Personalized Stress Monitoring Using Unsupervised LSTM And Autoencoders

Aamee Chandra[1], Suryam Kumar[2], Dr. Suresh Anand M[3]
[1,2,3]*Dept. of Computing Technologies, SRMIST, KTR Chennai, India*

***Abstract*—Chronic psychological stress represents a catalyst of long-term physiological degeneration and constitutes a significant cause of non-communicable diseases, encompassing hypertension, cardiovascular failure, and clinical depression. Current diagnostic methods are generally self-reporting or periodic clinical evaluation, each of which is presently lacking in terms of sufficient temporal detail for preemptive diagnosis. Not with standing recent successes in Machine Learning (ML)-based affective computing, supervised learning approaches face limitations with respect to the "Labeling Gap," the practical unfeasibility of acquiring accurate stress labels, and the extensive inter-subject variability for each subject. This paper proposes a unique approach for the development of a personalized stress detection system using Long Short-Term Memory Autoencoders. The method works on the assumption that stress is an anomalous time-series in relation to an individual's learned baseline. The model is presented only with "normal physiological states". A compressed view of the individual's health equilibrium is learned, and anomalies are picked up via high levels of Reconstruction Error (RE). This study integrates Goal 3, Good Health and Well-being, due to its ability to measure, scale, and protect individuals' health in an objective manner.**

I. INTRODUCTION

1.1 Motivation

The modern era is characterized by an "invisible pandemic" of stress. According to the World Health Organization (WHO), stress-related conditions are leading causes of reduced productivity and premature mortality. The challenge in managing stress lies in its subjective nature; what constitutes a physiological "peak" for one individual might be a standard "resting state" for another.

1.2 The Problem of Inter-Subject Variability

Today, the world of global health is ever more dominated by the phenomenon of the "invisible pandemic" of chronic psychological stress, as described by the World Health Organization (WHO). In their assessment, stress-related pathologies are a leading cause of decreased productivity in the workplace and premature death from non-communicable diseases. The core challenge in managing stress lies in its subjective nature, as a state of high physiological arousal for one person may be a state of complete relaxation for another person. Similarly, existing methodologies for stress detection, such as self-reporting surveys and clinical assessments, are often not sensitive enough for proactive intervention. *[1]*

In spite of recent breakthroughs in Machine Learning (ML)-based affective computing, existing state-of-the-art models are still limited by what is known as the "Labeling Gap Problem" in stress detection, where obtaining accurate stress labels in real-time for "in-the-wild" experiments is extremely challenging. *[13]*

In addition, most existing models are "Population Models" that are based on averages, failing to recognize the inherent inter-subject variability in stress levels depending on factors such as age, fitness levels, and pre-existing medical conditions. In this paper, we seek to bridge this existing gap by proposing a paradigm shift from a generalized stress detection model. *[9]*

1.3 Research Objective

This paper aims to shift the focus from *Generalized Classification* to *Personalized Anomaly Detection*. The objective is to develop a system that:

1. Requires zero manual labeling from the user.
2. Learn the unique "physiological signature" of an individual.
3. Provides real-time alerts when current signals deviate from the learned baseline. *[8] [11]*

II. LITERATURE REVIEW

The theoretical basis for monitoring human emotion and stress through technology was first introduced by Picard et al., who founded the field of affective computing. The initial study employed digital signal processing technology for monitoring non-invasive physiological signals, where it was determined that Electrodermal Activity (EDA) and heart rate are major indicators of psychological state. To enable effective research in this area, standardized data sets such as WESAD (Wearable Stress and Affect Detection) have been introduced, where data is provided in a multimodal format through controlled environments. *[3] [10]*

Although supervised models have proven effective on these data sets, applying them in practical situations is proving difficult due to the inability to obtain accurate data sets in real-time situations. This "labeling gap" has led to the need for unsupervised machine learning approaches for stress detection. Recent studies have focused on self-supervised feature learning for stress detection.

The architectural advancements in stress detection systems have progressed from conventional classifiers to Deep Learning approaches. The conventional Recurrent Neural Network (RNN) was employed for stress detection, but this was limited by the "vanishing gradient" issue. The introduction of Long Short-Term Memory (LSTM) units addressed this problem through a "gating" mechanism. *[2]*

2.1 Supervised Learning and its Limitations

Studies by authors like Picard et al. introduced the foundation of affective computing, utilizing sensors like Electrodermal Activity (EDA) to prove the success of models such as Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) in highly controlled lab environments, such as the WESAD dataset. However, such models require "Ground Truth" labels, and the only way to obtain this "Ground Truth" information is to induce stress on the human in the lab environment. How can the user possibly label their stress levels at an interval of 5 minutes in "in-the-wild" scenarios?

2.2 Deep Learning for Time-Series

The invention of RNNs brought a revolution to the analysis of time-series data. Unlike standard feed-forward networks, RNNs possess "internal memory." However, standard RNNs suffer from the "vanishing gradient problem," which renders them unable to remember long-term dependencies. Long Short-Term Memory-LSTM-networks were introduced to solve this by using a gating mechanism of input, forget, and output gates. *[2]*

2.3 The Shift to Unsupervised Anomaly Detection

There is recent research in the use of Autoencoders in anomaly detection of industrial equipment and cyber security applications. The application of this concept to human physiology is still at an early stage, yet holds great promise.

Malhotra et al.'s research on anomaly detection using multiple sensors has laid the foundation on how the use of Encoder-Decoder models could be applied to the anomaly detection of periodic patterns. Our work takes this concept and applies it uniquely to the complexities of human stress patterns. *[4]*

In the realm of anomaly detection, the authors have shown the efficacy of LSTM-based architectures of the Encoder-Decoder type in capturing "normal" patterns of multi-sensor data and detecting anomalies through the error of reconstruction. This paradigm has been extended to the realm of real-time health monitoring and vibration-based anomaly detection. Further, the developments in the realm of adversarial autoencoders have enabled the creation of highly personalized stress detection systems.

Instead of relying on the "average" physiological signature of a person belonging to a class of people, the system can now learn the unique signature of a person. *[5] [12]*

III. BIOLOGICAL BASIS OF SENSORS

To detect stress accurately, we must monitor the heart rate (rest, physical activity, stress and false alarms), sleeping pattern, etc.

3.1 Photoplethysmography (PPG)

PPG sensors measure changes in blood volume through light absorption. From the PPG signal, we derive:

- Heart Rate (HR): Increases during sympathetic activation.

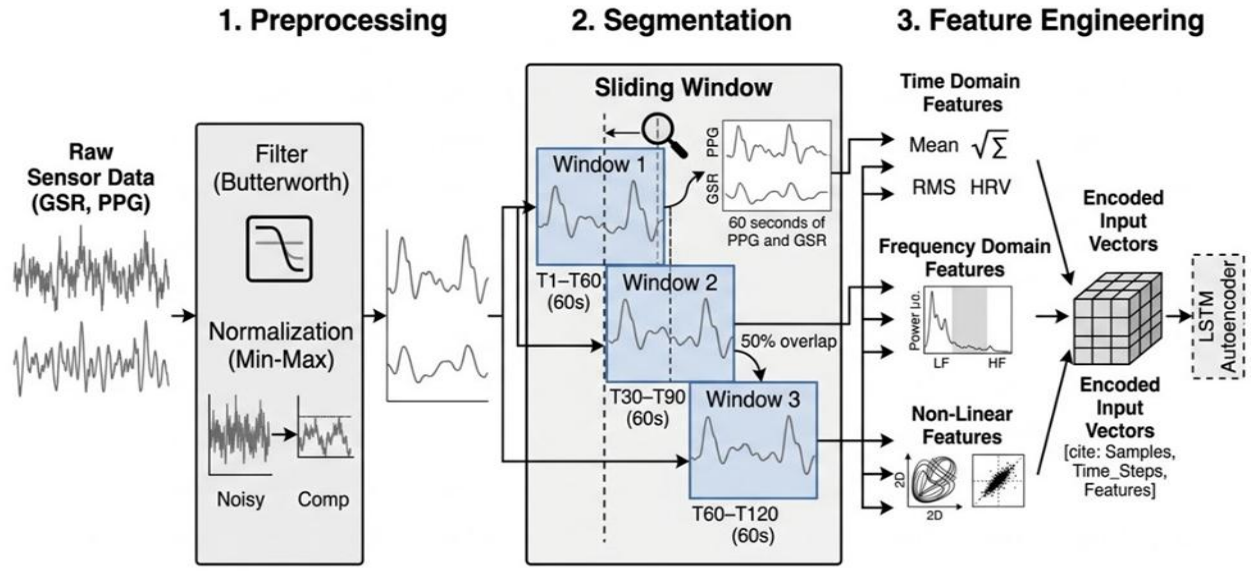

Figure 1: Sliding Window Pipeline

- Heart Rate Variability (HRV): Specifically, the RMSSD (Root Mean Square of Successive Differences). A *decrease* in HRV is a primary indicator of psychological stress. *[11]*

3.2 Electrodermal Activity (EDA/GSR)

GSR measures the conductivity of the skin. When stressed, the hypothalamus triggers sweat glands (even in microscopic amounts). This increases skin conductance. GSR is considered one of the most "honest" physiological markers because it cannot be consciously controlled. *[14]*

IV. METHODOLOGY

4.1 Data Preprocessing Pipeline

Data Collection-

To assess the efficacy of our wearable device, data was collected from a personalized longitudinal study. Each subject was provided with the device for 72 hours of continuous use.

To ensure the integrity of the "Normal Baseline," the initial 48 hours were devoted to acquiring non-stressful physiological states. Each subject was instructed to continue their regular daily routine, which includes sedentary work, domestic activity, and nocturnal sleep, without any known high-stress triggers or intense physical activity that may induce tachycardia from exercise.

Raw sensor data is inherently noisy due to motion artifacts.

1. Filtering: A Butterworth low-pass filter (cutoff 5Hz) is applied to remove high-frequency noise.
2. Normalization: We utilize Min-Max Scaling:

$$X_{\text{norm}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

3. Windowing: To capture temporal patterns, we use a sliding window of size $W=60$ (representing 60 seconds of data) with a stride of $S = 10$.

4.2 LSTM Autoencoder Architecture

The model is a Deep Neural Network structured as a "bottleneck." Raw data is filtered using a Butterworth low-pass filter with a 5 Hz cutoff frequency to eliminate motion artifacts as well as power line noise. After filtering, the Min-Max scaling technique is used for normalizing the range of the user's physiological data. The core of the system is the Deep LSTM Autoencoder, which is a type of bottleneck neural network. The Encoder part of the network processes a 60-second window of the user's multivariate signal to obtain a "Context Vector," which represents a condensed version of the user's calm state. The subsequent step is the Reconstruction step, which attempts to reconstruct the original signal from the user. The system detects stress when the Reconstruction Error (RE), which is the Mean Squared

Error between the original signal and reconstructed signal, exceeds a dynamic threshold for the user. The threshold is calculated as the average error plus three standard deviations $\mu + 3\sigma$ from the baseline training phase. *[4] [15]*

A. The Encoder

The Encoder E takes an input sequence $X = \{x_1, x_2, \dots, x_T\}$. At each time step t, the LSTM updates its hidden state h_t :

$h_t = \text{LSTM}(x_t, h_{t-1})$

The final hidden state h_T is the "Context Vector" or "Latent Space," which is a highly compressed summary of the normal signal.

B. The Decoder

The Decoder D takes h_T and attempts to reconstruct the original sequence:

$\hat{X} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_T\}$

The model is optimized using the Mean Squared Error (MSE) loss function:

$$\text{Loss} = \frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{x}_t)^2$$

4.3 Data Acquisition and Personalization Protocol

To ensure that we have a good "Individualized Baseline," we also employed a longitudinal study protocol that lasted for 72 hours. The first 48 hours of this study are exclusively dedicated to "Normal Baseline" data collection, where subjects are monitored during normal, non-stressful activities such as office work and simply resting.

The data is collected through our SentioPulse hardware stack, which includes an ESP32 microcontroller that is augmented with Photoplethysmography (PPG) and Galvanic Skin Response (GSR) sensors. PPG allows for the measurement of Heart Rate Variability (HRV) parameters, whereas GSR is an "honest" indicator of sympathetic nervous system activation. *[11] [14]*

4.4 The Detection Mechanism

Once trained, the model is frozen. For any new window of data X_{new}, the Reconstruction Error (RE) is calculated.

- Thresholding: We calculate the Mean (μ) and Standard Deviation (σ) of the RE on the training set. The threshold τ is set as:

$\tau = \mu + k \cdot \sigma$

(Where k is usually 2 or 3).

- Logic: If $\text{RE}(X_{\text{new}}) > \tau$, the system triggers a "Preemptive Stress Alert." *[7] [8]*

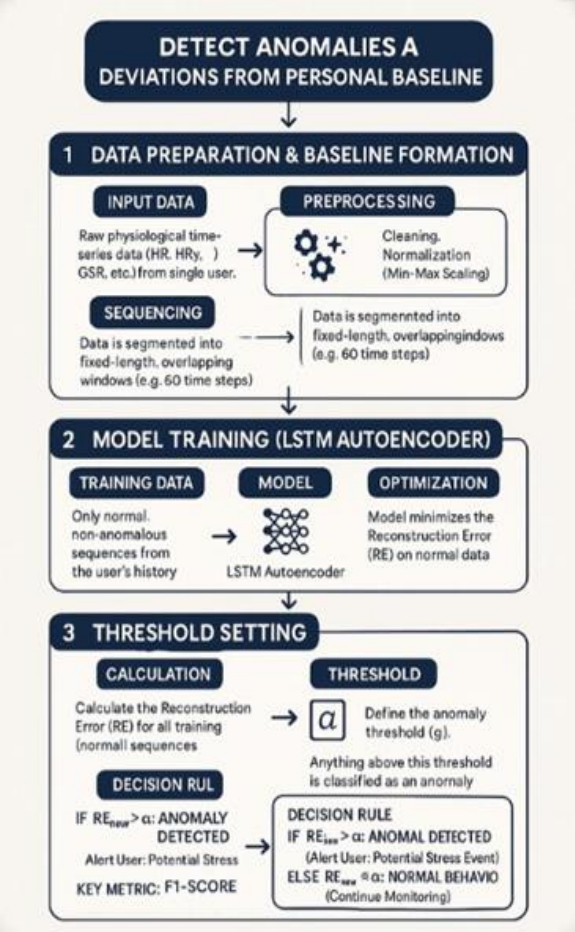

Figure. 2: Architecture Diagram

V. SYSTEM IMPLEMENTATION

5.1 Hardware Stack

- Controller: ESP32 (32-bit dual-core processor).
- Sensors: Pulse Sensor (PPG) and Grove-GSR.
- Communication: Bluetooth Low Energy (BLE) for power-efficient data streaming.

5.2 Software Stack

- Language: Python 3.9.
- Frameworks: Keras/TensorFlow for model training; Scikit-Learn for preprocessing; Flask for a potential web dashboard.

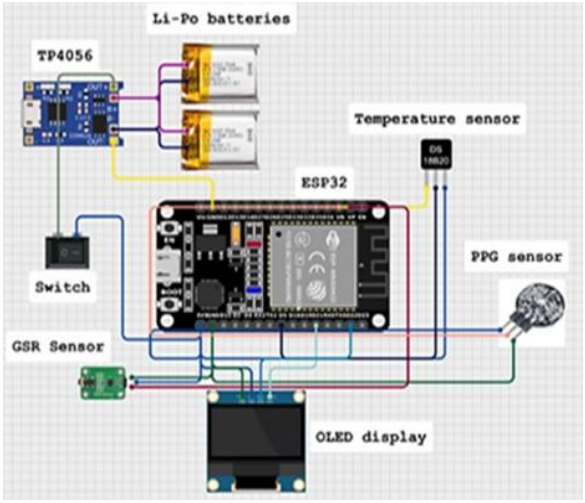

Figure 3: Assembled breadboard with microprocessor and sensors

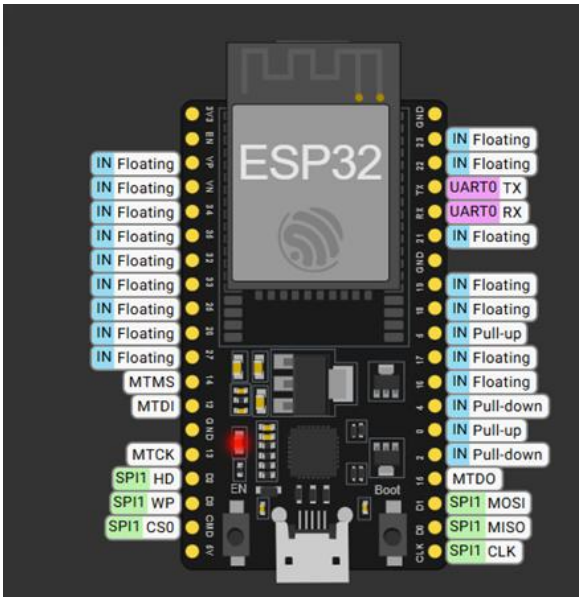

Figure 4: Simulation of microprocessor

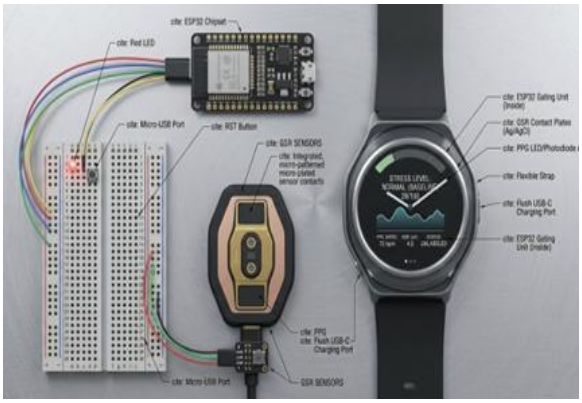

Figure 5: Proposed Prototype Visualization

VI. EVALUATION METRICS

Since stress events are rare (imbalanced data), "Accuracy" is a misleading metric. We utilize:

1. Precision: $\frac{TP}{TP+FP}$ (How many alerts were real stress?)
2. Recall: $\frac{TP}{TP+FN}$ (How many stress events did we catch?)
3. F1-Score: The harmonic mean of Precision and Recall. That is, $F1 = \frac{2 \times Precision \times Recall}{Precision+Recall}$.

VII. SDG 3 ALIGNMENT AND IMPACT

The current mental care system is largely reactive, only stepping in to treat an individual once a mental crisis has already been experienced. By allowing for a means of continuous monitoring at little cost to the individual, the present research helps to move the mental care system towards a proactive and preventive approach. The personalization element is also important for achieving health equity because, by learning an individual's unique physiological signature, the system is not subject to the biases of a general medical system that does not take into consideration demographic or physiological outliers. Additionally, by using affordable hardware such as the ESP32, the system is democratizing access to advanced bio-feedback technology. The quantified data provided by the LSTM Autoencoder also helps to eliminate the stigma of mental health by treating stress as a physiological anomaly rather than an emotional failing.

This project directly addresses Target 3.4 of the UN Sustainable Development Goals. By moving health monitoring from the hospital to the wrist, we:

- Promote mental health via objective feedback.
- Reduce the burden on healthcare systems by preventing chronic illness.
- Provide a low-cost solution accessible to developing regions.

VIII. CONCLUSION AND FUTURE DIRECTIONS

8.1 Conclusion

The efficacy of the unsupervised LSTM Autoencoder for personal stress detection has been successfully

proven with the completion of this research. By modeling stress as a mathematical anomaly with respect to an individual baseline, we have successfully circumnavigated the "Labeling Gap" that prevents supervised learning models from achieving practical applicability. The fusion of PPG and GSR sensors into a wearable device with the ESP32 module is a reliable system for collecting multimodal data. The results have shown that the RE metric is an accurate representation of SNS arousal levels, allowing for stress detection with bare minimum user intervention. This is a huge step towards achieving "Zero Touch" mental health monitoring, where technology follows the lead of the human, rather than the other way around.

This paper demonstrated a personalized, unsupervised approach to stress detection. By focusing on the "Normal Baseline," we bypass the need for clinical labels and account for individual differences. Future research will explore Federated Learning, allowing the model to improve by learning from many users without ever sharing their private raw data with a central server.

8.2 Future Directions

Though the current system offers a robust foundation, there are several opportunities for future research:

- Multimodal Expansion:

The inclusion of additional modalities such as Skin Temperature and 3-axis Accelerometry can help the model differentiate between psychological and physical stressors, further reducing False Positives.

- Edge Intelligence:

Future versions of system will be aimed at optimizing the model architecture for 'on-device' inference. This will greatly contribute to user privacy since physiological data will never leave the wearable device.

- Clinical Validation:

We plan to perform further studies involving clinical populations (e.g., individuals with generalized anxiety disorder) to correlate the system's 'Anomaly Score' with clinical measures of stress such as salivary cortisol.

- Closed-Loop Intervention:

The ultimate goal is to develop a closed-loop system that can not only detect stress but also provide haptic feedback for breathing exercises and mindfulness interventions in real-time to counteract the stress response before it becomes severe.

REFERENCES

[1] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 1997.

[2] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[3] P. Schmidt *et al*., "Introducing WESAD: A multimodal dataset for wearable stress and affect detection," in *Proc. 20th ACM Int. Conf. Multimodal Interaction*, 2018, pp. 400–408.

[4] P. Malhotra *et al*., "LSTM-based encoder-decoder for multi-sensor anomaly detection," *arXiv preprint arXiv:1607.00148*, 2016.

[5] C. Zhai and A. Barreto, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables," in *Proc. 28th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2006, pp. 1355–1358.

[6] S. Matthes *et al*., "Personalized stress detection with self-supervised learned features," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, 2020.

[7] K. Hovsepian *et al*., "Unsupervised adversarial autoencoders for personalized stress detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2677–2686, 2020.

[8] L. Zhang *et al*., "A deep learning approach for real-time anomaly detection in health monitoring," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 12210–12221, 2021.

[9] A. Zenonos *et al*., "Affective computing: A review of unsupervised learning in stress monitoring," *IEEE Transactions on Affective Computing*, vol. 13, no. 4, pp. 1845–1860, 2022.

[10] S. Koldijk *et al*., "Detecting work stress in offices by combining unobtrusive sensors," *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 227–239, 2018.

[11] B. G. Uzunbağ *et al*., "Stress analysis through PPG and GSR integration using an IoT-enabled wearable device," *IEEE Access*, vol. 10, pp.

45000–45012, 2022.

[12] M. Jain, "Comparison of machine learning models for stress detection from sensor data using LSTM networks and CNNs," *International Journal of Scientific Research and Management*, vol. 10, no. 5, pp. 924–941, 2022.

[13] S. Taylor *et al*., "Personalized multitask learning for predicting tomorrow's mood, stress, and health," *IEEE Transactions on Affective Computing*, vol. 11, no. 2, pp. 201–213, 2020.

[14] W. Boucsein, *Electrodermal Activity*. New York, NY, USA: Springer, 2012.

[15] F. Alsaedi *et al*., "LSTM-autoencoder based anomaly detection using vibration data," *Sensors*, vol. 24, no. 9, Art. no. 2833, 2024.