

AI-Based Stress and Fatigue Recognition Using Multimodal Wearable Physiological Signals

Jesper Diaz

School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
jesperdiaz588@uc.edu

Jean L. Walters

Department of Computer Science, University of Houston, Houston, TX, USA.
jeanwalters82@uh.edu

Sagar C. Verma

School of Computing, Clemson University, Clemson, SC, USA.
sagarwork@clemson.edu

Umesh D. Dutta

Department of Computer Science, University of New Hampshire, Durham, NH, USA.
dutta2000@unh.edu

Abstract

Stress and fatigue recognition from wearable physiological signals has emerged as a significant interdisciplinary challenge involving sensing hardware, machine learning, human factors, data governance, and healthcare policy. This paper presents a system-level examination of AI-based recognition architectures that integrate multimodal physiological data such as electrodermal activity, heart rate variability, skin temperature, and motion context. The discussion emphasizes structural trade-offs in wearable sensing infrastructure, including edge processing, cloud offloading, sensor fusion, and longitudinal deployment. Rather than focusing on a single algorithm or dataset, the paper analyzes how data representation, model architecture, privacy preservation, and fairness interact to determine system performance and societal acceptability. Architectural choices such as centralized training, federated learning, attention-based fusion, and context-aware inference are discussed in terms of robustness, generalization, energy consumption, and regulatory alignment. The analysis further explores the challenges of ambulatory sensing, motion artifacts, individual variability, and demographic bias. Governance and policy implications are examined in relation to worker monitoring, mental health applications, and continuous health surveillance. The paper concludes by outlining future directions for sustainable, fair, and explainable stress and fatigue recognition systems that can operate responsibly beyond controlled laboratory settings.

Keywords

wearable physiological sensing; stress recognition; fatigue detection; artificial intelligence; multimodal fusion; data governance; fairness; deployment sustainability.

1. Introduction

The capacity to recognize stress and fatigue from physiological signals has long been a central concern in affective computing and human factors research. Early work on machine

emotional intelligence demonstrated that patterns in electrodermal activity, heart rate, muscle tension, and respiration could be related to affective arousal and valence, although the signal-to-noise ratio and individual variability presented substantial challenges [1]. Subsequent experiments moved beyond constrained laboratory tasks to real-world driving settings, showing that physiological sensors could detect driver stress during naturalistic operations while also revealing the strong influence of environmental confounds and movement artifacts [2]. These foundational studies established that stress and fatigue are not singular physiological states but rather complex psychophysiological phenomena shaped by autonomic nervous system activity, cognitive workload, and emotional appraisal [3]. The need for automatic stress recognition in office environments further motivated multimodal measurement strategies because single-channel systems often failed to separate stress from physical activity, thermal comfort, or social context [4].

Contemporary wearable devices have transformed this research area by enabling continuous, minimally intrusive monitoring outside the laboratory. However, the transition from laboratory-grade instrumentation to consumer wearables introduces new structural problems related to data quality, sampling frequency, battery constraints, sensor placement, and long-term user adherence. Multimodal fusion provides one avenue for improving robustness, but it also increases system complexity and raises questions about how different signals should be weighted, synchronized, and interpreted across heterogeneous populations. This paper therefore adopts a system-level perspective, emphasizing architecture, infrastructure, governance, fairness, and deployment rather than a narrow algorithmic comparison. The aim is to clarify how technical design decisions are embedded within broader social, organizational, and regulatory environments.

2. Background and Related Work

A substantial body of work has developed around the construction of public datasets for stress and affect recognition using wearable sensors. One influential resource is the WESAD dataset, which provides synchronized electrocardiogram, electrodermal activity, electromyogram, respiration, temperature, and accelerometer data from wrist- and chest-worn devices, thereby allowing comparative evaluation of sensor placements and feature representations [5]. Systematic reviews of emotion recognition with wearables have identified that while many studies report high laboratory accuracy, performance drops considerably in ambulatory contexts due to missing data, motion artifacts, and heterogeneous devices [6]. The broader field of wearable biosensing has advanced rapidly through flexible electronics, low-power microcontrollers, and improved electrochemical sensors, enabling continuous monitoring beyond traditional clinical environments [7]. Surveys of wearable devices emphasize that wireless connectivity, energy supply, storage capacity, and user acceptability remain significant constraints for large-scale deployment [8]. Smart clothing and electronic textiles have extended sensing capabilities by embedding electrodes and conductive fabrics directly into garments, but these systems face durability, washability, and signal stability challenges [9]. Most importantly, comparative studies have shown that physiological stress detection in the wild differs fundamentally from laboratory measurement because uncontrolled physical activity, temperature changes, and social context introduce non-stationary noise [10].

These prior findings indicate that stress and fatigue recognition should not be treated solely as a supervised classification problem. Instead, it requires careful attention to the measurement infrastructure, the temporal dynamics of physiological responses, and the contextual variables that modulate sensor readings. The recognition task also intersects with mental health

screening, occupational safety, and personalized medicine, all of which demand different performance criteria and ethical safeguards. From an engineering standpoint, the challenge is not merely to achieve high accuracy on a single dataset but to design systems that remain reliable, interpretable, and fair across shifting deployment conditions.

3. Wearable Multimodal Physiological Sensing Infrastructure

The sensing layer of a stress and fatigue recognition system consists of multiple physiological transducers, signal conditioning circuits, microcontrollers, wireless radios, and power management units. Personal sensing frameworks have demonstrated that mobile and wearable sensors can provide dense longitudinal data relevant to mental health, yet they also create burdens related to data quality, participant burden, and privacy [11]. The integration of machine learning into medical practice has reinforced that the clinical utility of such systems depends less on raw sensor volume than on the reliability, representativeness, and calibration of the underlying data [12]. High-performance medicine increasingly expects wearable data to be integrated with electronic health records and decision support systems, but this integration can fail if sensor noise, missingness, and device heterogeneity are not addressed at the infrastructure level [13].

A central architectural decision is the placement of computational intelligence across the wearable, a nearby edge device, and the cloud. On-sensor processing reduces data transmission and improves privacy, but it is constrained by memory, processing speed, and energy. Cloud-based processing allows complex deep learning models and longitudinal aggregation, but it introduces latency, connectivity dependence, and centralized privacy risks. Edge processing offers an intermediate solution by enabling local feature extraction, adaptive calibration, and context-aware preprocessing while still allowing asynchronous synchronization with cloud services when connectivity is available. The choice among these deployments has structural implications for battery life, model update frequency, and the ability to detect acute stress episodes in real time.

Another critical dimension is sensor fusion. Multimodal systems can fuse electrodermal activity, heart rate variability, skin temperature, accelerometry, and sometimes electroencephalography or voice features. The assumption behind fusion is that each modality captures complementary aspects of stress and fatigue, such as sympathetic arousal, parasympathetic withdrawal, thermoregulatory change, and behavioral agitation. However, fusion also introduces synchronization problems, different sampling rates, and varying susceptibility to motion artifacts. Wearable systems must therefore implement robust timestamping, interpolation, and missing-data handling before model training. Without such infrastructure, high accuracy in controlled evaluations can become meaningless in operational settings.

4. AI Recognition Architectures and Multimodal Integration

The design of the learning architecture greatly influences how multimodal physiological streams are combined. One structural concern is that models trained on non-representative populations may encode biases that are not immediately visible in aggregate accuracy. Evidence from healthcare algorithms has shown that biased proxy variables and unequal data representation can produce systematic differences in model outputs across demographic groups, even when race is not explicitly included [14]. In stress and fatigue recognition, similar risks arise because baseline physiology, skin conductance responses, and heart rate variability differ across age, sex, medication use, cardiovascular fitness, and mental health

history. Architectural decisions such as feature normalization, personalization layers, and context conditioning can mitigate or amplify these disparities.

To address privacy and data governance challenges, federated learning has been proposed as an alternative to centralized training. Federated approaches allow models to be updated across distributed devices or institutions without sharing raw physiological data, thereby reducing privacy exposure and enabling broader collaboration across hospitals or workplaces [15]. In medicine, federated learning has shown promise for multi-institutional model development while maintaining local control over sensitive patient data [16]. For wearable stress recognition, federated architectures introduce additional challenges, including heterogeneous device capabilities, intermittent connectivity, statistical non-independence across users, and communication overhead. Nevertheless, they represent a structural alternative that aligns more naturally with data minimization principles.

Within local models, attention mechanisms have become increasingly important for selecting temporally relevant segments of physiological signals and for weighting multiple modalities. Recent work has shown that joint attention mechanism learning can facilitate opto-physiological monitoring during physical activity, when motion artifacts substantially degrade signal quality [17]. This line of research is relevant because fatigue and stress often occur during movement-rich occupational tasks, such as driving, manufacturing, or patient care. Attention-based architectures can learn to suppress unreliable sensor windows and amplify informative patterns, but they also increase computational complexity and require careful regularization to avoid overfitting. The trade-off between temporal convolutional networks, recurrent architectures, and attention-based models must be evaluated not only in terms of accuracy but also in terms of interpretability, robustness, and suitability for low-power devices.

5. Data Governance, Privacy, and Security

Stress and fatigue data are highly sensitive because they can reveal mental health states, emotional responses, and workplace behavior. The inclusion of neural signals such as electroencephalography further increases privacy concerns, because brain-derived data may be considered specially protected in some jurisdictions and can reveal information beyond the stated recognition task [18]. Studies comparing laboratory and wearable sensors have shown that although wearable systems improve ecological validity, they also generate continuous records of daily life that may be difficult to anonymize [19]. Even relatively simple heart rate monitoring, when combined with contextual information, can be used to infer stress, sleep, and health status, demonstrating that apparently innocuous physiological channels can become sensitive when aggregated over time [20].

A robust governance framework must therefore address consent, purpose limitation, data minimization, retention, and secondary use. In workplace deployment, employees may face implicit coercion to participate in fatigue monitoring, especially in safety-critical industries. Clear policies are needed to distinguish between aggregate safety analytics and individualized performance surveillance. Technical safeguards such as encryption, secure enclaves, differential privacy, and federated learning can reduce risks, but they cannot substitute for transparent governance and meaningful user autonomy. The system architecture must also support data subject rights, including access, correction, deletion, and portability, which are increasingly required under regulations such as the General Data Protection Regulation and analogous frameworks.

Security is equally important because wearable data streams are vulnerable to interception, device spoofing, and unauthorized inference. A stress recognition system that continuously transmits physiological data to a cloud server creates an attack surface that extends beyond the individual user to the organization. Secure device pairing, signed firmware updates, encrypted transport, and anomaly detection at the network layer are necessary components of a responsible deployment. Furthermore, model inversion and membership inference attacks can potentially reveal whether an individual's data were used in training, raising questions about how model updates should be logged and audited. Governance must therefore be integrated into the system architecture rather than treated as an afterthought.

6. Fairness, Robustness, and Generalization

Fairness in stress and fatigue recognition requires attention to physiological variability across populations. Heart rate variability is influenced not only by stress but also by age, cardiovascular fitness, medication, and workload history [21]. If training data are dominated by young, healthy adults, the resulting models may misclassify older workers or individuals with chronic conditions. Surveys of stress recognition techniques have documented that many models are evaluated on homogeneous samples with limited evidence of cross-population generalization [22]. Ambient intelligence systems in healthcare have likewise shown that sensor-based inference can introduce new forms of bias when trained in constrained environments and then applied to diverse real-world settings [23]. These issues are structural because they arise from the interaction between data collection, feature extraction, and model training rather than from a single algorithmic flaw.

Robustness is another serious concern. Wearable sensors are subject to motion artifacts, electrode displacement, sweat, temperature changes, and signal dropout. A model that performs well on cleaned laboratory data may fail when users engage in normal physical activity. Multimodal fusion can improve robustness because different modalities may fail under different conditions, but fusion can also propagate errors if one noisy channel dominates the learned representation. Temporal attention mechanisms can help by down-weighting unreliable segments, but their effectiveness depends on the availability of contextual labels such as activity type and sensor quality indicators. Context-aware modeling is therefore essential for real-world generalization.

Generalization also depends on the definition of the target state. Stress is not a single diagnostic category; it may be acute or chronic, task-induced or social, negatively or positively valenced. Fatigue may be physical, cognitive, or emotional. A system trained to recognize laboratory-induced cognitive stress may not generalize to occupational fatigue or clinical anxiety. Domain adaptation, personalization, and continuous recalibration can help address these shifts, but they introduce additional complexity. Model updates must be guarded to prevent catastrophic forgetting while still adapting to individual baselines. The system must also provide uncertainty estimates so that high-stakes decisions are not made on the basis of low-confidence predictions.

7. Deployment, Sustainability, and Policy Implications

Deployment of AI-based stress and fatigue recognition in operational environments has remained more difficult than prototype evaluations suggest. Wrist-worn devices using context information have demonstrated the feasibility of stress monitoring outside the laboratory, but they require careful integration with activity recognition to avoid confounding physical movement with stress responses [24]. Hybrid biosensing systems that combine chemical and

electrophysiological measurements illustrate the direction of more informative wearables, yet they also introduce new manufacturing, calibration, and disposal challenges [25]. Sustainable deployment must consider the environmental footprint of sensors, batteries, and frequent device replacement, especially as wearable health monitoring scales to large populations.

The policy implications extend across occupational health, mental health, insurance, and consumer wellness. In occupational settings, fatigue monitoring can improve safety in transportation, construction, and healthcare, but it may also enable continuous surveillance that harms worker dignity. Policies should promote worker participation, transparency about model limitations, and limits on automated disciplinary action. In mental health contexts, stress recognition may support early intervention, but it should not replace clinical judgment or contribute to diagnostic overreach. Regulatory frameworks for software as a medical device are still evolving, and many stress recognition systems occupy an ambiguous space between wellness applications and medical diagnostics. Clear guidance on validation, post-market monitoring, and adverse event reporting is needed.

System-level sustainability also involves algorithmic maintenance. Wearable models degrade over time as sensors drift, user behavior changes, and device firmware updates alter signal characteristics. Deployment pipelines must support continuous monitoring of data distributions and model performance. Feedback loops can enable periodic recalibration, but they must be designed to avoid reinforcing historical biases. Auditability, explainability, and human oversight are therefore essential attributes of responsible deployment. The goal is not to replace human decision-making but to provide meaningful physiological insights that can be interpreted by users, clinicians, and safety professionals.

8. Conclusion

AI-based stress and fatigue recognition using multimodal wearable physiological signals represents a promising but structurally demanding application of artificial intelligence. The technical challenges extend well beyond classifier design to include sensing infrastructure, data governance, fairness, robustness, and long-term sustainability. Multimodal fusion can improve recognition performance, but it introduces synchronization, privacy, and computational trade-offs that must be explicitly managed. Federated learning and attention-based architectures offer meaningful advantages, yet they also require careful evaluation in real-world deployment conditions. The field must move beyond narrow accuracy benchmarks and develop a broader evaluation framework that accounts for demographic diversity, motion artifacts, individual baselines, energy efficiency, and policy compliance. By integrating these considerations at the architecture and governance levels, future systems can provide safe, ethical, and effective support for stress and fatigue management across occupational, clinical, and everyday contexts.

References

1. Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1175–1191.
2. Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2), 156–166.

3. Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, 84(3), 394–421.
4. Alberdi, A., Aztiria, A., & Basarab, A. (2016). Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of Biomedical Informatics*, 59, 49–75.
5. Schmidt, P., Reiss, A., Duerichen, R., Marberger, C., & Van Laerhoven, K. (2018). Introducing WESAD, a multimodal dataset for wearable stress and affect detection. *Proceedings of the 2018 ACM International Conference on Multimodal Interaction*, 400–408.
6. Saganowski, S., Dutkowiak, A., Dziadek, A., Dziezyc, M., Komoszynska, J., Michalska, W., Polak, A., Ujma, M., & Kazienko, P. (2020). Emotion recognition using wearables: A systematic literature review—work-in-progress. *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 1–6.
7. Kim, J., Campbell, A. S., de Ávila, B. E.-F., & Wang, J. (2019). Wearable biosensors for healthcare monitoring. *Nature Biotechnology*, 37(4), 389–406.
8. Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., Hassan, M., & Seneviratne, A. (2017). A survey of wearable devices and challenges. *IEEE Communications Surveys & Tutorials*, 19(4), 2573–2620.
9. Chen, M., Ma, Y., Song, J., Lai, C.-F., & Hu, B. (2016). Smart clothing: Connecting human with clouds and big data for sustainable health monitoring. *Mobile Networks and Applications*, 21(5), 825–845.
10. Smets, E., De Raedt, W., & Van Hoof, C. (2018). Into the wild: The challenges of physiological stress detection in laboratory and ambulatory settings. *IEEE Journal of Biomedical and Health Informatics*, 23(2), 463–473.
11. Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology*, 13, 23–47.
12. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347–1358.
13. Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.
14. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453.
15. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R. M., Trask, A., Xu, D., Baust, M., & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, 3, 119.
16. Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchevskiy, V., Huang, S., Cohen, D., Pellow, C., Chen, L., Koutsoukos, S., Bilello, M., Maloney, E., & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Scientific Reports*, 10, 12598.

17. Zheng, X., Hu, S., Dwyer, V., Barrett, L., & Derakhshani, M. (2026). Joint attention mechanism learning to facilitate opto-physiological monitoring during physical activity. *Biomedical Signal Processing and Control*, 113, 108949.
18. Li, X., Zhang, Y., Tiwari, P., Song, D., Hu, B., Yang, M., Zhao, Z., Kumar, N., & Marttinen, P. (2020). EEG based emotion recognition: A tutorial and review. *ACM Computing Surveys*, 53(1), 1–57.
19. Ragot, M., Martin, N., Em, S., Pallamin, N., & Diverrez, J.-M. (2017). Emotion recognition using physiological signals: Laboratory vs. wearable sensors. In *Advances in Human Factors in Wearable Technologies and Game Design* (pp. 15–22). Springer.
20. Choi, J., & Gutierrez-Osuna, R. (2009). Using heart rate monitors to detect mental stress. In *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks* (pp. 219–223). IEEE.
21. Kaur, B., Durek, J. H., O'Kane, B. L., Tran, N., Moses, S., Luthra, M., & Ikonomidou, V. N. (2014). Heart rate variability and cognitive workload in operational environments: A review. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58(1), 218–222.
22. Sharma, N., & Gedeon, T. (2012). Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer Methods and Programs in Biomedicine*, 108(3), 1287–1301.
23. Haque, A., Milstein, A., & Fei-Fei, L. (2020). Illuminating the dark spaces of healthcare with ambient intelligence. *Nature*, 585(7824), 193–202.
24. Gjoreski, M., Luštrek, M., Gams, M., & Gjoreski, H. (2017). Monitoring stress with a wrist device using context. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 159–162.
25. Imani, S., Bandodkar, A. J., Mohan, A. M. V., Kumar, R., Yu, S., Wang, J., & Mercier, P. P. (2016). A wearable chemical–electrophysiological hybrid biosensing system for real-time health and fitness monitoring. *Nature Communications*, 7, 11650.