PHYSIOLOGICAL MONITORING OF FRONT-LINE CAREGIVERS FOR CV-19 SYMPTOMS: MULTI-RESOLUTION ANALYSIS & CONVOLUTIONAL-RECURRENT NETWORKS

Omid Dehzangi [a,b,c], Paria Jeihouni [b,c], Victor Finomore [a,b], Ali Rezai [a,b]
[a]Department of Neuroscience, [b]Rockefeller Neuroscience Institute,
[c]Lane Department of Computer Science and Engineering, West Virginia University

2021 IEEE International Conference on Image Processing (ICIP) | 978-1-6654-4115-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIP42928.2021.9506495

ABSTRACT

Due to easy transmission of the COVID-19, a crucial step is the effective screening of the front-line caregivers are one of the most vulnerable populations for early signs and symptoms, resembling the onset of the disease. Our aim in this paper is to track a combination of biomarkers in our ubiquitous experimental setup to monitor the human participants' operating system to predict the likelihood of the viral infection symptoms during the next 2 days using a mobile app, and an unobtrusive wearable ring to track their physiological indicators and self-reported symptoms. we propose a multi-resolution signal processing and modeling method to effectively characterize the changes in those physiological indicators. In this way, we decompose the 1-D input windowed time-series in multi-resolution (i.e. 2-D spectro-temporal) space. Then, we fitted our proposed deep learning architecture that combines recurrent neural network (RNN) and convolutional neural network (CNN) to incorporate and model the sequence of multi-resolution snapshots in 3-D time-series space. The CNN is used to objectify the underlying features in each of the 2D spectro-temporal snapshots, while the RNN is utilized to track the temporal dynamic behavior of the snapshot sequences to predict the patients' COVID-19 related symptoms. As the experimental results show, our proposed architecture with the best configuration achieves 87.53% and 95.12% average accuracy in predicting the COVID-19 related symptoms.

Index Terms— Ubiquitous physiological monitoring, COVID-19 Symptoms, Time-series prediction, CNN, RNN.

1. INTRODUCTION

COVID-19 has spread all over the world with a growing number of affected patients with more than 100 million confirmed cases as of January 25, 2021 (https://coronavirus.jhu.edu). The respiratory symptoms associated with the infection of this strain has lead to significantly high hospitalization [1], which have forced the medical workers into an increased amount of work. Due to its contagious nature of the virus, front-line medical workers are the most susceptible to be affected. It has been observed that the virus has an incubation period of 2 to 14 days [2], before the affected person starts developing or exhibiting symptoms. It is imperative to make sure the front-line caregivers are not affected, as it could render the entire medical facility in jeopardy. Therefore, effective and longitudinal screening of early signs and symptoms, resembling the onset of the disease is significant.

Several studies have attempted in utilizing predictive models to anticipate the occurrence of symptoms in various populations, such as [3–5]. Wearable consumer electronics have been proven to be effective in estimating the physiological and behavioral state pervasively and continuously [6, 7]. Though, little research has been conducted on COVID-related wearable monitoring in front-line caregivers. The goal of this paper is to investigate if early physiological biomarkers, in our ubiquitous recording setup, can indicate the onset of COVID-19 related symptoms. We aim to accurately identify the presence of viral infections on our participants and predict the development of symptoms during the next 2 days.

30 subjects participated in the study during the span of 4 months, where longitudinal data was acquired using a ring device and our in-house mobile app to track indicators at the Rockefeller Neuroscience Institute, West Virginia University (WVU). These sensor data indicators, including heart rate (HR) and heart rate variability (HRV) are used as the raw input to feed our signal processing and predictive time-series modeling system to anticipate the onset of the symptoms before they are physically manifested. Particularly, HRV is a significant measure representing the variation in time between each heartbeat. HRV is controlled by the autonomic nervous system (ANS), which is a primitive part of the nervous system that regulates HR, blood pressure, breathing, among others autonomously. We anticipate that HRV can be a potential predictor of COVID-19 related symptoms.

In the last few years, neural networks (NNs) have been widely employed to learn the underlying patterns in the data to enhance the predictive modeling performance. One of the most popular architectures employed for dealing with temporal features of the data is the recurrent neural networks (RNNs). Various versions of the RNNs such as the long short-term memory (LSTMs) have been developed in different applications such as precipitation forecasting [8], natural language processing [9], video segmentation [10], and even remaining useful life (RUL) estimation of a system [11]. The distinctive characteristic of the RNNs is a certain type of memory they possess to model time series data in time.

convolutional neural network (CNN) is another category, which is capable of detecting complex spatial features with different level of abstractions in the data in hand [12–14]. CNNs are merely able to extract the spatial features out of the data, regardless of the temporal properties, that is why they cannot be used for time-series data. Nevertheless, the integration of CNN and RNN seems to culminate in better perceiving the data, where CNN does the task of spatial feature extraction and RNN preserves the temporal properties, such as in [15, 16]. Also, authors in [17] could detect the QRS complex in ECG signal by feeding it to a CNN-LSTM architecture.

In this study, our aim is to investigate the relationship between the input physiological time-series patterns, and the COVID-19 related symptoms in front-line caregivers by leveraging the capabilities of our proposed architecture of jointly-trained CNN-RNN structure. We attempt to identify the future occurrence of symptoms such as fever, dry cough and difficulty breathing [18]. If identified successfully, that person can be isolated and tested for the presence of an infection by COVID-19 before the symptoms become apparent. In this way, it is plausible to prevent asymptomatic transmission [19] of the virus among the caregivers, as they can self isolate even before the symptoms manifest, thus reducing the probability of further transmission. To the best of our knowledge, this is the first study quantifying the wellness of COVID-19 front-line caregivers with respect to their physiological signal by using deep learning methods.

2. EXPERIMENTAL SETUP & DATA ACQUISITION

This study includes 30 front-line caregivers directly dealing with COVID-19 patients at the J.W. Ruby Memorial Hospital, WVU Medicine. Each subject was given an OURA ring, tracking their 24-hour HR and HRV during 120 days with the sampling rate of one per 3 mins. The data is synchronized using our specifically designed app and cloud. A graphic user interface was designed and installed on subjects' cell phones, on which they could report the presence of a set of symptoms, considered as labels, through a questionnaire that is filled twice daily to increase our label resolution. The outline of the experimental setup is depicted in Fig. 1.

3. DATA PREPARATION & FEATURE EXTRACTION

3.1. Data Framing

Let us assume that the physiological dataset collected for each subject is represented by $\mathbf{X} = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_D\end{bmatrix}$, where D denotes the number of days, and $\mathbf{x}_i$ is a vector of samples associated with the i^{th} day. Thus,

$$\mathbf{x}_i = \begin{bmatrix}{x_i}^1 & {x_i}^2 & {x_i}^3 & \cdots & {x_i}^N\end{bmatrix}^T, \tag{1}$$

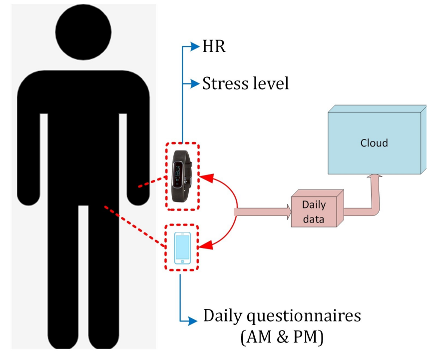

Fig. 1: The experimental setup for pervasive data collection.

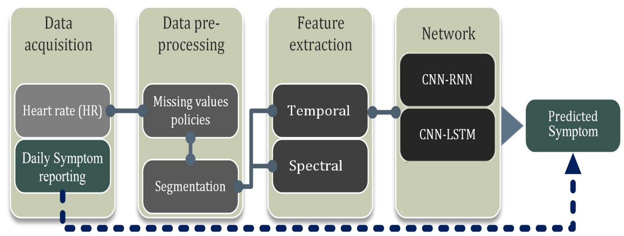

Fig. 2: Framework for the Symptom prediction.

where, $\mathbf{N}$ indicates the number of samples per day. Therefore, considering the sampling frequency to be equal to $f_s = 1/180^{Hz}$, the total number of samples per day would be $N = 24^{hours} \times 60^{mins} \times 60^{secs} \times 1/180^{Hz} = 480^{samples}$. Furthermore, to prepare high-resolution data, the samples associated with each day are segmented with a window of 5-hour duration, and the overlap of $o.l. = 90\%$ (30 mins increments) between two consecutive windows. Based on 5-hour duration for each window, the number of samples per window and the number of windows associated with each day would be 100 and 45, respectively.

3.2. Data Pre-processing

Prior to feeding the data to the network, two policies are chosen with the missing values: 1) those related to taking-off the ring are removed completely from the dataset, and 2) others missed due to the user's high activity during a much shorter amount of time are replaced by the median of the previous five samples. Finally, physiological samples per day are framed into a consecutive set of 5-hour long windows.

3.2.1. Labels

To assess the caregivers for COVID-related symptoms, a set of questions are asked twice daily via the provided mobile app, for which the caregivers provide their responses. The labels chosen here are various symptoms: fever, coughing, sneezing, sore throat, shortness of breath, sense of smell

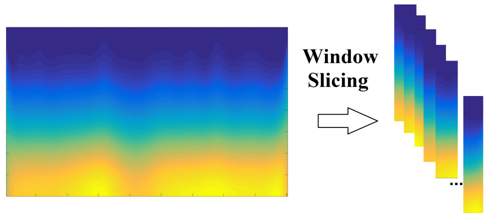

Fig. 3: SWVD multi-resolution respresentation of an HRV widow and the process of slicing.

change, where the presence of any of these symptoms was labelled as 1, and else as 0 for each 12-hr period and their corresponding windows of data. For the symptom class, the aggregation of all the COVID-19 related symptoms was considered in one positive class and we made sure that the positive class constituted 15% of the whole data for the performance measures to be meaningful.

3.3. Multi-Resolution Representation

Following the data segmentation, feature extraction is carried out to achieve the highest possible efficiency out of the network. It stands to reason that the spectral features are highly representative of the signal behavior. Therefore, we introduce the multi-resolution analysis, in which the signal is represented in joint time-frequency (TF) domains. Due to the non-stationary and multi-component nature of physiological signals, we use Wigner-Ville distribution that provides high spectral resolution and have good cross-term reduction capability [20]. Wigner-Ville distribution is a variant TF method, which incorporates smoothing by independent windows in time and frequency, namely $W(\tau)$ and $W(t)$, as below:

$$SWVD(t,\omega) = \int_{-\infty}^{+\infty} W_\omega(\tau)[\int_{-\infty}^{+\infty} W_t(u-t)x(u-\frac{\tau}{2}) x^*(u+\frac{\tau}{2})du]e^{j\omega\tau d\tau}. \tag{2}$$

Eq. 2, represents smoothed Wigner–Ville decomposition (SWVD), which provides great flexibility in the choice of time and frequency smoothing (illustrated in Figure 3).

4. METHODOLOGY

Using CNNs, we aim to conduct automatic feature extraction in the multi-resolution time series. Then, we use RNNs, whose mission is to preserve the temporal features of a time series in the sequential modeling task in-hand. We establish the methodology for the symptom prediction task based on a supervised learning paradigm, where the inputs are the 1-D physiological time series & the 2-D SWVD representation, and the labels are the self-reported symptoms (Fig. 2).

4.1. Convolutional Neural Network (CNN)

Each chunk of the data is taken in by the convolutional layer. With that being said, each chunk needs to be fed to the network separately, and then distribute the data on continuous time steps. In this study, each window of data is sliced into smaller pieces with 75% overlap (Fig. 3) to feed the Time scattered CNN separately and also augment the training data. Each slice is first passed to the convolutional layer, and then the entire windows go into a recurrent module one after another.

4.2. Recurrent Modules

RNN makes use of current input along with the history of information making its way forward from the previous time steps. The RNN layers are stacked on top of each other to extract more temporal dependencies between different parts of the sequence. In this study, we investigate the difference between the prediction results solely using the RNN modules given 1-D temporal data, and its combination in addition to the CNN transformation given the 2-D T-F SWVD. Fig. 4 visually explains our entire proposed modeling architecture designed for the COVID-19 symptom prediction.

5. EXPERIMENTS

In this section, the results achieved by RNN+1D (raw temporal data) and CNN-RNN+2D (SWVD spectro-temporal data) are compared to predict the existence of any symptoms during the next 1 (and 2) day(s). Also, the biomarkers of HR, HRV, and HR+HRV are compared as inputs to the deep learning models. The predictive models are trained and evaluated per individual and then averaged over subjects to get an overall accuracy and precision. The parameters set in the network are as follows: Learning rate= 0.001, loss function= *categorical cross-entropy*, optimizer function= *Adam*, batch size= 20, no. of epochs= 200 and the no. of recurrent layers= 4. Furthermore, we have divided the dataset into two parts, the train set and the validation set, where the percentage of the divisions are 80% and 20% (time-series fashion), respectively.

5.1. Results

The validation accuracies for both of the architectures are shown in the Table 1 and Table 2. It should be noted that the architecture is exactly the same as shown in Fig. 4, with and without the convolutional layer.

As can be seen in Table 1 and 2, the higher achieved results for both the next 1-day and 2-day prediction belong to the CNN-RNN architecture. The results demonstrate dramatic improvements compared to the results achieved by the

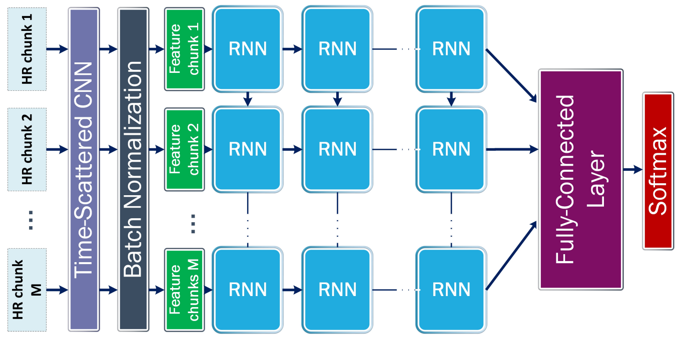

Fig. 4: Entire architecture for symptom prediction.

Table 1: Performance comparison for different architectures and different inputs for the next 1-day prediction.

Model		Acc (%)	Prec (%)
HR	RNN+1D Raw	84.85	79.80
HR	CNN-RNN + 2D SWVD	90.72	87.24
HRV	RNN+1D Raw	85.28	80.80
HRV	CNN-RNN + 2D SWVD	94.14	90.24
HR+HRV	RNN+1D Raw	87.11	83.60
HR+HRV	**CNN-RNN + 2D SWVD**	**95.12**	**91.03**

Table 2: Performance comparison for different architectures and different inputs for the next 2-day prediction.

Model		Acc (%)	Prec (%)
HR	RNN+1D Raw	76.53	71.68
HR	CNN-RNN + 2D SWVD	84.23	80.40
HRV	RNN+1D Raw	78.53	75.68
HRV	CNN-RNN + 2D SWVD	85.97	82.22
HR+HRV	RNN+1D Raw	79.72	77.13
HR+HRV	**CNN-RNN + 2D SWVD**	**87.53**	**83.62**

RNN only model, which shows the effectiveness of the joint CNN-RNN model optimization over the recurrent-based architecture alone. Table 1 and 2 also report the prediction results comparing the 1-D temporal representation (i.e., raw signal) vs. the 2-D multi-resolution representation (i.e., SWVD transformed). It is observed that the 2-D SWVD representation leads to better results compared to the raw temporal representation for the input frames, which demonstrates the profound effect of multi-resolution representation of data for predictive modeling. Also, the results in both Table 1 and 2 consistently show that the HRV measure provides a noticeably higher predictive capability to the models than the HR and their combination can moderately improve the results over the HRV measure alone. Due to the fact that there is a class imbalance (i.e. 79% for '0' & 21% for '1'), the precision measure is also provided in Table 1 and 2. As it can be seen, precision is following the prediction accuracy consistently, which is an evidence that the trained predictive models are not biased toward the dominant class.

6. CONCLUSION

In this paper, we have shown that by designing a deep learning method, specifically, using a CNN layer for feature extraction in multi-resolution jointly-trained with an memory-based RNN module, we can detect the underlying patterns in the HR HRV signals and predict the front-line caregiver subject population COVID-related symptoms in a longitudinal fashion. We empirically found that the "HR+HRV as the input + CNN-RNN + 2D SWVD representation" achieved the highest averaged prediction accuracy of 95.12% and 87.53% for the next 1-day and 2-day prediction, respectively. The results suggest that our developed system can serve as a critical decision-making tool to help contain the spread of the virus. This is significant as it shows the consumer electronics can be incorporated in medical diagnosis. As our future direction, the authors are working on improving the feature embedding module, designing novel modeling architectures, and employing other modalities of the wearable sensor recordings in our prediction and decision making on a larger population of subjects.

7. REFERENCES

[1] Robert Verity* et al., "Estimates of the severity of coronavirus disease 2019: a model-based analysis," *Th*, vol. 172, no. 9, pp. 577–582, 2020.

[2] Stephen A Lauer, "The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application," *Annals of internal medicine*, vol. 172, no. 9, pp. 577–582, 2020.

[3] Valdes A.M. Freidin M.B. et al. Menni, C., "Real-time tracking of self-reported symptoms to predict potential covid-19," *Nature Medicine*, vol. 26, pp. 037—1040, 2020.

[4] Ketko I Geva G Yosef R Toledo I Twig G Avramovitch E Gordon B Derazne E Fink N Furer A. Nitecki M, Taran B, "Self-reported symptoms in healthy young adults to predict potential covid-19 disease," *The Clinical Microbiology and Infection*, JAN, 2021.

[5] Nikos et al. Evangelou, "Self-diagnosed covid-19 in people with multiple sclerosis: a community-based cohort of the uk ms register," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 92, no. 1, pp. 107–109, 2021.

[6] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 1–12, Jan 2010.

[7] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321–1330, March 2015.

[8] SHI Xingjian, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," in *Advances in neural information processing systems*, 2015, pp. 802–810.

[9] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria, "Recent trends in deep learning based natural language processing," *IEEE Computational intelligence magazine*, vol. 13, no. 3, pp. 55–75, 2018.

[10] David Nilsson and Cristian Sminchisescu, "Semantic video segmentation by gated recurrent flow propagation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6819–6828.

[11] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long short-term memory network for remaining useful life estimation," in *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, June 2017, pp. 88–95.

[12] H. Chen, D. Ni, J. Qin, S. Li, X. Yang, T. Wang, and P. A. Heng, "Standard plane localization in fetal ultrasound via domain transferred deep neural networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1627–1636, Sep. 2015.

[13] A. Kumar, J. Kim, D. Lyndon, M. Fulham, and D. Feng, "An ensemble of fine-tuned convolutional neural networks for medical image classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 31–40, Jan 2017.

[14] S. Chen, J. Qin, X. Ji, B. Lei, T. Wang, D. Ni, and J. Cheng, "Automatic scoring of multiple semantic attributes with multi-task feature leverage: A study on pulmonary nodules in ct images," *IEEE Transactions on Medical Imaging*, vol. 36, no. 3, pp. 802–814, March 2017.

[15] Yong Xu, Qiuqiang Kong, Qiang Huang, Wenwu Wang, and Mark D Plumbley, "Convolutional gated recurrent neural network incorporating spatial features for audio tagging," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 3461–3466.

[16] Wootaek Lim, Daeyoung Jang, and Taejin Lee, "Speech emotion recognition using convolutional and recurrent neural networks," in *2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*. IEEE, 2016, pp. 1–4.

[17] Brosnan Yuen, Xiaodai Dong, and Tao Lu, "Inter-patient cnn-lstm for qrs complex detection in noisy ecg signals," *IEEE Access*, vol. 7, pp. 169359–169370, 2019.

[18] Min Cao, Dandan Zhang, Youhua Wang, Yunfei Lu, Xiangdong Zhu, Ying Li, Honghao Xue, Yunxiao Lin, Min Zhang, Yiguo Sun, et al., "Clinical features of patients infected with the 2019 novel coronavirus (covid-19) in shanghai, china," *MedRxiv*, 2020.

[19] Monica Gandhi, Deborah S Yokoe, and Diane V Havlir, "Asymptomatic transmission, the achilles' heel of current strategies to control covid-19," 2020.

[20] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal," *Algorithms and applications. Proc. IEEE*, vol. 80, pp. 540—-568, 1992.