for establishing wireless communication and allows data to be wirelessly transferred to the mobile application. This modular architecture enhances flexibility for future iterations and allows components to be

Fig. 4 Circuit diagram of hardware system created in [23]



easily updated. The MAX30102 sensor communicates with the NodeMCU ESP32 through the Inter-Integrated Circuits (I2C) protocol, and the wristband is carefully designed for comfortable and efficient system operation. The prototype design for the wristband in our system is shown in Fig. 5 below.

3.1.2 Comparing our wristband with existing devices

In our effort to assess the efficacy of our wristband, we perform a comparison between our wristband and the Huawei Watch 3 Pro, both adept at conducting measurements which are body temperature, heart rate, and SPO2. This comparative analysis serves as a litmus test to gauge the performance and accuracy of our wristband to a commercial device available in the market. By comparing both devices' functionalities, precision, and user experience, we aim to ascertain strengths and areas for improvement of our wristband, ensuring that it meets the expectations set by existing industry standards. We also compared it with the Fitbit smartwatch, which only measures heart rate. Through this rigorous examination, we endeavor to affirm our commitment to delivering cutting-edge technology that competes and innovates within the health wearable devices landscape.

3.2 Machine learning model

Our initial approach was to create a separate model for each datatype; a CNN model to classify the Cough Sounds recorded by the user, and a simple DNN model to classify the vital parameters sent from the wristband. Good results were achieved on both models, however, the diagnosing method of combining the results of two different models is not reliable. To create a single model for both data types, there should be a dataset that contains samples with Cough Sounds and Vital Parameters. Due to the unavailability of such a dataset, we create a new one by combining two existing datasets and training a single model to take both types of data and make a more efficient classification.

As explained in our work in [4], we convert the audio dataset (cough sounds) into mel-spectrograms. Since the dataset is unbalanced, we perform Pitch-Shifting and Spectral-data augmentation (Frequency and Time Masking) to create more samples that are slightly different from the original ones. We prefer this method over undersampling and oversampling since it prevents the model from overfitting on the data or under-learning.

In the initial approach (two separate models), there was no need to balance the samples of the two data types since each model would be trained separately. Augmentation was performed on the Cough dataset since the classes "Healthy" and "COVID-19" were unbalanced, and undersampling was performed on the "Healthy" class of the Vital Parameters dataset to match the number of samples in the "COVID-19" class.

Since in our second approach we are creating a new dataset by combining the COUGHVID Crowdsourcing Dataset and the Vital Parameters (Body Temperature, Heart Rate, and SpO2 Levels) from the COVID-19 Clinical Data Repository [22],

Fig. 5 Wristband prototype design



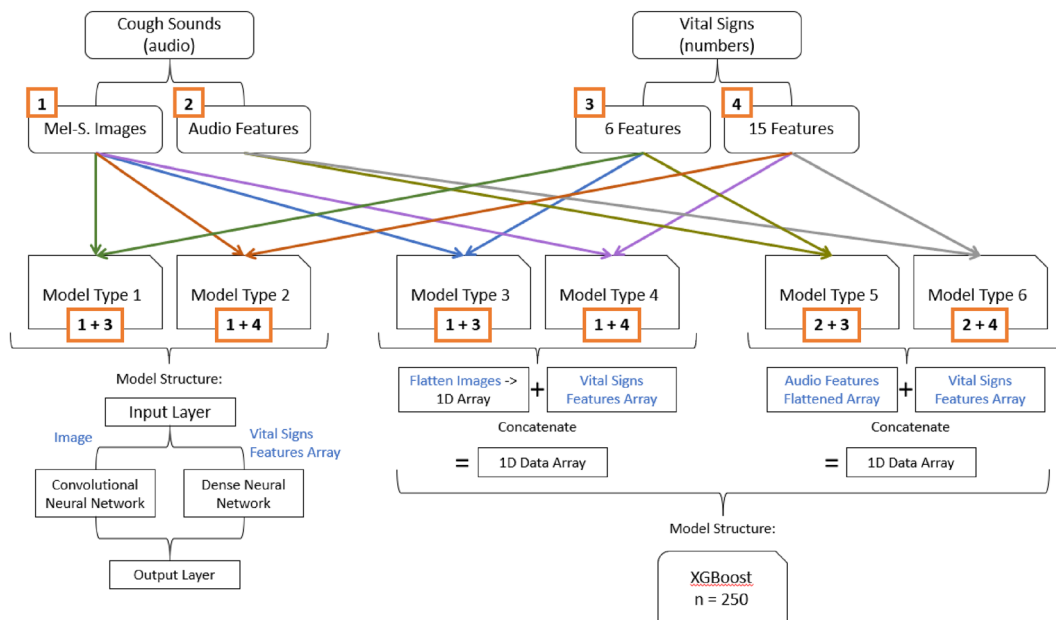


Fig. 6 Experiments conducted in this paper

Table 1 Cough dataset: images count before and after balancing

Class	Count before	Count after	Total
Positive	5,332	15,996	33,912
Negative	8,958	17,916	

Table 2 Cough dataset: audio count before and after balancing

Class	Count before	Count after	Total
Positive	5,332	15,996	33,912
Negative	8,958	17,916	

Table 3 Vital parameters dataset: vitals count before and after balancing

Class	Count before	Count after	Total
Positive	1,045	19,855	65,131
Negative	45,275	same	

more augmentation and balancing steps are needed. We perform 6 main experiments as can be seen in Fig. 6. Some of the experiments require the Image Representation of the Cough Sounds, while others require the original audio format of the data. Since the Data Augmentation was done on the Image representation, the Audio and Vital Parameters dataset are unbalanced. Augmentation is performed on the Positive Cough Sounds and Positive Vital Parameters. Details of the Datasets used for the 6 experiments are shown in Tables 1, 2, and 3.

3.2.1 Balancing the vital parameters dataset

The first step in our work is to balance the classes in the COVID-19 Clinical Data Repository. The Dataset provides many patient health data, including Test Results, Epi Factors, Comorbidities, Vitals, Symptoms, Radiological Findings, and more. In our work, we only need some values from the vital parameters, which are the Body Temperature, Heart Rate, and SPO2 levels.

After extracting the required data, the data was cleaned, and only the samples with the three values available (Body Temperature, Heart Rate, and SPO2 levels) are kept. The number of samples after cleaning are 1046 Positive samples and 45,275 Negative samples, as can be seen in Table 3. The Negative class has more samples than needed, however, the Positive class needed more samples, and therefore, augmentation was performed.

Two of the common symptoms of COVID-19 are (1) High Body Temperature, and (2) Low Oxygen Levels. Using this information, we created more versions of each Positive Sample by adjusting the Body Temperature, Heart Rate, and SPO2 levels. For the Temperature values, we created more samples by adding each of the following three values once [0.1, 0.2, 0.3]. The SPO2 levels, which are whole numbers, were augmented using the following three values [-1, -2, -3], while the Heart Rate was augmented with [1, -1]. We will call these values the Augmentation Values.

Combinations of the Augmentation Values are created, and are used to create more samples from each Positive Sample. 18 combinations can be made with the Augmentation Values, therefore, the number of samples after augmenting the Positive class is (18 + the original sample = 19) $1045 \times 19 = 19,855$. Since the total new size of the dataset is 65,131 samples, we already acquired enough samples to create combinations with the Cough Dataset, and no more balancing is needed.

3.2.2 Balancing the cough dataset-audio

The Image version of the Cough Dataset is already balanced from our previous work in [4], with 33,912 samples in total. As mentioned previously, Pitch Shifting on the Audio Data is performed, and Frequency and Time Masking are performed on the Mel-Spectrograms.

Since we are using the audio format of the data, we already have the first part of the augmentation done - the Pitch Shifting. However, more data augmentation is required to create more Positive samples. The **Count Before** column in Tables 1 and 2 show the number of samples in the classes after the Pitch Shifting was performed.

Frequency and Time masking can only be performed on spectrograms (Frequency and Time axes). Therefore, we apply Time Masking on the audio data by inserting a 1-second long silence at a random location in the audio file. The overall audio file length stays the same. For the each Negative sample, we perform the Time Masking only once, while we do it twice for each Positive sample. The total number of samples in the dataset is the same as the total number of samples in the Mel-Spectrogram Dataset, which is 33,912 samples.

3.2.3 Creating combination datasets

We create two main Combined Datasets, the Mel-Spectrogram—Vitals dataset, and the Audio Features—Vitals dataset. A script was created to pair every Positive Image/Audio file with a vital parameter sample randomly while checking that no duplicate combinations are created. The resulting datasets are the same size as the Mel-Spectrogram and Cough Audio datasets. No Cough sound, Mel-Spectrogram, or vital parameter was used twice.

The number of samples in the final Combined datasets are the same as the final samples in Tables 1 and 2.

3.2.4 Extracting additional statistical features from the vital Parameters

In [4], we extracted 6 main features from the vital parameters—Body Temperature, Heart Rate and SpO2, including: ['mean', 'minimum', 'maximum', 'rms', 'std', 'skew']. In this work, we train each model twice, once with the 6 features mentioned, and once with 9 additional features, including: ['range', 'median', 'variance', 'percentile_25', 'percentile_50', 'percentile_75', 'iqr', 'cv', 'kurt'], as can be seen in Fig. 6.

The features are not previously extracted and stored. During the model training process, the required features are extracted and fed into the model. Therefore, we have one copy of the vital parameters samples in the Combined Datasets, and we extract 6 or 15 features based on the experiment.

3.2.5 Explaining the data structures used in training

Three main experiments are performed in this work, and each experiment is conducted once with the six vital parameter features, and once with the 15 vital parameter features, as mentioned previously. The three main experiments can be explained as follows:

1. *Experiment 1:* [Model Type 1, Model Type 2] The Cough Mel-Spectrogram Combined Dataset is used in this experiment. The Images are kept in their Image format, and a Model is built by creating two sub models that have a common Input Layer and a common Output Layer. The input is split to the two models; the first being a CNN model to study the Images, and the second a Dense Neural Network to study the vital parameters. The output of the two sub models is concatenated and passed to the Output Layer.
2. *Experiment 2:* [Model Type 3, Model Type 4] The Cough Mel-Spectrogram Combined Dataset is used in this experiment. The Images are flattened, and the vital parameters are concatenated to the end of the Image array to create one long array of data. The Model created is an XGboost model.
3. *Experiment 3:* [Model Type 5, Model Type 6] The Cough Audio Combined Dataset is used in this experiment. Features are extracted from the audio files. The features matrix is flattened and the vital parameters data is concatenated to the end of the array. XGBoost model is used, similar to Experiment 2.

Table 4 shows the naming convention that we will be using in the rest of the paper.

3.2.6 Model structures

For the initial approach, the model used to classify the users' cough sounds is a CNN model. As explained in [4], the audio cough sounds are converted to Mel-Spectrograms, converting the audio dataset into an image dataset. A hybrid CNN-LSTM model is used, where the CNN part of the model is made of 4 Convolutional Layers with Average Pooling, Batch Normalization, ReLU Activation, and Dropout, while the LSTM section has 256 units with Dropout. The model was trained for 100 epochs with a batch-size of 256, a learning rate of 0.001, and an Adamax Optimizer. The model used for the vital parameters classification is a simple 2-Dense Layers model with ReLU Activation and Regularization, and was trained for 20 epochs. Details of the model structures used in the first approach can be found in Fig. 7.

As explained in Sect. 3.2.5, we have three main experiments in our second approach, with a total of six models created. The first experiment type Image-Vital Parameters was trained for 100 epochs with 0.01 learning rate, batch-size 128, and an Adamax Optimizer. In this experiment, the model was divided into two parts with common Input and Output Layers. The first part of the model used for training on the cough images has the same model structure as the one used in the Initial Approach explained before. The second part of the model for the vital parameters is a simple DNN model of three Dense Layers.

Experiment Flattened Images-Vital Parameters uses an XGBoost Classifier with 600 estimators and 5 as the maximum depth, while experiment Audio Features-Vital Parameters uses an XGBoost Classifier with 1000 estimators, 5 as the maximum depth, and no-limit for the number of leaves. For experiment Images-Vital Parameters, the data was divided into 75% Training, 10% Validation, and 15% Testing. As for experiments Flattened Images-Vital Parameters and Audio Features-Vital Parameters that use an XGBoost model, the data was divided into 85% Training and 15% Testing.

4 Results

In this section, we will show the results we received after comparing our wristband with existing products and the results from training our models.

4.1 Comparing wristband readings

We have generated three graphs that showcase the performance metrics and readings obtained from our wristband and Huawei watch. One of the figures also includes heart rate measurements obtained from a Fitbit smartwatch. We consider two metrics to compare the performance of our wristband to these commercial products: (i) The average absolute difference and (ii) the mean absolute error.

Given two sets of N sample values a_n and b_n , $n = 1, \dots, N$, the average absolute difference is given by:

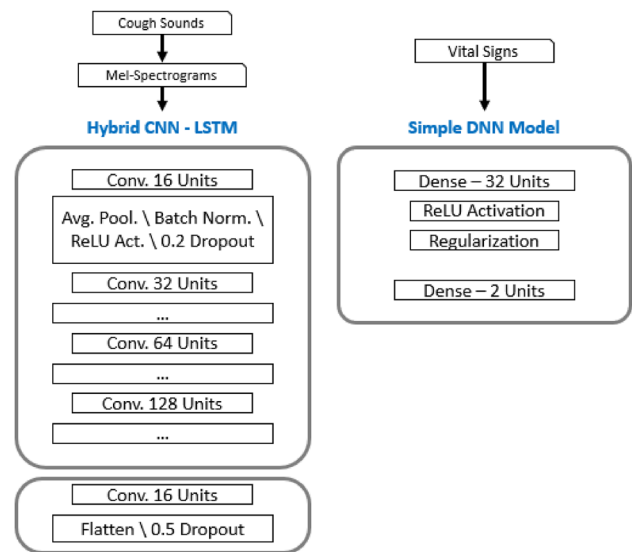
$$\text{AAD}(a, b) = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N |a_i - b_j| \quad (1)$$

Moreover, the mean absolute error is given by:

Table 4 Combined Datasets and Experiments Naming Convention

Type	Name	Description
Datasets	Cough Mel combined dataset	COVID-19 Crowdsourcing Dataset converted to Mel-Spectrograms, and combined with the vital parameters samples from the COVID-19 Clinical Repository.
	Cough audio combined dataset	COVID-19 Crowdsourcing Dataset kept in its audio ".wav" format and combined with the vital parameters samples from the COVID-19 Clinical Repository.
Experiments	Image-vital parameters	Cough Mel Combined Dataset used. The Mel-Spectrograms are kept in their Image format and passed to the model. Model has a common Input Layer, splits into CNN and DNN, then concatenates into a common Output Layer.
	Flattened images-vital parameters	Cough Mel Combined Dataset used. The Mel-Spectrograms are flattened, and vital parameters are concatenated to the end of the array. A single XGboost model is created.
	Audio features-vital parameters	Cough Audio Combined Dataset is used. MFCC features are extracted from the cough samples and flattened. Vital parameters samples are concatenated to the end of the array. A single XGBoost model is created.

Fig. 7 Model Structures for the Initial Approach



$$MAE(a, b) = \frac{1}{N} \sum_{n=1}^N |a_n - b_n| \tag{2}$$

Figure 8 shows a notable similarity between the curves, indicating consistent temperature readings. Both devices exhibited nearly identical curve shapes, providing accurate readings ranging from 33.5 to 35.8°C. It shows less than natural temperature measures because they are both wrist-worn devices. To delve deeper into our analysis, we calculated the average absolute difference between our device and the Huawei watch, which amounted to 0.5%. This indicates a minimal deviation between the two sets of readings, highlighting the reliability of our device in capturing temperature data. Furthermore, by computing the mean absolute error (MAE) across data points, we obtained a quantitative measure of the discrepancy between the devices' readings, which was calculated to be 0.2. This value confirms that our model works well and is reliable.

Examining the graph depicted in Fig. 9, we observe fluctuations in SpO2 readings captured by both our wristband and the Huawei watch. Our wristband occasionally registers slight decreases before stabilizing at levels comparable to the Huawei watch. However, unlike the Huawei watch, which requires manual readings and minute-by-minute monitoring, our device seamlessly provides continuous, automated readings. The average absolute difference between the measurements of the two devices being 2.9% indicates a relatively small deviation between their readings. Following this, we

Fig. 8 Temperature Comparison: Our Wristband vs. Huawei

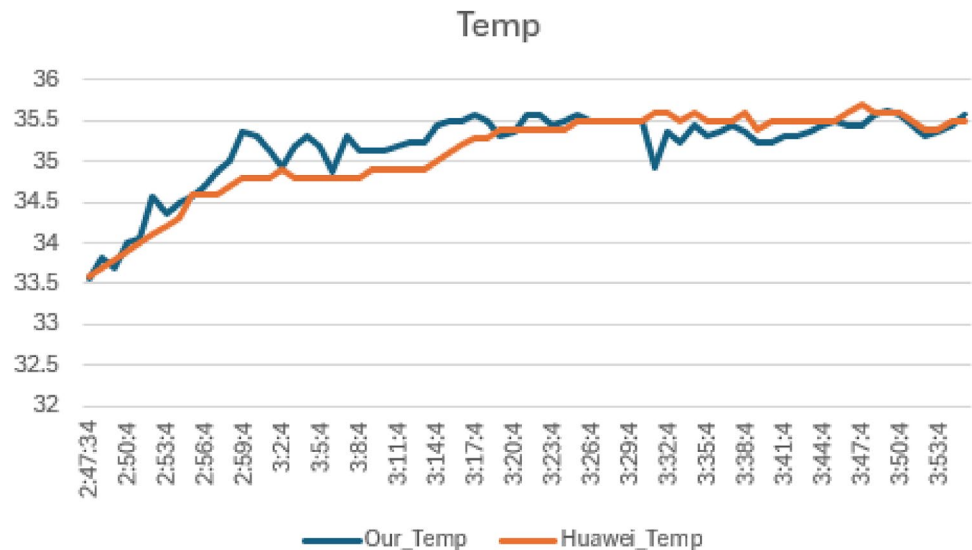
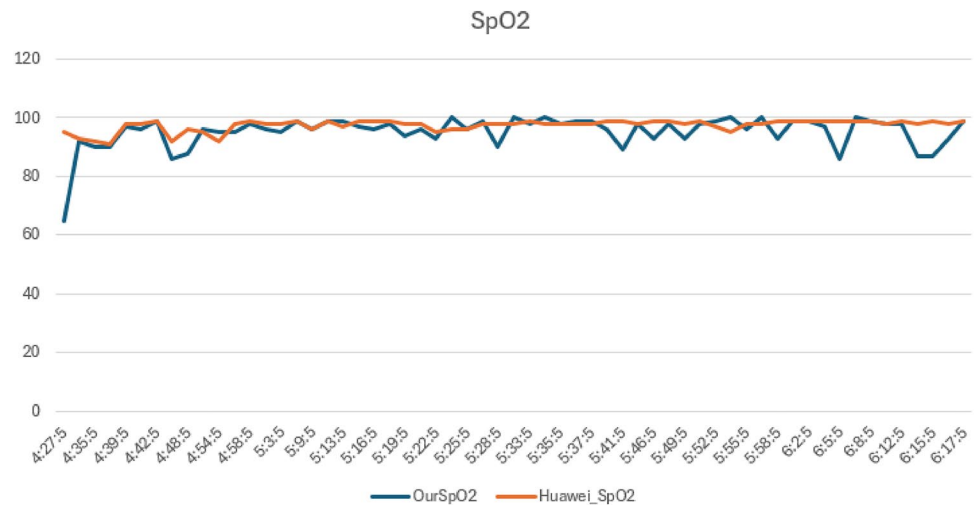


Fig. 9 SpO2 Comparison: Our Wristband vs. Huawei



calculated the MAE across all data points, which resulted in a calculated MAE value of 2.7. This value further emphasizes the consistency and accuracy of our wristband’s SpO2 measurements compared to the Huawei watch.

Finally, distinctive observations were made when comparing heart rates, as depicted in Figs. 10 and 11 below. In Fig. 10, we exclusively focused on the wristband and Huawei watch readings due to their automatic and continuous monitoring capabilities. Both devices’ curves exhibit a zigzag pattern, with simultaneous rises and falls, indicating similar trends.

In Fig. 11, readings from all three devices were captured for a shorter duration. This discrepancy arose because the Fitbit smartwatch necessitates manual readings and monitoring, unlike the other devices. Consequently, a significant discrepancy of 7.7% was observed between the readings obtained from Huawei and our wristband and 8.7% between the Fitbit and our wristband.

Subsequently, through the computation of the MAE, we derived a comprehensive assessment of the disparity between the readings recorded by the Huawei and Fitbit devices relative to our wristband. The resulting MAE values were 7.2 and 8.1, respectively. These figures signify a notable deviation in comparison to their prior measurements, primarily attributed to the reduced number of measurements and the necessity of manual readings on these devices.

Fig. 10 Heart Rate Comparison: Our Wristband vs. Huawei

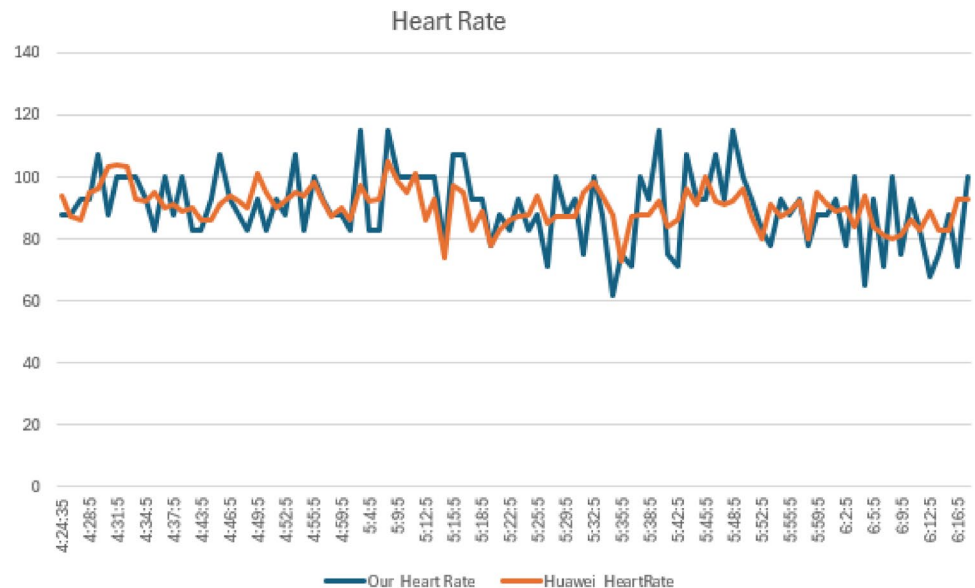
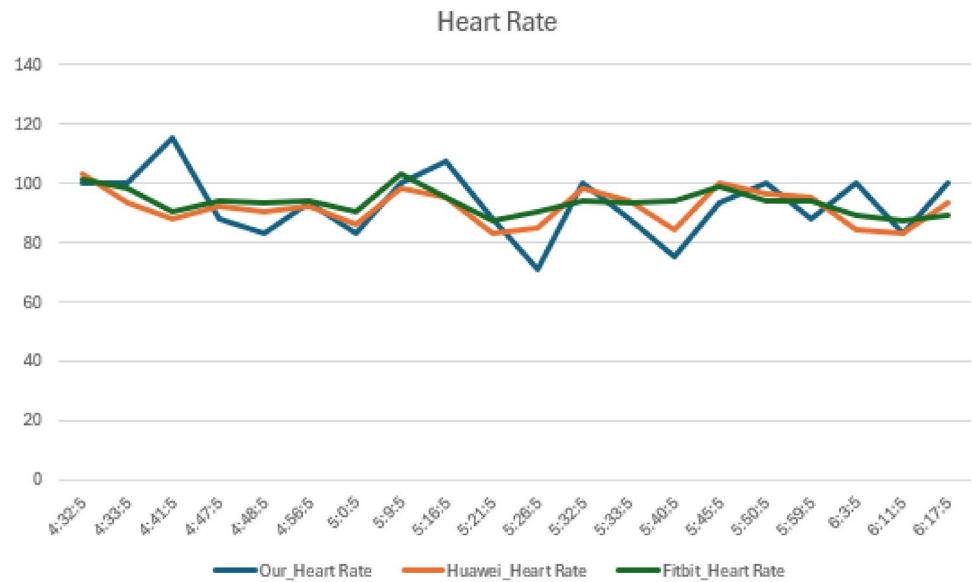


Fig. 11 Heart Rate Comparison: Our Wristband vs. Huawei vs. Fitbit



4.2 Model results

The testing results of the models are shown in Tables 5 and 6. Table 5 shows the results of the models created in our first approach. The Cough Model achieved high results of around 92% in Precision and around 91% in Recall and F1-Score. The vital parameters model achieved even higher results, at around 99% in Precision and F1-Score, and 100% in Recall.

From our second approach, the models with the best performances are the ones trained with the MFCC features from the audio format of the cough sounds. Our results are compared with the model results obtained by the researchers mentioned in Sect. 2. Generally, the best results obtained in our work are from the Audio Features - Vital Parameters 15 Features. Since we used an XGBoost Classifier, we cannot directly compare our results with the results from the other papers. However, if we are to compare the Experiment 3 models with the DNN model from [15], we used fewer cough features (only MFCC, while they included more features) and obtained better results with our method. Our models from

Table 5 Testing results for the models in the first approach

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Cough model	92.00	91.77	91.22	91.49
Vital parameters model	99.52	99.05	100.0	99.52

Table 6 Testing results for the models in the second approach

Model	Accuracy	Precision	Recall	F1-Score
<i>Experiment 1</i>				
Images-Vital Parameters 6 Features	90.00%	90.25%	89.34%	89.79%
Images-Vital Parameters 15 Features	88.00%	90.15%	85.27%	87.64%
<i>Experiment 2</i>				
Flattened Images-Vital Parameters 6 Features	90.00%	90.76%	87.51%	89.10%
Flattened Images-Vital Parameters 15 Features	90.00%	91.67%	86.94%	89.24%
<i>Experiment 3</i>				
Audio Features-Vital Parameters 6 Features	96.00%	96.96%	95.40%	96.18%
Audio Features-Vital Parameters 15 Features	97.00%	96.73%	96.08%	96.40%
[15] DNN	95.04%	90.92%	90.41%	90.09%
[16] CNN	79.00%	77.50%	69.00%	70.5%
[17] CNN	97.00%	–	98.50%	–
[19] SVM	–	–	68.00%	–
[18] SVM	93.75	90.00%	93.75%	88.75%

Experiment 1 also used a CNN model to classify the Image part of the data, and high results were achieved. The models from Experiment 2 achieved good results as well, however, the best results come from the models where the cough features are extracted directly from the audio files, and not converted to images first. In this case, Feature Extraction performed better than Representation Learning.

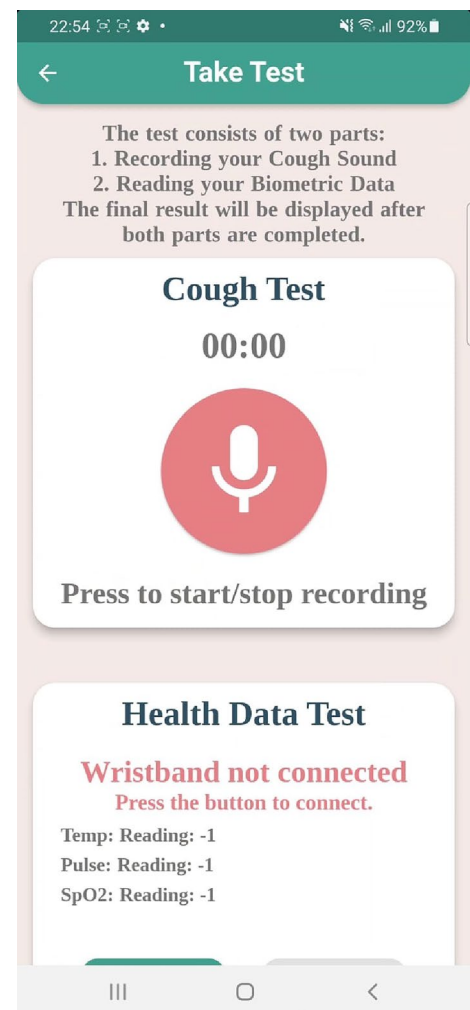
From Table 6, it can be seen that not all contributions in the literature consider a comprehensive evaluation of all of the metrics (i) accuracy, (ii) precision, (iii) recall, and (iv) F1-score. Those who do, e.g., [15, 16], and [18], are significantly outperformed by the proposed approach, especially the results of Experiment 3. Although [17] has a slight advantage in terms of recall, it should be noted that most of the previous works either consider cough sounds only (including [17]), or vital parameters only to detect infection. The few works that consider both, e.g., [19], achieve relatively low accuracy, as shown in Table 6. Thus, the proposed approach achieves excellent performance by combining cough sounds and vital parameters, while having reasonable complexity for the machine learning models.

4.3 Integrating the whole system with a mobile application

The final system created contains a Wristband, a Mobile Application, and a Server, as previously mentioned. The Mobile Application allows the users to monitor their health, and test for a COVID-19 infection by recording their cough and measuring their vital parameters. A full test can be performed using the "Take Test" functionality in the Mobile Application, as can be seen in Fig. 12.

The user records their cough and then connects the wristband to the application to read their vital parameters. The data recorded is sent to the Server to be classified using the Machine Learning model created. The Server will return the

Fig. 12 Take test screen in the mobile application

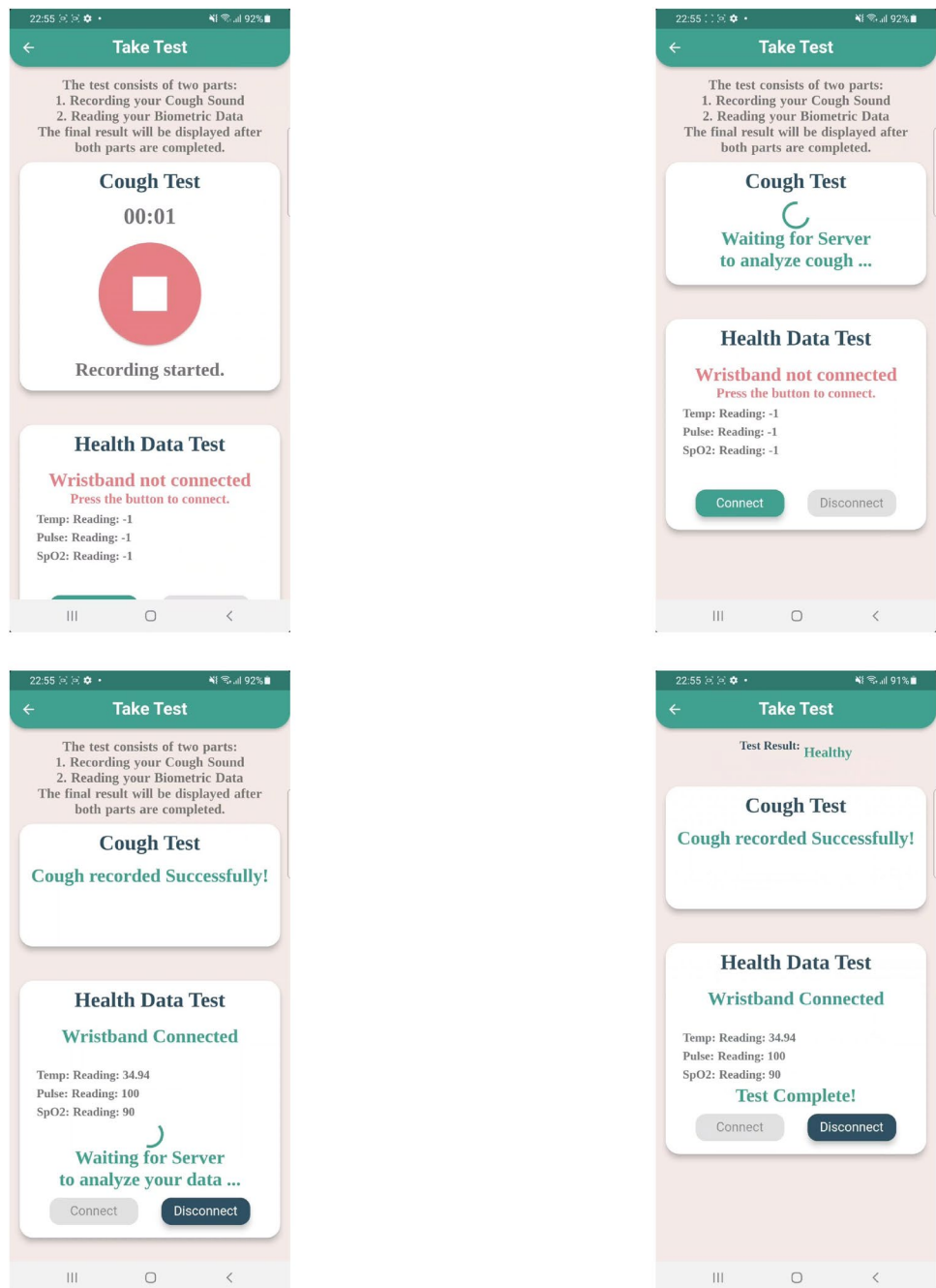


classification result to the Mobile Application, and it will be displayed to the user, as can be seen in the screenshots in Fig. 13.

4.4 Discussion

It should be noted that the wristband device could include many other sensors. First, the MAX30102 sensor could be replaced by a more expensive and accurate sensor. It was chosen for this preliminary prototype/proof of concept due to its wide availability and low cost. In addition, more sensors can be added to expand the applicability of the proposed approach. For example, an accelerometer could be used to detect falls and sudden movements. This is important in COVID situations where the patient is quarantined at home as it gives an indication that the patient needs immediate assistance (e.g., after falling due to the worsening of their condition). Moreover, a sensor for measuring blood sugar levels

Fig. 13 Full diagnostic test completed



could be added, as COVID was known to exacerbate the situation of diabetic patients [24]. Furthermore, a GPS sensor can be used (or the location can be used from the mobile app, not necessarily by adding an extra sensor to the bracelet; both options could work). If the mobile app is integrated with a national health system, this could allow the authorities to track the pandemic spread, identify the most infected areas (by using the locations of the infected users and evaluating their density per area), and even check for quarantine violations (in case the patient moves for large distances while they should stay at home). This could help complement the works that model pandemic spread and try to analyze user behavior in times of pandemics, such as [25–27], for example.

Moreover, the proposed approach is adaptable, since it can be expanded to diagnose other respiratory diseases, or re-trained for future pandemics. However, significant challenges need to be surmounted in this case. Before training for any future pandemic, the features of the pandemic need to be quickly identified, and the medical teams need to perform a significant number of tests to build datasets that can be used for training appropriate machine learning models. It is only then that such models can be integrated with the proposed approach. Particularly, in case of respiratory pandemics, cough sounds need to be distinguished between various diseases. The sounds used in this paper, based on the common datasets in the literature, e.g., [21], are classified as either “Covid” or “Non-Covid”. Similar datasets need to be derived for future pandemics (say, “Pandemic X”), to distinguish cough sounds between “Pandemic X” or “Non-Pandemic X”. However, this might not be sufficient in the prevalence of multiple serious pandemics. Elaborate datasets containing cough sounds from multiple pandemics need to be derived, so that training can be done to have accurate classification between “Pandemic 1”, “Pandemic 2”, etc., or “Healthy”. A possible way to deal with this is to have multi-stage classification, such as “Healthy” versus “Non-Healthy” first, then classify the “Non-Healthy” cases between the various candidate pandemics. Despite these huge challenges, once they are overcome, the proposed system can be upgraded mainly through a software update. The user would still perform a cough sound and have their vital parameters collected by the wristband, but the machine learning model used in the application would be more advanced to handle the more complex situation.

5 Conclusion

Our research demonstrates the potential of machine learning for early respiratory disease detection, particularly in the context of COVID-19. Through the analysis of common infection symptoms, we have developed a wristband device designed to accurately detect respiratory symptoms. Our wristband prototype demonstrates remarkable accuracy, with temperature, heart rate, and SPO2 measurements exhibiting an 99.7%, 82.0%, and 95.3% accuracy, respectively, compared to existing commercial products. This accuracy is evident through comparisons with widely-used wearable devices like the Huawei watch and Fitbit smartwatch. The continuous and automated monitoring capabilities of our wristband empower individuals to actively manage their health, enabling early intervention and personalized healthcare strategies. Additionally leveraging machine learning techniques, including several ML models utilizing cough sounds and vital parameters measurements, our approach has demonstrated remarkable accuracy. In fact, our proposed models have outperformed other state-of-the-art models in the literature, solidifying their efficacy in accurately diagnosing respiratory illnesses. Moreover, our approach is adaptable, since it can be expanded to diagnose other respiratory diseases, or re-trained for future pandemics. In this way, we mitigate the spread of the disease and contribute to an enhanced healthcare system.

Looking ahead, our research sets the stage for numerous future opportunities, including further enhancing our wristband device by integrating additional sensors and functionalities to improve diagnostic capabilities. Furthermore, we anticipate exploring future extensions such as implementing analogous methodologies for addressing other pandemics and refining our system to distinguish cough sounds associated with multiple diseases. These endeavors promise to broaden the scope of our research and its practical implications, positioning our technology as a versatile tool for healthcare monitoring and management.

Acknowledgements This publication was supported by Qatar University grant no. QUCD-IRCC-CENG-24-349. The findings achieved herein are solely the responsibility of the authors.

Authors' contributions Conceptualization Ayah Abdel-Ghani, Amira Abdalla, Zaineh Abughazzah, Mahnoor Akhund, Elias Yaacoub; Methodology Ayah Abdel-Ghani, Amira Abdalla, Zaineh Abughazzah, Mahnoor Akhund, Khalid Abualsaud, Elias Yaacoub; Formal analysis and investigation Ayah Abdel-Ghani, Amira Abdalla, Zaineh Abughazzah, Mahnoor Akhund; Writing-original draft preparation Ayah Abdel-Ghani, Amira Abdalla; Writing-review and editing Khalid Abualsaud and Elias Yaacoub; Funding acquisition Khalid Abualsaud and Elias Yaacoub; Supervision Elias Yaacoub.

Funding This publication was supported by Qatar University grant no. QUCD-IRCC-CENG-24-349.

Availability of data and materials The COUGHVID dataset is available at <https://doi.org/10.1038/s41597-021-00937-4>; The vital parameters dataset is available at <https://covidclinicaldata.org/>.

Code availability Software code available upon request. Please contact the first author.

Declarations

Ethics approval and consent to participate This research did not involve experiments on human subjects and thus did not seek IRB approval. For the machine learning model, we used existing online datasets available from other studies, as indicated in the Data Availability statement. As for the readings from the wristband and their comparison to commercial devices, they were performed on one member of the research team who volunteered to wear the wristband, for proof-of-concept validation.

Informed consent No. Please check Ethics Statement above.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Karesh W, et al. Wildlife trade and global disease emergence. *Emerg Infect Dis*. 2005;11:1000–2. <https://doi.org/10.3201/eid1107.050194>.
2. Mathieu E. et al. Coronavirus pandemic (COVID-19). In: *Our world in data*; 2020. <https://ourworldindata.org/coronavirus>.
3. Abid H, Mohd J, Ibrahim HK. Current status and applications of artificial intelligence (AI) in medical field: an overview. *Curr Med Res Pract*. 2019;9(6):231–7.
4. Abdel-Ghani A. et al. Efficient pandemic infection detection using wearable sensors and machine learning. In: *2023 international wireless communications and mobile computing (IWCMC)*; 2023. p. 1562–1567. <https://doi.org/10.1109/IWCMC58020.2023.10182781>.
5. Xiao L, et al. Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. *PLoS Biol*. 2017;15(1):1–30. <https://doi.org/10.1371/journal.pbio.2001402>.
6. Salvatore T, et al. Validity evaluation of the fitbit charge2 and the Garmin vivosmart HR+ in free-living environments in an older adult cohort. *JMIR mHealth uHealth*. 2019;7(6): e13084.
7. Paola K, Rasik M, Saba Z. The use of wearable ECG devices in the clinical setting: a review. *Curr Emerg Hosp Me R*. 2022;10:1–6. <https://doi.org/10.1007/s40138-022-00248-x>.
8. Chong Kimberly PL, et al. Consumer perceptions of wearable technology devices: retrospective review and analysis. *JMIR mHealth uHealth*. 2020;8(4): e17544. <https://doi.org/10.2196/17544>.
9. Majumder S, Mondal T, Deen MJ. Wearable sensors for remote health monitoring. *Sensors*. 2017;17:130. <https://doi.org/10.3390/s17010130>.
10. NHS choices. Pulse oximetry to detect early deterioration of patients with COVID-19 in primary and community care settings. Accessed: 9-Sep-2022, from <https://www.england.nhs.uk/coronavirus/publication/pulse-oximetry-to-detect-early-deterioration-of-patients-with-covid-19-in-primary-and-community-care-settings/>. June 2020.
11. Mendelson Y, Pujary C. Measurement site and photodetector size considerations in optimizing power consumption of a wearable reflectance pulse oximeter. In: *Proceedings of the 25th annual international conference of the IEEE engineering in medicine and biology society (IEEE Cat. No.03CH37439)*. Vol. 4. 2003, 3016–3019 Vol.4. <https://doi.org/10.1109/IEMBS.2003.1280775>.
12. Caleb P. et al. WristO2: reliable peripheral oxygen saturation readings from wrist-worn pulse oximeters. In: *2021 IEEE international conference on pervasive computing and communications workshops and other affiliated events (PerCom Workshops)*. 2021, 623–629. <https://doi.org/10.1109/PerComWorkshops51409.2021.9430986>.
13. Yang L. A heart rate finger ring and its smartphone APP through customized NFC. PhD thesis. Rose-Hulman Institute of Technology, 2015.
14. Alqudaihi Kawther S, et al. Cough sound detection and diagnosis using artificial intelligence techniques: challenges and opportunities. *IEEE Access*. 2021;9:102327–44. <https://doi.org/10.1109/ACCESS.2021.3097559>.
15. Pal A, Sankarasubbu M. Pay attention to the cough: early diagnosis of COVID-19 using interpretable symptoms embeddings with cough sound signal processing. In: *CoRR abs/2010.02417* (2020). [arXiv: 2010.02417](https://arxiv.org/abs/2010.02417).
16. Tammuz D. Signal analysis and classification of audio samples from individuals diagnosed with COVID-19. San Diego: University of California; 2020.

17. Jordi L, Ferran H, Brian S. COVID-19 artificial intelligence diagnosis using only cough recordings. *IEEE Open J Eng Med Biol.* 2020;1:275–81. <https://doi.org/10.1109/OJEMB.2020.3026928>.
18. Sudaroli Vijayakumar D, Sneha M. Low cost Covid-19 preliminary diagnosis utilizing cough samples and keenly intellectual deep learning approaches. *Alexandria Eng J.* 2021;60(1):549–57. <https://doi.org/10.1016/j.aej.2020.09.032>.
19. Han J. et al. Exploring automatic COVID-19 diagnosis via voice and symptoms from crowdsourced data. In: *CoRR abs/2102.05225* (2021). [arXiv: 2102.05225](https://arxiv.org/abs/2102.05225).
20. Skander H, et al. Attention-based hybrid CNN-LSTM and spectral data augmentation for COVID-19 diagnosis from cough sound. *J Intell Inf Syst.* 2022;59(2):367–89.
21. Orlandic L, Teijeiro T, Atienza D. The COUGHVID crowdsourcing dataset, a corpus for the study of large-scale cough analysis algorithms. *Sci Data.* 2021. <https://doi.org/10.1038/s41597-021-00937-4>.
22. Carbon Health and Braid Health. Coronavirus Disease 2019 (COVID-19) Clinical data repository. Version 10-20-2020. Accessed: February 15, 2024, from <https://covidclinicaldata.org/>. 2020.
23. Fritzing: open-source software for creating electronic schematics and PCB layouts. Accessed: May 31, 2024, on <https://fritzing.org/>. 2021.
24. Joshi AM, Shukla UP, Mohanty SP. Smart healthcare for diabetes during COVID-19. *IEEE Consum Electron Magazine.* 2021;10(1):66–71.
25. Mishra R, Gupta HP, Dutta T. Analysis, modeling, and representation of COVID-19 spread: a case study on India. *IEEE Trans Comput Soc Syst.* 2021;8(4):964–73. <https://doi.org/10.1109/TCSS.2021.3077701>.
26. Narayani V, Debasish G. The impact of mobility and interventions on the spread of diseases. *IEEE Trans Comput Soc Syst.* 2022;10(5):2291–311.
27. Gouissem A, Abualsaud K, Yaacoub E, Khattab T, Guizani M. A novel pandemic tracking map: from theory to implementation. *IEEE Access.* 2021;9:51106–20.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.