HHS Public Access

Author manuscript

Annu Int Conf IEEE Eng Med Biol Soc. Author manuscript; available in PMC 2022 November 01.

Published in final edited form as:

Annu Int Conf IEEE Eng Med Biol Soc. 2021 November ; 2021: 2140–2143. doi:10.1109/EMBC46164.2021.9629794.

Detection of COVID-19 Using Heart Rate and Blood Pressure: Lessons Learned from Patients with ARDS

Milad Asgari Mehrabadi[1], **Seyed Amir Hossein Aqajari**[1], **Iman Azimi**[2], **Charles A. Downs**[3], **Nikil Dutt**[4], **Amir M. Rahmani**[4,5,6]

[1]Department of Electrical Engineering and Computer Science, University of California Irvine, USA.

[2]Department of Computing, University of Turku, Turku, Finland.

[3]School of Nursing and Health Studies, University of Miami, USA.

[4]Department of Computer Science, University of California, Irvine, USA.

[5]School of Nursing, University of California, Irvine, USA.

[6]Institute for Future Health, University of California, Irvine, USA

Abstract

The world has been affected by COVID-19 coronavirus. At the time of this study, the number of infected people in the United States is the highest globally (31.2 million infections). Within the infected population, patients diagnosed with acute respiratory distress syndrome (ARDS) are in more life-threatening circumstances, resulting in severe respiratory system failure. Various studies have investigated the infections to COVID-19 and ARDS by monitoring laboratory metrics and symptoms. Unfortunately, these methods are merely limited to clinical settings, and symptom-based methods are shown to be ineffective. In contrast, vital signs (e.g., heart rate) have been utilized to early-detect different respiratory diseases in ubiquitous health monitoring. We posit that such biomarkers are informative in identifying ARDS patients infected with COVID-19. In this study, we investigate the behavior of COVID-19 on ARDS patients by utilizing simple vital signs. We analyze the long-term daily logs of blood pressure (BP) and heart rate (HR) associated with 150 ARDS patients admitted to five University of California academic health centers (containing 77,972 samples for each vital sign) to distinguish subjects with COVID-19 positive and negative test results. In addition to the statistical analysis, we develop a deep neural network model to extract features from the longitudinal data. Our deep learning model is able to achieve 0.81 area under the curve (AUC) to classify the vital signs of ARDS patients infected with COVID19 versus other ARDS diagnosed patients. Since our proposed model uses only the BP and HR, it would be possible to review data prior to the first reported cases in the U.S. to validate the presence or absence of COVID-19 in our communities prior to January 2020. In addition, by utilizing wearable devices, and monitoring vital signs of subjects in everyday settings it is possible to early-detect COVID-19 without visiting a hospital or a care site.

Corresponding Author: Milad Asgari Mehrabadi, Department of Electrical Engineering and Computer Science, University of California Irvine, Irvine, CA, USA, masgarim@uci.edu.

I. INTRODUCTION

Acute respiratory distress syndrome (ARDS) is a potential life-threatening consequence of infection with SARS-CoV2, the novel coronavirus that causes COVID-19 [1]. ARDS is characterized by an overwhelming immune response and non-cardiogenic pulmonary edema that compromise gas exchange, resulting in severe respiratory failure. ARDS mortality ranges from 40%−60%; however, it is unclear if the mortality rate is substantially higher if associated with COVID-19 infection, as it varies from 28.8%−62% [1], [2]. Currently, more than 136 million people worldwide have been infected with SARS-CoV-2 [3]. In the United States, 31.2 million people have been infected with over 562,000 deaths [3]. The impact of the COVID-19 pandemic is considerable and efforts to mitigate its spread through early detection cannot be over-emphasized.

Infections to COVID-19 have been conventionally investigated in clinical settings by monitoring laboratory metrics and symptoms [4], [5]. These studies have focused on a large amount of subjective questionnaires and invasive laboratory test results. For example, Jehi et al. [4] use a large number of features extracted from demographics, comorbidities, immunization history, symptoms, travel history, laboratory variables, and medications to predict the infection with COVID-19. Li et al. [1] show that the oxygenation index and respiratory system compliance could be leveraged to study ARDS patients infected with COVID-19. Force et al. [6] propose that ARDS caused by factors rather than COVID19 results in reduced lung compliance. However, reduced lung compliance in ARDS is typical of the disease [1].

Such diagnostics are the gold standard methods to investigate COVID-19 and ARDS patients; however, they are limited to hospitals and clinical settings. Moreover, subjective symptom-based analyses were shown to be an ineffective strategy to qualify an individual's likelihood of contracting COVID-19 [5]. In contrast, various studies showed that vital signs such as heart rate and blood pressure could be exploited for early detection of infections and respiratory diseases [7], [8], [9]. We posit that such biomarkers are informative in identifying ARDS patients infected with COVID-19. These biomarkers can be collected continuously and remotely due to the recent advancements in wearable electronics and Internet-of-Things-based devices (e.g., Omron® HeartGuide wrist-band [10]). Therefore, the effectiveness of these biomarkers in early COVID-19 detection extends the monitoring services to remote settings.

Recognition of COVID-19 infections using big sensory data necessitates novel modeling and analysis techniques. The state-of-the-art studies often use traditional statistical models to predict COVID-19 infections. These studies have mostly studied the linear statistical relationship and association between the health parameters or extracted features from the subject's demographics, symptoms, laboratory tests, and medications [4], [5]. For example, a full multi-variate logistic model is constructed in [4] to predict COVID-19 using extracted features. However, such data with complex intensive longitudinal structure and temporal characteristics need to be investigated using nonlinear and advanced methods. Machine learning algorithms, including Artificial Neural Networks, can be tailored in this regard to extract linear/nonlinear correlations in the data throughout the health monitoring. In this

paper, we investigate the behavior of COVID-19 on ARDS patients by proposing a deep neural network (DNN) model which utilizes three longitudinal features from the University of California COVID Research Data Set (UCCORDS) [11]: systolic and diastolic blood pressure and heart rate. We compare individuals who developed ARDS with and without COVID-19 to assess potential markers that could be used in early detection and prevention strategies. Moreover, we utilize statistical features and neural networks to distinguish between ARDS caused by COVID-19 and other factors.

II. METHODS

A. Data Set

UC-CORDS data set provides comprehensive, structured information of patients admitted to the hospital at the University of California's five academic health centers (i.e., UC Davis Health, UC San Diego Health, UC Irvine Health, UCLA Health, and UCSF Health). This data set provides a wide range of information, including different observations, measurements and COVID-19 test results of patients.

Notably, the vital signs, including heart rate, systolic and diastolic blood pressure, are recorded continuously every 30 minute. Since the data set is fully anonymized, it is not possible to access actual dates. However, we only considered hospitalized patients who were diagnosed with ARDS (IDs 4195694 and 4191650 from SNOMED vocabulary [12]) and we included data after their first COVID-19 test. Since the number of observations with negative COVID-19 test results is more than positives, we randomly selected fewer patients with negative test results based on the age distribution. This re-sampling resulted in a more balanced data set (i.e., 39,802 data points for each feature in the positive group and 38,170 samples in the negative group). As of April 1st 2021, this led to 150 participants for the positive and negative test groups (i.e., 75 participants for each group). Table I shows the age distribution of the patients per each COVID-19 test result.

In addition, another valuable aspect of this data set is the longitudinal monitoring of the vital signs. The data set contains, on average, 136.6 and 57.4 days for the negative and positive test groups, respectively.

B. Ethics

The data was jointly reviewed by the Institutional Review Boards of all UC Health campuses and was determined to be non-human subjects research. Moreover, UC-CORDS does not contain any patient identifier such as name and phone number. As such, UC-CORDS is a HIPAA limited data set.

C. Statistical Analyses

To show the correlation of features (i.e., blood pressures and heart rates) and COVID-19 test results, statistical features have been extracted. We measure basic features, including mean, minimum (min), maximum (max), and standard deviation (std) of DBP, SBP, and HR for each subject. Besides, we utilize the Point Biserial correlation between the proposed features and COVID-19 test results. This correlation, which is similar to Pearson's correlation, is

used when one of the variables is binary, and the other variable is a continuous number [13]. In other words, this measurement indicates the difference between the categorical groups' distribution.

D. Neural Networks

In this study, we are interested in the COVID-19 detection using longitudinal heart rate and blood pressure monitoring. To perform the detection, we propose a DNN architecture combining convolutional neural network (CNN), and a long short-term memory (LSTM). Such a model is utilized to leverage the embedded structure of longitudinal data. We have considered three channels of vital signs, i.e., heart rate (HR), systolic and diastolic blood pressure (SBP and DBP), as the inputs and the COVID-19 test result for the network's output. Table II summarizes the detailed structure of the proposed network. It consists of CNNs (capturing the spatial information), followed by a max-pooling layer, a LSTM layer (capturing the order in time series data) [14], and finally two fully connected layers (extracting the embeddings). We use grid search to tune the hyperparameters (e.g., number of filters and neurons) of the DNN. We randomly select 75% (112) of patients as train, the rest as the test data, and accuracy, f1 score and AUC metrics were chosen for the performance evaluation.

We label positive COVID-19 test results with '1' (26,984 samples in the train, 46.66%, and 8,018 samples in the test data, 75.97%) and the negative ones with '0'. For the learning task, TensorFlow package has been utilized.

Besides, to assess the detection's effectiveness, we test our model on different time intervals on test subjects. In other words, we are interested in the possibility of COVID-19 test result detection by only using a limited number of samples (in days). We evaluate the model with different interval sizes, which is extracted including N days ($N \in \{2,4,\ldots,60\}$) of subject's data. This evaluates the model's performance by looking only at a limited number of days.

Finally, for visualization purposes, the t-SNE method [15] was used over the dense layer's output to reduce the feature space's dimension to two.

III. RESULTS

A. Statistical Observations

We measure basic statistical features of BP and HR and compare them with COVID-19 test results. Table III shows the Point Biserial correlation between these features and age with COVID-19 test results, and Table IV represents 95% confidence interval (CI) of these biomarkers for each COVID-19 test group.

Table III shows significant positive correlations between the mean of HR/DBP, std of HR/DBP, maximum value of HR, minimum value of SBP and COVID-19 test results. Fig. 1 illustrates the difference in the distribution of average HR (mean hr) between each age group for the positive and negative test results. Although (mean hr) shows a significant correlation with COVID-19 test results, there is an overlap in the distribution of such a feature between

positive and negative results. This visualization further supports the fact that using only statistical features to detect people infected with COVID-19 is challenging.

B. Deep Learning

Due to the longitudinal aspect of the data, we consider a DNN architecture to detect COVID-19 test results by only looking at BP and HR. The accuracy of this model reached as high as 0.79, 0.87 precision, 0.84 recall, 0.85 f1 score and 0.81 area under the curve (AUC) for the entire test data. Besides, Fig. 2a shows the accuracy of our model with respect to the first 60 days of data. Fig. 2b illustrates the corresponding area under the curve (AUC) with given days while Fig. 2c shows the f1 score. Fig. 2a shows an increase in the model's accuracy at the beginning, starting from 0.52 and reaching as high as 0.78% on day 12th. The small drop in the measurements after day 12th is because of the increase in the number of false negatives compared to true positives. By observing more data for both groups, the performance of the model constantly increases (day 20th).

To visualize the extracted features using DNN, we used t-SNE method [15] to reduce the feature space dimension to two. We performed this method on the output of the dense layer with 100 neurons. Fig. 3 shows that using extracted features by the DNN, the positive and negative cases are almost separated.

IV. DISCUSSION & CONCLUSIONS

A few of our observations warrant additional discussion. First, monitoring of blood pressure and heart rate may provide a useful strategy for individuals living in collective communities, such as nursing homes or rehabilitation facilities, as well as for healthy community-dwelling adults. The potential impact could be to mitigate the spread of COVID-19, as well as allowing early detection of complications associated with infection, such as those at greater risk for ARDS.

Second, we assessed for the presence of comorbidities in COVID-19 positive patient with ARDS, and reported that comorbid diagnoses such as type 2 Diabetes Mellitus, hyperglycemia, chronic obstructive pulmonary disease, elevated transaminase, and lactic acid dehydrogenase, brady-cardia, acute ST segment elevation myocardial infarction, and metabolic derangements were more prevalent (data not shown). This observation is in-line with other reports [1], [2], [16] demonstrating increased vulnerability among those with chronic health conditions, as well as reported metabolic derangements observed with COVID-19 infection, especially among adults over 60 years of age.

Third, there are other potential applications in modeling COVID-19. Specifically, there has been a discussion of how early COVID-19 arrived in the United States; the first cases were reported in California. It would be possible to review data prior to the first reported cases in the U.S. to validate the presence or absence of COVID-19 in our communities prior to January 2020. This is of importance as the viral genome sequence was confirmed in late January 2020, which allowed for the use of polymerase chain reaction to detect viral genetic material [17]. Antibody testing, which has been shown to be inconsistent, was used in the preceding months, raising the question of how early was COVID-19 in the United States.

Moreover, the related works focus on laboratory measurements and symptoms to detect the infected patients or severe positive cases of COVID-19 using statistical methods [4], [5]. In contrast, we considered two easily accessible features as well as utilizing a deep learning method to capture the short- and long-term dependencies in the time series data. There is a correlation between the simple statistical features and COVID-19 test results. However, simple logistic regression models are insufficient due to the overlap in the feature space. Leveraging the nonlinear features extracted from our proposed neural network, we distinguished negative and positive COVID-19 test results with the AUC as high as 0.81 by using only blood pressure and heart rate values.

Although our findings are only based on ARDS population, these achievements could potentially lead future directions of our research to investigate the aforementioned vital signs for COVID-19 detection tasks with other populations.

In conclusion, we proposed a DNN-based model to investigate the non-linear patterns in simple vital signs, namely, blood pressure and heart rate, which can be easily and reliably measured without the need for skilled medical professionals, in ARDS patients with positive and negative COVID-19 test results. Our proposed model achieved 0.79 accuracy, 0.85 f1 score and 0.81 AUC. Using wearable devices, it is possible to monitor vital signs of subjects in everyday settings without visiting a hospital or a care site. Utilizing the proposed model allows early detection of COVID-19 cases in free-living conditions.

ACKNOWLEDGMENT

Biomedical computing facilities are supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health (Grant UL1 TR001414). C. A. Downs is also supported by NR016957. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

REFERENCES

[1]. Li X et al. , “Acute respiratory failure in covid-19: is it “typical” ards?” Critical Care, vol. 24, pp. 1–5, 2020. [PubMed: 31898531]

[2]. Tang X et al. , “Comparison of hospitalized patients with ards caused by covid-19 and h1n1 [published online ahead of print, 2020 mar 26],” Chest, pp. 30 558–4, 2020.

[3]. “COVID-19 dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University,” https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6, [Online; accessed 31-Aug-2020].

[4]. Jehi L et al. , “Individualizing risk prediction for positive covid-19 testing: results from 11,672 patients.” Chest, 2020.

[5]. Callahan A et al. , “Estimating the efficacy of symptom-based screening for covid-19,” NPJ digital medicine, vol. 3, no. 1, pp. 1–3, 2020. [PubMed: 31934645]

[6]. Force ADT et al. , “Acute respiratory distress syndrome,” Jama, vol. 307, no. 23, pp. 2526–2533, 2012. [PubMed: 22797452]

[7]. Shashikumar SP et al. , “Early sepsis detection in critical care patients using multiscale blood pressure and heart rate dynamics,” Journal of electrocardiology, vol. 50, no. 6, pp. 739–743, 2017. [PubMed: 28916175]

[8]. Matsumura K et al. , “Comparison of the clinical course of covid-19 pneumonia and acute respiratory distress syndrome in 2 passengers from the cruise ship diamond princess in february 2020,” The American journal of case reports, vol. 21, pp. e926 835–1, 2020.

[9]. Gattinoni L et al., “Covid-19 pneumonia: Ards or not?” 2020.

[10]. “Wearable Blood Pressure Monitor and Watch, HeartGuide by OMRON,” https://omronhealthcare.com/, [Online; accessed 31-Aug-2020].

[11]. “University of California Health creates centralized data set to accelerate COVID-19 research,” https://www.universityofcalifornia.edu/pressroom/university-california-health-creates-centralized-data-set-accelerate-covid-19-research, 2020, [Online; accessed 15-Aug-2020].

[12]. “SNOMED Clinical Terms,” https://www.snomed.org/, [Online; accessed 15-Aug-2020].

[13]. Sheskin DJ, Handbook of parametric and nonparametric statistical procedures. crc Press, 2020.

[14]. Hochreiter S and Schmidhuber J, “Long short-term memory,” Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997. [PubMed: 9377276]

[15]. Maaten L. v. d. et al. , “Visualizing data using t-sne,” Journal of machine learning research, vol. 9, no. Nov, pp. 2579–2605, 2008.

[16]. Wu C et al. , “Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in wuhan, china,” JAMA internal medicine, 2020.

[17]. “Whole genome of novel coronavirus, 2019-nCoV, sequenced,” www.sciencedaily.com/releases/2020/01/200131114748.htm, 31 January 2020, [Online; accessed 31-Aug-2020]

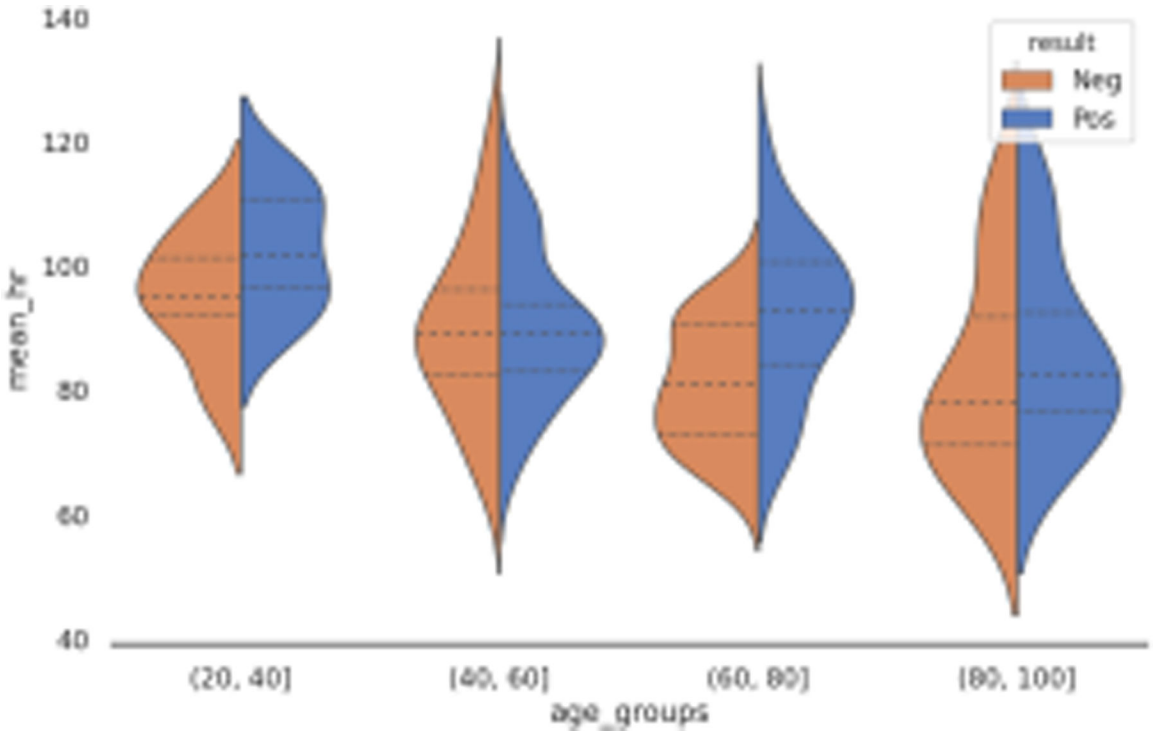

Fig. 1:

The distribution comparison of average HR between the positive and negative test results for each age group. The dashed lines represent the quartiles.

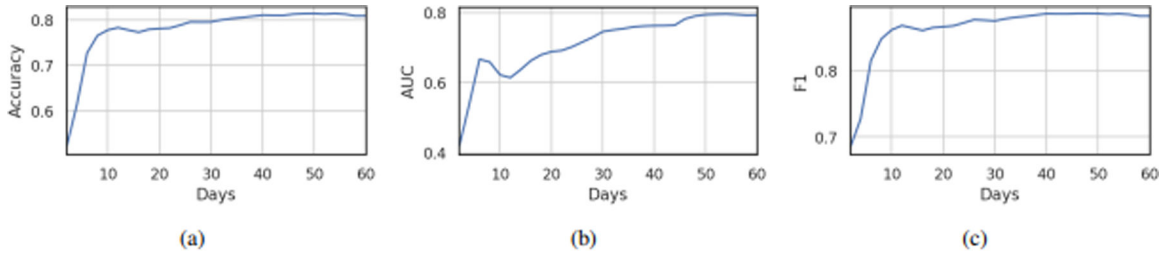

Fig. 2:

The performance of the model in terms of accuracy (a), AUC (b) and f1 score (c) using test data with respect to the number of included days.

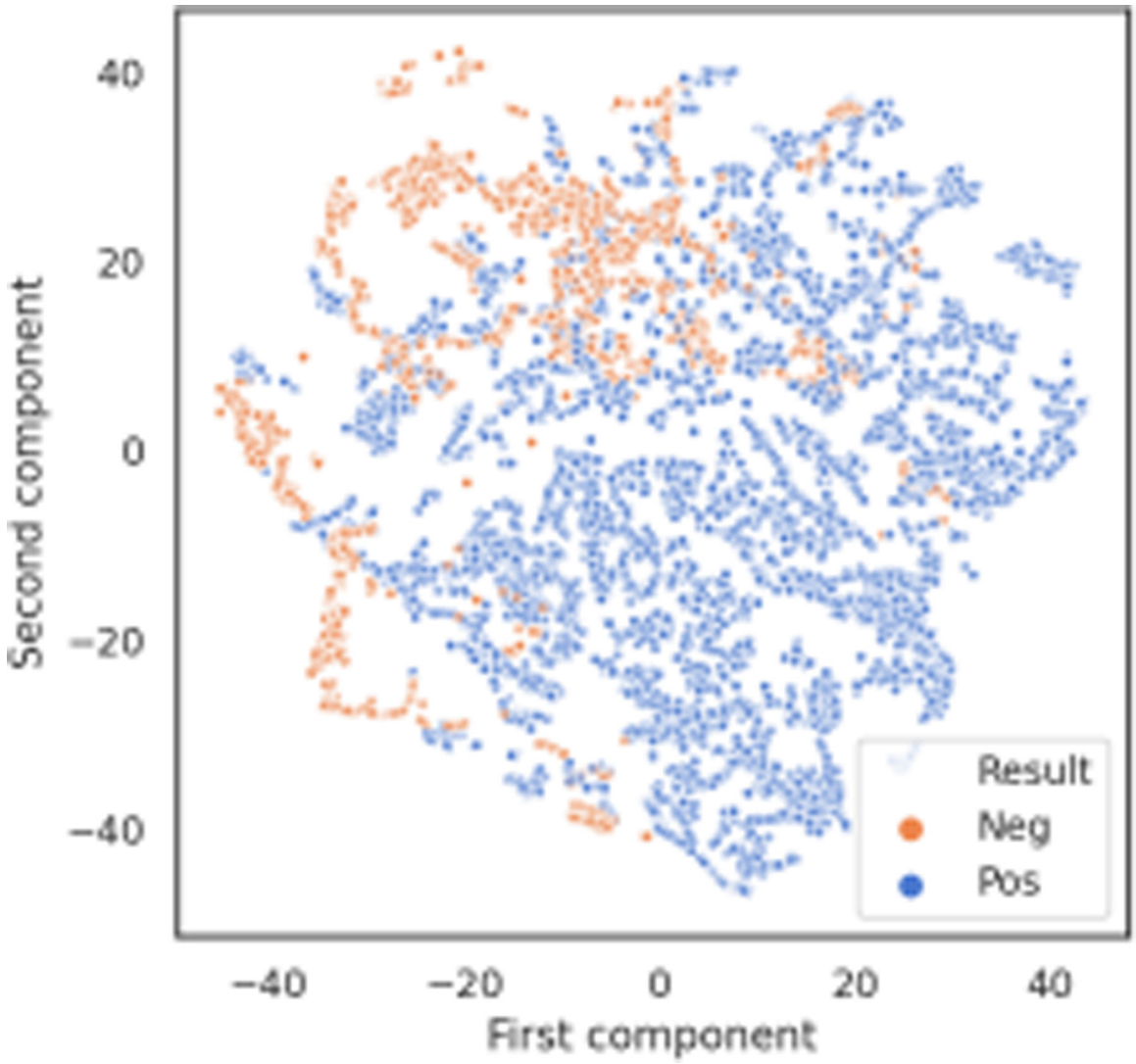

Fig. 3:
The 2-dimensional representation of test data using t-SNE considering the entire test data.

TABLE I:

Age distribution of subject with different COVID19 test results.

	Number of Participants		Number of Samples	
Age range	**Negative**	**Positive**	**Negative**	**Positive**
20 – 40	10	10	5724	5291
41 – 60	27	30	13995	16068
61 – 80	27	24	16395	15020
80 – 100	11	11	2096	3423

TABLE II:

The architecture of the proposed neural network.

Layer	Output Shape
1D CNN	(None, 4, 14, 64)
Max Fooling	(None, 4, 7, 64)
1D CNN	(None, 4, 5, 32)
Max Fooling	(None, 4, 2, 32)
Flatten	(None, 4, 64)
LSTM	(None, 64)
Flatten	(None, 64)
Dense	(None, 100)
Dense	(None, 1)

TABLE III:

Point Biserial correlation of statistical features and COVID-19 test results (* shows significant correlation).

		Correlation	*P*-value
HR	mean	0.20	.012*
	std	0.21	.009*
	min	−0.10	.22
	max	0.18	.020*
DBP	mean	−0.19	.015*
	std	−0.19	.016*
	min	−0.10	.20
	max	−0.14	.075
SBP	mean	−0.15	.063
	std	0.08	.32
	min	−0.17	.031*
	max	0.13	.10
Age		−0.05	0.51

TABLE IV:

95% confidence intervals of HR, DBP and SBP for each COVID-19 test result group.

Biomaker	95% CI (Positive)	95% CI (Negative)
HR	95.97–96.35	84.87–85.24
DBP	94.95–65.20	70.83–71.13
SBP	117.66–118.09	127.20–127.72