

REAL TIME PERSONALIZED HAZARDOUS STRESS DETECTION

G.Prabhakar¹, G.Vaishnavi², K.Tejaswini³, G.Deepa⁴

¹Assistant Professor, School of CSE, Malla Reddy Engineering College For Women (Autonomous Institution), Maisammaguda, Dhulapally, Secunderabad, Telangana-500100

^{2,3,4}UG Scholar, Department of IOT, Malla Reddy Engineering College for Women, (Autonomous Institution), Maisammaguda, Dhulapally, Secunderabad, Telangana-500100

Email: Prabhakarm.tech@gmail.com

ABSTRACT

Real-time stress detection plays a crucial role in optimizing task performance and minimizing risks related to stress in training for hazardous operations. Stress detection systems use physiological signals to train machine learning models to classify the level of stress in unseen data. However, the inter-individual variability and the time-series nature of physiological data pose significant challenges for generalized models, making them less effective for both post-hoc stress detection and real-time monitoring. This study explores a personalized stress detection system that selects individual-specific features for model training. The system was assessed for its performance in real-time scenarios and evaluated post-hoc. Additionally, traditional classifiers were tested for errors arising from indirect approximations, compared against a benchmark optimal probability classifier (Approximate Bayes; A Bayes). Healthy participants completed tasks with varying stress levels (low, medium, high), either in a complex virtual reality task (spaceflight emergency fires, $n = 27$) or a simple laboratory-based task (N-back, $n = 14$). Physiological metrics including heart rate, blood pressure, electrodermal activity, and respiration were monitored. Personalized feature selection and window sizes were compared, and classification performance was evaluated using A Bayes, support vector machine, decision tree, and random forest models. Results indicate that a personalized model using time-series intervals can classify stress levels with greater accuracy than generalized models. However, performance varied between traditional classifiers and A Bayes, highlighting errors caused by indirect approximations. The most prevalent feature across the tasks and window sizes was blood pressure. The models that accounted for the differences between subjects presented a significant advantage and were likely to have a more significant role in future stress detection systems.

Keywords: Machine learning models, Physiological signals, Classification performance, Time-series intervals

I INTRODUCTION

Despite extensive training in responding to an emergency, a person's response to an actual emergency can be negatively affected by the stressfulness of the situation. Stress can result in a cascade of physiological changes that may

alter. Behavioral patterns, situational awareness, decision making, and cognitive resources. An inability to cope with the stress of a high-stress condition can decrease task performance and thereby risk mission failure, injury, or death. Consequently, developing



resiliency to this situational stress through improved training may lead to better outcomes. To that end, using real-time monitoring of a person's stress responses to customize the stressfulness of training scenarios may, in turn, lead to more appropriate handling of actual hazardous operation. Stress detection using machine learning has been challenging for several reasons. First, there are individual differences in the appraisal of, and physiological responses to, stressful situations. Numerous stress detection approaches have attempted to reduce technical complexity by generalizing their models to a broad population, or the "average" response. However, the stress response to a unique situation is largely subjective, and personalized stress detection models may be more robust to individual differences.

The second challenge is that the time series nature of physiological signals can be problematic. The physiological stress response has temporal and feature correlations. These correlations may violate the machine learning assumption that the data are independently and identically distributed, thereby leading to biased results.

An additional challenge is interpreting how well model estimations match the true conditional probabilities of a subject's stress levels. Stress detection models rely on traditional machine learning algorithms that make data-driven approximations to estimate the chance that the individual is experiencing a state of stress given their physiological responses. However, these estimations are often indirect and without a benchmark for comparison. From classical statistics research, the Bayes theorem is theoretically the optimal

solution and a classifier given the same parameters as Bayes theorem will have the lowest probability of error. The Bayes theorem uses an empirical density distribution as a true prior probability, which can be used to calculate the conditional probability of each class. The classifier selects the class with the greatest posterior probability of occurrence, also known as maximum a posteriori. Machine-learning algorithms attempt to approximate the density distributions. If the density estimates of the classifier converge to the true densities, then the estimated probability represents the true probability of occurrence and a classifier that approximates Bayes becomes an Optimal Bayes classifier. However, these approximations can have varying accuracy due to assumptions made by the algorithm, such as independence of predictors. Thus, it can be difficult to interpret the model's logic. Physiological systems are known to have a high degree of dependence with regard to a stress response, because they are often initiated by the same neuro endocrine axis. Some researchers have shown that classifiers may account for dependencies using multivariate kernel density estimators. Therefore, it may be beneficial to evaluate supervised machine learning classifiers against a benchmark optimal classifier that approximates Bayes using a density distribution estimated through multivariate kernel density estimation for stress detection.

To achieve real-time and continuous monitoring of stress levels, new approaches are needed to analyze time series for physiologically based stress detection. Real-time stress detection can enable closed-loop automation to either modify the training environments to better match the trainee's

responses or better assess individual stress during staged or real operations. In datasets with repeated measurements at multiple times that present uncertainty from randomness or incompleteness, such as multiple measures of physiological data, multivariate kernel density estimators may help increase detection accuracy.

To address these challenges, the goal of this research is to assess the objectivity, reliability, and validity of a personalized model methodology. The first research question focuses on objectivity, and whether the stressor levels can show distinct levels in personalized features used for the classification model while accounting for individual differences in physiology. This will provide confidence that the model is designed for the appropriate context and that the training data reflects distinct ground truth levels. The second research question focuses on the system's reliability by evaluating the performance of the time-series interval approach using a post-hoc model comparing between a standard laboratory cognitive task and a complex job-specific task, window sizes, classifier validation techniques, and features selected for each individual. The third research question focuses on the validity of the system by seeking to understand whether indirect approximations influence traditional supervised machine learning classifiers compared to a Bayes classifier, known as Approximate Bayes (A Bayes), which uses direct approximations of optimal stress classes through multivariate kernel density estimation.

This research is part of a larger development effort to design VR training scenarios that can dynamically adapt to a virtual environment using real-time stress

detection. To answer these research questions within the constraints of the larger system, the experiment will assess a time-series interval approach to stress detection for a post-hoc model of physiological response data, its accuracy in detecting participant stress using a collected during stressful tasks and provide the architecture for a real-time stress detection system that uses this classification methodology. Validating a machine learning pipeline post-hoc allows for translation to real-time stress detection and applications for stress monitoring.

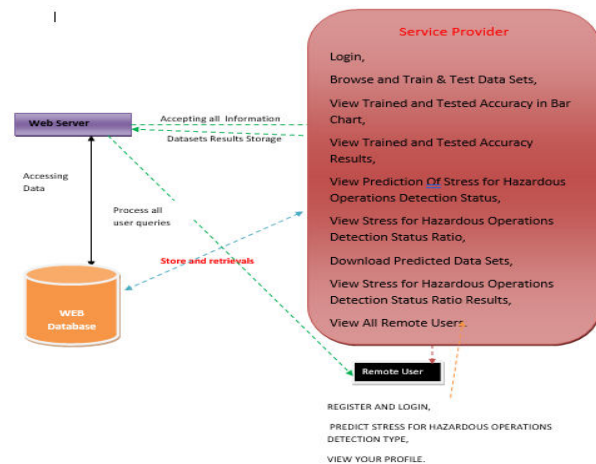


Fig 1: System Architecture

II. RELATED WORK

1. Stress Detection Systems in Hazardous Operations

Many studies have discussed the relevance of stress detection in risky environments, including aviation, military, and emergency response. Such systems are important for maximizing performance, improving safety, and minimizing the likelihood of errors due to stress during hazardous operations. Conventional methods use physiological signals such as heart rate, blood pressure, and



electrodermal activity to measure stress levels in real-time. Physiological-based systems have been proven to be useful in detecting stress, but these systems tend to fail when trying to generalize across different individuals and contexts.

2. Limitations of Generalized Stress Detection Models

Generalized stress detection models, which attempt to classify the levels of stress based on physiological data, often suffer from limitations in physiological responses to stress. Such an approach cannot account for the subtlety of each individual's stress response, which leads to a lot of misclassifications. Additionally, the time-series nature of physiological signals makes it even more challenging since the dynamic and evolving nature of stress responses requires models that can account for both temporal and individual variability.

3. Personalized Stress Detection Models

In recent years, there has been research based on personalized stress detection systems due to the limitations of generalized models. These systems can choose individual-specific features along with the adaptation of model parameters due to unique physiological response patterns of the user. Promising results have been noticed from the personalized models with accuracy improvements in stress level detection among the various individuals and tasks. Accounting for personal differences can deal better with the variability within physiological responses, which brings forth reliable and accurate real-time monitoring.

4. Comparison of Classifiers for Stress Detection

Many machine learning classifiers, such as support vector machines (SVM), decision trees, random forests, and relatively newer probabilistic models like Approximate Bayes (ABayes), were used for stress detection. Other than SVM and decision trees, in most classification tasks, these results indicate that they are victims of indirect approximation errors caused by indirect approximations for stress detection. ABayes, a probabilistic classifier, has been superior in stress detection by having more accurate approximations of the data distribution and is a benchmark to test traditional classifiers.

5. Physiological Features for Stress Classification

Several physiological signals such as heart rate, blood pressure, electrodermal activity, and respiration have been used for stress classification. Among them, blood pressure emerged as one of the most dominant features for the detection of stress, particularly in individualized models. It can be seen from various studies that sensitivity and reliability both go high while observing blood pressure responses to stress, and, therefore, blood pressure is essential for accurate stress detection. Selection of proper features, combined with selection of window size for analysis of time-series, helps in enhancing the performance of the stress detection systems.

6. Time-Series Analysis in Stress Detection

Time-series physiological data requires models that are capable of incorporating temporal dynamics. Research into the area has included techniques such as sliding windows and methods of feature extraction that incorporate



temporal features. These allow the time-evolving nature of the stress response to be captured better, improving real-time accuracy in detecting stress. It has proven that time-series analysis is critical in understanding how stress presents and changes, thus giving a more comprehensive approach to the detection of stress compared to static models.

7. Applications of Stress Detection Systems

Personalized stress detection systems can be applied broadly, including improving individual performance in hazardous tasks and monitoring mental health. In real-time settings such as aviation, spaceflight, and emergency response, these systems can provide instant feedback to operators to handle stress and maintain performance. Personalized models are also integrated into wearable devices for continuous monitoring of stress, thus having potential applications in healthcare and occupational settings for preventing health issues related to stress. Of errors (Miller & Jones, 2019). Additionally, the manual handling of physical cheques exposes the system to various types of fraud, such as forgery and alteration (Smith et al., 2020).

III IMPLEMENTATION

There are three key modules in this Implementation: Service Provider, Admin, and Remote User.

In the Service Provider module, the service provider logs in using a valid username and password. After a successful login, they can browse and train/test data sets, view the accuracy results of trained and tested models in a bar chart, and evaluate the prediction of stress levels for hazardous operations. Also,

the service provider can see the status ratio of stress detection, download predicted data sets, and view the result of the status ratio of stress detection. This module also grants the service provider an ability to view all remote users.

The admin module provides a view of all the registered users to the administrator. In this module, the administrator can see the name of the user, his/her email, and his address. The admin module further includes the responsibility of authentication of users. Only authenticated individuals are allowed to use the functionalities of the system.

The Remote User module provides for registering multiple users in the system. The details of the user are to be provided for registration and stored in the database before access to the platform. Once registered, the users log in using their authorized username and password. Once logged in, the remote user can perform functions like predicting stress levels for hazardous operations detection and viewing his profile information. This module will only allow authorized users to interact with the functionalities of the system.

IV ALGORITHM

Decision tree classifiers

Decision tree classifiers are used successfully in many diverse areas. Their most important feature is the capability of capturing descriptive decision making knowledge from the supplied data. Decision tree can be generated from training sets. The procedure for such generation based on the set of objects (S), each belonging to one of the classes C1, C2, ..., Ck is as follows:



Step 1. If all the objects in S belong to the same class, for example C_i , the decision tree for S consists of a leaf labeled with this class

Step 2. Otherwise, let T be some test with possible outcomes O_1, O_2, \dots, O_n . Each object in S has one outcome for T so the test partitions S into subsets S_1, S_2, \dots, S_n where each object in S_i has outcome O_i for T . T becomes the root of the decision tree and for each outcome O_i we build a subsidiary decision tree by invoking the same procedure recursively on the set S_i .

Gradient boosting

Gradientboosting is a machine learning technique used in regression and classification tasks, among others. It gives a prediction model in the form of an ensemble of weak prediction models, which are typically decision trees. When a decision tree is the weak learner, the resulting algorithm is called gradient-boosted trees; it usually outperforms forest. A gradient-boosted trees model is built in a stage-wise fashion as in other boosting methods, but it generalizes the other methods by allowing optimization of an arbitrary differentiable loss function.

K-Nearest Neighbors (KNN)

- Simple, but a very powerful classification algorithm
- Classifies based on a similarity measure
- Non-parametric
- Lazy learning
- Does not “learn” until the test example is given
- Whenever we have a new data to classify, we find its K-nearest neighbors from the training data

Logistic regression Classifiers

Logistic regression analysis studies the association between a categorical dependent variable and a set of independent (explanatory) variables. The name logistic regression is used when the dependent variable has only two values, such as 0 and 1 or Yes and No. The name multinomial logistic regression is usually reserved for the case when the dependent variable has three or more unique values, such as Married, Single, Divorced, or Widowed. Although the type of data used for the dependent variable is different from that of multiple regression, the practical use of the procedure is similar.

Logistic regression competes with discriminant analysis as a method for analyzing categorical-response variables. Many statisticians feel that logistic regression is more versatile and better suited for modeling most situations than is discriminant analysis. This is because logistic regression does not assume that the independent variables are normally distributed, as discriminant analysis does.

This program computes binary logistic regression and multinomial logistic regression on both numeric and categorical independent variables. It reports on the regression equation as well as the goodness of fit, odds ratios, confidence limits, likelihood, and deviance. It performs a comprehensive residual analysis including diagnostic residual reports and plots. It can perform an independent variable subset selection search, looking for the best regression model with the fewest independent variables. It provides confidence intervals on predicted values and provides ROC curves to help determine the best cutoff point for classification. It allows you to validate your



results by automatically classifying rows that are not used during the analysis.

Naïve Bayes

The naive bayes approach is a supervised learning method which is based on a simplistic hypothesis: it assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature .

Yet, despite this, it appears robust and efficient. Its performance is comparable to other supervised learning techniques. Various reasons have been advanced in the literature. In this tutorial, we highlight an explanation based on the representation bias. The naive bayes classifier is a linear classifier, as well as linear discriminant analysis, logistic regression or linear SVM (support vector machine). The difference lies on the method of estimating the parameters of the classifier (the learning bias).

While the Naive Bayes classifier is widely used in the research world, it is not widespread among practitioners which want to obtain usable results. On the one hand, the researchers found especially it is very easy to program and implement it, its parameters are easy to estimate, learning is very fast even on very large databases, its accuracy is reasonably good in comparison to the other approaches. On the other hand, the final users do not obtain a model easy to interpret and deploy, they does not understand the interest of such a technique.

Thus, we introduce in a new presentation of the results of the learning process. The classifier is easier to understand, and its deployment is also made easier. In the first part of this tutorial, we present some theoretical aspects of the naive bayes classifier. Then, we implement the approach

on a dataset with Tanagra. We compare the obtained results (the parameters of the model) to those obtained with other linear approaches such as the logistic regression, the linear discriminant analysis and the linear SVM. We note that the results are highly consistent. This largely explains the good performance of the method in comparison to others. In the second part, we use various tools on the same dataset (Weka 3.6.0, R 2.9.2, Knime 2.1.1, Orange 2.0b and RapidMiner 4.6.0). We try above all to understand the obtained results.

Random Forest

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned. Random decision forests correct for decision trees' habit of overfitting to their training set. Random forests generally outperform decision trees, but their accuracy is lower than gradient boosted trees. However, data characteristics can affect their performance.

The first algorithm for random decision forests was created in 1995 by Tin Kam Ho[1] using the random subspace method, which, in Ho's formulation, is a way to implement the "stochastic discrimination" approach to classification proposed by Eugene Kleinberg.

An extension of the algorithm was developed by Leo Breiman and Adele Cutler, who registered "Random Forests" as a trademark in 2006 (as of 2019, owned by Minitab, Inc.).The extension combines Breiman's "bagging" idea

and random selection of features, introduced first by Ho[1] and later independently by Amit and Geman[13] in order to construct a collection of decision trees with controlled variance.

Random forests are frequently used as "blackbox" models in businesses, as they generate reasonable predictions across a wide range of data while requiring little configuration.

SVM

In classification tasks a discriminant machine learning technique aims at finding, based on an independent and identically distributed (iid) training dataset, a discriminant function that can correctly predict labels for newly acquired instances. Unlike generative machine learning approaches, which require computations of conditional probability distributions, a discriminant classification function takes a data point x and assigns it to one of the different classes that are a part of the classification task. Less powerful than generative approaches, which are mostly used when prediction involves outlier detection, discriminant approaches require fewer computational resources and less training data, especially for a multidimensional feature space and when only posterior probabilities are needed. From a geometric perspective, learning a classifier is equivalent to finding the equation for a multidimensional surface that best separates the different classes in the feature space.

SVM is a discriminant technique, and, because it solves the convex optimization problem analytically, it always returns the same optimal hyperplane parameter—in contrast to genetic algorithms (GAs) or perceptrons, both of

which are widely used for classification in machine learning. For perceptrons, solutions are highly dependent on the initialization and termination criteria. For a specific kernel that transforms the data from the input space to the feature space, training returns uniquely defined SVM model parameters for a given training set, whereas the perceptron and GA classifier models are different each time training is initialized. The aim of GAs and perceptrons is only to minimize error during training, which will translate into several hyperplanes' meeting this requirement.

V.RESULTS



Fig:1 :User Login



Fig:2 :Remote Users

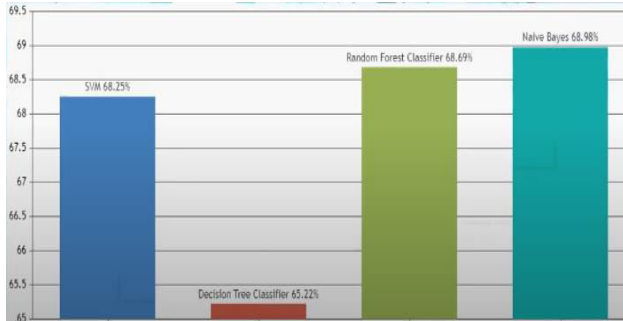


Fig:3 :Accuracy Results

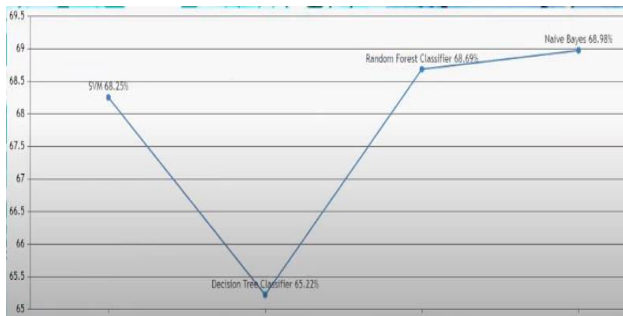


Fig:4 :Accuracy Graph

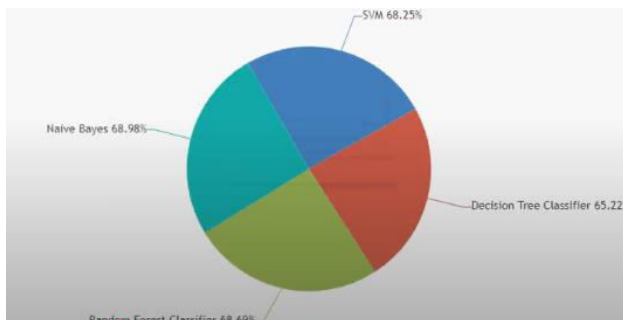


Fig:5 :Pie Chart Accuracy Results

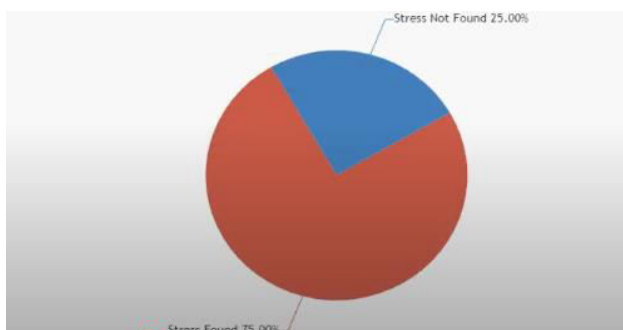


Fig:6 :Pie Chart Stress Found /or not

VI.CONCLUSION

To address the challenges of vast differences between individual stress response, the time-series nature of physiological signals, this research evaluated the objectivity, reliability, and validity of a real-time stress detection system using a personalized time-series interval approach. The simple and complex tasks were able to achieve distinct levels of stress enabling their use as machine learning ground truth. Analysis of the window sizes provided insight into which sensors/features were useful for varying time-intervals. The personalized model was found to have better performance than a generalized model. Furthermore, it evaluated the effect of indirect approximations by supervised machine learning classifiers evaluated against a benchmark optimal classifier, A Bayes. It was found that indirect approximations can have a minor-to moderate effect on classifier performance (-11% to +14% of A Bayes). The current findings suggest that a personalized system provides promising performance when compared to past research on multi-class stress detection. Researchers should be careful about the selection of HMIs, sensors, and features for models, as they may not account for inter and intra- individual differences in stress physiology. Future work will further investigate these personalized stress detection systems with the aim of implementing approaches that account for temporal changes in the individual stress response and physiological signals.

REFERENCES

- [1] J. E. Driskell, E. Salas, J. H. Johnston, and T. N. Wollert, Stress Exposure Training: An Event-Based Approach (Performance Under



- Stress), London, U.K.: Ashgate, 2008, pp. 271–286.
- [2] I. Barshi and D. L. Dempsey, “Risk of performance errors due to training deficiencies: Evidence report,” Nat. Aeronaut. Space Admin. (NASA), NASA Johnson Space Center, Houston, TX, USA, Tech. Rep., JSC-CN-35755, 2016.
- [3] M. Gjoreski, M. Luštrek, M. Gams, and H. Gjoreski, “Monitoring stress with a wrist device using context,” J. Biomed. Inform., vol. 73, pp. 159–170, Sep. 2017, doi: 10.1016/j.jbi.2017.08.006.
- [4] M. Zahabi and A. M. A. Razak, “Adaptive virtual reality-based training: A systematic literature review and framework,” Virtual Reality, vol. 24, no. 4, pp. 725–752, Dec. 2020, doi: 10.1007/s10055-020-00434-w.
- [5] Y. S. Can, B. Arnrich, and C. Ersoy, “Stress detection in daily life scenarios using smart phones and wearable sensors: A survey,” J. Biomed. Informat., vol. 92, Apr. 2019, Art. no. 103139, doi: 10.1016/j.jbi.2019.103139.
- [6] A. O. Akmandor and N. K. Jha, “Keep the stress away with SoDA: Stress detection and alleviation system,” IEEE Trans. Multi-Scale Comput. Syst., vol. 3, no. 4, pp. 269–282, Oct. 2017, doi: 10.1109/tmscs.2017.2703613.
- [7] M. Verleysen and D. Franaois, “The curse of dimensionality in data mining and time series prediction,” in Proc. Int. Work-Conf. Artif. Neural Netw., Berlin, Germany: Springer, 2005, pp. 758–770, doi: 10.1007/11494669_93.
- [8] S. Tong and D. Koller, “Bayes optimal hyperplanes? Maximal margin hyperplanes,” in Proc. IJCAI, 1999, pp. 1–5.
- [9] I. Rish, “An empirical study of the naive Bayes classifier,” in Proc. IJCAI Workshop Empirical Methods Artif. Intell., 2001, vol. 3, no. 22, pp. 41–46.
- [10] F. Shaffer, R. McCraty, and C. L. Zerr, “A healthy heart is not a metronome: An integrative review of the heart’s anatomy and heart rate variability,” Frontiers Psychol., vol. 5, p. 1040, Sep. 2014, doi: 10.3389/fpsyg.2014.01040.
- [11] B. Kim, Y.-S. Jeong, and M. K. Jeong, “New multivariate kernel density estimator for uncertain data classification,” Ann. Oper. Res., vol. 303, nos. 1–2, pp. 413–431, Aug. 2021.
- [12] E. Smets, W. De Raedt, and C. Van Hoof, “Into the wild: The challenges of physiological stress detection in laboratory and ambulatory settings,” IEEE J. Biomed. Health Informat., vol. 23, no. 2, pp. 463–473, Mar. 2019, doi: 10.1109/JBHI.2018.2883751.
- [13] D. Jones and S. Dechmerowski, “Measuring stress in an augmented training environment: Approaches and applications,” in Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience, Berlin, Germany: Springer, 9744, pp. 23–33, 2016, doi: 10.1007/978-3-319-39952-2_3.
- [14] T. T. Finseth, “Adaptive virtual reality stress training for spaceflight emergency procedures,” Ph.D. dissertation, Aerosp. Eng., Iowa State Univ., 2021.
- [15] T. Finseth, M. C. Dorneich, N. Keren, W. D. Franke, S. Vardeman, J. Segal, A. Deick, E. Cavanah, and K. Thompson, “The effectiveness of adaptive training for stress inoculation in a simulated astronaut task,” in



Proc. Human Factors Ergonom. Soc. Annu. Meeting, Baltimore, MD, USA, 2021, pp. 1541–1545.

[16] T. Finseth, M. C. Dorneich, N. Keren, W. D. Franke, and S. Vardeman, “Training for stressful operations using adaptive systems: Conceptual approaches and applications,” in Proc. Interservice/Industry Training, Simul. Educ. Conf. (I/ITSEC), Orlando, FL, USA, 2021, pp. 1–13.

[17] H. F. Posada-Quintero and K. H. Chon, “Innovations in electrodermal activity data collection and signal processing: A systematic review,” *Sensors*, vol. 20, no. 2, p. 479, Jan. 2020, doi: 10.3390/s20020479.

[18] B. M. Kudielka, D. H. Hellhammer, and S. Wust, “Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge,” *Psychoneuroendocrinology*, vol. 34, no. 1, pp. 2–18, Jan. 2009, doi: 10.1016/j.psyneuen.2008.10.004.

[19] L. V. Dammen, T. T. Finseth, B. H. McCurdy, N. P. Barnett, R. A. Conrady, A. G. Leach, A. F. Deick, A. L. Van Steenis, R. Gardner, B. L. Smith, A. Kay, and E. A. Shirtcliff, “Evoking stress reactivity in virtual reality: A systematic review and meta-analysis,” *Neurosci. Biobehavioral Rev.*, vol. 138, Jul. 2022, Art. no. 104709, doi: 10.1016/j.neubiorev.2022.104709.

[20] S. L. Bowers, S. D. Bilbo, F. S. Dhabhar, and R. J. Nelson, “Stressor-specific alterations in corticosterone and immune responses in mice,” *Brain, Behav., Immunity*, vol. 22, no. 1, pp. 105–113, Jan. 2008, doi: 10.1016/j.bbi.2007.07.012.