Digital Interventions to Reduce Distress Among Frontline Health Care Providers: Analysis of Self-Perceived Stress

Binh Nguyen[1,*], Andrei Torres[2], Alice Rueda[3], Walter Sim[3], Douglas M Campbell[4], Wendy Lou[5], Bill Kapralos[3], Lindsay Beavers[4], Adam Dubrowski[2] , Venkat Bhat[3,**], and Sridhar Krishnan[1,**]

[1]Dept of Electrical, Computer, and Biomedical Engineering, Toronto Metropolitan Univerity, Toronto, ON M5B 2K3
[2]maxSIMhealth, Ontario Tech University, Oshawa, ON, L1H 7K4
[3]Interventional Psychiatry Program, St. Michael's Hospital, Toronto M5B 1W8
[4]Allan Waters Family Simulation Program, St Michael's Hospital, Unity Health Toronto, Toronto M5B 1W8
[5]Dalla Lana School of Public Health, University of Toronto, Toronto M5T 1P8
[*]Corresponding author: Binh Nguyen, binh.nguyen@ryerson.ca
[**]Corresponding Co-senior authors

***Abstract*— Due to the constraints of the COVID-19 pandemic, healthcare workers have reported behaving in ways that are contrary to their values, which may result in distress and injury. This work is the first of its kind to evaluate the presence of stress in the *COVID-19 VR Healthcare Simulation for Distress* dataset. The dataset collected passive physiological signals and active mental health questionnaires. This paper focuses on correlating electrocardiogram, respiration, photoplethysmography, and galvanic skin response with the Perceived Stress Scale (PSS)-10 questionnaire. The analysis involved data-driven techniques for a robust evaluation of stress among participants. Low-complexity pre-processing and feature extraction techniques were applied and support vector machine and decision tree models were created to predict the PSS-10 scores of users. Imbalanced data classification techniques were used to further enhance our understanding of the results. Decision tree with oversampling through Synthetic Minority Oversampling Technique achieved an accuracy, precision, recall, and F1 of 93.50%, 93.41%, 93.31%, and 93.35%, respectively. Our findings offer novel results and clinically valuable insights for stress detection and potential for translation to edge computing applications to enhance privacy, longitudinal monitoring, and simplify device requirements.**

I. Introduction

Healthcare workers (HCW) during the COVID-19 pandemic have been exposed to high levels of stress, anxiety, distress, and depression. They have reported behaving in ways that are contrary to their moral values and professional commitments that degrade their integrity [1]. Moral distress is thought to occur when a participant is aware of the ethical and correct choice but is powerless due to the constraints they are encountered with and is derived from the onset of prolonged stress, severe stress, or both [2].

Currently, there are various available techniques to understand, monitor, and reduce underlying mental disorders such as the use of VR to treat major depressive disorder (MDD), anxiety, post-traumatic stress disorder and phobias [3]. Physiological biomarkers and self-reports are other methods to objectively evaluate underlying psychiatric disorders. The autonomic nervous system is the control system that acts largely unconsciously and regulates bodily functions, including heart rate and respiratory rate. This system is the primary mechanism in control of the fight-or-flight response. It is expected that users undergoing stressful situations will be reflective in their physiological biomarkers [4]. Self-reports, such as the Perceived Stress Scale (PSS)-10, are surveys with a set number of questions to evaluate their self-perceived mental health.

Digital devices such as smartphones and wearables have the capability of collecting passive and active data to monitor the potential onsets of distressful experiences. Preventative measures for mental health concerns, such as the use of digital interventions and longitudinal monitoring, aim to reduce the costs of treatment. These digital devices, including smartphones and wearables, are capable of collecting passive and active data to monitor the potential onset of distressful experiences [5].

In a previous study, we collected the *COVID-19 VR Healthcare Simulation for Distress* (COVID-19 VR HSD) dataset which included collecting various physiological signals and questionnaires from HCW during a VR simulation that elicited stressful situations similar to COVID-19 [6]. We hypothesize that we can classify user distress through the use of physiological signals during a controlled simulation. The motivation for using low-complex analysis is to offer a clinically valuable tool that can be applied for longitudinal monitoring through the use of edge computing applications.

This paper will be the first of its kind to evaluate the self-perceived stress of users by correlating the PSS-10 responses to the collected physiological signals. Data-driven analysis will be applied as we hypothesize that the stressful situations in the VR scenarios will be reflected in the physiological signals.

II. Related Works

In a study by Paul et al. [7], VR was used to administer behavioral activation to a clinical population. This work successfully evaluated the feasibility, acceptability, and tolerability for participants diagnosed with MDD to be exhibited in VR interventions.

The use of physiological signals is a common parameter used to monitor a person's mental health. In a paper by Schmidt et al. [4], the state-of-the-art research of wearable technology that used physiological signals to evaluate affect states was reviewed and outlined. A broad overview was conducted of the methods and theories employed to evaluate user affect from wearable technology. Moreover, Arza et al. [8] created a multivariate approach by using skin temperature, heart rate, and pulse wave signals to evaluate physiological response. The results supported the use of digital biomarkers to monitor stress within participants. An article by Smets et al. [9] proposes the future focus of research to move from laboratory experiments to ambulant stress detection.

Prior work has examined the use of complex computations such as neural networks [10] or multi-layer perceptron [11] to evaluate stress in participants. These proposed models and analyses can achieve high accuracies but lack interpretability and potential translation to edge computing [12].

Edge computing is local computation and storage of data that offers an increase in privacy and the possibility for longitudinal monitoring. Edge computing offers a number of advantages compared to traditional cloud computing, including increased privacy, reduced form factor, and reduced power consumption [13]. With edge computing, data is processed locally before being sent to the cloud, which reduces the number of computations and the amount of power needed. This can make edge computing particularly useful for wearables and other devices that are used for long-term data collection in the wild. The increased privacy, reduced form factor, and reduced power consumption of edge computing can make it more feasible to use wearables for extended periods of time, enabling the collection of more accurate and detailed data.

To the authors' knowledge, no prior research has been conducted to explore the feasibility of VR interventions to examine self-perceived stress in the context of COVID-19. Given this reasoning, we collected the COVID-19 VR HSD dataset [6] to attempt to classify distress in physiological signals. This paper aims to develop a low-complexity analysis to evaluate the separability of the stress severities of the PSS-10 questionnaire. To do so, we empirically evaluate the physiological response of individuals for the identification of potential factors for long-term psychological trauma and impairment associated with ethical conflicts and moral violations.

III. Dataset

This study is the first to analyze perceived stress in the COVID-19 VR HSD dataset [6], which is composed of the collection of physiological data and responses from mental health questionnaires from 15 HCW while undergoing a VR simulation scenario which included a re-enaction of a HCW working in a clinical environment during the COVID-19 pandemic. A flow diagram of the experiment can be seen in Figure 1. The pre-brief was held to inform the participant of the purpose of the session, to collect pre-survey data, and obtain written consent. During the pre- and post-test components, the VR experiment was conducted to identify circumstances that have the potential to cause distress. The intervention was composed of a short video that educated the user based on the Moral Injury Guide [14]. Finally, the debrief was hosted after the experiment and was composed of a semistructured interview, open-ended questions, and an exit survey.

During the experiment, physiological signals were continuously collected, whereas mental health questionnaires were administered at the beginning and end. This paper focuses on respiration (RESP), electrocardiogram (ECG), galvanic skin response (GSR), photoplethsmography (PPG), PSS-10, and information extracted from the signals which include ECG pulse rate (PR), ECG RR interval, respiration rate elevated, and respiration rate. PSS-10 is a questionnaire composed of 10 questions to evaluate the severity of the participant's perceived stress. We have decided to adopt the PSS-10 based on recommendation from the subject matter experts involved in the project, that have vast experience in the fields of psychology. The PSS-10 scale has a 0-40 range with diagnosis of low, moderate and high perceived stress for score ranges of 0-13, 14-26, and 27-40, respectively.

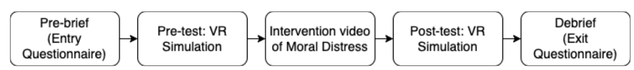

Fig. 1: Diagram of experimental components [6]

IV. Methods

This study is the first to analyze the dataset using a low-complexity pre-processing, feature extraction, and classification pipeline. During pre-processing, physiological data is normalized, segmented, and windowed for analysis. The PSS-10 scores are then clustered into groups of severities. Next, statistical and heart rate variability (HRV) features are extracted for classification using support vector machine (SVM) and decision tree (DT).

A. *Pre-processing*

In order to properly analyze the data, the physiological signals were normalized using min-max normalization [15]. The data was then segmented to before and after the intervention. Segment 1 includes data from the pre-brief, pre-test and the entry questionnaire, and can be considered as the user's baseline data. Segment 2 includes the data from post-test, debrief and the exit questionnaire, and can be considered the affect state post-intervention. To better analyze the data, 20 second windows were used for Segment 1 and 2. The fixed window size was chosen for computational simplicity and experimentally achieved the best results for the separability of user distress.

To pre-process the active data, PSS-10 was encoded to its respective severities and was represented as class 1, 2, and 3, as low, moderate, and high perceived stress, respectively.

B. Feature Analysis

Our preliminary findings and previous studies have influenced the use of statistical and HRV features to be prominent in the detection of stress [8]. Statistical features calculated the mean and variance of the window of each signal. HRV features included low frequency (LF), high frequency (HF), very low frequency (VLF), LF/HF, and root mean square of successive differences (RMSSD). Table I lists the features that were collected alongside their respective descriptions.

TABLE I: Description of features extracted

Feature	Description
Mean	Average value of the window
Variance	Variance value of the window
LF	Relative power of low frequency band of HRV (0.04–0.15 Hz)
HF	Relative power of low frequency band of HRV (0.15–0.4 Hz)
VLF	Relative power of low frequency band of HRV (0.0033–0.04 Hz)
LF/HF	Ratio of Low and High frequency
RMSSD	Root mean square of successive RR interval differences

C. Classification and validation

Classification methodology evaluated the separability between the PSS-10 severities among the users. SVM and DT classifiers are supervised machine learning algorithms and our work utilized a one-vs-all approach in conjunction with a linear SVM, and a binary classification for DT. SVM and DT offer performance metrics such as accuracy, precision, recall, and F1-score. Additionally, the chosen models offer high explainability, that is to understand the decision or output while at the same time offering a low complexity. Having said that, SVM and DT were chosen due to their capabilities for performance, explainability, and complexity given the dataset size.

Stress is an underlying psychological disorder that requires intensive repetition to cause severe stress. Due to the limited time in the VR experiment, causing severe stress would not be ethical or achievable. Therefore, there may be an imbalance in the dataset. To handle the imbalanced dataset, we apply three tests, which include weighted classification, downsampling, and Synthetic Minority Oversampling Technique (SMOTE).

Weighted classification is a method to assign different weights to different classes in a dataset. The weights are inversely proportional to the class frequency and can be calculated using Eq. 1, where w represents the weight for the respective class, C is the number of classes, n_C is the number of samples per class C and N represents the total number of samples.

$$w_C = \frac{N}{C \times n_C} \tag{1}$$

The SMOTE algorithm is an oversampling approach to rebalance the training set. The main purpose of the algorithm is to introduce synthetic samples rather than replication of samples. To do so, the synthetic samples are created by interpolating among the minority class instance within a defined neighborhood [16].

V. Results

After pre-processing, the mean and standard deviation length of the experiments (n=15) were 47.12 and 6.41 minutes, respectively. The average length of Segment 1 and 2 were 22.16 and 24.96 minutes, respectively. A sample visualization of the collected physiological signals can be seen in Fig. 2. Each participant contains two segments before and after the intervention, each segment was windowed with a 20-second window to achieve a total of 2144 windows for 15 participants. Prior to clustering, the histogram of the PSS-10 scores is visualized in Fig. 3. As we can see, the PSS-10 scores are sporadic, therefore clustering was used to organize the group values into their respective severity cases for efficient analysis. The histogram can be seen in Fig. 4 where class 1, 2, and 3, are low, moderate and high perceived stress, respectively. Class 1, 2, and 3 contained 447, 1299, and 398 samples, respectively.

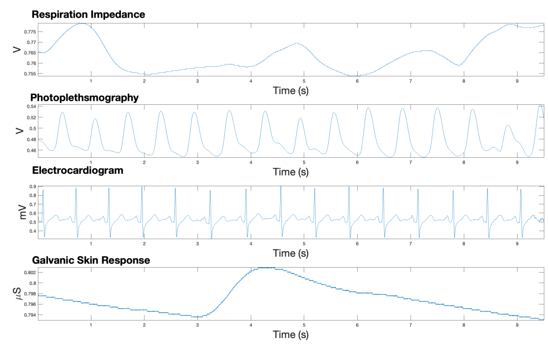

Fig. 2: Physiological signal visualization [6]

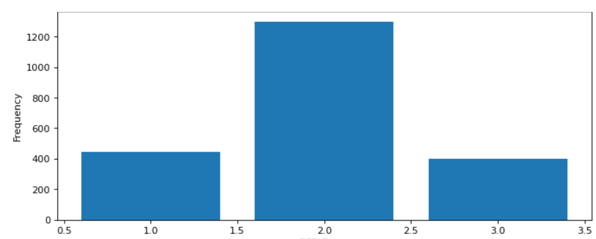

Fig. 3: Histogram for PSS-10 scores

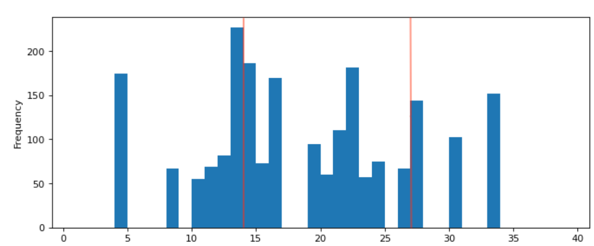

Fig. 4: Histogram for clustered PSS-10 scores

After clustering, the data was split into training and testing sets. DT and SVM achieved an accuracy, precision, recall, and F1 of 82.92%, 80.34%, 80.46%, 80.40% and 68.17%, 64.51%, 71.51%, 65.8%, respectively.

There is a large imbalance within the dataset as seen in Fig. 4. This is expected as experiencing extreme stress is a rare occurrence that requires continuous and extreme exposure to distress. Therefore, we evaluated three tests for imbalanced classification. The results from weighted classification, downsampling, and SMOTE can be seen in Table II. The weighted classification weights for class 1, 2, and 3 were 1.56, 0.55, and 1.79, respectively. It can be seen that the usage of DT achieved the highest classifications through oversampling from SMOTE.

TABLE II: Imbalanced classification results

DT	Weighted Classification	Downsampling	SMOTE
Accuracy (%)	82.23	78.55	**93.50**
Precision (%)	79.79	78.11	**93.41**
Recall (%)	79.94	78.36	**93.31**
F1 (%)	79.82	78.14	**93.35**
SVM			
Accuracy (%)	87.42	70.75	77.69
Precision (%)	84.67	72.76	78.08
Recall (%)	85.65	70.48	77.03
F1 (%)	85.11	70.11	76.92

VI. Discussion and Conclusion

In this study, we evaluated the performance of a low-complexity algorithm clinical tool to detect stress. Our findings support literature to evaluate mental health through the use of physiological data [4], [8]. Additionally, the use of the low-complexity algorithm supports future applications for ambulant stress detection in real-world applications [9]

Advantages of using a low-complexity algorithm are that it allows for potential deployment on edge computing devices, which enables the collection of longitudinal data over extended periods of time, potentially providing valuable insights into individual and population-level patterns of physiological responses. The use of continuous and passive data collection can improve the ecological validity of the collected data, as it reflects real-world situations and behaviors rather than artificially controlled laboratory environments. This can provide a more accurate representation of the physiological responses of individuals in their daily life and can help to improve the generalizability of the results.

However, it is important to consider the potential limitations and ethical concerns such as the collection of large amounts of data over extended periods of time on devices may present challenges for data storage and analysis. Additionally, it is crucial to obtain informed consent from participants and to adhere to ethical guidelines in the collection and use of physiological data.

Overall, our results suggest that the low-complexity algorithm used in this study has potential for deployment in continuous and passive data collection of users in their daily life. Further research is needed to evaluate the performance of the algorithm in real-world settings and to address the potential limitations and ethical concerns of continuous and passive data collection.

Acknowledgment

The authors would like to thank NSERC, OGS, and Canada's DND IDeAS Program for funding the project. We would like to acknowledge and thank the Unity Health Toronto Simulation Program.

References

[1] M. Travers, "COVID-19 Poses A Major Quandary For The Field Of Moral Psychology." https://www.forbes.com/sites/traversmark/2020/04/15/covid-19-poses-a-major-quandary-for-the-field-of-moral-psychology/. 2022-02-26.
[2] E. G. Epstein and S. Delgado, "Understanding and addressing moral distress," *Online Journal of Issues in Nursing*, vol. 15, sep 2010.
[3] M. H. Hatta, H. Sidi, C. S. Koon, N. A. C. Roos, S. Sharip, F. D. A. Samad, O. W. Xi, S. Das, and S. M. Saini, "Virtual Reality (VR) Technology for Treatment of Mental Health Problems during COVID-19: A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 9, p. 5389, 2022.
[4] P. Schmidt, A. Reiss, R. Dürichen, and K. V. Laerhoven, "Wearable-based affect recognition—a review," *Sensors (Switzerland)*, vol. 19, no. 19, pp. 1–42, 2019.
[5] Y. National Research Council (US) and Institute of Medicine (US) Committee on the Prevention of Mental Disorders and Substance Abuse Among Children, Y. A. R. Advances, P. Interventions, M. E. O'Connell, T. Boat, and K. E. Warner, "Benefits and Costs of Prevention," 2009.
[6] B. Nguyen, A. Torres, W. Sim, D. Kenny, D. M. Campbell, L. Beavers, W. Lou, B. Kapralos, E. Peter, A. Dubrowski, S. Krishnan, and V. Bhat, "Digital Interventions to Reduce Distress Among Health Care Providers at the Frontline: Protocol for a Feasibility Trial," *JMIR Res Protoc*, vol. 11, p. e32240, feb 2022.
[7] M. Paul, K. Bullock, and J. Bailenson, "Virtual Reality Behavioral Activation as an Intervention for Major Depressive Disorder: Case Report," *JMIR Ment Health 2020;7(11):e24331 https://mental.jmir.org/2020/11/e24331*, vol. 7, p. e24331, nov 2020.
[8] A. Arza, J. M. Garzón-Rey, J. Lázaro, E. Gil, R. Lopez-Anton, C. de la Camara, P. Laguna, R. Bailon, and J. Aguiló, "Measuring acute stress response through physiological signals: towards a quantitative assessment of stress," *Medical and Biological Engineering and Computing*, vol. 57, pp. 271–287, jan 2019.
[9] E. Smets, W. De Raedt, and C. Van Hoof, "Into the Wild: The Challenges of Physiological Stress Detection in Laboratory and Ambulatory Settings," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, pp. 463–473, mar 2019.
[10] Q. Xu, T. L. Nwe, and C. Guan, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, pp. 275–281, jan 2015.
[11] A. Arsalan, M. Majid, S. M. Anwar, and U. Bagci, "Classification of Perceived Human Stress using Physiological Signals," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 1247–1250, jul 2019.
[12] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining Explanations: An Overview of Interpretability of Machine Learning," *Proceedings - 2018 IEEE 5th International Conference on DS and Advanced Analytics, DSAA 2018*, pp. 80–89, may 2018.
[13] F. Qureshi and S. Krishnan, "Wearable Hardware Design for the Internet of Medical Things (IoMT)," *Sensors 2018, Vol. 18, Page 3812*, vol. 18, p. 3812, nov 2018.
[14] "Phoenix Australia – Centre for Posttraumatic Mental Health and the Canadian Centre of Excellence – PTSD," *Phoenix Australia – Centre for Posttraumatic Mental Health and the Canadian Centre of Excellence – PTSD moralinjuryguide.ca*.
[15] S. Krishnan, "Biomedical signal analysis for connected healthcare," *Elsevier*, 2021.
[16] A. Fernández, S. García, F. Herrera, and N. V. Chawla, "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary," *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905, apr 2018.