

Review

Not peer-reviewed version

Review of ML Approaches in Anxiety Detection from In-Lab to In-the-Wild

[Maxine He](#)[†], [Abdulrahman Alkurdi](#)[†], [Jean L. Clore](#), [Richard B. Sowers](#), [Elizabeth T. Hsiao-Wecksler](#), [Manuel E. Hernandez](#)^{*}

Posted Date: 28 July 2025

doi: 10.20944/preprints202507.2265.v1

Keywords: anxiety; deep learning; machine learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Review of ML Approaches in Anxiety Detection from In-Lab to In-the-Wild

Maxine He ^{1,†}, Abdulrahman Alkurdi ^{2,†}, Jean L. Clore ³, Richard B. Sowers ⁴, Elizabeth T. Hsiao-Weckler ^{1,2,5} and Manuel E. Hernandez ^{1,5,6,7,8*}

¹ Neuroscience Program, University of Illinois Urbana-Champaign (UIUC)

² Department of Mechanical Science and Engineering, UIUC

³ Department of Clinical Psychiatry, University of Illinois College of Medicine at Peoria, University of Illinois Chicago

⁴ Department of Industrial and Enterprise Systems Engineering, UIUC

⁵ Department of Biomedical and Translational Sciences, Carle Illinois College of Medicine, UIUC

⁶ Department of Health and Kinesiology, UIUC

⁷ Department of Bioengineering, UIUC

⁸ Beckman Institute, UIUC

* Correspondence: mhermand@illinois.edu

† Authors with equal contributions.

Abstract

This review provides a detailed exploration of the current state of anxiety detection using machine learning (ML), with a focus on both feature-based and end-to-end models. The field has experienced rapid growth, with a significant increase in academic output necessitating an updated review of commonly used ML models and their performance, anxiety-inducing methodologies, data collection conditions, and dataset utilization. Feature-based ML models, such as Support Vector Machines, have been extensively employed due to their interpretability and simplicity. However, these models require manual feature engineering, which can be labor-intensive and potentially biased. End-to-end deep learning models have emerged as powerful alternatives, capable of utilizing raw signal directly and handling large datasets. Additionally, this review categorizes stressors into six distinct types – social, mental, physical, emotional, driving, and daily-life stressors – to provide a better overview of methodologies used to elicit anxiety. These are further explored based on whether data were collected under controlled in-lab conditions or real world in-the-wild conditions. This review underscores the need for further exploration into model architecture and their suitability for different types of data, advocating for a more nuanced and personalized approach to anxiety detection using machine learning.

Keywords: anxiety; deep learning; machine learning

1. Introduction

Mental health includes cognitive wellness and the ability to cope with adverse events in everyday life to function productively. Appropriate amounts of anxiety are essential for survival and motivation. Prolonged anxiety, however, can lead to mental illness and disorders that negatively impact cognition, emotion, behavior, and perception of self and the world [1]. The global COVID-19 pandemic significantly worsened the quality of life, with isolation, quarantine, and social avoidance further exacerbating mental health issues, leading to increased reports of symptoms of stress, anxiety, and depression [2]. Anxiety disorders, particularly, have become a major global health challenge due to their high prevalence and substantial economic impact, yet access to mental health services

remains limited [3,4]. Anxiety disorders are also associated with increased physical health conditions like cardiovascular diseases and weakened immunity, which could further impair decision making and cognitive performance [5,6]. These issues highlight the need for the development of effective diagnostic and therapeutic tools to improve treatment access and support mental health care systems [7,8].

The detection and monitoring of anxiety through machine learning (ML) has become an increasingly crucial area of research in affective computing. Significant advancements have been made with traditional feature-based (FB) models. Such models are popular for their interpretability and ease of implementation. They rely on manually engineered features derived from physiological or behavioral signals, enabling domain knowledge to play a critical role in their performance. However, this manual feature engineering process can be labor-intensive and may introduce biases or fail to capture subtle, complex patterns within the data [9,10].

End-to-end (E2E) approaches eliminate the need for explicit feature extraction by learning representations directly from raw data. These models, often include deep learning models, excel in handling complex, non-linear relationships and temporal dependencies in time-series signals [11,12]. A comparative study [11] highlights the dependency of E2E models on dataset size and quality, while also noting that FB models retain their relevance in many affective computing applications. Despite their computational intensity and less transparent operation, these models have demonstrated comparable performance in anxiety detection tasks.

As the volume of academic work in this field continues to grow, there is a pressing need to review existing model architectures, methodologies, data collection conditions, and dataset utilization. This review provides a comprehensive overview and summary of current state-of-the-art research in the ML application to anxiety detection. It also examines the challenges of deploying these models in real-world scenarios and explores future research directions for advancing the field toward the future of robust ML models and personalized healthcare.

1.1 Defining Anxiety

Stress and anxiety are closely related concepts often used interchangeably in literature, which can obscure definitions and reduce clarity. Stress is typically characterized as a current stimulus that triggers temporary physiological and psychological responses [13]. Anxiety, a negative affective and cognitive state that contributes to mental health disorders, consists of excessive worrying, rumination, or fear of future threats [14]. Spielberger's theory further differentiates anxiety into state and trait anxiety [15–17]. State anxiety refers to a temporary and transient negative emotional reaction accompanied by physiological responses to stressful stimuli. Trait anxiety, on the other hand, denotes an individual's inherent tendency to experience state anxiety when encountering stress [18].

While stress and anxiety share overlapping physiological responses mediated by the hypothalamic-pituitary-adrenal (HPA) axis and sympathetic nervous system (e.g., cortisol release, increased heart rate) [14], [19]—this review focuses on 'state anxiety' to describe the transient and stimuli-driven physiological responses. Many studies use the term 'stress' to describe these responses, while their investigations often align closely with what would be categorized as 'state anxiety', involving detecting physiological markers and psychological response to transient threats [20–24]. Several studies also used State-Trait Anxiety Inventory (STAI) [17] scores to confirm the presence of 'stress' elicited by anxiety-inducing stimuli [20,21], [25]. Furthermore, although the terms 'stress' and 'stressors' are sometimes used interchangeably to refer to both the cause and responses to adverse stimuli [26], we reserve the term 'stressor' specifically for the adverse stimuli presented to participants to evoke anxiety.

Throughout this review, we will use 'anxiety' as a simplified alternative term for state anxiety, capturing both subjective emotional states and physiological responses for understanding the impact of stressors. This approach, which distinguishes stressors (external triggers) from anxiety (cognitive, emotional, and physiological responses), aligns with Bystritsky and Kronemyer's transdiagnostic

model of the 'stress/anxiety complex' [26]. Here, stress and anxiety are conceptualized as interdependent phenomena on a continuum, where stressors elicit anxiety and anxiety exacerbate stress reactivity through shared neurobiological pathways (e.g., HPA axis dysregulation) [26]. While their framework underscores the interdependence of these phenomena, the persistent conflation of 'stress' and 'anxiety' in literature necessitated the inclusion of 'stress' in our search strategy to ensure comprehensive review of the current field.

1.2 Measuring Anxiety

three interconnected domains: subjective, behavioral, and physiological systems [27]. The subjective system captures an individual's self-reported experiences of anxiety, often through questionnaires or clinical interviews. The behavioral system encompasses observable expressions, such as facial expressions, body language, and other non-verbal cues. Finally, the physiological system reflects the biological and autonomic processes underlying anxiety.

i. Traditional Measures

Traditional anxiety assessment tools include standardized questionnaires such as the STAI [17] and the Beck Anxiety Inventory [28]. The subjective measures are straightforward to administer, though these methods are limited by recall bias, a systematic error due to inaccuracies in remembering past events, and their inability to monitor changes in real-time across different contexts [9], [29]. Clinical interviews conducted by trained professionals include assessment of anxiety symptoms, history of mental illness, and stressful events exposure. It offers a more precise diagnosis and cause of anxiety disorders than self-reports but are costly and time-consuming [30].

ii. Behavioral Measures

Behavioral observations focus on non-verbal aspects like facial expressions and body language and are useful for those unable to complete questionnaires or discuss their emotions. However, these observations can be subjective and may not accurately reflect anxiety in individuals who may suppress their symptoms [31], [32].

ii. Physiological Measures

The exploration of anxiety detection has significantly benefited from physiological measures, notably through the analysis of brain and cardiac activities, electromyography (EMG), respiration, and electrodermal activity (EDA). Brain activity measurement such as electroencephalogram (EEG), functional near-infrared spectroscopy and magnetic resonance imaging offer insights into neurological responses to anxiety, despite challenges such as spatial resolution, susceptibility to noise, and cost [33]. Cardiac activity, particularly heart rate variability (HRV) interpolated from electrocardiogram (ECG) or photoplethysmography (PPG), has been established as a robust indicator of anxiety, with advancements in wearable technologies enhancing real-time, non-invasive monitoring [34,35]. EMG [36] also provide peripheral insights into anxiety levels through muscle tension and autonomic nervous system activity [37–39]. EDA [40] and other measures, including pupil dilation [41] and skin temperature changes [42], have been incorporated into ML algorithms for anxiety detection, demonstrating the potential of non-invasive techniques [35,43]. Behavioral measures, leveraging observable physical and interactive markers, offer a complementary approach to understanding anxiety through the analysis of physical activity, voice characteristics, or other behavioral patterns, despite the need for adaptation to individual differences and real-world applicability [44,45]. These multifaceted approaches underscore the complexity of anxiety detection and the necessity for diverse measurement techniques to capture its nuanced manifestations.

1.3. Inducing Anxiety

Anxiety can be triggered in laboratory environments through the presentation of different stressor types. In-lab environments are specifically designed to conduct experiments under standardized conditions. They allow researchers to carefully manipulate variables, administer stressors, and supervise the study protocol while minimizing noise and artifacts, external influences, and confounding factors. The consistent and controlled atmosphere of in-lab conditions enables the collection of high-quality data to ensure that any observed effects on anxiety are attributable to the experimental stressor rather than random variables. On the other hand, in-the-wild conditions passively monitor participants' responses in real-world scenarios with minimal intervention. Here we characterize these stressors into six main categories: social, mental, physical, emotional, driving, and daily-life stressors.

Social stressors, such as the Trier Social Stress Test (TSST) [111], are designed to induce anxiety through social evaluation and performance pressure. The TSST typically involves public speaking and mental arithmetic tasks. Participants are asked to perform these tasks in front of an audience or evaluators, who often maintain neutral or critical expressions to trigger stress responses associated with social evaluative threat and fear of negative judgment.

Mental stressors typically involve tasks that demand high levels of cognitive processing, thus creating conditions of mental overload known to induce stress. Commonly used tasks include the Stroop Color-Word Test (SCWT), mental arithmetic, and n -back tasks. The SCWT is a classic psychological task used to evaluate the cognitive inference, attention, and selective inhibition response via the Stroop effect [46] to elicit state anxiety and acute stress responses [47]. The SCWT typically involves congruent and incongruent conditions, in which participants are required to name the color of presented words. The word itself might be colored by the same (congruent) or different color (incongruent; e.g., the word 'blue' printed in red font). In incongruent conditions, the conflict between the word's meaning and the font color creates cognitive interference, requiring the brain to inhibit automatic responses, leading to increased mental effort and stress [46]. Mental arithmetic tasks are another common type of mental stressor. These tasks require quick mental calculations and induce anxiety by taxing cognitive resources, such as working memory and processing speed, as well as frustration or stress due to their difficulty, the potential for errors, or time pressures [48]. The n -back working memory task is another popular mental task that requires participants to monitor a sequence of stimuli and indicate when the current stimulus matches the one presented n steps earlier [130], [131]. This task increases in difficulty as the value of n increases to put greater demands on the participant's working memory and cognitive control.

Physical stressors, such as the cold pressor test [48], elicit anxiety responses through physical exertion, discomfort, or pain. During the cold pressor test, participants immerse their hand in ice-cold water for a set period, activating the sympathetic nervous system due to the discomfort associated with cold exposure.

Emotional stressors involve exposing participants to stressful images or videos, such as those from the International Affective Picture System (IAPS) [49]. These stimuli are designed to elicit negative emotional states like fear, disgust, or sadness for anxiety responses.

Driving stressors is listed as a separate category because driving encompasses a combination of cognitive, emotional and physical demands that requires attention, decision making and coordination of motor movements [50,51]. It is also a common real-world activity, and responses induced in this context are highly relevant for studying anxiety detection in daily functioning [51]. Driving tasks usually involve simulated driving tasks that place participants in virtual or controlled real-world driving environments mimicking challenging or demanding driving conditions. These scenarios induce anxiety due to the need for constant attention, quick decision-making, and the potential for simulated hazards, reflecting real-world stressors faced by drivers. Such stressors are particularly relevant for studying anxiety in contexts where performance and safety are critical.

Daily-life stressors are a category that we propose. It refers to anxiety-inducing situations encountered in naturalistic settings, such as workplace challenges and academic examinations. There

is a growing interest in real-world ('in-the-wild') conditions, where data collection occurs in naturalistic settings as participants conduct their routine daily activities with little or no interference or supervision. These scenarios provide a more ecologically valid context for studying anxiety [52,53].

1.4. Detecting Anxiety

ML-based anxiety detection efforts predominantly utilized FB models, which required extensive domain-specific knowledge for manual feature extraction and engineering [12,54]. These models often depended on handcrafted features derived from physiological signals or behavioral data. The typical process involves the following steps: 1) Data Collection, where physiological signals are gathered from participants; 2) Data Preprocessing, which includes removing noise and artifacts; 3) Feature Extraction, which involves the derivation of relevant features; 4) Feature Selection using techniques to retain most informative features; 5) Model Training, where selected features are used to train classification or prediction models, with optional hyperparameter fine-tuning to optimize performance; 6) Model Validation employing methods such as test datasets, k-fold cross-validation, or leave-one-participant-out validation; and 7) Model Evaluation using performance metrics like accuracy, precision, recall, and F1-scores to assess effectiveness (Figure 1).

FB models often allow easier interpretation of results and identification of features contributing most to the classification or prediction, and experts with domain knowledge can incorporate features known to be relevant to improve the model performance. FB models generally require less computation power and can be suitable for small datasets. However, relying on feature extraction and selection could potentially introduce bias and information loss. For instance, important subtle indicators of anxiety might be overlooked if they are not part of the selected features. Generalizability may also be limited because handcrafted features may not generalize well to new datasets or populations, and FB models may struggle with nonlinear nature and complexities in physiological signals. In contrast, the field has seen a significant shift toward E2E deep learning models like CNN and LSTM. These models can automatically learn hierarchical representations from raw data, reducing the need for manual feature extraction and potentially capturing more complex and subtle patterns associated with anxiety [12]. Advancements in computational power, the availability of large datasets, and improved algorithms have facilitated this transition, enabling models that can generalize better across different populations and settings.

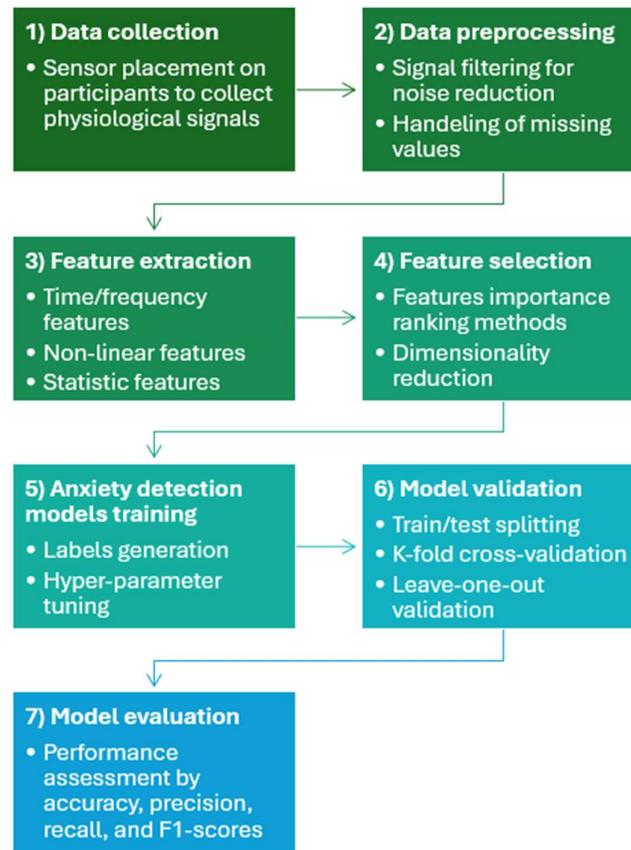


Figure 1. Typical process of developing feature-based anxiety detection models. End-to-end models bypass feature extraction and selection steps (Step 3 and 4) by directly utilizing raw signals.

2. Review of Anxiety Detection using ML

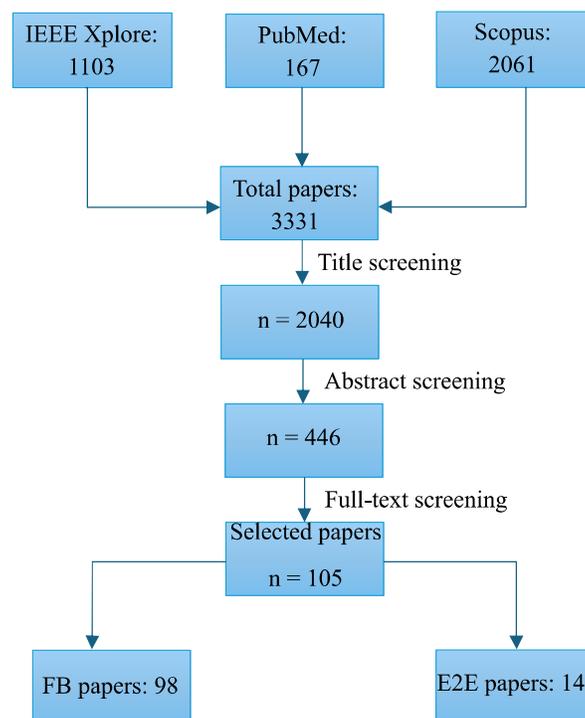
2.1 Methods

The literature review methodology included a thorough search across Scopus, IEEE Xplore, and PubMed databases, covering publications from 2010 to November 2024. The search was conducted using keyword combinations, such as ‘stress or anxiety or mental workload detection,’ and ‘AI or machine learning or deep learning’ (Table I). A systematic screening process was implemented, involving title screening, abstract evaluation, and full-text analysis. Inclusion criteria focused on original articles that use one or more stressors to induce anxiety and use physiological signals to measure anxiety. Review papers were excluded from consideration.

The initial search yielded 3331 articles, i.e., 2061, 1103, and 167 articles from Scopus, IEEE Xplore, and PubMed, respectively (Figure 2). Title screening narrowed the field to 2040. Further abstract evaluation reduced this number to 446, and a full-text review culminated in 105 relevant papers. Data extracted from these articles included ML architectures, ground truth questionnaires, environmental conditions, tasks for inducing anxiety, model performance, physiological signals, and datasets utilized. A more detailed summary of each paper, categorized based on stressor types, can be found in Supplemental Material Tables I- VI. These articles were further segregated between ML approaches that used FB models (98 articles) or E2E models (14 articles) (Tables II and III). Seven papers overlapped and used both FB and E2E models [24], [56]–[61] (note [55] used traditional ML models for E2E).

Table 1. Search string used for each database.

Database	Search String
PubMed	("machine learning") AND "anxiety" NOT ("depression" OR "Autism" OR Stroke OR "depressive" OR phobia)
IEEE Xplore	("machine learning") AND (("psychological stress" OR "mental stress" OR "emotional stress" OR "mental workload" OR "stressful" OR "anxiety"))
Scopus	TITLE-ABS-KEY ("machine learning" AND ("psychological stress" OR "mental stress" OR "emotional stress" OR "mental workload" OR "cognitive workload" OR "Cognitive stress" OR "anxiety")) AND NOT TITLE-ABS (review OR survey OR scoping OR autism OR autistic OR diabetic) AND NOT TITLE (treatment OR suicide OR surgery OR depression OR depressed OR "anxiety disorders" OR vaccine OR child OR children OR cells OR glycemia OR tumor OR tremor OR gender OR wealth OR "mental illness" OR disorder OR "management system" OR "intelligence" OR disease) AND (LIMIT-TO (PUBSTAGE, "final")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp"))

**Figure 2.** Flowchart of database results and screening methods for articles, segregated based on machine learning approaches used: feature-based (FB) and end-to-end (E2E). Seven selected papers used both FB and E2E models.

2.2 ML models and architectures

2.2.1. FB Models

Early studies in anxiety detection predominantly utilized FB models including traditional ML and deep learning models (Table II, alphabetized by last names). These models depend on engineered features specifically chosen for their relevance to the state of interest and are often supervised by manually created labels or ground truth. Their advantages include interpretability, simplicity, and lower computational demands, making them particularly effective when working with limited datasets. However, the performance of FB models may be dependent on the quality of features, which requires domain expertise. Deep learning methods within the FB paradigm, while also requiring

engineered features as input, process these features through networks inspired by the human brain's structure, consisting of interconnected nodes organized into layers: one input layer, one or more hidden layers, and an output layer [56]. Artificial Neural Networks (ANNs) serve as the foundation for more complex architectures, with feed-forward neural networks being the simplest form where information flows only in one direction from input to output. Building upon ANNs, specialized architectures like CNN which originally designed to process images, have become increasingly popular for processing engineered features in anxiety detection [12]. However, regardless of whether traditional ML or deep learning is used, the performance of FB models remains dependent on the quality of the engineered features, which requires domain expertise [56].

Table 2. Articles that used feature based models

FB category	Models	References
Traditional models	SVM	[57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [53], [76], [77], [78], [79], [25] [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90],[91], [92], [93], [94], [21], [95], [96], [97], [98], [22], [99],[100], [101], [102], [103], [104] [105], [106], [107], [108], [109] [24], [110], [111]
	RF	[57], [59], [60], [62], [65], [112], [68], [69], [55], [113], [114], [71], [72], [115], [116], [53], [78], [117], [79], [83], [84], [82], [87], [118], [88], [92,93], [94], [95], [119], [120], [121], [122], [23], [101], [102], [103,105], [24], [111]
	kNN	[57], [59], [62], [63], [67], [55], [71], [73], [74], [75], [53], [76], [77], [79], [80] [123], [82], [83], [85], [123], [42], [87], [89], [91], [124], [97], [21], [100], [121], [101], [122], [102], [103], [106], [108,109] [125], [111]
	Naïve Bayes	[63], [64], [65], [113], [69], [71], [73], [75], [53], [77], [126], [80], [85], [100], [101], [103], [105], [125], [111]
	Boost	[57], [59], [60], [62], [67], [113], [72], [74], [78], [79], [127], [75], [53], [78], [87], [118], [21], [95], [96], [121], [122], [101], [128], [104], [106], [107], [24]
	DT	[58], [60], [62], [63], [64], [65], [67], [113], [69], [53], [79], [127], [80], [82], [83], [86], [96], [97], [121], [122], [100], [101], [102],[106], [108,109], [24]
	LR	[62], [65], [68], [129], [69], [117], [81], [82], [83], [84], [87], [118], [100], [101], [102], [103], [104] [105], [107], [24], [111]
	LDA	[59], [63], [64], [67], [73], [76], [79], [80], [85], [87], [89], [130], [121], [122], [99], [108,109], [24]
	Ensemble	[53], [85], [131], [52], [97], [106]
Deep learning models	CNN	[132], [57], [133], [134], [62], [135], [136], [69], [74], [84], [90], [91], [137], [23], [138]
	ANN	[67], [113], [139], [81], [42], [131], [91], [140], [101], [103], [106], [141]
	MLP	[62], [69], [71], [78], [81], [83], [102], [142]
	LSTM	[62], [139], [143], [52], [140], [144]
	ResNet	[78], [145]
	RNN	[143]
	CNN-LSTM	[62]
GCN	[132]	

SVM: support vector machine; RF: random forest; kNN: k-nearest neighbors; DT: decision tree; LDA: linear discriminant analysis; LR: linear regression; CNN: convolutional neural network; ANN: artificial neural network; RNN: recurrent neural network; MLP: multilayer perceptron; LSTM: long short-term memory, GCN: graph convolutional network.

2.2.2 E2E Models

E2E models, typically based on deep learning, operate directly on raw data without the need for manual feature engineering (Table III, alphabetized by last names). These models, like CNN and Long Short-Term Memory (LSTM), can capture intricate details through complex architecture, and automatically learn features to capture patterns and relationships within the data. LSTM is a specialized deep learning architecture, as a subset of recurrent neural networks, designed to handle sequential data by maintaining a memory cell that can selectively remember or forget information over time, making it ideal for time-series physiological signals [146]. For example, Onim & Thapliyal [143] achieved 96% accuracy using LSTM to analyze EDA, BVP, and temperature data from wearable devices for detecting stress in elderly individuals. However, these models are often criticized for being ‘black box’ models due to a lack of interpretability and are computationally intensive, requiring substantial data to prevent overfitting [116].

Table 3. Articles that used end-to-end models.

E2E models	References
CNN	[57], [147], [148], [69], [11], [149], [150]
FCN	[151], [11]
Inception Time	[11]
LSTM	[152], [153], [154], [152]
Multi-ResNet	[11]
ResNet	[87], [151]
Encoder	[11]
Time CNN	[11]
CNN-LSTM	[147], [153], [11]
MLP	[153], [11], [151], [130], [124]
MLP-LSTM	[153]
RF and kNN	[55]

CNN: convolutional neural network; FCN: fully convolutional neural network; MLP: multilayer perceptron; LSTM: long short-term memory.

3. Results

3.1. ML Techniques and Performances

The review of the literature revealed a diverse range of FB machine learning approaches applied to anxiety detection using physiological signals (Table II). Deep learning algorithms also gained popularity in recent studies using extracted features, suggesting a transitional phase before fully E2E models are more broadly adopted [132,135]. Traditional machine learning models using engineered features have demonstrated consistently strong performance across multiple studies. Among classical algorithms, SVM was the most prevalent, used in 60 studies, followed by RF in 40 studies, and kNN in 39 studies. SVM has been particularly effective, as shown by Schmidt et al. [122], publisher of WEASAD dataset (see next section), who achieved 93.12% accuracy for binary classification. RF have also shown impressive results, with [114] reporting accuracy as high as 99.5%

using HRV features from the same WESAD dataset [122]. These classical FB models perform well with smaller datasets and provide interpretable results through feature importance analysis [100].

E2E approaches have shown varying degrees of success, with performance ranging significantly across studies. Fan et al. [149] demonstrated strong results using a CNN with Convolutional Attention Block Module, achieving 97.5% accuracy in three-class classification using the WESAD dataset. However, other studies show more modest results - Dziezyc et al. [11] reported 79% accuracy using fully convolutional networks on raw physiological signals. This variability in performance is particularly evident in real-world applications, as shown by [107], where accuracy dropped significantly from lab conditions (76.4%) to in-the-wild conditions (64.7-70.1% F1 score).

The performance of these ML models, as measured typically by accuracy, also varied depending on the nature of the task used to induce stress, the environment of the data collection, and the signals utilized. For instance, laboratory studies generally reported higher accuracy, with models achieving up to 99% accuracy [63], [114]. In contrast, studies that attempted to detect stress in wild or semi-wild environments faced more challenges in that they often showed a decrease in model performance due to the increased noise and variability in the data [118], [119].

Also, most of the studies focused on binary classification (e.g., no anxiety vs. anxiety), while a few also moved on to multiclass classification (e.g., low vs. moderate vs. high anxiety) for added granularity in anxiety detection. For example, Schmidt et al. [122] and Bobade et al. [67] compared binary classification with three-class classification (baseline vs. amusement vs. stress) using the WESAD dataset. However, it is noted that the multiclass classification often yielded lower performance due to the intrinsic complexity in multiclass problems and subtle difference in physiological responses across tasks [67], [122], [60,62]. There has also been a trend of using ensemble methods to incorporate multiple models and improve results [53], [85], [131], [52], [97], [106]

3.2. Open Datasets for Anxiety Detection

Several widely used datasets are employed to validate model performance on various stressor types for anxiety detection.

Social Stressors: The WESAD (Wearable Stress and Affect Detection) dataset includes data from 15 participants exposed to the TSST alongside baseline and amusement conditions [122]. It provides multimodal measurements using ECG, EDA, PPG, respiration, temperature, and acceleration.

Mental Stressors: The SWELL-KN (Smart Reasoning for Well-being at Home and at Work - Knowledge Work) dataset examines mental workload via HRV in 25 participants performing simulated office tasks, such as writing reports and reading emails, under no stress, time pressure, or interruptions [155]. The CLAS (Cognitive Load and Stress) dataset focuses on cognitive stress, featuring ECG, EDA, and respiration data collected during tasks with varying mental stressors (math problems, logic problems and SCWT). The MAUS (Multimodal Affect and Understanding Stress) dataset used *n*-back test for mental stress and provides ECG, EDA, and PPG [156]. The CogLoad dataset captures ECG, EDA, and body temperature during six mental tasks [157]. Similarly, the EEG During Mental Arithmetic Tasks Dataset provides EEG data to analyze neural activity using mental stressors [158]. The SAM40 (Stress and Anxiety Monitoring) dataset offers EEG with peripheral signals like ECG, EDA, and respiration data collected from 40 participants during SCWT and mirror image recognition tasks [159].

Driving Stressors: The SRAD (Stress Recognition in Automobile Drivers) dataset contains multimodal data of 17 healthy drivers completing driving tasks in a semi-controlled wild condition of driving around Boston, MA, USA. [160]. ECG, EMG, foot and hand EDA, and respiration data were collected. The protocol included a resting state, driving in the city, and driving on the highway. The AffectiveROAD dataset emphasizes affective states during driving and includes ECG, EDA, and respiration to also provide multimodal understanding of driving stress [161].

Emotional Stressors: The DEAP (Dataset for Emotion Analysis using Physiological Signals) includes EEG, ECG, EDA, respiration, and BVP signals from 32 participants watching music videos designed to evoke specific emotional responses [162]. It features self-reported measures of arousal,

valence, and dominance to annotate emotional states. The CASE (Continuously Annotated Signals of Emotion) dataset has physiological signals (ECG, BVP, EMG, EDA, respiration and skin temperature) from 30 participants viewing video clips that contained different emotions [163].

Daily-life Stressors: The Multimodal Dataset for Nurses provides ECG, EDA, and skin temperature data collected from nurses during physically demanding work shifts, with observational data on workload and patient interactions adding context to the stress measurements [164].

Mixed types of stressors: The Non-EEG Dataset for Neural Status Assessment captures non-EEG signals and used multiple types of stressors. ECG, EDA, and respiration were collected during multiple anxiety-inducing conditions, including physical stressor (physical exercise), mental stressor (SCWT), and emotional stressor (watching a horror movie), to evaluate neurological states [165].

3.3. Model Performances Based on Stressor types

Table 4 separates the tasks used into different stressor categories, namely social, mental, emotional, driving, physical, and daily-life stressors. Seventy-nine studies used a single stressor in the protocol, while 26 studies used a combination of different types of stressors. Ninety-three studies were conducted in a lab-controlled environment, while 9 studies focused on natural settings for real-world anxiety detection, 5 of which included both lab and in-the-wild conditions. Eight papers using the SRAD dataset were considered as 'semi-wild environment', combining real-world variability with predefined protocol, to monitor driver's physiological states [160].

Table 4. Articles grouped by stressors used to induce anxiety.

Stressor Type	References
Social Stressors	[132], [57], [59], [135], [67], [68], [136], [113], [129], [114], [11], [149], [75], [154], [78], [151], [79], [127], [131], [118], [21], [95], [143], [119], [96], [140], [120], [152], [121], [150], [98], [22], [122], [99], [100], [166], [145], [107], [24], [111]
Mental Stressors	[132], [58], [133], [134], [63,64], [65], [148], [167], [112], [68], [55], [129], [69], [72], [139], [168], [74], [75], [116], [53], [154], [77], [117], [169], [126], [80], [81], [123], [82], [83], [85], [86], [87], [170], [90], [124], [118], [137], [130], [95], [119], [96], [97], [144], [99], [101], [103], [128], [104], [105], [106], [107], [108,109], [111]
Physical Stressors	[57], [61], [72], [154], [77], [84], [171], [22], [107], [111]
Emotional Stressors	[134], [168], [73], [78], [115], [75], [77], [110], [141]
Driving Stressors	[147], [172], [62], [70], [71], [87], [88], [92], [93], [94], [102], [138]
Daily-life Stressors	[60], [153], [116], [53], [77], [52], [96], [23], [107]

3.3.1. Social Stressors

TSST is mostly used to induce social-evaluative threat and state anxiety, and the WESAD dataset is a common benchmark for comparing and evaluating model performance. Most studies used study protocol design of the dataset as the ground truth, while a few also used STAI scores or clustering to label anxiety states [24,100,127,136]. ECG, BVP via PPG, EDA, and respiration were common physiological signals used. Studies often combined these signals to capture a more multimodal view of physiological changes associated with anxiety, as each signal provides unique insights into autonomic nervous system activity. Commonly extracted features included statistical measures (mean, standard deviation, min, max), HRV metrics such as root mean square of successive differences (RMSSD), low-frequency (LF) to high-frequency (HF) ratios, and EDA-derived metrics like skin conductance peaks and rise time. Notably, studies using feature selection often found that EDA- and ECG-derived features were the most predictive of anxiety levels [113,127,166].

Traditional models dominate in anxiety detection research. SVM was frequently used for binary classification tasks and often demonstrated high accuracy. For example, Akella et al. [59] reported SVM accuracy ranging from 24% to 91% using EEG signals, and Sandulescu et al. [98] achieved SVM accuracy between 73.26% and 83.08% using PPG and EDA signals. RF, AdaBoost, and kNN are also common FB ML models showing high accuracy in anxiety detection, and cardiovascular signals and EDA signals are popular candidates as biomarkers of stress responses. Moreover, it was observed that binary classification of stress tended to achieve higher accuracy compared to multiclass classification, where distinguishing between different types of stressors or different levels can be more challenging.

More recent studies adopt deep learning models, and their performance is comparable to traditional FB models. Studies report that CNN and LSTM had comparable, sometimes higher, performance compared to traditional ML models, and sometimes achieving accuracies above 90% in binary classification tasks [131,132,135,136,143]. Bobade et al. [67] found that ANN outperformed other classifiers, achieving 95.2% accuracy in binary classification, while Chatterjee et al. [136] demonstrated that CNN-based models could achieve accuracies ranging from 90.3% to 94.2%. Studies such as [11], [149], [151,154] reported promising results with deep learning and E2E architectures, indicating that these models can learn effective feature representations from raw signals. This advantage reduces the bias introduced by manual feature selection, enhancing model generalizability.

3.3.2. Mental Stressors

This category is the most represented across studies, with 56 studies focusing on tasks that induce anxiety through mental cognitive tasks. Sample sizes ranged widely in number, from as few as 5 to as many as 90 participants, and included diverse groups such as teenagers, college students, and older adults. However, many studies did not include detailed demographics like race, ethnicity, or gender. Besides the aforementioned common mental stressors in Section I c, some studies also used more unique tasks such as gaming or virtual reality exercises [139], [89], [170], [115]. EEG, HRV derived from ECG or PPG, and EDA were commonly used for anxiety detection induced by mental stressors. Some studies also reported their preprocessing steps to ensure data quality, using techniques like independent component analysis and bandpass filtering to remove noise and artifacts for EEG [89], [105]. Other methods like template matching, threshold filters, outlier removal and empirical mode decomposition, and baseline correction were usually used for ECG and PPG signals for reliable HRV extraction [87,95,96,119].

While deep learning models like CNNs and LSTMs achieved high accuracy, traditional machine learning models such as SVM and RF still demonstrated strong performance, particularly in binary tasks. For example, in the study by Badr et al. [63], SVMs achieved accuracies as high as 99.58% in detecting stress using EEG data during the SCWT. Benchekroun et al. [112] reported RF accuracies between 71% and 84% for stress detection, and Campanella et al. [68] found similar results using a combination of PPG and EDA features.

Deep learning methods were increasingly applied, especially in studies using EEG data. For example, Appriou et al. [134] achieved a 70% accuracy in classifying mental workload with CNNs, although the small dataset limited their effectiveness. LSTM networks, known for their ability to analyze sequential data, achieved a 95% accuracy in stress detection during gaming tasks in the study by Dhaouadi et al. [139]. Ensemble methods further enhanced performance by combining predictions from multiple models for final classification [85,106,173].

Similar features to those used in social stressors from ECG and EDA signals also show potential in classifying anxiety levels. However, generalization across different stress conditions and tasks posed a significant challenge in these studies. Several researchers examined how well models trained under one set of conditions would perform under different stress-inducing scenarios. Mamdouh et al. [89] tested models trained on mental arithmetic tasks against data from virtual-reality-induced

stress, finding that kNN slightly outperformed SVM and LDA, with an accuracy of 87.1%. This finding underscored the challenges in transferring models between different stress contexts.

3.3.3. Physical Stressors

Physical stressors in these studies included tasks such as physical exercise using a cycle-ergometer or walking exercises, or thermal stress induced by physical activity at high temperatures or by cold ice water at low temperatures (cold pressor test [48]). These stressors were used to evaluate how physical demand can trigger significant anxiety responses. In the study by Delmastro et al. [72], physical anxiety was induced using a cycle-ergometer exercise in conjunction with the SCWT to examine anxiety responses in older adults. The study utilized ECG and EDA signals to detect changes in anxiety levels, with RF and AdaBoost models achieving accuracies between 85.4% and 88.2%. These findings suggest that physical activity, while inducing a low level of anxiety, can positively impact cognitive performance during stressful tasks.

Han et al. [77] explored anxiety detection during both in-lab and real-world conditions, using a variety of stressors including physical exercises such as plank exercises and the TSST. The study recorded ECG, PPG, and EDA signals to monitor anxiety responses. The kNN model showed high accuracy, with up to 94.55% accuracy for in-lab data and 100% for certain real-world conditions when physical activities were excluded due to motion artifacts. This result highlights the challenges of detecting anxiety in dynamic, real-world environments where physical activity can introduce significant noise. In another study [119], physical anxiety was induced by the cold pressor test and walking exercises. The study aimed to compare the performance of different heart rate monitoring devices in detecting anxiety. Using ECG and PPG signals, RF achieved accuracies between 81% to 85%, demonstrating that slight differences in heart rate readings across devices did not significantly impact the accuracy of anxiety detection. Lastly, Sandulescu et al. [22], [98], examined the feasibility of using a customized wearable system to detect thermal and mental anxiety. The study involved physical activity at 40°C to induce thermal stress and used the TSST for mental stress. However, the results for thermal anxiety detection were not clearly reported, leaving some uncertainty about the system's effectiveness under thermal stress conditions.

3.3.4. Emotional Stressors

This category focused on tasks that evoked emotional reactions and challenges through emotionally challenging stimuli such as viewing video clips with negative arousing. Three studies, [168], [106], [118], [147], used emotional visual stimuli from the IAPS [148], which provided images that induced negative, positive or neutral emotions. For example, Ding et al. [168] used negative visual stimuli from IAPS and a stop-signal task to induce emotional stress, with physiological signals like ECG and EDA used to predict state anxiety with regression models. The correlation coefficients between predicted and actual STAI scores using LASSO regression ranged from 0.4748 to 0.5528, suggesting moderate accuracy in capturing emotional responses. Han et al. [77] utilized ECG, PPG, and EDA signals, with kNN and SVM models achieving high accuracy in lab settings (94.55%) where motion artifacts were minimal.

Other studies, e.g., [137], [149], used video clips containing emotionally charged content to elicit anxiety. Another study by Henry et al. [78] explored the generalizability of ML models focusing solely on cardiac signals using two datasets (CASE and WESAD), one with emotional stress by watching video clips with negative emotions and the other used a social stressor, respectively. SVM models performed slightly better than other models, and WESAD had higher accuracy compared to CASE. Using ECG signals also suggested better performance than PPG due to higher signal quality. Gazi et al. [115] examined anxiety detection in individuals with arachnophobia by exposing them to spiders in a virtual environment. This study used ECG, EDA, and respiration signals, achieving an accuracy of 88% with an RF model when adding respiratory features as complementary information.

Common physiological signals include ECG and EDA, and some researchers also utilized pupil diameter, and even facial cues to detect emotional stress responses. Giannakakis et al. [75]

investigated the use of facial cues and camera-based PPG to detect stress during tasks such as social exposure, emotional recall, and viewing stressful images. The study reported accuracy between 85.54% and 91.68%, with kNN and AdaBoost models having the highest performance. The inclusion of facial features like eye blinks, head movement and changes in pupil diameter, alongside heart rate data, proved to be effective in classifying stress responses to emotionally charged stimuli. Using pupil diameter features alone, Erkus et al. [73] reported accuracies ranging between 44.4% and 67.9% using LDA and coarse tree. Two studies explored brain activity for anxiety detection. For ECG, time and frequency domain and nonlinear HRV features such as RMSSD, LF/HF ratios, and standard deviations from Poincaré plot were commonly used [168], [77,115], while for EDA, features like mean skin conductance response and mean amplitude [147], [151] also contributed to anxiety detection as an indicator of the sympathetic nervous system activation.

3.3.5. Driving Stressors

Eight out of 12 studies in Table IV utilized the SRAD dataset, and four studies [87,92–94] used simulated driving tasks in lab-controlled condition. ECG, EMG, foot and hand GSR, and respiration data were collected. The protocol included a resting state (low stress), driving in the city (medium stress), and driving on the highway (high stress). Additionally, self-reported stress and anxiety levels were collected from participants.

The primary physiological signals used across these studies included ECG and EDA, with some studies [102,147,172] also incorporating additional signals such as respiration and EMG. These signals were processed to extract various features, including HRV features, skin conductance response from EDA, and ECG-derived respiration. To ensure accurate stress detection, noise reduction methods like template matching for ECG R peak detection, z-scores for removing unreliable RR intervals, baseline correction, and wavelet decomposition were commonly applied.

RF consistently showed high performance across multiple studies, particularly excelling in both binary and three-class stress classification tasks. RF achieved an accuracy of 98.2% in a study [102] and area under the curve values of 91.5% in [88], demonstrating its robustness in handling stress detection using physiological signals. SVM also performed well, with varying levels of success depending on the study. Cruz et al. [70] reported an impressive accuracy of 96.3% using a tree-optimized SVM model for its effectiveness in stress detection during driving tasks using ECG as the single modality. Other traditional models like kNN, DT, and LR were used with similar success. For example, kNN achieved an accuracy of 91.2% in Siam et al. [102] and 80.05% in Dalmeida et al. [71], suggesting that these models can still be valuable in stress detection, albeit with typically lower performance compared to RF and SVM.

Arya et al. [62] reported an accuracy of 95.67% using LSTM for binary classification and 88.70% for ternary classification, CNN and CNN-LSTM hybrids were also effective, with CNN achieving 94.2% accuracy in binary classification and 85.60% in three-class classification [62]. However, the CNN-LSTM hybrid did not significantly outperform the standalone LSTM model. A notable approach in some studies was the use of automated pipeline optimization for ML, such as the Tree-based Pipeline Optimization Tool. Liu et al. [88] reported that this optimization method, with an extra tree classifier, improved the AUC to 93.4%, indicating the potential of automated ML in refining model performance.

Across studies, key features for stress detection included heart rate, HRV, and skin conductance features. The consistent importance of these features across different models and datasets illustrates their potential as reliable biomarkers for anxiety detection. While deep learning models, particularly LSTM, generally outperformed traditional ML models, RF consistently showed strong performance, making it a reliable choice for stress classification tasks. These results suggest that while deep learning models offer advantages, traditional models like RF can still be highly effective, especially in cases with limited data.

In addition, Dalmeida et al. [71] explored the generalization of models trained on the SRAD dataset and tested on the AffectiveRoad dataset; although they did not report performance metrics. This work highlights the ongoing challenge of ensuring that anxiety detection models can generalize well across different datasets and real-world scenarios. The high accuracy observed in studies like Siam et al. [102] also suggest the feasibility of applying anxiety classification in real-world scenarios, such as monitoring drivers in near-real-time with short recording periods (e.g., 1-minute segments). This practical application is crucial for developing effective anxiety detection systems for high-stress environments like driving.

3.3.6. Daily-life stressors

The studies using daily-life stressors are listed in Table V. Model performances on daily-life stressors performance varies largely, ranging from 31.3% [23] to 100% [77] in classification accuracy. Studies incorporating lab-based and real-world data emphasized the challenges in maintaining high model accuracy outside the lab [53,77], [116], [107]. Models trained on lab data often showed reduced performance when applied in real-world scenarios [77,107], highlighting the importance of robust preprocessing and the inclusion of context-aware algorithms to handle the variability found in real-life situations. In several studies [53,60,107,116,153], they also mentioned the imbalanced samples in the data collected in the real-world (more non-anxious events than anxious events), and Başaran et al. [153] proposed semi-supervised learning to address the intensive labeling in massive in-the-wild dataset.

Table 5. Articles grouped by the experimental environment.

Stressor Type	References
Social Stressors	[132], [57], [59], [135], [67], [68], [136], [113], [129], [114], [11], [149], [75], [154], [78], [151], [79], [127], [131], [118], [21], [95], [143], [119], [96], [140], [120], [152], [121], [150], [98], [22], [122], [99], [100], [166], [145], [107], [24], [111]
Mental Stressors	[132], [58], [133], [134], [63,64], [65], [148], [167], [112], [68], [55], [129], [69], [72], [139], [168], [74], [75], [116], [53], [154], [77], [117], [169], [126], [80], [81], [123], [82], [83], [85], [86], [87], [170], [90], [124], [118], [137], [130], [95], [119], [96], [97], [144], [99], [101], [103], [128], [104], [105], [106], [107], [108,109], [111]
Physical Stressors	[57], [61], [72], [154], [77], [84], [171], [22], [107], [111]
Emotional Stressors	[134], [168], [73], [78], [115], [75], [77], [110], [141]
Driving Stressors	[147], [172], [62], [70], [71], [87], [88], [92], [93], [94], [102], [138]
Daily-life Stressors	[60], [153], [116], [53], [77], [52], [96], [23], [107]

Data collected in laboratory settings typically resulted in higher accuracies compared to data collected in natural settings [23], [53,116], [77], [96], while adding context information seemed to improve the model performance [53,116]. This difference underscores the challenges of applying lab-trained models to real-world environments, where unpredictable factors such as ambient noise, physical movements, and uncontrolled psychological states reduce the quality of signal detection [52,85,131].

4. Discussion

4.1 Summary

The review confirms the widespread application of traditional ML approaches in anxiety detection, with SVM being the most prevalent due to its ability to handle high-dimensional datasets and achieve high accuracy. Ensemble methods showed competitive results, indicating their

effectiveness in combining multiple weak learners for robust predictions. Deep learning methods are increasingly popular for both FB and E2E approaches, achieving comparable or superior performance to traditional ML models. Widely used datasets like WESAD, SWELL-KN, and SRAD have enabled benchmarking and comparative analyses across ML models. Binary classification remains the most common approach, often achieving higher accuracies than multiclass classification due to the complexity of distinguishing subtle physiological differences across multiple states. However, the variations in features employed and model architecture across different studies make it challenging to compare across studies and identify the most reliable features or robust models for anxiety detection. This observation was supported by reviews in the relevant field, which highlight that no single model or feature sets universally outperform others due to diverse experimental conditions and participant characteristics [14], [174].

Social stressors, particularly those induced by the TSST, dominate the field, as seen in studies using the WESAD dataset. These studies often achieve high accuracy because of controlled experimental designs and multimodal data collection. Mental stressors are the most represented category, with studies employing tasks like SCWT to elicit cognitive stress. Driving stressors provide valuable insights into stress detection in semi-wild environments using datasets like SRAD and AffectiveROAD. Physical and emotional stressors further expand the scope, with studies focusing on tasks like physical exercise and emotional challenge to understand diverse anxiety responses. Lastly, there has been increasing interest in translating lab findings and models to in-the-wild data, highlighting future directions in anxiety detection for real-life.

4.2. Key observations

Comparison between FB and E2E models: The effectiveness of both approaches appears to be heavily influenced by dataset size and signal type. For ECG/PPG signals, feature-based methods often show superior performance, as demonstrated by Jahanjoo et al. [79], achieving 95.55% accuracy using SVM with HRV features. In contrast, E2E methods show promise with EEG data that [124] achieved 99.56% accuracy using kNN with PCA-reduced EEG raw data.

Real-world applications present unique challenges for both approaches. Gjoreski et al. [117] showed that while SVM performed well in laboratory conditions (67-71% accuracy), performance varied significantly in real-world settings. The study by Naegelin et al. [95] is particularly insightful, showing that combining physiological signals with contextual features (like mouse and keyboard activity) improved real-world performance, achieving F1 scores of 62.5%. This suggests that hybrid approaches, combining both feature engineering and deep learning techniques, might be most effective for practical applications.

FB models demonstrate several key advantages. Their primary strength lies in robustness and consistency across different datasets, achieving over 90% accuracy with traditional ML models. They allow clear interpretability through feature importance analysis. These models also perform well with limited data and require less computational resources. However, FB models face limitations in requiring domain expertise for feature engineering. They may also miss complex patterns in raw data and need feature redesign for different sensor types, as shown in [112] where model performance dropped significantly when transferring between ECG and PPG signals.

E2E models offer distinct advantages in their ability to automatically learn features from raw data. They excel at capturing complex temporal patterns and show remarkable adaptability to different input data types, evidenced by Ragav et al. [152] reaching 99.86% accuracy across various physiological signals. However, E2E typically require substantial training data, who observed poor CNN performance (40-48%) with smaller datasets [134]. E2E models also demand greater computational resources and often lack interpretability, as noted by Dziezyc et al. [11]. Lastly, Toshnazarov et al. [107] demonstrated their tendency to show performance degradation in real-world settings, with accuracy dropping from 76.4% in laboratory conditions to 64.7-70.1% in practical applications.

Stressors differentiation and generalization: While many reviewed studies incorporated multiple stressors as part of their study protocols, it is important to note that majority of these models were not specifically designed to evaluate the distinct impact of unique stressors on physiological signals. Instead, multiple stressors were grouped together, and these studies did not fine-tune their analyses to distinguish between different types of stressors. Understanding the specific effects of each stressor could be crucial, as different stressors can elicit varied physiological responses, and accurately identifying which stressor is at play could lead to more precise anxiety detection models and better-targeted interventions. For example, physical stressors and psychosocial stressors may induce different patterns of physiological changes [175] that are important for fine-grained anxiety detection and personalization of stress management interventions. By distinguishing between stressors, researchers can better understand how different individuals react to specific stress conditions, leading to more personalized approaches to anxiety management. Meanwhile, it is still essential to recognize the significance of the generalization of anxiety responses across a range of conditions to ensure model robustness to detect anxiety in diverse real-world scenarios, where multiple types of stressors might be present simultaneously.

Real-world challenges and model robustness: A significant challenge in anxiety detection research is translating laboratory findings to practical, real-world applications. Although laboratory-based studies often report high accuracy rates, the ML performances generally decline when applied to real-world settings [23,53,77,96]. The majority of studies focused on distinguishing between baseline and anxious states, which, while effective in controlled lab settings, may not translate well to real-world scenarios where multiple affective states are present [60,153]. This aspect of model robustness in the presence of diverse affective states has not been thoroughly tested, representing a considerable gap in the research for evaluating the testing environment in influencing model selection and performance.

The data quality particularly affects the quality and reliability of in-the-wild data, where factors like physical activities, user compliance to protocol, and proper devices handling [176] can significantly degrade the ML performance. Wearable sensors could vary in their signal quality and are subject to measurement errors, especially during physical activity [177]. Physiological data are prone to contamination by artifacts resulting from movement, external noise, and other non-physiological factors (e.g., fit of wearable devices, ambient temperature and humidity, and device maintenance) [178,179]. Effective methods for filtering these artifacts are crucial for separating signals from noises to ensure data integrity. Addressing this issue could involve two primary strategies: 1) noise reduction by developing methods to remove or minimize noise from the data, and 2) noise-robust detection by developing methodologies that are inherently robust to the presence of noise. Efforts have also been made to discriminate between psychological stressors and physical stressors, for instance, by using accelerometer data to categorize the physical activity state during anxiety detection [77], [119].

In-the-wild data also suffers from imbalanced samples due to the nature of data collection that results in more non-anxious period compared to anxious events, potentially leading to reduced performance [53,60,107,116,153]. Also, the continuous monitoring often resulted in huge dataset with thousands of data points and relies on intensive labeling using self-reports for ground truth, while semi-supervised learning showed promising for circumvent this dependence on the ground truth [153]. Future research should address the lack of in-depth, comprehensive studies in wild settings, consider a wider variety of affective states, and focus on increasing the size and diversity of participant groups to enhance the generalizability and reliability of anxiety detection models.

Sample size and generalizability across population: Research often involves relatively homogenous participant groups, which may not well represent the diverse population, especially for underrepresented minority groups. Many papers emphasized model enhancement, and they relied on data from a limited number of participants, often fewer than 20, which restricts the robustness and generalizability of their findings. For instance, the study by Han et al. [77], which was limited to three participants, yielded an 81% accuracy rate, and two studies by Gjoreski et al. [53,116] were

limited to 5 participants each for the in-the-wild conditions. While these papers claimed to boost the model performance by considering physical activity as context information, the small sample sizes could limit the generalizability to new, unseen data due to overfitting. Most studies in this field predominantly involved young, healthy adults, often from similar cultural and ethnic backgrounds. This homogeneity fails to capture the variability in physiological and psychological responses to stress across different groups. Ages and ethnic factors can influence stress perception and physiological responses, potentially impacting the model's performance in a diverse population [180]. Likewise, gender differences can play a role in stress responses, with research indicating that males and females may exhibit different physiological and psychological reactions to stressors [181]. Consequently, models trained predominantly on one gender may perform less effectively when applied to the other.

Individual variability in stress responses: The variability in how anxiety manifests across different individuals poses a significant challenge. Genetic factors, personal health history, and previous experiences contribute to how one responds to stress [182], [183], making it difficult to establish a generalized, one-size-fits-all model for detection. This variability means that physiological measures can differ widely among individuals due to factors like trait personality and resilience [31], [184]. Additionally, individuals' physiological responses to stressors may change over time as they develop coping mechanisms or due to repeated exposures [185]. Studies need to consider these adaptive responses and potentially track changes over extended periods to understand and predict anxiety reactions accurately. Anxiety response varies not only across individuals but also within an individual over time, depending on health, mood, and external circumstances. This intra-individual variability can complicate the interpretation of physiological data. Future research should focus on developing adaptive models that learn and adjust to individual baselines dynamically. Incorporating longitudinal tracking to continuously update the model based on individual's changing physiological and emotional state could provide more accuracy.

Toward anxiety prediction and intervention: The future of anxiety detection lies in moving beyond mere detection toward predicting anxiety onset and developing intervention strategies. There is a notable gap in methodologies for predicting anxiety onset. However, promising applications in real-time detection, such as virtual reality therapy and HRV biofeedback training [186,187], offer avenues to ameliorate anxiety levels. Leveraging existing anxiety detection frameworks could significantly enhance the evaluation and improvement of anxiety therapies, facilitating coping mechanisms for negative emotions.

5. Conclusion

This review paper examined the current state of machine learning methods in anxiety detection, revealing the significant advancements as well as the ongoing challenges in the field. The extensive review highlights a diverse range of methodologies, from traditional FB models to advanced E2E approaches, each with their own set of strengths and limitations. A key observation was the progression from traditional FB models toward more advanced FB deep learning models, reflecting the research interest toward more sophisticated, data-driven approaches and model architectures. The results have shown high accuracy of these models in controlled laboratory settings, while also underscoring the need for expanded research in real-world environments to assess and enhance their practical applicability. The insights from studies conducted in-the-wild, though limited, have revealed a crucial gap in the research and a pressing need for models that can effectively operate amidst the complexities of real-life scenarios. The review also highlights the importance of understanding the physiological and psychological responses elicited by different types of stressors, including social, mental, physical, emotional, driving and daily-life stressors. While many studies group stressors into broad binary categories, distinguishing between stressor types can enhance the precision of anxiety detection models and support the development of more personalized interventions. As the field progresses, future research should focus on improving the robustness, versatility, and generalizability of anxiety detection models. Emphasizing real-world applicability,

enhancing noise resilience, exploring innovative architectures, and expanding the understanding of affective state variability will be key in advancing the field for mental health promotion.

Author contributions: Conceptualization, A.A., J.L.C., R.B.S., E.T.H.-W., M.H., and M.E.H.; methodology, A.A. and M.H.; data curation, A.A. and M.H.; writing—original draft preparation, A.A., J.L.C., R.B.S., E.T.H.-W., M.H., and M.E.H.; writing—review and editing, A.A., J.L.C., R.B.S., E.T.H.-W., M.H., and M.E.H.; visualization, A.A. and M.H.; supervision, E.T.H.-W., and M.E.H.; project administration, M.E.H.; funding acquisition, J.L.C., R.B.S., E.T.H.-W., and M.E.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Jump ARCHES endowment through the Health Care Engineering Systems Center.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data is publicly available.

Acknowledgement: We thank Dr. Rebecca Smith for her expertise and assistance in the literature review process.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
CASE	Continuously Annotated Signals of Emotion
CNN	Convolutional Neural Network
DT	Decision Tree
E2E	End-to-End
ECG	Electrocardiogram
EDA	Electrodermal Activity
EEG	Electroencephalogram
EMG	Electromyography
FB	Feature Based
FCN	Fully Convolutional Neural Network
GCN	Graph Convolutional Network
HF	High Frequency
HPA	Hypothalamic–Pituitary–Adrenal axis
HRV	Heart Rate Variability
IAPS	International Affective Picture System
kNN	k-Nearest Neighbors
LDA	Linear Discriminant Analysis
LF	Low Frequency
LR	Linear Regression
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multi-Layer Perceptron
PPG	Photoplethysmography
RESP	Respiration
RF	Random Forest
RMSSD	Root Mean Square of Successive Differences
RNN	Recurrent Neural Network
SCWT	Stroop Color and Word Test
STAI	State-Trait Anxiety Inventory
SVM	Support Vector Machine
SWELL-KN	Smart Reasoning for Well-being at Home and at Work – Knowledge Work
TEMP	Temperature

TSST Trier Social Stress Test
WESAD Wearable Stress and Affect Detection

References

- Manderscheid, R.W.; Ryff, C.D.; Freeman, E.J.; McKnight-Eily, L.R.; Dhingra, S.; Strine, T.W. Evolving Definitions of Mental Illness and Wellness. *Prev. Chronic Dis.* **2010**, *7*, A19.
- Salari, N.; Hosseinian-Far, A.; Jalali, R.; Vaisi-Raygani, A.; Rasoulpoor, S.; Mohammadi, M.; Rasoulpoor, S.; Khaledi-Paveh, B. Prevalence of Stress, Anxiety, Depression among the General Population during the COVID-19 Pandemic: A Systematic Review and Meta-Analysis. *Global. Health* **2020**, *16*.
- Canersals, J.; Voltas, N.; Hernández-Martínez, C.; Cosi, S.; Arija, V. Prevalence of DSM-5 Anxiety Disorders, Comorbidity, and Persistence of Symptoms in Spanish Early Adolescents. *Eur. Child Adolesc. Psychiatry* **2019**, *28*, 131–143, doi:10.1007/s00787-018-1207-z.
- Wittchen, H.U.; Jacobi, F.; Rehm, J.; Gustavsson, A.; Svensson, M.; Jönsson, B.; Olesen, J.; Allgulander, C.; Alonso, J.; Faravelli, C.; et al. The Size and Burden of Mental Disorders and Other Disorders of the Brain in Europe 2010. *Eur. Neuropsychopharmacol.* **2011**, *21*, 655–679, doi:10.1016/j.euroneuro.2011.07.018.
- Celano, C.M.; Daunis, D.J.; Lokko, H.N.; Campbell, K.A.; Huffman, J.C. Anxiety Disorders and Cardiovascular Disease. *Curr. Psychiatry Rep.* **2016**, *18*, doi:10.1007/s11920-016-0739-5.
- Segerstrom, S.C.; Miller, G.E. Psychological Stress and the Human Immune System: A Meta-Analytic Study of 30 Years of Inquiry. *Psychol. Bull.* **2004**, *130*, 601–630, doi:10.1037/0033-2909.130.4.601.
- Thomas, K.C.; Ellis, A.R.; Konrad, T.R.; Holzer, C.E.; Morrissey, J.P. County-Level Estimates of Mental Health Professional Shortage in the United States. *Psychiatr. Serv.* **2009**, *60*, 1323–1328, doi:10.1176/PS.2009.60.10.1323.
- Satiani, A.; Niedermier, J.; Satiani, B.; Svendsen, D.P. Projected Workforce of Psychiatrists in the United States: A Population Analysis. *Psychiatr. Serv.* **2018**, *69*, 710–713, doi:10.1176/appi.ps.201700344.
- Althubaiti, A. Information Bias in Health Research: Definition, Pitfalls, and Adjustment Methods. *J. Multidiscip. Healthc.* **2016**, *9*, 211–217, doi:10.2147/JMDH.S104807.
- Julian, L.J. Measures of Anxiety. *Arthritis Care* **2011**, *63*, 1–11, doi:10.1002/acr.20561.Measures.
- Dziezyc, M.; Gjoreski, M.; Kazienko, P.; Saganowski, S.; Gams, M. Can We Ditch Feature Engineering? End-to-End Deep Learning for Affect Recognition from Physiological Sensor Data. *Sensors (Switzerland)* **2020**, *20*, 1–21, doi:10.3390/s20226535.
- Mentis, A.F.A.; Lee, D.; Roussos, P. Applications of Artificial Intelligence–machine Learning for Detection of Stress: A Critical Overview. *Mol. Psychiatry* **2023**, 1–13, doi:10.1038/s41380-023-02047-6.
- Spielberger, C.D. *Theory and Research on Anxiety*; Spielberger, C.D., Ed.; Academic Press Inc.: Oxford, England, 1966; ISBN 9781483258362.
- Daviu, N.; Bruchas, M.R.; Moghaddam, B.; Sandi, C.; Beyeler, A. Neurobiological Links between Stress and Anxiety. *Neurobiol. Stress* **2019**, *11*.
- Spielberger, C.D. Theory and Research on Anxiety. In *Anxiety and Behavior*; Elsevier, 1966; pp. 3–20 ISBN 9781483258362.
- Spielberger, C.D. Notes and Comments Trait-State Anxiety and Motor Behavior. *J. Mot. Behav.* **1971**, *3*, 265–279, doi:10.1080/00222895.1971.10734907.
- Spielberger, C.D.; Gonzalez-Reigosa, F.; Martinez-Urrutia, A.; Natalicio, L.F.S.; Natalicio, D.S. The State-Trait Anxiety Inventory. *Rev. Interam. Psicol. J. Psychol.* **1971**, *5*, 3–4, doi:10.30849/RIP/IJP.V5I3.
- Wiedemann K. Anxiety and Anxiety Disorders, International Encyclopedia of the Social & Behavioral Sciences. **2001**, 560–567.
- Duval, E.R.; Javanbakht, A.; Liberzon, I. Neural Circuits in Anxiety and Stress Disorders: A Focused Review. *Ther. Clin. Risk Manag.* **2015**, *11*, 115–126.
- Ding, Y.; Cao, Y.; Duffy, V.G.; Wang, Y.; Zhang, X. Measurement and Identification of Mental Workload during Simulated Computer Tasks with Multimodal Methods and Machine Learning. *Ergonomics* **2020**, *63*, 896–908, doi:10.1080/00140139.2020.1759699.

21. Mozos, O.M.; Sandulescu, V.; Andrews, S.; Ellis, D.; Bellotto, N.; Dobrescu, R.; Ferrandez, J.M. Stress Detection Using Wearable Physiological and Sociometric Sensors. *Int. J. Neural Syst.* **2017**, *27*, 1650041, doi:10.1142/S0129065716500416.
22. Sandulescu, V.; Dobrescu, R. Wearable System for Stress Monitoring of Firefighters in Special Missions. In Proceedings of the 2015 E-Health and Bioengineering Conference, EHB 2015; IEEE, November 2016; pp. 1–4.
23. Schmidt, P.; Dürichen, R.; Reiss, A.; Van Laerhoven, K.; Plötz, T. Multi-Target Affect Detection in the Wild: An Exploratory Study. In Proceedings of the Proceedings - International Symposium on Wearable Computers, ISWC; 2019; pp. 211–219.
24. Vaz, M.; Summavielle, T.; Sebastião, R.; Ribeiro, R.P. Multimodal Classification of Anxiety Based on Physiological Signals. *Appl. Sci.* **2023**, *13*, doi:10.3390/app13116368.
25. Jiao, Y.; Wang, X.; Liu, C.; Du, G.; Zhao, L.; Dong, H.; Zhao, S.; Liu, Y. Feasibility Study for Detection of Mental Stress and Depression Using Pulse Rate Variability Metrics via Various Durations. *Biomed. Signal Process. Control* **2023**, *79*, doi:10.1016/j.bspc.2022.104145.
26. Bystritsky, A.; Kronemyer, D. Stress and Anxiety: Counterpart Elements of the Stress/Anxiety Complex. *Psychiatr. Clin. North Am.* **2014**, *37*, 489–518, doi:10.1016/j.psc.2014.08.002.
27. Taschereau-Dumouchel, V.; Michel, M.; Lau, H.; Hofmann, S.G.; LeDoux, J.E. Putting the “Mental” Back in “Mental Disorders”: A Perspective from Research on Fear and Anxiety. *Mol. Psychiatry* **2022**, *27*, 1322–1330, doi:10.1038/s41380-021-01395-5.
28. Beck, A.T.; Epstein, N.; Brown, G.; Steer, R.A. An Inventory for Measuring Clinical Anxiety: Psychometric Properties. *J. Consult. Clin. Psychol.* **1988**, *56*, 893–897, doi:10.1037/0022-006X.56.6.893.
29. Demetriou, C.; Ozer, B.U.; Essau, C.A. Self-Report Questionnaires. *Encycl. Clin. Psychol.* **2015**, 1–6, doi:10.1002/9781118625392.WBCEP507.
30. Arikian, S.R.; German, J.M. A Review of the Diagnosis, Pharmacologic Treatment, and Economic Aspects of Anxiety Disorders. *Prim. Care Companion J. Clin. Psychiatry* **2001**, *3*, 110–117, doi:10.4088/pcc.v03n0302.
31. Weinberger, D.A.; Schwartz, G.E.; Davidson, R.J. Low-Anxious, High-Anxious, and Repressive Coping Styles: Psychometric Patterns and Behavioral and Physiological Responses to Stress. *J. Abnorm. Psychol.* **1979**, *88*, 369–380.
32. Kaplan, R.; Saccuzzo, D. *Psychological Testing: Principles, Applications, and Issues*; Cengage Learning: Boston, MA, 1982; ISBN 9781337517065.
33. Ancillon, L.; Elgendi, M.; Menon, C. Machine Learning for Anxiety Detection Using Biosignals: A Review. *Diagnostics* **2022**, *12*, 1794, doi:10.3390/diagnostics12081794.
34. Kim, H.G.; Cheon, E.J.; Bai, D.S.; Lee, Y.H.; Koo, B.H. Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry Investig.* **2018**, *15*, 235–245, doi:10.30773/pi.2017.08.17.
35. Giannakakis, G.; Grigoriadis, D.; Giannakaki, K.; Simantiraki, O.; Roniotis, A.; Tsiknakis, M. Review on Psychological Stress Detection Using Biosignals. *IEEE Trans. Affect. Comput.* **2022**, *13*, 440–460, doi:10.1109/TAFFC.2019.2927337.
36. Merletti, R.; Avenaggiato, M.; Botter, A.; Holobar, A.; Marateb, H.; Vieira, T.M.M. Advances in Surface EMG: Recent Progress in Detection and Processing Techniques. *Crit. Rev. Biomed. Eng.* **2010**, *38*, 305–345, doi:10.1615/CritRevBiomedEng.v38.i4.10.
37. Hidaka, O.; Yanagi, M.; Takada, K. Mental Stress-Induced Physiological Changes in the Human Masseter Muscle. *J. Dent. Res.* **2004**, *83*, 227–231, doi:10.1177/154405910408300308.
38. Wijsman, J.; Grundlehner, B.; Penders, J.; Hermens, H. Trapezius Muscle EMG as Predictor of Mental Stress. *Trans. Embed. Comput. Syst.* **2013**, *12*, doi:10.1145/2485984.2485987.
39. Tsai, C.M.; Chou, S.L.; Gale, E.N.; McCall, W.D. Human Masticatory Muscle Activity and Jaw Position under Experimental Stress. *J. Oral Rehabil.* **2002**, *29*, 44–51, doi:10.1046/j.1365-2842.2002.00810.x.
40. Turpin, G.; Grandfield, T. Electrodermal Activity. *Encycl. Stress* **2007**, 899–902, doi:10.1016/B978-012373947-6.00139-2.
41. Ren, P.; Barreto, A.; Huang, J.; Gao, Y.; Ortega, F.R.; Adjouadi, M. Off-Line and on-Line Stress Detection through Processing of the Pupil Diameter Signal. *Ann. Biomed. Eng.* **2014**, *42*, 162–176, doi:10.1007/s10439-013-0880-9.

42. Palanisamy, K.; Murugappan, M.; Yaacob, S. Descriptive Analysis of Skin Temperature Variability of Sympathetic Nervous System Activity in Stress. *J. Phys. Ther. Sci.* **2012**, *24*, 1341–1344, doi:10.1589/jpts.24.1341.
43. Long, N.; Lei, Y.; Peng, L.; Xu, P.; Mao, P. A Scoping Review on Monitoring Mental Health Using Smart Wearable Devices. *Math. Biosci. Eng.* **2022**, *19*, 7899–7919, doi:10.3934/mbe.2022369.
44. Gedam, S.; Paul, S. A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques. *IEEE Access* **2021**, *9*, 84045–84066, doi:10.1109/ACCESS.2021.3085502.
45. Vizer, L.M.; Zhou, L.; Sears, A. Automated Stress Detection Using Keystroke and Linguistic Features: An Exploratory Study. *Int. J. Hum. Comput. Stud.* **2009**, *67*, 870–886, doi:10.1016/j.IJHCS.2009.07.005.
46. Scarpina, F.; Tagini, S. The Stroop Color and Word Test. *Front. Psychol.* **2017**, *8*, 1–8, doi:10.3389/fpsyg.2017.00557.
47. Tulen, J.H.M.; Moleman, P.; van Steenis, H.G.; Boomsma, F. Characterization of Stress Reactions to the Stroop Color Word Test. *Pharmacol. Biochem. Behav.* **1989**, *32*, 9–15, doi:10.1016/0091-3057(89)90204-9.
48. Lovallo, W. The Cold Pressor Test and Autonomic Function: A Review and Integration. *Psychophysiology* **1975**, *12*, 268–282, doi:10.1111/j.1469-8986.1975.tb01289.x.
49. Lang, P.J.; Bradley, M.M.; Cuthbert, B.N. International Affective Picture System (IAPS): Technical Manual and Affective Ratings. *NIMH Cent. Study Emot. Atten.* **1997**, 39–58.
50. Magaña, V.C.; Pañeda, X.G.; Garcia, R.; Paiva, S.; Pozueco, L. Beside and behind the Wheel: Factors That Influence Driving Stress and Driving Behavior. *Sustain.* **2021**, *13*, 18–21, doi:10.3390/su13094775.
51. Chung, W.-Y.; Chong, T.-W.; Lee, B.-G. Methods to Detect and Reduce Driver Stress: A Review. *Int. J. Automot. Technol.* **2019**, *20*, 1051–1063, doi:10.1007/s12239-019-0099-3.
52. Pasha, S.T.; Halder, N.; Badrul, T.; Setu, J.H.; Islam, A.; Alam, M.Z. Physiological Signal Data-Driven Workplace Stress Detection Among Healthcare Professionals Using BiLSTM-AM and Ensemble Stacking Models. *Adv. Sci. Eng. Technol. Int. Conf. ASET* **2024**, 1–10, doi:10.1109/ASET60340.2024.10708646.
53. Gjoreski, M.; Luštrek, M.; Gams, M.; Gjoreski, H. Monitoring Stress with a Wrist Device Using Context. *J. Biomed. Inform.* **2017**, *73*, 159–170, doi:10.1016/j.jbi.2017.08.006.
54. Gedam, S.; Paul, S. A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques. *IEEE Access* **2021**, *9*, 84045–84066, doi:10.1109/ACCESS.2021.3085502.
55. Chandra, V.; Sethia, D. Machine Learning-Based Stress Classification System Using Wearable Sensor Devices. *IAES Int. J. Artif. Intell.* **2024**, *13*, 337–347, doi:10.11591/ijai.v13.i1.pp337-347.
56. Shrestha, A.; Mahmood, A. Review of Deep Learning Algorithms and Architectures. *IEEE Access* **2019**, *7*, 53040–53065, doi:10.1109/ACCESS.2019.2912200.
57. Ahmad, Z.; Rabbani, S.; Zafar, M.R.; Ishaque, S.; Krishnan, S.; Khan, N. Multilevel Stress Assessment from ECG in a Virtual Reality Environment Using Multimodal Fusion. *IEEE Sens. J.* **2023**, *23*, 29559–29570, doi:10.1109/JSEN.2023.3323290.
58. Agarwal, S.; Sharma, S.; Faisal, K.N.; Sharma, R.R. Induced Stress Identification Using EEG: A Framework Based on MVMD and Machine Learning. *2024 IEEE Int. Students' Conf. Electr. Electron. Comput. Sci. SCEECS 2024* **2024**, 1–5, doi:10.1109/SCEECS61402.2024.10481868.
59. Akella, A.; Singh, A.K.; Leong, D.; Lal, S.; Newton, P.; Clifton-Bligh, R.; McLachlan, C.S.; Gustin, S.M.; Maharaj, S.; Lees, T.; et al. Classifying Multi-Level Stress Responses from Brain Cortical EEG in Nurses and Non-Health Professionals Using Machine Learning Auto Encoder. *IEEE J. Transl. Eng. Heal. Med.* **2021**, *9*, 1–9, doi:10.1109/JTEHM.2021.3077760.
60. Abd Al-Alim, M.; Mubarak, R.; M. Salem, N.; Sadek, I. A Machine-Learning Approach for Stress Detection Using Wearable Sensors in Free-Living Environments. *Comput. Biol. Med.* **2024**, *179*, 108918, doi:10.1016/j.combiomed.2024.108918.
61. AlShorman, O.; Masadeh, M.; Heyat, M.B. Bin; Akhtar, F.; Almahasneh, H.; Ashraf, G.M.; Alexiou, A. Frontal Lobe Real-Time EEG Analysis Using Machine Learning Techniques for Mental Stress Detection. *J. Integr. Neurosci.* **2022**, *21*, 020, doi:10.31083/jjin2101020.
62. Arya, L.; Chowdhary, H.; Agrawal, I.; Sreedevi, I. Towards Accurate Stress Classification: Combining Advanced Feature Selection and Deep Learning. In Proceedings of the 2023 3rd IEEE International

- Conference on Software Engineering and Artificial Intelligence, SEAI 2023; Institute of Electrical and Electronics Engineers Inc., 2023; pp. 47–52.
63. Badr, Y.; Al-Shargie, F.; Tariq, U.; Babiloni, F.; Al-Mughairbi, F.; Al-Nashash, H. Classification of Mental Stress Using Dry EEG Electrodes and Machine Learning. In Proceedings of the 2023 Advances in Science and Engineering Technology International Conferences, ASET 2023; Institute of Electrical and Electronics Engineers Inc., 2023.
 64. Badr, Y.; Al-Shargie, F.; Tariq, U.; Babiloni, F.; Al-Mughairbi, F.; Al-Nashash, H. Mental Stress Detection and Mitigation Using Machine Learning and Binaural Beat Stimulation. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2023**, 1–5, doi:10.1109/EMBC40787.2023.10340673.
 65. Bahameish, M.; Stockman, T.; Requena Carrión, J. Strategies for Reliable Stress Recognition: A Machine Learning Approach Using Heart Rate Variability Features. *Sensors* **2024**, *24*, 1–21, doi:10.3390/s24103210.
 66. Beh, W.-K.; Wu, Y.-H.; Wu, A.-Y. Robust PPG-Based Mental Workload Assessment System Using Wearable Devices. *IEEE J. Biomed. Heal. Informatics* **2023**, *27*, 2323–2333, doi:10.1109/JBHI.2021.3138639.
 67. Bobade, P.; Vani, M. Stress Detection with Machine Learning and Deep Learning Using Multimodal Physiological Data. *Proc. 2nd Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2020* **2020**, 51–57, doi:10.1109/ICIRCA48905.2020.9183244.
 68. Campanella, S.; Altaleb, A.; Belli, A.; Pierleoni, P.; Palma, L. A Method for Stress Detection Using Empatica E4 Bracelet and Machine-Learning Techniques. *Sensors* **2023**, *Vol. 23, Page 3565* **2023**, *23*, 3565, doi:10.3390/S23073565.
 69. Cui, Z.; Ma, Y.; Ma, M.; Huang, R.; Du, B. Towards a Lightweight Stress Prediction Model: A Study on Dimension Reduction and Individual Models in HRV Analysis. *Proc. Int. Conf. Parallel Distrib. Syst. - ICPADS* **2023**, 1709–1716, doi:10.1109/ICPADS60453.2023.00238.
 70. Cruz, A.P.; Pradeep, A.; Sivasankar, K.R.; Krishnaveni, K.. A Decision Tree Optimised SVM Model for Stress Detection Using Biosignals. In Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP); IEEE, July 2020; pp. 0841–0845.
 71. Dalmeida, K.M.; Masala, G.L. Hrv Features as Viable Physiological Markers for Stress Detection Using Wearable Devices. *Sensors* **2021**, *21*, 2873, doi:10.3390/s21082873.
 72. Delmastro, F.; Martino, F. Di; Dolciotti, C. Cognitive Training and Stress Detection in MCI Frail Older People through Wearable Sensors and Machine Learning. *IEEE Access* **2020**, *8*, 65573–65590, doi:10.1109/ACCESS.2020.2985301.
 73. Erkus, E.C.; Purutcuoglu, V.; Ari, F.; Gokcay, D. Comparison of Several Machine Learning Classifiers for Arousal Classification: A Preliminary Study. In Proceedings of the 2020 Medical Technologies Congress (TIPTEKNO); IEEE, November 19 2020; pp. 1–7.
 74. Fernandez, J.; Martínez, R.; Innocenti, B.; López, B. Contribution of EEG Signals for Students' Stress Detection. *IEEE Trans. Affect. Comput.* **2024**, 1–12, doi:10.1109/TAFFC.2024.3503995.
 75. Giannakakis, G.; Pediaditis, M.; Manousos, D.; Kazantzaki, E.; Chiarugi, F.; Simos, P.G.; Marias, K.; Tsiknakis, M. Stress and Anxiety Detection Using Facial Cues from Videos. *Biomed. Signal Process. Control* **2017**, *31*, 89–101, doi:10.1016/j.bspc.2016.06.020.
 76. Hag, A.; Al-Shargie, F.; Handayani, D.; Asadi, H. Mental Stress Classification Based on Selected Electroencephalography Channels Using Correlation Coefficient of Hjorth Parameters. *Brain Sci.* **2023**, *13*, doi:10.3390/brainsci13091340.
 77. Han, H.J.; Labbaf, S.; Borelli, J.L.; Dutt, N.; Rahmani, A.M. Objective Stress Monitoring Based on Wearable Sensors in Everyday Settings. *J. Med. Eng. Technol.* **2020**, *44*, 177–189, doi:10.1080/03091902.2020.1759707.
 78. Henry, J.; Lloyd, H.; Turner, M.; Kendrick, C. On the Robustness of Machine Learning Models for Stress and Anxiety Recognition from Heart Activity Signals. *IEEE Sens. J.* **2023**, *23*, 14428–14436, doi:10.1109/JSEN.2023.3276413.
 79. Jahanjoo, A.; Taherinejad, N.; Aminifar, A. High-Accuracy Stress Detection Using Wrist-Worn PPG Sensors. *Proc. - IEEE Int. Symp. Circuits Syst.* **2024**, 1–5, doi:10.1109/ISCAS58744.2024.10558012.
 80. Abdul Kader, L.; Al-Shargie, F.; Tariq, U.; Al-Nashash, H. One-Channel Wearable Mental Stress State Monitoring System. *Sensors* **2024**, *24*, doi:10.3390/s24165373.

81. Kalra, P.; Sharma, V. Mental Stress Assessment Using PPG Signal a Deep Neural Network Approach. *IETE J. Res.* **2023**, *69*, 879–885, doi:10.1080/03772063.2020.1844068.
82. Kim, N.; Seo, W.; Kim, S.; Park, S.M. Electrogastragram: Demonstrating Feasibility in Mental Stress Assessment Using Sensor Fusion. *IEEE Sens. J.* **2021**, *21*, 14503–14514, doi:10.1109/JSEN.2020.3026717.
83. Kim, H.; Kim, M.; Park, K.; Kim, J.; Yoon, D.; Kim, W.; Park, C.H. Machine Learning-Based Classification Analysis of Knowledge Worker Mental Stress. *Front. Public Heal.* **2023**, *11*, 1–6, doi:10.3389/fpubh.2023.1302794.
84. Kim, N.; Lee, S.; Kim, J.; Choi, S.Y.; Park, S.M. Shuffled ECA-Net for Stress Detection from Multimodal Wearable Sensor Data. *Comput. Biol. Med.* **2024**, *183*, 109217, doi:10.1016/j.compbiomed.2024.109217.
85. Konar, D.; De, S.; Mukherjee, P.; Roy, A.H. A Novel Human Stress Level Detection Technique Using EEG. *2023 Int. Conf. Network, Multimed. Inf. Technol. NMITCON 2023* **2023**, 1–6, doi:10.1109/NMITCON58196.2023.10276159.
86. Kurniawan, H.; Maslov, A. V.; Pechenizkiy, M. Stress Detection from Speech and Galvanic Skin Response Signals. In Proceedings of the Proceedings of CBMS 2013 - 26th IEEE International Symposium on Computer-Based Medical Systems; IEEE, June 2013; pp. 209–214.
87. Lingelbach, K.; Bui, M.; Diederichs, F.; Vukelic, M. Exploring Conventional, Automated and Deep Machine Learning for Electrodermal Activity-Based Drivers' Stress Recognition. In Proceedings of the 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC); IEEE, October 17 2021; pp. 1339–1344.
88. Liu, Y.; Li, H.; Wang, J.; Zhang, H.; Zheng, X. Psychological Stress Detection Based on Heart Rate Variability. In Proceedings of the International Conference on Electronic Information Engineering and Computer Science (EIECS 2022); Yue, Y., Ed.; SPIE, April 21 2023; Vol. 12602, p. 150.
89. Mamdouh, M.; Mahmoud, R.; Attallah, O.; Al-Kabbany, A. Stress Detection in the Wild: On the Impact of Cross-Training on Mental State Detection. **2023**, 150–158, doi:10.1109/NRSC58893.2023.10153050.
90. Marthinsen, A.J.; Galtung, I.T.; Cheema, A.; Sletten, C.M.; Andreassen, I.M.; Sletta, Ø.; Soler, A.; Molinas, M. Psychological Stress Detection with Optimally Selected EEG Channel Using Machine Learning Techniques. *CEUR Workshop Proc.* **2023**, *3576*, 53–68.
91. Mevlevioğlu, D.; Tabirca, S.; Murphy, D. Real-Time Classification of Anxiety in Virtual Reality Therapy Using Biosensors and a Convolutional Neural Network. *Biosensors* **2024**, *14*, doi:10.3390/bios14030131.
92. Meteier, Q.; Capallera, M.; Ruffieux, S.; Angelini, L.; Abou Khaled, O.; Mugellini, E.; Widmer, M.; Sonderegger, A. Classification of Drivers' Workload Using Physiological Signals in Conditional Automation. *Front. Psychol.* **2021**, *12*, doi:10.3389/fpsyg.2021.596038.
93. Meteier, Q.; De Salis, E.; Capallera, M.; Widmer, M.; Angelini, L.; Abou Khaled, O.; Sonderegger, A.; Mugellini, E. Relevant Physiological Indicators for Assessing Workload in Conditionally Automated Driving, Through Three-Class Classification and Regression. *Front. Comput. Sci.* **2022**, *3*, doi:10.3389/fcomp.2021.775282.
94. Meteier, Q.; Capallera, M.; de Salis, E.; Angelini, L.; Carrino, S.; Widmer, M.; Abou Khaled, O.; Mugellini, E.; Sonderegger, A. A Dataset on the Physiological State and Behavior of Drivers in Conditionally Automated Driving. *Data Br.* **2023**, *47*, 109027, doi:10.1016/J.DIB.2023.109027.
95. Naegelin, M.; Weibel, R.P.; Kerr, J.I.; Schinazi, V.R.; La Marca, R.; von Wangenheim, F.; Hoelscher, C.; Ferrario, A. An Interpretable Machine Learning Approach to Multimodal Stress Detection in a Simulated Office Environment. *J. Biomed. Inform.* **2023**, *139*, 104299, doi:10.1016/J.JBI.2023.104299.
96. Plarre, K.; Raij, A.; Hossain, S.M.; Ali, A.A.; Nakajima, M.; Al'Absi, M.; Ertin, E.; Kamarck, T.; Kumar, S.; Scott, M.; et al. Continuous Inference of Psychological Stress from Sensory Measurements Collected in the Natural Environment. In Proceedings of the Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN'11; 2011; pp. 97–108.
97. Rajendran, V.G.; Jayalalitha, S.; Adalarasu, K.; Thalamalaichamy, M. Analysis and Classification of EEG Data When Playing Video Games and Relax Using EEG Biomarkers. *AIP Conf. Proc.* **2024**, *3180*, doi:10.1063/5.0224678.
98. Sandulescu, V.; Andrews, S.; Ellis, D.; Bellotto, N.; Mozos, O.M. Stress Detection Using Wearable Physiological Sensors. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **2015**, *9107*, 526–532.

99. Setz, C.; Arnrich, B.; Schumm, J.; Marca, R. La; Tr, G.; Ehlert, U. Discriminating Stress From Cognitive Load Using a Wearable EDA Device. *Technology* **2010**, *14*, 410–417.
100. Sharisha Shanbhog, M.; Medikonda, J.; Rai, S. Unsupervised Machine Learning Approach for Stress Level Classification Using Electrodermal Activity Signals. *Proc. CONECCT 2024 - 10th IEEE Int. Conf. Electron. Comput. Commun. Technol.* **2024**, 1–6, doi:10.1109/CONECCT62155.2024.10677181.
101. Shaposhnyk, O.; Yanushkevich, S.; Babenko, V.; Chernykh, M.; Nastenka, I. Inferring Cognitive Load Level from Physiological and Personality Traits. In Proceedings of the 2023 International Conference on Information and Digital Technologies (IDT); IEEE, June 22 2023; pp. 233–242.
102. Siam, A.I.; Gamel, S.A.; Talaat, F.M. Automatic Stress Detection in Car Drivers Based on Non-Invasive Physiological Signals Using Machine Learning Techniques. *Neural Comput. Appl.* **2023**, *35*, 12891–12904, doi:10.1007/s00521-023-08428-w.
103. Silva, E.; Aguiar, J.; Reis, L.P.; Sá, J.O. e.; Gonçalves, J.; Carvalho, V. Stress among Portuguese Medical Students: The EuStress Solution. *J. Med. Syst.* **2020**, *44*, doi:10.1007/s10916-019-1520-1.
104. Souchet, A.D.; Lamarana Diallo, M.; Lourdeaux, D. Acute Stress Classification with a Stroop Task and In-Office Biophilic Relaxation in Virtual Reality Based on Behavioral and Physiological Data. *Proc. - 2023 IEEE Int. Symp. Mix. Augment. Real. Adjunct, ISMAR-Adjunct 2023* **2023**, 537–542, doi:10.1109/ISMAR-Adjunct60411.2023.00115.
105. Subhani, A.R.; Mumtaz, W.; Saad, M.N.B.M.; Kamel, N.; Malik, A.S. Machine Learning Framework for the Detection of Mental Stress at Multiple Levels. *IEEE Access* **2017**, *5*, 13545–13556, doi:10.1109/ACCESS.2017.2723622.
106. Swapnil, S.S.; Nuhi-Alamin, M.; Rahman, K.M.; Sarkar, A.K.; Siam, M.Z.H. An Ensemble Approach to Classify Mental Stress Using EEG Based Time-Frequency and Non-Linear Features. *2024 3rd Int. Conf. Adv. Electr. Electron. Eng. ICAEEE 2024* **2024**, 1–6, doi:10.1109/ICAEEE62219.2024.10561805.
107. Toshnazarov, K.; Lee, U.; Kim, B.H.; Mishra, V.; Najarro, L.A.C.; Noh, Y. SOSW: Stress Sensing with Off-the-Shelf Smartwatches in the Wild. *IEEE Internet Things J.* **2024**, *11*, 21527–21545, doi:10.1109/JIOT.2024.3375299.
108. Troyee, T.G.; Chowdhury, M.H.; Khondakar, M.F.K.; Hasan, M.; Hossain, M.A.; Hossain, Q.D.; Ali Akber Dewan, M. Stress Detection and Audio-Visual Stimuli Classification from Electroencephalogram. *IEEE Access* **2024**, *12*, 145417–145427, doi:10.1109/ACCESS.2024.3471590.
109. Troyee, T.G.; Karim Khondakar, M.F.; Hasan, M.; Chowdhury, M.H. A Comparative Analysis of Different Preprocessing Pipelines for EEG-Based Mental Stress Detection. *Proc. - 6th Int. Conf. Electr. Eng. Inf. Commun. Technol. ICEEICT 2024* **2024**, 370–375, doi:10.1109/ICEEICT62016.2024.10534380.
110. Xing, M.; Fitzgerald, J.M.; Klumpp, H. Classification of Social Anxiety Disorder With Support Vector Machine Analysis Using Neural Correlates of Social Signals of Threat. *Front. Psychiatry* **2020**, *11*, doi:10.3389/fpsy.2020.00144.
111. Zhu, L.; Spachos, P.; Ng, P.C.; Yu, Y.; Wang, Y.; Plataniotis, K.; Hatzinakos, D. Stress Detection Through Wrist-Based Electrodermal Activity Monitoring and Machine Learning. *IEEE J. Biomed. Heal. Informatics* **2023**, *27*, 2155–2165, doi:10.1109/JBHI.2023.3239305.
112. Benchekroun, M.; Chevallier, B.; Beouiss, H.; Istrate, D.; Zalc, V.; Khalil, M.; Lenne, D. Comparison of Stress Detection through ECG and PPG Signals Using a Random Forest-Based Algorithm. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2022**, 2022-July, 3150–3153, doi:10.1109/EMBC48229.2022.9870984.
113. Chauhan, A.R.; Akhil; Kumar, S. Analysing Effectiveness of Different Physiological Biomarkers in Detecting Stress. *Proc. - 2023 IEEE World Conf. Appl. Intell. Comput. AIC 2023* **2023**, 71–75, doi:10.1109/AIC57670.2023.10263957.
114. Dahal, K.; Bogue-Jimenez, B.; Doblaz, A. Global Stress Detection Framework Combining a Reduced Set of HRV Features and Random Forest Model. *Sensors* **2023**, *23*, 5220, doi:10.3390/S23115220/S1.
115. Gazi, A.H.; Lis, P.; Mohseni, A.; Ompi, C.; Giuste, F.O.; Shi, W.; Inan, O.T.; Wang, M.D. Respiratory Markers Significantly Enhance Anxiety Detection Using Multimodal Physiological Sensing. In Proceedings of the BHI 2021 - 2021 IEEE EMBS International Conference on Biomedical and Health Informatics, Proceedings; 2021; pp. 1–4.

116. Gjoreski, M.; Gjoreski, H.; Luštrek, M.; Gams, M. Continuous Stress Detection Using a Wrist Device - in Laboratory and Real Life. In Proceedings of the UbiComp 2016 Adjunct - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2016; pp. 1185–1193.
117. Iyer, G.G.; Udhayakumar, R.; Gopakumar, S.; Karmakar, C. Optimizing Temporal Segmentation of Multi-Modal Non-EEG Signals for Human Stress Analysis. *6th IEEE Int. Conf. Artif. Intell. Eng. Technol. IICAIET 2024* **2024**, 494–499, doi:10.1109/IICAIET62352.2024.10730226.
118. Morshed, M. Bin; Rahman, M.M.; Nathan, V.; Zhu, L.; Bae, J.; Rosa, C.; Mendes, W.B.; Kuang, J.; Gao, A. Core Body Temperature and Its Role in Detecting Acute Stress: A Feasibility Study. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* **2024**, 1606–1610, doi:10.1109/ICASSP48485.2024.10447599.
119. Pinge, A.; Bandyopadhyay, S.; Ghosh, S.; Sen, S. A Comparative Study between ECG-Based and PPG-Based Heart Rate Monitors for Stress Detection. *2022 14th Int. Conf. Commun. Syst. NETWORKS, COMSNETS 2022* **2022**, 84–89, doi:10.1109/COMSNETS53615.2022.9668342.
120. Quadir, M.A.; Bhardwaj, S.; Verma, N.; Sivaraman, A.K.; Tee, K.F. IoT-Based Mental Health Monitoring System Using Machine Learning Stress Prediction Algorithm in Real-Time Application. *Lect. Notes Electr. Eng.* **2023**, 1021 LNEE, 249–263, doi:10.1007/978-981-99-1051-9_16.
121. Rashid, N.; Mortlock, T.; Faruque, M.A. AI Stress Detection Using Context-Aware Sensor Fusion from Wearable Devices. *IEEE Internet Things J.* **2023**, doi:10.1109/JIOT.2023.3265768.
122. Schmidt, P.; Reiss, A.; Duerichen, R.; Van Laerhoven, K. Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection. *ICMI 2018 - Proc. 2018 Int. Conf. Multimodal Interact.* **2018**, 400–408, doi:10.1145/3242969.3242985.
123. Karthikeyan, P.; Murugappan, M.; Yaacob, S. EMG Signal Based Human Stress Level Classification Using Wavelet Packet Transform. In *Communications in Computer and Information Science*; 2012; Vol. 330 CCIS, pp. 236–243 ISBN 9783642351969.
124. Mazlan, M.R. Bin; Sukor, A.S.B.A.; Adom, A.H. Bin; Jamaluddin, R.B.; Awang, S.A.B. Investigation of Different Classifiers for Stress Level Classification Using PCA-Based Machine Learning Method. *2023 19th IEEE Int. Colloq. Signal Process. Its Appl. CSPA 2023 - Conf. Proc.* **2023**, 168–173, doi:10.1109/CSPA57446.2023.10087367.
125. Wijsman, J.; Grundlehner, B.; Liu, H.; Hermens, H.; Penders, J. Towards Mental Stress Detection Using Wearable Physiological Sensors. In Proceedings of the Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS; IEEE, August 2011; pp. 1798–1801.
126. Kader, L.A.; Yahya, F.; Tariq, U.; Al-Nashash, H. Mental Stress Assessment Using Low in Cost Single Channel EEG System. In Proceedings of the 2023 Advances in Science and Engineering Technology International Conferences, ASET 2023; Institute of Electrical and Electronics Engineers Inc., 2023.
127. Jain, A.; Kumar, R. Machine Learning Based Anxiety Detection Using Physiological Signals and Context Features. *Proc. - 2nd Int. Conf. Adv. Comput. Comput. Technol. InCACCT 2024* **2024**, 116–121, doi:10.1109/InCACCT61598.2024.10551003.
128. Sim, D.Y.Y.; Chong, C.K. Effects of Dimension Reduction Methods on Boosting Algorithms for Better Prediction Accuracies on Classifications of Stress EEGs. *2023 6th Int. Conf. Electron. Electr. Eng. Technol.* **2024**, 49–54, doi:10.1109/eeet61723.2023.00031.
129. Choi, J.; Ahmed, B.; Gutierrez-Osuna, R. Development and Evaluation of an Ambulatory Stress Monitor Based on Wearable Sensors. *IEEE Trans. Inf. Technol. Biomed.* **2012**, 16, 279–286, doi:10.1109/TITB.2011.2169804.
130. Mozafari, M.; Goubran, R.; Green, J.R. A Fusion Model for Cross-Subject Stress Level Detection Based on Transfer Learning. In Proceedings of the 2021 IEEE Sensors Applications Symposium, SAS 2021 - Proceedings; 2021; pp. 1–6.
131. Masrur, N.; Halder, N.; Rashid, S.; Setu, J.H.; Islam, A.; Ahmed, T. Performance Analysis of Ensemble and DNN Models for Decoding Mental Stress Utilizing ECG-Based Wearable Data Fusion. *2024 IEEE Int. Black Sea Conf. Commun. Networking, BlackSeaCom 2024* **2024**, 276–279, doi:10.1109/BlackSeaCom61746.2024.10646297.

132. Adarsh, V.; Gangadharan, G.R. Mental Stress Detection from Ultra-Short Heart Rate Variability Using Explainable Graph Convolutional Network with Network Pruning and Quantisation. *Mach. Learn.* **2024**, *113*, 5467–5494, doi:10.1007/s10994-023-06504-9.
133. Al-Shargie, F.; Badr, Y.; Tariq, U.; Babiloni, F.; Al-Mughairbi, F.; Al-Nashash, H. Classification of Mental Stress Levels Using EEG Connectivity and Convolutional Neural Networks. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2023**, 1–5, doi:10.1109/EMBC40787.2023.10340398.
134. Appriou, A.; Cichocki, A.; Lotte, F. Modern Machine-Learning Algorithms: For Classifying Cognitive and Affective States From Electroencephalography Signals. *IEEE Syst. Man, Cybern. Mag.* **2020**, *6*, 29–38, doi:10.1109/MSMC.2020.2968638.
135. Shirley Benita, D.; Shamila Ebenezer, A.; Susmitha, L.; Subathra, M.S.P.; Jeba Priya, S. Stress Detection Using CNN on the WESAD Dataset. *ESIC 2024 - 4th Int. Conf. Emerg. Syst. Intell. Comput. Proc.* **2024**, 308–313, doi:10.1109/ESIC60604.2024.10481604.
136. Chatterjee, D.; Dutta, S.; Shaikh, R.; Saha, S.K. A Lightweight Deep Neural Network for Detection of Mental States from Physiological Signals. *Innov. Syst. Softw. Eng.* **2022**, 1–8, doi:10.1007/s11334-022-00470-6.
137. Mortensen, J.A.; Mollov, M.E.; Chatterjee, A.; Ghose, D.; Li, F.Y. Multi-Class Stress Detection Through Heart Rate Variability: A Deep Neural Network Based Study. *IEEE Access* **2023**, *11*, 57470–57480, doi:10.1109/ACCESS.2023.3274478.
138. Zontone, P.; Affanni, A.; Piras, A.; Rinaldo, R. Convolutional Neural Networks Using Scalograms for Stress Recognition in Drivers. *Eur. Signal Process. Conf.* **2023**, 1185–1189, doi:10.23919/EUSIPCO58844.2023.10290079.
139. Dhaouadi, S.; Ben Khelifa, M.M. A Multimodal Physiological-Based Stress Recognition: Deep Learning Models' Evaluation in Gamers' Monitoring Application. In Proceedings of the 2020 International Conference on Advanced Technologies for Signal and Image Processing, ATSIP 2020; IEEE, September 2020; pp. 1–6.
140. Praveenkumar, S.; Karthick, T. Automatic Stress Recognition System with Deep Learning Using Multimodal Psychological Data. In Proceedings of the Proceedings of the 2022 International Conference on Electronic Systems and Intelligent Computing, ICESIC 2022; IEEE, April 22 2022; pp. 122–127.
141. Uddin, J. An Autoencoder Based Emotional Stress State Detection Approach Using Electroencephalography Signals. *J. Inf. Syst. Telecommun.* **2023**, *11*, 24–30, doi:10.52547/jist.32267.11.41.24.
142. Eisenbarth, H.; Chang, L.J.; Wager, T.D. Multivariate Brain Prediction of Heart Rate and Skin Conductance Responses to Social Threat. *J. Neurosci.* **2016**, *36*, 11987–11998, doi:10.1523/JNEUROSCI.3672-15.2016.
143. Onim, M.S.H.; Thapliyal, H. Predicting Stress in Older Adults with RNN and LSTM from Time Series Sensor Data and Cortisol. *Proc. IEEE Comput. Soc. Annu. Symp. VLSI, ISVLSI* **2024**, 300–306, doi:10.1109/ISVLSI61997.2024.00062.
144. Rashmi, C.R.; Shantala, C.P. Cognitive Stress Recognition During Mathematical Task and EEG Changes Following Audio-Visual Stimuli for Relaxation. *Int. Conf. Sustain. Commun. Networks Appl. ICSCNA 2023 - Proc.* **2023**, 612–617, doi:10.1109/ICSCNA58489.2023.10370715.
145. Tigranyan, S.; Martirosyan, A. Breaking Barriers in Stress Detection: An Inter-Subject Approach Using ECG Signals. *2024 IEEE 48th Annu. Comput. Software, Appl. Conf.* **2024**, 1850–1855, doi:10.1109/compsac61105.2024.00293.
146. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput.* **2019**, *31*, 1235–1270, doi:10.1162/neco_a_01199.
147. Amin, M.; Ullah, K.; Asif, M.; Shah, H.; Mehmood, A.; Khan, M.A. Real-World Driver Stress Recognition and Diagnosis Based on Multimodal Deep Learning and Fuzzy EDAS Approaches. *Diagnostics* **2023**, *13*, 1897, doi:10.3390/diagnostics13111897.
148. Barki, H.; Chung, W.Y. Mental Stress Detection Using a Wearable In-Ear Plethysmography. *Biosensors* **2023**, *13*, 397, doi:10.3390/bios13030397.
149. Fan, T.; Qiu, S.; Wang, Z.; Zhao, H.; Jiang, J.; Wang, Y.; Xu, J.; Sun, T.; Jiang, N. A New Deep Convolutional Neural Network Incorporating Attentional Mechanisms for ECG Emotion Recognition. *Comput. Biol. Med.* **2023**, *159*, 106938, doi:10.1016/J.COMPBIOMED.2023.106938.

150. Sah, R.K.; Cleveland, M.J.; Habibi, A.; Ghasemzadeh, H. Stressalyzer: Convolutional Neural Network Framework for Personalized Stress Classification. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2022**, *2022-July*, 4658–4663, doi:10.1109/EMBC48229.2022.9871842.
151. Huynh, L.; Nguyen, T.; Nguyen, T.; Pirttikangas, S.; Siirtola, P. StressNAS: Affect State and Stress Detection Using Neural Architecture Search. In Proceedings of the Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers; ACM: New York, NY, USA, September 21 2021; pp. 121–125.
152. Ragav, A.; Krishna, N.H.; Narayanan, N.; Thelly, K.; Vijayaraghavan, V. Scalable Deep Learning for Stress and Affect Detection on Resource-Constrained Devices. In Proceedings of the Proceedings - 18th IEEE International Conference on Machine Learning and Applications, ICMLA 2019; 2019; pp. 1585–1592.
153. Başaran, O.T.; Can, Y.S.; André, E.; Ersoy, C. Relieving the Burden of Intensive Labeling for Stress Monitoring in the Wild by Using Semi-Supervised Learning. *Front. Psychol.* **2023**, *14*, doi:10.3389/fpsyg.2023.1293513.
154. Halder, N.; Setu, J.H.; Rafid, L.; Islam, A.; Amin, M.A. Smartwatch-Based Human Stress Diagnosis Utilizing Physiological Signals and LSTM-Driven Machine Intelligence. *Adv. Sci. Eng. Technol. Int. Conf. ASET* **2024**, 1–8, doi:10.1109/ASET60340.2024.10708730.
155. Koldijk, S.; Sappelli, M.; Verberne, S.; Neerinx, M.A.; Kraaij, W. The Swell Knowledge Work Dataset for Stress and User Modeling Research. *ICMI 2014 - Proc. 2014 Int. Conf. Multimodal Interact.* **2014**, 291–298, doi:10.1145/2663204.2663257.
156. Beh, W.-K.; Wu, Y.-H.; An-Yeu; Wu MAUS: A Dataset for Mental Workload Assessment on N-Back Task Using Wearable Sensor. **2021**, 1–9.
157. Gjoreski, M.; Kolenik, T.; Knez, T.; Luštrek, M.; Gams, M.; Gjoreski, H.; Pejović, V. Datasets for Cognitive Load Inference Using Wearable Sensors and Psychological Traits. *Appl. Sci.* **2020**, *Vol. 10*, Page 3843 **2020**, *10*, 3843, doi:10.3390/AP10113843.
158. Zyma, I.; Tukaev, S.; Seleznev, I.; Kiyono, K.; Popov, A.; Chernykh, M.; Shpenkov, O. Electroencephalograms during Mental Arithmetic Task Performance. *Data* **2019**, *4*, 2–7, doi:10.3390/data4010014.
159. Ghosh, R.; Deb, N.; Sengupta, K.; Phukan, A.; Choudhury, N.; Kashyap, S.; Phadikar, S.; Saha, R.; Das, P.; Sinha, N.; et al. SAM 40: Dataset of 40 Subject EEG Recordings to Monitor the Induced-Stress While Performing Stroop Color-Word Test, Arithmetic Task, and Mirror Image Recognition Task. *Data Br.* **2022**, *40*, 107772, doi:10.1016/j.dib.2021.107772.
160. Healey, J.A.; Picard, R.W. Detecting Stress during Real-World Driving Tasks Using Physiological Sensors. *IEEE Trans. Intell. Transp. Syst.* **2005**, *6*, 156–166, doi:10.1109/TITS.2005.848368.
161. Haouij, N. El; Poggi, J.M.; Sevestre-Ghalila, S.; Ghozi, R.; Jadane, M. AffectiveROAD System and Database to Assess Driver's Attention. *Proc. ACM Symp. Appl. Comput.* **2018**, 800–803, doi:10.1145/3167132.3167395.
162. Koelstra, S.; Muhl, C.; Soleymani, M.; Jong-Seok Lee; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. DEAP: A Database for Emotion Analysis Using Physiological Signals. *IEEE Trans. Affect. Comput.* **2012**, *3*, 18–31, doi:10.1109/T-AFFC.2011.15.
163. Sharma, K.; Castellini, C.; van den Broek, E.L.; Albu-Schaeffer, A.; Schwenker, F. A Dataset of Continuous Affect Annotations and Physiological Signals for Emotion Analysis. *Sci. Data* **2019**, *6*, 1–13, doi:10.1038/s41597-019-0209-0.
164. Hosseini, S.; Gottumukkala, R.; Katragadda, S.; Bhupatiraju, R.T.; Ashkar, Z.; Borst, C.W.; Cochran, K. A Multimodal Sensor Dataset for Continuous Stress Detection of Nurses in a Hospital. *Sci. Data* **2022**, *9*, 1–13, doi:10.1038/s41597-022-01361-y.
165. Birjandtalab, J.; Cogan, D.; Pouyan, M.B.; Nourani, M. A Non-EEG Biosignals Dataset for Assessment and Visualization of Neurological Status. *IEEE Work. Signal Process. Syst. SiPS Des. Implement.* **2016**, 110–114, doi:10.1109/SiPS.2016.27.
166. Tanwar, R.; Singh, G.; Pal, P.K. FuSeR: Fusion of Wearables Data for StrEss Recognition Using Explainable Artificial Intelligence Models. *2023 14th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2023* **2023**, 1–6, doi:10.1109/ICCCNT56998.2023.10307589.

167. Beh, W.K.; Wu, Y.H.; Wu, A.Y. Robust PPG-Based Mental Workload Assessment System Using Wearable Devices. *IEEE J. Biomed. Heal. Informatics* **2023**, *27*, 2323–2333, doi:10.1109/JBHI.2021.3138639.
168. Ding, Y.; Liu, J.; Zhang, X.; Yang, Z. Dynamic Tracking of State Anxiety via Multi-Modal Data and Machine Learning. *Front. Psychiatry* **2022**, *13*, doi:10.3389/fpsy.2022.757961.
169. Jiao, Y.; Wang, X.; Zhao, L.; Dong, H.; Du, G.; Zhao, S.; Liu, Y.; Liu, C.; Wang, D.; Liang, W. An Improved Sequence Coding-Based Gray Level Co-Occurrence Matrix for Mild Stress Assessment. *Biomed. Signal Process. Control* **2024**, *95*, 106357, doi:10.1016/j.bspc.2024.106357.
170. Mamdouh, M.; Mahmoud, R.; Attallah, O.; Al-Kabbany, A. Stress Detection in the Wild: On the Impact of Cross-Training on Mental State Detection. *Natl. Radio Sci. Conf. NRSC, Proc.* **2023**, *2023-May*, 150–158, doi:10.1109/NRSC58893.2023.10153050.
171. Ribeiro, G.; Postolache, O.; Ferrero, F. A New Intelligent Approach for Automatic Stress Level Assessment Based on Multiple Physiological Parameters Monitoring. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 1–14, doi:10.1109/TIM.2023.3342218.
172. Akbas, A. Evaluation of the Physiological Data Indicating the Dynamic Stress Level of Drivers. *Sci. Res. Essays* **2011**, *6*, 430–439, doi:10.5897/SRE10.943.
173. Bh, S.; Neelima, K.; Deepanjali, C.; Bhuvanashree, P.; Duraipandian, K.; Rajan, S.; Sathiyarayanan, M. Mental Health Analysis of Employees Using Machine Learning Techniques. In Proceedings of the 2022 14th International Conference on Communication Systems and NETWORKS, COMSNETS 2022; 2022; pp. 1–6.
174. Saganowski, S.; Perz, B.; Polak, A.G.; Kazienko, P. Emotion Recognition for Everyday Life Using Physiological Signals From Wearables: A Systematic Literature Review. *IEEE Trans. Affect. Comput.* **2023**, *14*, 1876–1897, doi:10.1109/TAFFC.2022.3176135.
175. Dedovic, K.; Duchesne, A.; Andrews, J.; Engert, V.; Pruessner, J.C. The Brain and the Stress Axis: The Neural Correlates of Cortisol Regulation in Response to Stress. *Neuroimage* **2009**, *47*, 864–871, doi:10.1016/j.neuroimage.2009.05.074.
176. Stuart, T.; Hanna, J.; Gutruf, P. Wearable Devices for Continuous Monitoring of Biosignals: Challenges and Opportunities. *APL Bioeng.* **2022**, *6*, doi:10.1063/5.0086935.
177. Boudreaux, B.D.; Hebert, E.P.; Hollander, D.B.; Williams, B.M.; Cormier, C.L.; Naquin, M.R.; Gillan, W.W.; Gusew, E.E.; Kraemer, R.R. *Validity of Wearable Activity Monitors during Cycling and Resistance Exercise*; 2018; Vol. 50; ISBN 0000000000.
178. Smets, E.; De Raedt, W.; Van Hoof, C. Into the Wild: The Challenges of Physiological Stress Detection in Laboratory and Ambulatory Settings. *IEEE J. Biomed. Heal. Informatics* **2019**, *23*, 463–473, doi:10.1109/JBHI.2018.2883751.
179. Can, Y.S.; Gokay, D.; Kılıç, D.R.; Ekiz, D.; Chalabianloo, N.; Ersoy, C. How Laboratory Experiments Can Be Exploited for Monitoring Stress in the Wild: A Bridge between Laboratory and Daily Life. *Sensors (Switzerland)* **2020**, *20*, 838, doi:10.3390/s20030838.
180. Choi, J.B.; Hong, S.; Nelesen, R.; Bardwell, W.A.; Natarajan, L.; Schubert, C.; Dimsdale, J.E. Age and Ethnicity Differences in Short-Term Heart-Rate Variability. *Psychosom. Med.* **2006**, *68*, 421–426, doi:10.1097/01.psy.0000221378.09239.6a.
181. Graves, B.S.; Hall, M.E.; Dias-Karch, C.; Haischer, M.H.; Apter, C. Gender Differences in Perceived Stress and Coping among College Students. *PLoS One* **2021**, *16*, 1–12, doi:10.1371/journal.pone.0255634.
182. Mueller, A.; Strahler, J.; Armbruster, D.; Lesch, K.P.; Brocke, B.; Kirschbaum, C. Genetic Contributions to Acute Autonomic Stress Responsiveness in Children. *Int. J. Psychophysiol.* **2012**, *83*, 302–308, doi:10.1016/j.ijpsycho.2011.11.007.
183. Ellis, B.J.; Jackson, J.J.; Boyce, W.T. The Stress Response Systems: Universality and Adaptive Individual Differences. *Dev. Rev.* **2006**, *26*, 175–212, doi:10.1016/j.dr.2006.02.004.
184. McEwen, B.S. The Neurobiology of Stress: From Serendipity to Clinical Relevance. *Brain Res.* **2000**, *886*, 172–189, doi:10.1016/S0006-8993(00)02950-4.
185. Grissom, N.; Bhatnagar, S. Habituation to Repeated Stress: Get Used to It. *Neurobiol. Learn. Mem.* **2009**, *92*, 215–224, doi:10.1016/j.nlm.2008.07.001.

186. Kothgassner, O.D.; Goreis, A.; Bauda, I.; Ziegenaus, A.; Glenk, L.M.; Felnhofer, A. Virtual Reality Biofeedback Interventions for Treating Anxiety: A Systematic Review, Meta-Analysis and Future Perspective. *Wien. Klin. Wochenschr.* **2022**, *134*, 49, doi:10.1007/S00508-021-01991-Z.
187. Gradl, S.; Wirth, M.; Zillig, T.; Eskofier, B.M. Visualization of Heart Activity in Virtual Reality: A Biofeedback Application Using Wearable Sensors. In Proceedings of the 2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2018; IEEE, March 2018; Vol. 2018-Janua, pp. 152–155.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.