International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Human Mental Stress Detection Using Machine Learning: A Comprehensive Review

Jitendra Kumar[1], Dr. Sweety Maniar[2]

[1]Research Scholar, IT & Computer Science, Swaminarayan University, Kalol—382725, Gujarat, India
[2]Phd Supervisor

Abstract
Mental stress, a pervasive global health concern, significantly impacts individual well-being, economic productivity, and contributes to a spectrum of physical and psychological disorders. The concurrent proliferation of wearable sensors, ubiquitous smart devices, and online social platforms has generated unprecedented volumes of multimodal data, creating a fertile ground for the objective, continuous, and automated detection of human stress. This paper provides a comprehensive and in-depth review of the application of machine learning (ML) techniques for mental stress detection. We survey the diverse landscape of data modalities, from direct physiological signals such as Electroencephalography (EEG) and Heart Rate Variability (HRV), to indirect behavioral cues derived from video and speech analysis, and rich textual data mined from social media. A detailed analysis of the machine learning paradigms employed is presented, covering classical models like Support Vector Machines (SVM) and Ensemble Learning, as well as advanced Deep Learning architectures including Convolutional and Recurrent Neural Networks (CNNs, RNNs) and state-of-the-art transformer-based models for Natural Language Processing (NLP). Finally, we discuss promising future directions poised to overcome these hurdles, including the development of sophisticated multimodal fusion techniques, the creation of closed-loop systems for real-time stress mitigation, and the integration of more advanced, context-aware artificial intelligence.

Keywords: Mental Stress, Stress Detection, Machine Learning, Deep Learning, Physiological Signals, Wearable Sensors, Affective Computing, Electroencephalography (EEG), Heart Rate Variability (HRV), Natural Language Processing (NLP), Multimodal Fusion, Interpretability, Mental Health Technology.

1. Introduction

Mental stress, defined as the physiological and psychological response to perceived threats or demands—known as stressors—is an inescapable component of modern life. While acute, short-term stress can be adaptive, enhancing focus and performance, chronic or excessive stress is maladaptive. Traditional methods for assessing stress predominantly rely on self-report instruments like the Perceived Stress Scale (PSS) and clinical interviews. Although valuable, these methods are inherently subjective, prone to recall bias, sporadic, and fail to capture the dynamic, moment-to-moment fluctuations of stress experienced in naturalistic, real-world settings. This significant methodological gap has spurred a concerted search for objective, continuous, and non-intrusive methods for stress detection.

The convergence of ubiquitous computing—particularly through wearable technology and smartphones—and artificial intelligence has catalyzed a paradigm shift in mental health monitoring.

Machine learning (ML), a field of AI that endows systems with the ability to learn complex patterns from data without explicit programming, is at the epicentre of this revolution [1][4] By analyzing vast streams of data from diverse sources, ML models can identify subtle yet consistent physiological, behavioral, and linguistic patterns that are indicative of stress. This capability offers a powerful toolkit for early detection, continuous monitoring, and personalized intervention, moving mental healthcare from a reactive to a proactive model.

This review provides a structured and comprehensive overview of this rapidly evolving domain. We outline the detailed methodological pipeline, from data acquisition to model deployment, and elucidate the underlying conceptual framework that connects the neurophysiological stress response to a computationally tractable machine-learning problem. An expanded synthesis of the literature is presented, followed by a rigorous comparative analysis of different approaches and a deep discussion of their implications.

2. Methodology

The methodologies reviewed in the literature adhere to this conceptual framework, demonstrating significant diversity in both data acquisition techniques and the computational models applied.

2.1. Data Acquisition and Modalities: The choice of data modality is fundamental to the system's design, balancing factors like intrusiveness, accuracy, and scalability. A diagram could effectively illustrate these diverse data streams flowing into a central processing unit.

- **Physiological:** This category represents the most direct measurement of the stress response. Data is captured using various sensors. EEG data is collected with headsets ranging from multi-channel, clinical-grade caps to more portable, consumer-grade devices. ECG is often captured with chest straps for high-fidelity R-peak detection, while photoplethysmography (PPG) sensors in smartwatches and fitness trackers provide a more convenient, albeit sometimes less accurate, alternative for estimating HRV. GSR/EDA sensors are typically integrated into wrist-worn devices or research-grade equipment.
- **Behavioral:** These modalities capture the external manifestations of an individual's internal state. Video-based systems use cameras to record facial expressions, with algorithms then identifying the activation of specific Facial Action Units (AUs) associated with negative affect, such as brow lowering (AU4) or lip corner depression (AU15) [8]. Audio-based systems analyze vocal prosody, extracting acoustic features like pitch (fundamental frequency), jitter (pitch perturbation), shimmer (amplitude perturbation), and Mel-frequency cepstral coefficients (MFCCs) [9]. Actigraphy, typically from accelerometers in wearables, is used to monitor sleep quality, restlessness, and general activity levels, all of which are modulated by stress [3].
- **Textual:** Leveraging the digital footprint of individuals, this modality involves scraping textual data from social media platforms like Reddit or Twitter. This process requires careful handling of data anonymization to protect user privacy. NLP techniques are then applied to these large corpora to find linguistic markers correlated with stress [2] [7].

2.2. Machine Learning and Deep Learning Models The analytical core of these systems consists of a wide range of algorithms:

- **Classical ML Models:** These models are often favored when feature sets are well-understood and interpretability is a priority. Support Vector Machines (SVM) remain highly popular for their effectiveness in finding optimal separating hyperplanes in high-dimensional feature spaces [5].

Ensemble methods like Random Forests and Gradient Boosting are also widely used due to their robustness and ability to handle complex feature interactions [6].

- **Deep Learning (DL) Models:** DL models have become dominant for their ability to perform automatic feature learning from raw, high-dimensional data. CNNs, for example, are adept at learning spatial patterns and can be applied to 2D representations of time-series data, like spectrograms from audio or EEG signals arranged in a topographical map. RNNs and LSTMs are inherently suited for sequential data, as their internal memory allows them to capture temporal dependencies in physiological time series or the sequential structure of language [10] [9].
- **NLP Models:** In the textual domain, the field has rapidly advanced from using bag-of-words models to sophisticated transformer architectures like BERT. Transformers utilize a self-attention mechanism, allowing the model to weigh the importance of different words in a sentence dynamically, leading to a much deeper contextual understanding. This enables the detection of nuanced expressions of stress that were previously intractable [11] [12].

3. Comparison Tables

Stress Detection in Humans through EEG

Research Title	Dataset Employed	Method Used	Strength of the Technique (Inferred)	Potential for Enhancement (Inferred)
R. Katmah, F. Al-Shargie, U. Tariq, F. Babiloni, F. Al-Mughairbi, and H. Al-Nashash, “A review on mental stress assessment methods using eeg signals,” Sensors, vol. 21, no. 15. MDPI AG, Aug. 01, 2021. doi: 10.3390/s21155043.	Not specified (review paper)	Review of EEG-based methods	Comprehensive overview of existing methods.	Can guide future research by identifying gaps.
R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel, and A. S. Malik, “Machine learning framework for the detection of mental stress at multiple levels,” IEEE Access, vol. 5, pp. 13545–13556, Jul. 2017, doi: 10.1109/ACCESS.2017.2723622.	Not specified	Machine learning framework	Explores multi-level detection of stress.	Further validation on diverse datasets.
G. Jun and S. K. G., “EEG based stress level identification,” in 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016, pp. 3270–3274. doi: 10.1109/SMC.2016.7844738.	Not specified	EEG-based stress level identification	Direct approach to stress level identification.	Integration with other biosignals for robustness.
G. Giorgos, G. Dimitris, and T. Manolis, “Detection of stress/anxiety state from EEG features during video watching,” Annu Int Conf IEEE Eng	Not specified	EEG feature analysis during video watching	Contextual stress detection.	Real-time application and broader scenarios.

Med Biol Soc., vol. 7, no. 1, pp. 6037–6041, 2015, doi: 10.1109/EMBC.2015.7319767.				

Stress Detection Using Speech Signal

Research Title	Dataset Employed	Method Used	Strength of the Technique (Inferred)	Potential for Enhancement (Inferred)
Mustaqeem and Soonil Kwon "A CNN-Assisted Enhanced Audio Signal Processing for Speech Emotion Recognition" MDPI Journal 28 December 2019.	Not specified	CNN-Assisted Audio Signal Processing	Combines CNN with audio processing for emotion.	Further enhancements in audio feature extraction.
Issa, D.; Fatih Demirci, M.; Yazici, A. Speech emotion recognition with deep convolutional neural networks. Biomed. Signal Process. Control 2020, 59, 101894.	Not specified	Deep convolutional neural networks	Deep learning for speech emotion recognition.	Optimization of network architectures.
Zamil, A.A.A.; Hasan, S.; Baki, S.M.J.; Adam, J.M.; Zaman, I. Emotion Detection from Speech Signals using Voting Mechanism on Classified Frames. In Proceedings of the 2019 International Conference on Robotics,	Not specified	Voting Mechanism on Classified Frames	Ensemble method for robust emotion detection.	Testing with a wider range of datasets.

Recognition of Stress via Audio Visual information

Research Title	Dataset Employed	Method Used	Strength of the Technique (Inferred)	Potential for Enhancement (Inferred)
Luna-Jiménez, C.; Kleinlein, R.; Griol, D.; Callejas, Z.; Montero, J.M.; Fernández-Martínez, F. A Proposal for Multimodal Emotion Recognition Using Aural Transformers and Action Units on RAVDESS Dataset. Appl. Sci. 2022, 12, 327. https://doi.org/10.3390/app12010327	RAVDESS Dataset	Aural Transformers and Action Units	Utilizes a specific public dataset and advanced techniques.	Exploring fusion strategies and real-time performance.
Mansouri-Benssassi, E.; Ye, J. Speech Emotion Recognition With Early Visual	Not specified	Spiking Neural	Novel neural	Efficiency and biological

Cross-modal Enhancement Using Spiking Neural Networks. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.		Networks with Cross-modal Enhancement	network approach with cross-modal fusion.	plausibility.
Luna-Jiménez, C.; Griol, D.; Callejas, Z.; Kleinlein, R.; Montero, J.M.; Fernández-Martínez, F. Multimodal Emotion Recognition on RAVDESS Dataset Using Transfer Learning. Sensors 2021, 21, 7665. https://doi.org/10.3390/s21227665	RAVDESS Dataset	Transfer Learning	Efficiently leverages pre-trained models.	Fine-tuning for specific stress recognition tasks.

Towards Deep Learning-Based Facial Recognition

Research Title	**Dataset Employed**	**Method Used**	**Strength of the Technique (Inferred)**	**Potential for Enhancement (Inferred)**
Naga, P.; Marri, S.D.; Borreo, R. Facial emotion recognition methods, datasets and technologies: A literature survey. Mater. Today Proc. 2021.	Not specified (literature survey)	Literature survey on methods, datasets, technologies	Comprehensive overview of the field.	Identifies trends and future research directions.
Ashraf, A.; Gunawan, T.; Rahman, F.; Kartiwi, M. A Summarization of Image and Video Databases for Emotion Recognition. In Recent Trends in Mechatronics Towards Industry 4.0. Lecture Notes in Electrical Engineering; Springer: Singapore, 2022; Volume 730, pp. 669–680.	Various Image and Video Databases	Summarization and analysis of databases	Provides insights into available resources.	Helps in selecting appropriate datasets for research.

4. Literature Review

Stress Detection in Humans through EEG

Electroencephalography (EEG) signals, representing the brain's electrical activity, offer a direct physiological window into stress states, making them a primary modality for objective stress assessment. Comprehensive reviews, such as that by Instrumental in charting the landscape of EEG-based stress detection methods. These reviews not only catalog existing techniques but critically identify gaps, suggesting that while the field is robust, there remains a significant need for more robust, generalized, and clinically validated solutions. [13]

Demonstrated a machine learning framework for multi-level mental stress detection, highlighting a move towards more granular, severity-based assessment rather than simple binary stress/non-stress classification.[14]
Explored direct EEG-based stress level identification. While straightforward, such approaches often benefit from multimodal integration with other biosignals (e.g., heart rate, skin conductance) to enhance reliability and provide a more holistic stress profile, mitigating the inherent variability of EEG.[15]
Contextual understanding is crucial for practical applications.
Pioneered stress/anxiety state detection from EEG features during video watching, illustrating the potential for applications in interactive and dynamic environments. The strength here lies in linking brain activity to specific stimuli.[16]

Stress Detection Using Speech Signal
Speech signals offer a highly accessible and non-invasive modality for stress and emotion detection, with significant potential for real-world applications. Early foundational work, exemplified by
Introduced a CNN-assisted approach to enhance audio signal processing for speech emotion recognition. Convolutional Neural Networks (CNNs) excel at learning hierarchical features, which can significantly improve the extraction of relevant acoustic patterns.[17]
Leveraged deep convolutional neural networks directly for speech emotion recognition, indicating a strong trend towards end-to-end deep learning solutions that can learn features directly from raw audio.[18]
Employed a voting mechanism based on classified frames for emotion detection. This ensemble method enhances robustness by combining multiple decisions, a valuable strategy for improving overall accuracy, but requires extensive validation across diverse datasets to prove its generalizability.[19]

Recognition of Stress via Audio Visual Information
Combining audio and visual cues for stress and emotion recognition offers a richer, more robust dataset than single modalities, leading to potentially higher accuracy and reliability. This multimodal approach is particularly relevant for applications requiring a comprehensive understanding of human affective states.
Made a significant contribution by proposing multimodal emotion recognition using Aural Transformers and Action Units on the RAVDESS Dataset. The strength of this work lies in its utilization of a well-known public dataset, facilitating reproducibility and comparison, combined with advanced transformer architectures for audio and Action Units for facial expressions. Future directions include exploring more sophisticated fusion strategies (e.g., early, late, or hybrid fusion) and optimizing models for real-time performance in unconstrained settings. [20]
Introduced Spiking Neural Networks with early visual cross-modal enhancement for speech emotion recognition. This indicates a promising move towards biologically plausible and potentially more energy-efficient models. Future work could focus on scalability and hardware implementation for real-time applications. Finally,[21]
Further explored multimodal emotion recognition on the RAVDESS Dataset using Transfer Learning. The power of transfer learning lies in efficiently leveraging pre-trained models, reducing the need for massive domain-specific datasets. Fine-tuning these models for specific stress recognition tasks, rather

than general emotion recognition, and assessing their performance in diverse real-world scenarios are important next steps.[22]

Towards Deep Learning-Based Facial Recognition

Facial recognition has become a cornerstone of emotion and stress detection, especially with the transformative power of deep learning. This field benefits immensely from comprehensive reviews that map its rapid evolution.

Provided a valuable literature survey on facial emotion recognition methods, datasets, and technologies. This survey's strength lies in its ability to consolidate diverse research, helping to identify current trends (e.g., the dominance of deep learning, the emergence of 3D facial models) and highlighting critical areas for future work, such as improving robustness to occlusions, varying lighting conditions, and diverse ethnic facial features. Similarly,[23]

Contributed by summarizing existing image and video databases specifically curated for emotion recognition. Future research will likely focus on synthesizing more diverse and realistic facial datasets, potentially leveraging generative adversarial networks (GANs), and developing standardized benchmarks for cross-dataset evaluation. Furthermore, the integration of explainable AI (XAI) techniques will be crucial for understanding how deep learning models interpret facial cues and ensuring fair and transparent decision-making.[24]

5. Future Directions and Research Gaps

Addressing the current challenges effectively illuminates the path forward, pointing to several promising and critical future directions for research.

- **Real-time Stress Detection and Mitigation:** The ultimate vision extends beyond mere detection to intervention. The future lies in creating closed-loop biofeedback systems that not only identify stress in real-time but also deliver personalized, just-in-time interventions. The work by [5] on using binaural beat stimulation is an early prototype. Future systems could integrate a wide array of interventions, such as triggering guided breathing exercises on a smartwatch, suggesting a short walk, or adapting the user's music playlist to be more calming.
- **Advanced Wearable Sensors:** The field's progress is inextricably linked to sensor innovation. The next generation of wearables will likely be more comfortable, more accurate, and more energy-efficient. We can anticipate the integration of novel sensors capable of non-invasively measuring biomarkers previously confined to the lab, such as cortisol levels in sweat, providing a more direct window into the HPA axis activity.
- **Advanced Learning Paradigms:** To break the dependency on large, labeled datasets, the field must embrace more advanced learning paradigms. Federated learning holds promise for training models on decentralized data (e.g., on users' phones) without compromising raw data privacy. Unsupervised and semi-supervised learning techniques are crucial for leveraging the vast amounts of unlabeled data being generated. Furthermore, reinforcement learning could be used to train agents that learn optimal, personalized intervention strategies over time based on user feedback and physiological responses.
- **Personalized and Context-Aware Models:** One-size-fits-all models are doomed to fail. The future requires deeply personalized models that learn an individual's unique physiological and behavioral stress signature. Crucially, these models must be context-aware. A system must be able to

differentiate a high heart rate caused by exercise from one caused by an argument. This requires integrating multimodal sensor data with contextual information from smartphone calendars, GPS, ambient sound, and app usage to build a rich, holistic understanding of the user's life and the triggers of their stress.

6. Conclusion

The application of machine learning to the detection of human mental stress has rapidly evolved from a niche academic pursuit into a vibrant and impactful field of research. This review has charted its significant progress, detailing the impressive breadth of data modalities being harnessed—from the electrical whispers of the brain to the overt language of social media—and the growing sophistication of the machine learning techniques being employed. From interpretable classical algorithms to powerful deep learning architectures, these models have demonstrated a remarkable and promising ability to identify the complex, multi-faceted signature of stress across a variety of meaningful contexts.
However, the journey from research prototype to reliable, real-world tool is far from complete. The path is obstructed by formidable challenges, chief among them being the development of models that can generalize across diverse contexts and individuals. The critical needs for algorithmic transparency and interpretability, alongside the pressing ethical imperatives of data privacy and security, must be placed at the forefront of the research agenda.

7. References

1. 1 Akhtar, F. et al. (2020). "Role of Machine Learning in Human Stress: A Review." In Proc. Int. Comput. Conf. Wavelet Active Media Technol. Inf. Process. (ICCWAMTIP).
2. 2 Alnajjar, F. et al. (2023). "Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing." J. Soc. & Hum. Anal., 3, pp. 80–91.
3. 6 Bharath, B. et al. (2024). "Comparative Analysis of Machine Learning Algorithms for Detection of the Stress of Humans During Sleep." Int. J. Intell. Syst. Appl. Eng., 12(4), pp. 1491–1500.
4. 9 Gedam, F., & Paul, S. (2021). "A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques." IEEE Access, 9, pp. 84045–84066.
5. 13 Halim, Z., & Rehan, M. (2019). "SVM-Based Detection of Driving-Induced Stress Patterns in EEG." Appl. Soft Comput., 85, art. 105819.
6. 19 Jayawickrama, J. G., & Rupasingha, R. A. H. M. (2022). "Ensemble Learning Approach to Human Stress Detection Based on Behaviours During Sleep." In Proc. Int. Conf. Adv. Comput.
7. 28 Lin, M. et al. (2022). "Deep Learning Model for Psychological Stress Detection in Social Media." IEEE Trans. Neural Netw. Learn. Syst., 33(9), pp. 4324–4333.
8. 29 Llanes, R. G., & Reyes, R. (2022). "Stress Detection in Video Feed: Utilizing Facial Action Units as Indicators." In Proc. IEEE Eurasia Conf. Biomed. Eng. Healthc. Sustain. (ECBIOS).
9. 32 Mirsamadi, S. et al. (2017). "Automatic Speech Emotion Recognition Using Recurrent Neural Networks." In Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), pp. 2227–2231.
10. 37 Paul, A. et al. (2023). "Hybrid Deep Learning Approach for Stress Detection Using Decomposed EEG Signals." Diagnostics, 13(11), p. 1936.
11. 46 Turcan, E., & McKeown, K. (2021). "BERT-Based Explainable Framework for Stress Detection." In Proc. Conf. Empir. Methods Nat. Lang. Process. (EMNLP), pp. 616–627.

12. 49 Yang, H. et al. (2023). "KC-Net: Knowledge-Enhanced Contextual Network for Stress Detection." IEEE Trans. Affect. Comput., 14(1), pp. 582–593.
13. R. Katmah, F. Al-Shargie, U. Tariq, F. Babiloni, F. Al-Mughairbi, and H. Al-Nashash, “A review on mental stress assessment methods using eeg signals,” Sensors, vol. 21, no. 15. MDPI AG, Aug. 01, 2021. doi: 10.3390/s21155043.
14. R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel, and A. S. Malik, “Machine learning framework for the detection of mental stress at multiple levels,” IEEE Access, vol. 5, pp. 13545–13556, Jul. 2017, doi: 10.1109/ACCESS.2017.2723622.
15. G. Jun and S. K. G., “EEG based stress level identification,” in 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016, pp. 3270–3274. doi: 10.1109/SMC.2016.7844738.
16. G. Giorgos, G. Dimitris, and T. Manolis, “Detection of stress/anxiety state from EEG features during video watching,” Annu Int Conf IEEE Eng Med Biol Soc., vol. 7, no. 1, pp. 6037–6041, 2015, doi: 10.1109/EMBC.2015.7319767.
17. Mustaqeem and Soonil Kwon “A CNN-Assisted Enhanced Audio Signal Processing for Speech Emotion Recognition” MDPI Journal 28 December 2019.
18. Issa, D.; Fatih Demirci, M.; Yazici, A. Speech emotion recognition with deep convolutional neural networks. Biomed. Signal Process. Control 2020, 59, 101894.
19. Zamil, A.A.A.; Hasan, S.; Baki, S.M.J.; Adam, J.M.; Zaman, I. Emotion Detection from Speech Signals using Voting Mechanism on Classified Frames. In Proceedings of the 2019 International Conference on Robotics.
20. Luna-Jiménez, C.; Kleinlein, R.; Griol, D.; Callejas, Z.; Montero, J.M.; Fernández-Martínez, F. A Proposal for Multimodal Emotion Recognition Using Aural Transformers and Action Units on RAVDESS Dataset. Appl. Sci. 2022, 12, 327. https://doi.org/10.3390/app 12010327.
21. Mansouri-Benssassi, E.; Ye, J. Speech Emotion Recognition With Early Visual Cross-modal Enhancement Using Spiking Neural Networks. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.
22. Luna-Jiménez, C.; Griol, D.; Callejas, Z.; Kleinlein, R.; Montero, J.M.; Fernández-Martínez, F. Multimodal Emotion Recognition on RAVDESS Dataset Using Transfer Learning. Sensors 2021, 21, 7665. https://doi.org/10.3390/s21227665.
23. Naga, P.; Marri, S.D.; Borreo, R. Facial emotion recognition methods, datasets and technologies: A literature survey. Mater. Today Proc. 2021.
24. Ashraf, A.; Gunawan, T.; Rahman, F.; Kartiwi, M. A Summarization of Image and Video Databases for Emotion Recognition. In Recent Trends in Mechatronics Towards Industry 4.0. Lecture Notes in Electrical Engineering; Springer: Singapore, 2022; Volume 730, pp. 669–680.