HealthSense: Unobtrusive Continuous Stress Monitoring Using a Novel Dual ECG-PPG Patch

Glenn Fernandes[12], Boyang Wei[12], Christopher Romano[2], Deniz Ulusel[1], Henry K Dambanemuya[1], Yang Gao [2], Roozbeh Ghaffari [3] John Rogers[3], and Nabil Alshurafa[12]
[1]Department of Computer Science,[2]Department of Preventive Medicine, [3]Department of Biomedical Engineering, Northwestern University, Evanston, IL, USA

***Abstract*—Stress, a significant risk factor for chronic disease, manifests as changes in heart rate, respiration rate, and blood pressure. Non-invasive wearables like smartwatches can continuously track these physiological indicators to predict stress, enabling clinicians to develop and test interventions. However, most current devices are rigid and lack skin conformity, resulting in suboptimal signal quality and adherence during extended use. Furthermore, existing flexible sensors employ either electrocardiogram (ECG) or photoplethysmography (PPG), but not both, which is useful for calculating pulse arrival time (PAT) – known to correlate with stress. Addressing these challenges, we introduce HealthSense, a novel, flexible, and skin-conformable device that integrates ECG, PPG, and Inertial Measurement Unit (IMU) sensors into a single wearable. We assessed the comfort of wearing HealthSense and the feasibility of stress prediction by conducting a stress-induction study with 11 participants. Participants rated the comfort level of wearing the device on a Likert scale of 1-5, with 80% rating it as a 5 (most comfortable). Using statistical features, heart rate variability (HRV) related features, and PAT from our sensor data, we trained machine learning (ML) models to predict minute-level perceived and physiological stress with F1-scores of 85.5% and 87.7%, respectively. Additionally, using SHAP values, we identified PAT, systolic time, and pulse as the most significant contributors to the predictions. These findings enhance the understanding of physiological manifestations of stress and lays the groundwork for future stress-reduction interventions.**

I. Introduction

Over 60% of adults globally are afflicted by stress due to health, financial, and societal factors, presenting a pervasive public health challenge warranting clinical intervention [1]. Stress can lead to severe physiological consequences, including elevated blood pressure, atherosclerosis, myocardial infarction, and hypertension. Stress adversely affects the endocrine, immune, and reproductive systems and is linked to premature birth and developmental disorders in children [2]. Prior research suggests that women are more affected by stress than men, experiencing higher psychological distress and greater susceptibility to psychiatric illness [3]. Additionally, stress is associated with depression and unhealthy behaviors such as overeating, smoking, and substance abuse. To prevent the downstream effects of stress, identifying biomarkers of prolonged stress is crucial for developing effective tools and interventions.

Traditional stress monitoring techniques, such as self-reports, cortisol measurements, and blood pressure response, face significant limitations in accuracy, convenience, and practicality for continuous monitoring. Self-reporting involves tracking stress levels through logging or smartphone apps, which can be burdensome, easily missed and is not objective (i.e., capturing physiological manifestations of stress known to impact health). Cortisol measurement from saliva samples provides more objective measures but requires lab visits for accurate assessment, and poorly correlates with stress. Blood pressure and stress are correlated, but blood pressure measurements, while accurate, follow strict protocols and can even induce stress during the collection process (white coat syndrome). Although traditional stress monitoring techniques are accurate for short-term assessments, their limitations make them unsuitable for long-term, continuous monitoring of stress biomarkers, highlighting the need for more objective and practical solutions.

Researchers are developing continuous monitoring methods that leverage the ubiquity of wireless and mobile sensors for low-cost, unobtrusive solutions that do not interfere with daily activities. One example is Glabella, a wearable device in the form of glasses that estimates blood pressure using PPG sensors at multiple head locations [4]. The authors utilized PPG data to measure pulse transit time (the time a pulse wave takes to travel between two arterial sites), which is shown to correlate with blood pressure and, by extension, stress. While Glabella provides a novel sensor in an eye-glass form factor, it is constrained by inconsistent sensor contact and has not been shown to estimate stress. In this paper, we focus on extracting PAT (time interval between ECG R-peak and PPG peak), which, according to recent studies [5], is a better correlate to blood pressure and, hence, a more robust proxy to stress. Recent studies [6] demonstrate the potential of chest-worn devices combining ECG, PPG, and accelerometers for continuous, cuffless blood pressure estimation, crucial for continuous stress monitoring over extended periods. However, these devices are rigid and not skin-conformable. In contrast, King et al. [7] found that the flexible and hidden wearables, such as the Biostamp RC, was favored over wrist-worn devices for comfort. However, since it only uses ECG sensors, calculating PAT remains a challenge, which can improve our ability to predict and estimate objective physiological stressors.

To address the challenges of poor skin conformity, discomfort during extended use due to device rigidity, and the need for simultaneous ECG and PPG measurements for calculating PAT, we detail the following contributions in this paper:

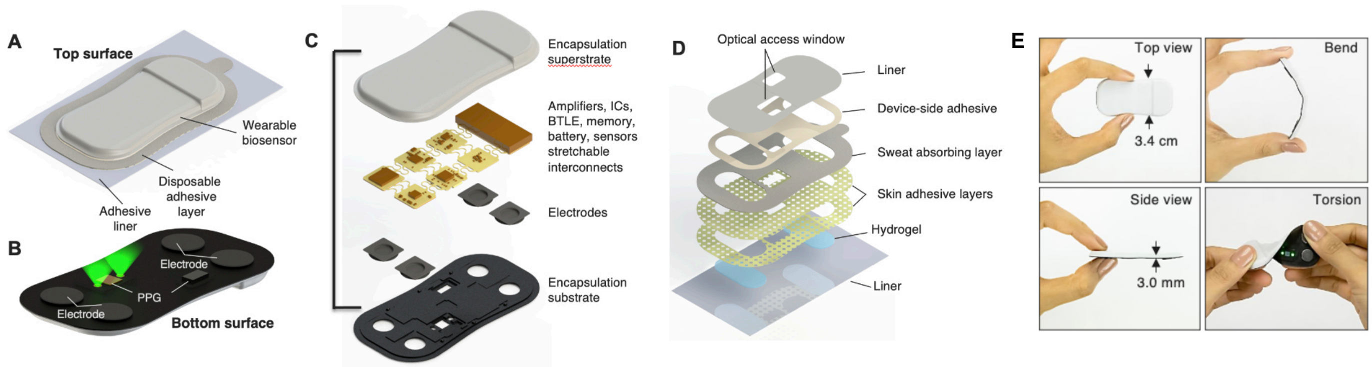

Fig. 1. Schematic and visual representations of the HealthSense wearable sensor and its adhesive layers. (A) Top surface of the HealthSense device, showing the biosensor patch, adhesive layer, and liner. (B) Bottom surface with electrodes and optical sensors. (C) Exploded view of the HealthSense patch, highlighting four polymeric and electronic layers. (D) Exploded view of the disposable adhesive, showing five material layers for a soft, breathable interface. (E) HealthSense patch images: top view (top left), side view (bottom left), under bend (top right), and twist (bottom right) deformations. The HealthSense patch integrates ECG/PPG sensors for single-device PAT and stress measurements.

1) **HealthSense**: A novel, flexible, skin-conformable device integrating ECG, PPG, and IMU sensors for continuous, unobtrusive monitoring with a three-day battery life.
2) **Stress-Induction User Study**: A study with 11 participants demonstrating high comfort levels for continuous wear, with 80% rating it 5 on a Likert scale out of 5.
3) **Stress Detection Models**: Our ML models extracted features from ECG, PPG, and PAT data to predict perceived and physiological stress, achieving best weighted F1-scores of 85.5% and 87.7%, respectively. We also identified and explained critical features that contributed to the predictions.

II. Method

A. Device and Sensing Modality

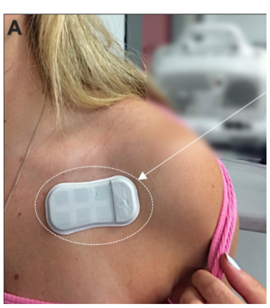

Fig. 2. Person wearing HealthSense device

We designed the HealthSense (Fig. 1) device as a compact, wearable patch offering 24-hour skin-friendly adhesion for easy application and secure wear (Fig. 2). The rechargeable battery supports up to three days of continuous sensing with all sensors active, making it ideal for prolonged stress measurement. The device features a 3-axis accelerometer and a gyroscope to capture six degrees of freedom in inertial sensing. It also includes two pairs of electrodes for ECG sensing and PPG sensors that utilize red or green visible light. The device is capable of collecting raw data at a max sampling rate of 400 Hz.

B. User Study

Building on prior research [7], we conducted our study using the HealthSense device with 11 female participants ages 24 to 64 (M = 38.36, SD = 16.00). Physiological stress was defined as the state during induced stress activities, while perceived stress was measured through participants' responses to micro-EMA questions on a Likert scale after each in-lab activity. Stress was induced by having participants perform a series of activities listed in Table I. On conclusion of each activity, participants completed the micro-EMA for the activity.

TABLE I
STRESS INDUCTION ACTIVITIES

Activity	Description
Rest	10 min: Relaxation period
PASAT	10 min: 4 rounds of 60 trials at 4000ms, 2400ms, 2000ms, 1600ms, 1200ms.
Rest	5 min: Relaxation period
Video Game	5 min: 2-Player Mortal Kombat Game
Conversation	5 min: Unstructured conversation with researcher
Rest	5 min: Relaxation period
Crying Baby	3:30 min: Listening to a 45-sec recording of infant crying with two 1-min rest periods
Rest	10 min: Relaxation period
Cold Pressor	$\approx$ 1 min: Inserting dominant hand in an ice bucket for as long as the participant can.

For the first activity, we utilized a well-established psychological stressor, the Paced Auditory Serial Addition Test (PASAT), using Inquisit's Software Tool. This non-verbal computer-based test presented participants with an audible sequence of single-digit numbers (1-9) at a constant speed, requiring them to mentally sum up the last two digits. Participants then selected the current sum from a circle of numbers (1-18) on a computer screen by clicking the appropriate number. The other stressful activities included a two-player Mortal Kombat video game, a crying infant audio, and the cold pressor test, where participants immersed their hands in ice water. The duration of the cold pressor varied by participant, as each was instructed to leave their hand in ice water for as long as possible, up to one minute. We randomized the order of all stressful activities to counterbalance potential cross-over effects, and no single stressor was dominant. The activities alternated between stressful and non-stressful activities. Non-stressful activities included conversing with the researcher or an activity-free relaxation period. Since participants respond differently to various stressful situations, we administered a micro-EMA after each activity, allowing participants to self-report their stress levels.

C. Wearability Questionnaire

At the end of the lab session, we asked each participant to complete a device wearability survey to assess the comfort and feasibility of the device in a real-world setting. Participants completed a short survey to answer the following questions about the HealthSense device:

1) How comfortable was the device? (On a scale from 1 - 5, 5 being the most comfortable)
2) During the activity, were you aware of the device?
3) Would you consider wearing the device all day?
4) Would you consider wearing the device while sleeping?
5) If paid $100, would you be willing to wear the sensor for 30 days?

III. Data Preprocessing

In this section, we build upon the data preprocessing by King et al. [7] and present our preprocessing steps applied to the raw ECG and PPG signals. We used the Python HeartPy library for signal preprocessing. First, raw ECG data were segmented into 1-minute sliding windows with 50% overlap. For each 1-minute segment, a band-pass filter between 0.67 Hz and 124 Hz isolated the QRS complexes for R-peaks detection, followed by min-max normalization. For each 1-minute segment, we split the signal into 0.6-second intervals. We then used an autoencoder-based noise detection model to filter nosiy ECG segments. Within each 1-minute sliding window, we classified the 0.6-second segmented intervals as clean or noisy. We extracted R-peaks only from sequences that contained a minimum of three consecutive clean 0.6-second intervals. For the PPG signal, we applied the same 1-minute segmentation approach as used for the ECG signal. We applied a Notch filter with a cutoff frequency of 0.05 Hz to remove baseline wander, followed by a bandpass Butterworth filter to process signals between 0.67 Hz and 124 Hz. Based on the noise model results for identifying noisy ECG signals, we removed the corresponding noisy timepoints from the PPG signal as well.

A. Feature Extraction from ECG

After applying the preprocessing steps mentioned above on the raw ECG signal, we extracted minute-level ECG features based on the clean ECG intervals. To identify valid R-R Inter-Beat Intervals (IBIs), we used a two-moving average algorithm to extract peaks from clean ECG intervals. The detected IBIs were passed into a Criterion Beat Difference function [8] to filter invalid R-peaks. 15% of the final IBIs were randomly selected for manual inspection to ensure quality. Both minute-level heart-rate variability (HRV) and statistical features were extracted using valid IBIs.

B. Feature Extraction from PPG

We implemented feature point detection, which involved detecting the wave crest, trough point, inflection point, and extreme point in clean 1-minute PPG segments. We extracted the Cardiac Period (CP), the Systolic Upstroke Time (ST), and the Diastolic Time (DT). Referencing the R peaks detected in the ECG segment, we calculated PAT as the time delay between the ECG R peak and the PPG fiducial point (maximum value of the first derivative of the PPG waveform). For each PAT caluclated, we set the normal range of 300 to 340 and exluded values outside of this range. We calculated the min, max, mean, median, range, and standard deviation for each valid minute-level PAT value.

IV. Stress Detection Models

We trained and evaluated four supervised machine learning (ML) models—random forest, decision tree, adaptive boosting, and gradient boosting machines—on data from 11 participants to predict perceived and physiological stress. The features extracted from ECG and PPG were used as input to predict perceived and physiological stress. We used the Python sci-kit-learn library to train and evaluate the ML models. To evaluate the stress prediction models, we used stratified 5-fold cross-validation. To avoid redundancy and high feature correlations, we stratified the data within each fold by participant (non-overlapped) into 80% for the training set and 20% for the validation set. We then applied correlation-based feature selection (CFS) to the training set's 34 features (2 PPG-based, 31 ECG-based, and 1 combined). We trained and validated the ML models after identifying the best features and optimized the hyperparameters using Bayesian search over three iterations in the training set. For each fold, we reported the precision, recall, and weighted F1 score on the hold-out set using the model trained with the optimized parameters. To further understand the impact of each feature extracted from different sensing modalities on physiological stress, we applied SHAP on the best-performed model, analyzing the 9 most common features selected by CFS from the entire set of 34 features.

V. Results

A. Wearability

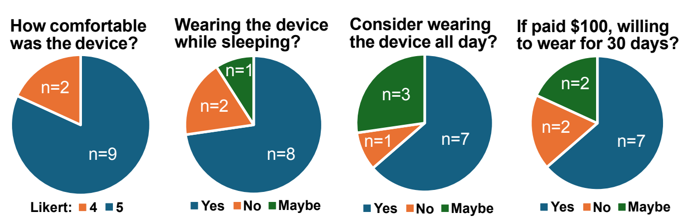

Fig. 3. Wearability survey resuls

All participants rated the device as comfortable ($\geq$ 4 on Likert scale), with 81.82% indicating they would consider wearing the device while sleeping, and 90.91% open to wearing it all day 3. When incentivized with a $100 reward, 81.82% of participants were willing to wear the device for

30 days, suggesting a high level of overall acceptability and feasibility for long-term use.

B. Model Performance

Our dataset comprised a total of 1,005 minutes of activities. The autoencoder noise model identified 389 minutes as noisy, leaving 616 minutes of clean data. Of these, 200 minutes were stress-induced activities, while 416 minutes were from non-stressful rest sessions. The gradient boost machine (GBM) outperformed other models, achieving 85.5% F1 for perceived stress and 87.7% F1 for physiological stress (Table II).

TABLE II
STRESS PREDICTION RESULTS BY 5-FOLD CROSS-VALIDATION

Model	Perceived Stress			Physiological Stress		
	Precision	Recall	F-1	Precision	Recall	F-1
GBM	**0.855**	**0.859**	**0.855**	**0.840**	**0.738**	**0.877**
DT	0.809	0.806	0.805	0.756	0.727	0.836
ADA	0.683	0.747	0.695	0.814	0.649	0.822
RDF	0.804	0.797	0.766	0.800	0.500	0.783

GBM: Gradient Boost Machine, DT: Decision Tree, ADA: Adaptive Boosting, RDF: Random Forest. All F1 scores are weighted.

C. Feature Importance

We further analyzed the feature importance for predictions of physiological stress, which is most germane to the health outcomes of concern. During correlation-based feature subset selection (CFS), the following common features were identified for each fold: pulse, mean, kurtosis, 80th percentile, 40th percentile, RMS, count from minute-level ECG signal, systolic time from the PPG signal, and average pulse arrival time (PAT) from both signals combined.

The SHAP summary plot (Fig. 4), revealed that PAT (average), systolic time, and pulse were the most influential features in predicting physiological stress. Systolic time demonstrated an inverse relationship with physiological stress, with shorter intervals indicating higher stress levels. Conversely, pulse rate was positively associated with physiological stress, with higher rates reflecting increased stress levels.

As observed in the SHAP figure, average PAT emerged as the most important feature. However, the SHAP values for average PAT are spread across both positive and negative ranges. This distribution suggests that PAT has a complex relationship with stress prediction, depending on its specific value and interactions with other features. We believe this complexity arises from the interactions between the selected features. Further research is necessary to fully understand the independent impact of each feature.

VI. CONCLUSION AND FUTURE WORK

Our study demonstrates the potential of HealthSense, a novel flexible, skin-conformable wearable integrating ECG, PPG, and IMU sensors, in predicting stress with high accuracy. The device was rated highly comfortable by participants, and our ML models achieved notable performance in predicting both perceived and physiological stress. Future work will focus on expanding the sample size, improving model accuracy, and integrating real-time feedback mechanisms to provide personalized stress management interventions. Further research will also explore the long-term usability and effectiveness of HealthSense in diverse populations and real-world settings.

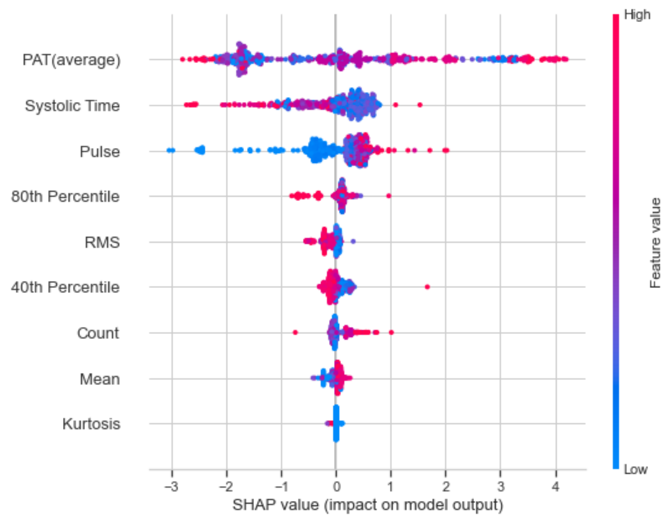

Fig. 4. SHAP summary plot for feature importance in detecting physiological stress.

VII. ACKNOWLEDGEMENT

This material is based upon work supported by the National Institute of Health (NIH) under award numbers K25DK113242, R03DK127128, R21EB030305, and R01DK129843.

REFERENCES

[1] "apa.org," https://www.apa.org/news/press/releases/stress/2023/collective-trauma-recovery, [Accessed 16-05-2024].

[2] S. H. Massey, A. E. Hatcher, C. A. C. Clark, J. L. Burns, D. S. Pine, A. D. Skol, D. K. Mroczek, K. A. Espy, D. Goldman, E. Cook, and L. S. Wakschlag, "Does MAOA increase susceptibility to prenatal stress in young children?" *Neurotoxicology and Teratology*, vol. 61, pp. 82–91, May 2017.

[3] I. Sandanger, J. Nygård, T. Sørensen, and T. Moum, "Is women's mental health more susceptible than men's to the influence of surrounding stress?" *Social Psychiatry and Psychiatric Epidemiology*, vol. 39, pp. 177–184, 2004.

[4] C. Holz and E. J. Wang, "Glabella: Continuously Sensing Blood Pressure Behavior using an Unobtrusive Wearable Device," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 58:1–58:23, Sep. 2017.

[5] M. Di Rienzo, A. Avolio, G. Rizzo, Z. M. I. Zeybek, and L. Cucugliato, "Multi-site pulse transit times, beat-to-beat blood pressure, and isovolumic contraction time at rest and under stressors," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 2, pp. 561–571, 2021.

[6] J. Park, S. Yang, J. Sohn, J. Lee, S. Lee, Y. Ku, and H. C. Kim, "Cuffless and continuous blood pressure monitoring using a single chest-worn device," *IEEE Access*, vol. 7, pp. 135 231–135 246, 2019.

[7] Z. D. King, J. Moskowitz, B. Egilmez, S. Zhang, L. Zhang, M. Bass, J. Rogers, R. Ghaffari, L. Wakschlag, and N. Alshurafa, "micro-stress ema: A passive sensing framework for detecting in-the-wild stress in pregnant mothers," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–22, 2019.

[8] K. Hovsepian, M. Al'Absi, E. Ertin, T. Kamarck, M. Nakajima, and S. Kumar, "cstress: towards a gold standard for continuous stress assessment in the mobile environment," in *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 2015, pp. 493–504.