Recent Advances in Stress Detection with Machine and Deep Learning

Literature Survey (2024–2025)

J.Krithika

Introduction

- **Objective:** Review of 10 recent peer-reviewed studies on stress detection using machine and deep learning
- **Focus Areas:**Title/Authors
- Year/Source
- Main Objective
- Data Type
- Models/Methods
- Methodology Summary
- Key Results
- Limitations/Future Work

Study Overview

Study & Year	Data Type	ML/DL Model(s)	Key Performance
Kumar & Ankayarkanni (2025)	PPG (Pulse-rate variability)	Hybrid CNN + XGBoost	98.87% train acc, 93.28% val acc (F1 97.25%)
Al-Alim *et al.* (2024)	ECG, Skin Temp, Skin Cond.	KNN, SVC, DT, RF, XGBoost	RF: 98.29% acc (binary, no SMOTE); XGBoost: 98.98% (3-class, with SMOTE)
Afify *et al.* (2025)	EEG (32-channels)	VGGish-CNN	99.25% accuracy (5-fold CV)
Fontes *et al.* (2024)	rPPG (facial video)	LSTM, GRU, 1D-CNN (hybrid)	95.83% accuracy
Xiang *et al.* (2025)	Accel, EDA, HR, Temp	CNN (time/freq fusion)	91.0% accuracy (F1 0.91)
Nechyporenko *et al.* (2024)	GSR (EDA), PPG	KNN (best)	83.3% accuracy
Moser *et al.* (2024)	Wearable (Empatica E4: EDA, HR)	LSTM + GAN ensemble (LSTM-DGE)	+3.0% recall, +7.18% precision (vs baseline)

Table of Studies:

Study & Year	Data Type	ML/DL Model(s)	Key Performance
Motaman *et al.* (2025)	PPG (WESAD wristband)	Dilated CNN	93.56% accuracy, AUC 96.52%
Kim *et al.* (2024)	EEG + GSR (VR scenario)	CNN (ResNet), Vision Transformer	AUROC 0.954 (multi-modal ResNet-50) Abdelfattah *et al.* (2025)
Abdelfattah *et al.* (2025)	Multimodal (ACC, ECG, BVP, TEMP, RESP, EMG, EDA)	RNN (best cross-subj.), RF (best within-subj.)	RNN: F1 93% (cross-subject); RF: F1 ~99% (same-subj)

Leveraging Conv-XGBoost Algorithm for Perceived Mental Stress Detection Using Photoplethysmography

- **Authors:** Geethu S. Kumar, B. Ankayarkanni.
- **Year/Source:** 2025, *Intelligence-Based Medicine*p
- **Objective:** Improve stress detection accuracy from PPG signals using a hybrid model.
- **Data:** PPG (pulse-rate variability) signals from a wearable dataset.
- **Models:** A novel hybrid "Conv-XGBoost" combining Convolutional Neural Network (CNN) feature extraction with eXtreme Gradient Boosting (XGBoost) classifier
- **Methodology:** PPG time-series were converted to spectrogram images; a CNN processed the spectrograms to extract features, and XGBoost took those features as input. The model was trained/validated on a benchmark PPG stress dataset.
- **Results:** The Conv-XGBoost model achieved very high performance: training accuracy 98.87%, validation accuracy 93.28%, and F1-score 97.25%, substantially outperforming conventional classifiers. It showed robustness to noise/variability in PPG. **Limitations/Future:** The study suggests extending the approach to integrate multiple physiological modalities with advanced deep architectures for even better generalizationpapers.ssrn.com.

A Machine-Learning Approach for Stress Detection Using Wearable Sensors in Free-Living Environments

Authors: Mohamed Abd Al-Alim *et al.*

Year/Source: 2024, *Computers in Biology and Medicine*

Objective: Detect stress in real-life (free-living) conditions using wearable data.

Data: SWEET dataset from 240 subjects, including **ECG**, skin temperature (ST), and skin conductance (SC) recordings

Models: Classical ML classifiers: K-Nearest Neighbors (KNN), Support Vector Classification (SVC), Decision Tree (DT), Random Forest (RF), and XGBoost

Methodology: The authors trained and tested these models under various setups: binary stress/non-stress classification and three-level stress classification, each with and without SMOTE resampling to handle class imbalancepThey evaluated accuracy and F1 for all scenarios.

Results: Random Forest performed best for binary classification without SMOTE (98.29% accuracy, F1 97.89%) With SMOTE, KNN was best for binary (95.70% acc). For three-class stress, RF achieved 97.98% (F1 97.22%) without SMOTE, while XGBoost gave 98.98% accuracy with SMOTEThese high accuracies demonstrate that simple ML on wearable ECG/EDA data can be very effective for stress detection.

Limitations/Future: The study was limited to pre-collected lab data (SWEET) and focused on classical ML; future work could explore deep models, real-time deployment, and other physiological signals.

Stress Detection Based on EEG Under Varying Cognitive Tasks Using Convolution Neural Network

Authors: Heba M. Afify, Kamel K. Mohammed, Aboul Ella Hassanien.

Year/Source: 2025, *Neural Computing and Applications*l

Objective: Measure short-term stress levels during mental tasks using EEG.

Data: EEG signals from the **SAM40 dataset** (32-channel EEG from 40 subjects performing four cognitive tasks)

Models: Deep CNN using **VGGish** for feature extraction and a custom CNN classifier.

Methodology: EEG signals were preprocessed and converted into spectrogram representations. A five-stage pipeline (preprocessing, segmentation, filtering, spectrogram generation, classification) was used. The VGGish network (typically audio feature extractor) was adapted to pull features from EEG spectrograms. Then a CNN classifier (VGGish-CNN) was trained. They compared the hybrid VGGish-CNN to a baseline VGGish-alone approach.

Results: The VGGish-CNN model achieved extremely high accuracy: **99.25%** on the SAM40 stress classification taskl. K-fold cross-validation confirmed stability. This suggests the proposed CNN effectively captures EEG stress patterns.

Limitations/Future: The authors note this was a preliminary study; future work will explore EEG-based stress detection in contexts such as neurorehabilitation and real-world settings. They also imply extending the model beyond 32 channels or different stressors.

Enhancing Stress Detection: A Comprehensive Approach through rPPG Analysis and Deep Learning Techniques

Authors: Laura Fontes *et al.*

Year/Source: 2024, *Sensors* (Open Access)mdpi.commdpi.com.

Objective: Develop a contactless (remote) stress detection method using facial video.

Data: UBFC-Phys dataset – videos of 56 subjects (after excluding 12) with remote photoplethysmography (rPPG) signals extracted from face video

Models: Hybrid deep networks including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) networks, and 1D Convolutional Neural Networks (1D-CNN)

Methodology: Faces in video were used to estimate rPPG (i.e., pulse and physiology remotely). The approach included signal processing to extract features (e.g. spectrogram of rPPG), followed by training several DL models. The authors performed hyperparameter optimization and data augmentation. They tested each model and hybrid combinations of LSTM/GRU/1D-CNN on the binary stress classification task.

Results: The best hybrid DL model achieved **95.83% accuracy** on the UBFC-Phys dataset. This demonstrates that rPPG-based, video-derived signals can be used for accurate stress detection with hybrid DL.

Limitations/Future: The study highlights potential for contactless monitoring but notes real-world conditions (lighting, movement) still pose challenges. Future work may include fusing this with other modalities and testing on more diverse environments.

A Multi-Modal Deep Learning Approach for Stress Detection Using Physiological Signals: Integrating Time and Frequency Domain Features

Authors: Jun-Zhi Xiang *et al.*

Year/Source: 2025, *Frontiers in Physiology*frontiersin.orgfrontiersin.org.

Objective: Detect stress in an occupational (nursing) context by fusing diverse physiological signals.

Data: Wearable physiological signals collected from nurses during work: accelerometer data (x,y,z), electrodermal activity (EDA), heart rate (HR), and skin temperature

Models: Customized deep learning framework with two parallel CNN branches – one for time-domain features and one for frequency-domain (via FFT) features, followed by fully connected layers .They used data augmentation (sliding windows, jittering) and SMOTE to balance classesfrontie

Methodology: Time-series data were segmented into windows. Statistical features (mean, variance, etc.) from raw signals and frequency-domain features from FFT were extracted. Two CNNs processed these feature sets separately. Their outputs were fused and fed to dense layers for classification into "stress" vs "no-stress." The model was trained and tested on an occupational dataset with labels of stress level.

Results: The multimodal CNN achieved **91.00% accuracy** and F1-score 0.91 for stress detection. This outperformed traditional ML classifiers (logistic regression, RF, XGBoost). An ablation study showed combining time and frequency features was critical.

Limitations/Future: The authors note future work should explore adding more modalities, enabling real-time detection, and improving generalization for wider applicability

Galvanic Skin Response and Photoplethysmography for Stress Recognition Using Machine Learning and Wearable Sensors

Authors: Alina Nechyporenko *et al.*

Year/Source: 2024, *Applied Sciences* (MDPI)mdpi.com.

Objective: Recognize acute stress induced by loud sirens, using wearable biosignals.

Data: Experimental study with 37 participants exposed to a simulated air-raid siren. During baseline, "no-stress" (silence), stress (siren sound), and recovery, **GSR (EDA)** and **PPG** signals were recorded using Shimmer wearable sensors.

Models: Several classical ML algorithms (k-Nearest Neighbors (KNN), Naive Bayes, Random Forest, SVM, etc.).

Methodology: Signals were preprocessed (artifact removal). Features were extracted from GSR (skin conductance) and PPG (heart rate variability indices). Feature selection identified the most stress-sensitive features (e.g. HRV parameters plus GSR). Different classifiers were trained to distinguish "stress" (siren) vs "no-stress."

Results: A kNN classifier performed best, achieving **83.3% accuracy** for stress vs no-stress. The authors report this performance using a combination of novel HRV and GSR features.

Limitations/Future: The scenario was a controlled lab simulation with a specific stressor. Real-world stress is more varied; future work should test in more naturalistic high-stress settings. The sample size (n=37) is moderate

An Explainable Deep Learning Approach for Stress Detection in Wearable Sensor Measurements

Authors: Martin Karl Moser, Maximilian Ehrhart, Bernd Resch.

Year/Source: 2024, *Sensors*mdpi.com.

Objective: Develop an interpretable DL model for real-time stress detection from wrist-worn sensors.

Data: Physiological data from the Empatica E4 wearable (likely including EDA, heart rate, etc.) recorded in a city-driving stress experiment (details in paper)

Models: A novel LSTM-based model augmented by a **Deep Generative Ensemble (DGE)** of conditional GANs (termed LSTM-DGE), plus an explainability layer using Integrated Gradients (IG)

Methodology: The LSTM-DGE was trained on sparsely labeled sensor data by using GANs to generate more training examples (addressing low-data regimes). The IG method was applied to highlight which input features (e.g. EDA, HR features) most influenced the stress prediction. Performance was compared against a state-of-the-art expert-rule stress detector.

Results: The proposed LSTM-DGE model improved over the baseline: +3.0 percentage points recall and +7.18 points precision .Crucially, IG explanations aligned with known stress markers (e.g. EDA peaks), validating the model.

Limitations/Future: While demonstrating better performance and interpretability, the study suggests further work on more varied data sources. Also, as an acute-stress setting (driving), chronic stress was not addressed.

A Dilated CNN-Based Model for Stress Detection Using Raw PPG Signals

Authors: Koorosh Motaman *et al.*

Year/Source: 2025, *IET Wireless Sensor Systems*researchgate.netresearchgate.net.

Objective: Build a simple DL model for stress detection from raw PPG data.

Data: PPG signals from the Empatica E4 wristband, specifically using the WESAD public dataset. WESAD contains multi-modal data labeled with “stress” and “non-stress” states.

Models: A deep **Dilated Convolutional Neural Network (Dilated CNN)** tailored to time-series PPG.

Methodology: Unlike approaches that hand-craft features, the authors fed raw PPG sequences into the dilated CNN without explicit denoising or preprocessing. The dilations allow capturing longer temporal context. The network was trained to classify each window as “stressed” or “not stressed.”

Results: The Dilated CNN achieved **93.56% test accuracy** and **AUC 96.52%** on the WESAD stress classification task. These results are notable given the minimal preprocessing. The model’s simplicity suggests feasibility for real-time wearable use.

Limitations/Future: The work used one dataset (WESAD); future work should test cross-dataset generalization. Also, multi-class or continuous stress regression was not explored.

Deep Learning-Based Stress Detection for Daily Life Use Using Single-Channel EEG and GSR in a Virtual Reality Interview Paradigm

Authors: Hun-gyeom Kim, Solwoong Song, Baek Hwan Cho, Dong Pyo Jang.

Year/Source: 2024, *PLOS ONE*journals.plos.org

Objective: Detect social/anticipatory stress during an interview using EEG and skin conductance.

Data: Controlled VR interview scenario with 30 participants. Collected **single-channel EEG (behind the ear)** and **GSR (electrodermal activity)** during stress-inducing interviews. Participants self-reported stress (VAS) to validate the induction.

Models: Several deep models – five CNN architectures (including ResNet variants) and a Vision Transformer (ViT). They also tested a “multi-column” model that fuses EEG and GSR features.

Methodology: Single EEG channel and GSR signal were processed separately or jointly. In single-modality models (“single-column”), each modality was fed to a network (e.g. ResNet-152 for GSR). In the multi-column model, EEG and GSR features were combined (e.g. via concatenation) and classified (ResNet-50 based). Performance was measured by AUROC.

Results: For single modalities, **ResNet-152 on GSR** achieved AUROC 0.944±0.027, and **ViT on EEG** reached 0.886±0.069j. The multi-column (fused) model gave the best result: **AUROC 0.954** (±0.018). This shows that combining EEG with GSR improves stress classification.

Limitations/Future: The study used a lab VR setting; future work should test longer-term or real-world use. Only one EEG channel was used – more channels could be explored. Also, they note ear-EEG is user-friendly but may limit signal richness.

Machine and Deep Learning Models for Stress Detection Using Multimodal Physiological Data

- **Authors:** Eman Abdelfattah, Shreehar Joshi, Shreekar Tiwari.
- **Year/Source:** 2025, *IEEE Access*researchgate.net.
- **Objective:** Compare various ML and DL methods on a multimodal stress dataset.
- **Data:** WESAD wearable sensor data (15 subjects) covering 4 states (baseline, stress, amusement, meditation). Signals used: tri-axial **acceleration (ACC)**, **ECG**, **blood volume pulse (BVP)**, body **temperature**, **respiration**, **EMG**, and **EDA**researchgate.net.
- **Models:** Seven classical classifiers (Logistic Regression, Naive Bayes, AdaBoost, XGBoost, Decision Tree, Extra Trees, Random Forest) and three DL models (fully-connected DNN, CNN, and RNN)researchgate.net.
- **Methodology:** All models were trained to classify the 4 states, using two schemes: (a) same-subject train/test splits, and (b) cross-subject (leave-some-subjects-out) splits. Features were preprocessed to "processed metrics" (e.g. HRV features from ECG, etc.) for ML models; DL models likely used raw or minimally processed inputs.
- **Results:** Traditional ML (RF, ExtraTrees, XGB) achieved very high F1-scores (~99%) when trained and tested on the same subjectsresearchgate.net. In the cross-subject setting, the RNN achieved an F1 of 93%. (Other DL models had similar high performance but RNN was highlighted.) Models generally performed better on chest-worn data in the same-subject case, and better on wrist data in the cross-subject caseresearchgate.net.
- **Limitations/Future:** This study is comprehensive but focused on a single dataset (WESAD, which is an induced-stress lab dataset). Future work could explore generalization to other datasets or deployment in less controlled environments.

Key Findings

- **Predominant Use of Physiological Signals:**
  - Reliance on biosignals (ECG, PPG, EEG, GSR/EDA).
  - Combination of modalities to enhance accuracy.
- **Deep Learning and Hybrid Models:**
  - Use of CNNs, LSTMs, and hybrid architectures.
  - High reported accuracies (>90%) on controlled datasets.

Trends

- **Predominant Use of Physiological Signals:** All studies relied on biosignals (ECG, PPG, EEG, GSR/EDA, respiration, etc.) from wearables or contactless sensorsHeart-related signals (PPG, ECG) and skin conductance dominate, reflecting the underlying stress response. A few studies combined modalities (e.g. EEG+GSR, or multi-sensor fusion) to boost accuracy.
- **Deep Learning and Hybrid Models:** Most recent papers leverage deep neural networks (CNNs, LSTMs, Transformers) or hybrid combinations (CNN + gradient boosting, CNN+LSTM, etc.). For example, Conv-XGBoostand Dilated CNN represent novel DL architectures, while others compare deep models to classical ML. Hybrid time/frequency-domain networks are also popular

contd

- **High Reported Accuracy:** Reported classification accuracies are very high (often >90%), typically on well-known datasets (WESAD, SWEET, UBFC-Phys) or controlled experiments. This indicates ML methods can capture stress signatures when trained/tested on similar data.
- **Explainability and Realism:** A few works explicitly address interpretability (e.g. Integrated Gradients in Moser et al. to highlight stress features) and realistic settings (e.g. free-living wearables, VR experiments, remote (camera) monitoring). However, many approaches still assume lab conditions and ample labeled data.

GAPS

- Despite progress, gaps remain. Many studies use small or homogeneous datasets (e.g. 15–56 subjects) and evaluate models on data split from the same subjects. Cross-dataset or long-term generalization is rarely tested. Only one study tackled free-living data. Real-time deployment and on-device constraints are under-explored (though Motaman et al. note their model suits real-time use). Interpretability is seldom addressed aside from one study. Also, stress is often treated as binary or categorical; continuous stress level prediction is less studied. Finally, textual or behavioral data (speech, social media) remain largely unexamined in peer-reviewed work (the focus is on physiology).

Future Research Directions

- **Multimodal Integration:** Combine physiological and contextual data.
- **Real-world Deployment:** Focus on longitudinal monitoring and diverse datasets.
- **Transfer and Personalization:** Enhance model generalization across users.
- **Explainable and Lightweight Models:** Improve interpretability and efficiency.
- **Continuous Stress Metrics:** Shift towards continuous stress level estimation.
- **New Modalities and Hybrid Approaches:** Explore additional data sources and hybrid algorithms.

Conclusion

- **Summary:** Recent advances in stress detection using machine and deep learning show promising results but highlight the need for more diverse datasets and real-world applications.
- **Call to Action:** Encourage further exploration of multimodal approaches and real-time implementations for effective stress monitoring.

References

1. **Kumar & Ankayarkanni (2025)**
   Title: Leveraging Conv-XGBoost algorithm for perceived mental stress detection using Photoplethysmography
   Link: ScienceDirect
2. **Al-Alim et al. (2024)**
   Title: A machine-learning approach for stress detection using wearable sensors in free-living environments
   Link: ScienceDirect
3. **Afify et al. (2025)**
   Title: Stress detection based EEG under varying cognitive tasks using convolution neural network
   Link: SpringerLink
4. **Fontes et al. (2024)**
   Title: Enhancing Stress Detection: A Comprehensive Approach through rPPG Analysis and Deep Learning Techniques
   Link: PubMed
5. **Xiang et al. (2025)**
   Title: An Explainable Deep Learning Approach for Stress Detection in Wearable Sensor Measurements
   Link: PMC

Reference

6. **Nechyporenko et al. (2024)**
   Title: Galvanic Skin Response and Photoplethysmography for Stress Recognition Using Machine Learning and Wearable Sensors
   Link: MDPI
7. **Moser et al. (2024)**
   Title: An Explainable Deep Learning Approach for Stress Detection in Wearable Sensor Measurements
   Link: PMC
8. **Motaman et al. (2025)**
   Title: A Dilated CNN-Based Model for Stress Detection Using Raw PPG Signals
   Link: IET Research Journals
9. **Kim et al. (2024)**
   Title: Enhancing Stress Detection: A Comprehensive Approach through rPPG Analysis and Deep Learning Techniques
   Link: PubMed
10. **Abdelfattah et al. (2025)**
    Title: An Explainable Deep Learning Approach for Stress Detection in Wearable Sensor Measurements
    Link: PMC


Thank you