



Machine Learning-Based Classification of Anxiety-Related Physiological Arousal Using ECG, EDA, and Respiration Signals

Sabrina Aghniya Ilmi, Achmad Rizal^{ID}, Rita Magdalena^{*ID}, Ayu Sekar Safitri^{ID}

School of Electrical Engineering, Telkom University, Bandung 40257, Indonesia

Corresponding Author Email: ritamagdalenat@telkomuniversity.ac.id

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310412>

ABSTRACT

Received: 4 December 2025

Revised: 26 January 2026

Accepted: 8 February 2026

Available online: 30 April 2026

Keywords:

anxiety detection, physiological arousal, electrocardiogram, electrodermal activity, respiration, machine learning, classification

Anxiety can significantly affect mental and physical health, yet traditional detection methods relying on self-report questionnaires are often biased. Biosignal-based approaches provide a more objective way to characterize anxiety-related physiological responses. This study develops a multimodal physiological arousal classification framework using electrocardiogram (ECG), electrodermal activity (EDA), and respiration (RSP) signals. The signals were preprocessed, normalized, and features were extracted from time, frequency, and nonlinear domains. Two physiological arousal labeling schemes derived from heart rate (HR) and skin conductance response (SCR) were used to define low, medium, and high arousal levels associated with anxiety-eliciting stimuli. Five machine learning classifiers were evaluated: Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression, and XGBoost. The best performance was achieved by XGBoost using combined ECG, EDA, and RSP features with HR-derived labels, reaching an accuracy of 92%. The inclusion of RSP consistently improved classification performance compared to ECG and EDA alone. Feature analysis indicated that HRV from ECG and SCR characteristics from EDA contributed strongly to physiological arousal classification. These results demonstrate that multimodal biosignals can effectively model anxiety-related autonomic activation and support the development of non-invasive wearable systems for monitoring physiological responses under anxiety-provoking conditions.

1. INTRODUCTION

Anxiety disorders are among the most prevalent mental health conditions and significantly affect quality of life, work productivity, and cognitive performance [1-4]. Global studies report an increasing prevalence of anxiety-related conditions, such as social phobia and generalized anxiety, particularly among university students and young adults due to academic stress, lifestyle changes, and social pressure [5, 6]. Anxiety often co-occurs with other physiological and psychological conditions, including gastrointestinal problems and sleep disorders, which complicates accurate diagnosis [7, 8].

Conventional diagnostic approaches, such as self-report questionnaires and structured clinical interviews, are inherently subjective and prone to recall bias, limiting their suitability for continuous or real-time monitoring [9, 10]. These limitations have motivated the development of objective approaches based on physiological signals that reflect autonomic nervous system activity associated with anxiety and stress responses [11-14].

Recent advances in wearable sensing and artificial intelligence (AI) have enabled the continuous monitoring of emotional and stress-related physiological states using biosignals such as electrocardiogram (ECG), electrodermal activity (EDA), and respiration (RSP) [6, 15-17]. These signals capture complementary aspects of autonomic nervous

system activity: ECG reflects cardiovascular regulation through heart rate and heart rate variability (HRV), EDA reflects sympathetic activation through skin conductance response (SCR), and RSP captures breathing dynamics that are sensitive to emotional arousal [13, 18]. These biosignals have therefore been widely used to characterize anxiety-related and stress-related physiological responses [19-21].

In experimental fear-elicitation paradigms, anxiety is typically operationalized as physiological arousal induced by emotionally salient stimuli, rather than as a clinical diagnosis. Datasets such as the Spider ECG-EDA-RSP dataset [22, 23] are designed according to this paradigm, where participants are exposed to fear-inducing stimuli while autonomic responses are recorded. In this context, HRV, SCR, and respiratory patterns provide objective measures of anxiety-related autonomic activation during baseline, exposure, and recovery phases.

Machine learning (ML) methods, including Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression, and deep learning models, have demonstrated strong capability in modeling these physiological patterns [7, 20, 22]. For example, Tzevelekakis et al. [9] developed a lightweight convolutional neural network for stress classification using ECG, while Ancillon et al. [15] reviewed the effectiveness of biosignal-based approaches for anxiety and stress monitoring. Househ et al. [23] further

emphasized the role of multimodal biosignals and wearable AI for monitoring mental health related physiological responses.

Although prior studies have demonstrated the feasibility of biosignal-based anxiety and fear recognition, their methodological scope remains limited when examined in the context of multimodal fear-elicitation datasets. Petrescu et al. [11] proposed a protocol for anxiety estimation in virtual-reality exposure using ECG and EDA features, but the analysis was restricted to a fixed feature set and did not evaluate respiratory signals or the effect of alternative classifiers. Moreover, anxiety levels were estimated through a nonlinear physiological function rather than a systematic comparison of machine-learning models.

Vulpe-Grigorași et al. [24] focused exclusively on ECG-derived ultra-short HRV features and trained a one-dimensional convolutional neural network for anxiety detection in arachnophobic individuals. While their results confirmed the relevance of HRV for anxiety-related arousal, the study did not incorporate EDA or respiration, which limits the representation of sympathetic nervous system activity and multimodal physiological dynamics.

Ihmig et al. [18] employed ECG, EDA, and RSP signals in a randomized controlled trial with spider-fearful individuals and investigated both two-level and three-level anxiety classification. However, their final models were primarily based on a fixed classifier type (bagged trees), and the specific contribution of respiration to classification performance was not systematically examined. In addition, although multiple labeling strategies were proposed, direct comparisons between heart-rate-based and skin-conductance-based physiological arousal definitions were not the central focus of their evaluation.

Consequently, despite the availability of multimodal biosignal datasets, it remains unclear how respiration contributes to anxiety-related physiological arousal modeling, how different machine-learning classifiers perform under the same conditions, and how alternative physiological arousal definitions influence classification outcomes.

To address these challenges, this study proposes a multimodal physiological arousal classification framework that integrates ECG, EDA, and RSP signals with five ML classifiers: RF, SVM, KNN, Logistic Regression, and XGBoost [9, 18, 25-27]. The system includes preprocessing (filtering and normalization), feature extraction (time-domain, frequency-domain, and non-linear features), and HR- and SCR-derived arousal labeling to classify physiological responses into low, medium, and high anxiety-related arousal levels. The primary aim of this research is to evaluate how effectively multimodal biosignals capture autonomic activation under anxiety-eliciting stimuli and to identify the most suitable machine-learning model for this task.

The contributions of this paper are summarized as follows:

- 1) Development of a multimodal biosignal pipeline (ECG, EDA, and RSP) for anxiety-related physiological arousal classification.
- 2) Comprehensive feature extraction, including HRV, SCR, and respiration-based metrics.
- 3) Comparative evaluation of five machine-learning classifiers (RF, SVM, KNN, Logistic Regression, and XGBoost).
- 4) Demonstration of improved performance (up to 92%) through multimodal feature fusion and XGBoost optimization.

The remainder of this paper is organized as follows. Section

2 describes the dataset, preprocessing, and methodology. Section 3 presents the experimental results and evaluation metrics. Section 4 discusses our findings and their implications. Section 5 concludes the paper and outlines potential directions for future research.

2. METHOD

The system design in this study provides a structured overview of the workflow used to detect anxiety levels from physiological signals. Figure 1 illustrates the complete process from acquiring secondary biosignals to conducting classification analysis. The dataset consists of ECG, EDA, and RSP signals, which are first processed through filtering and segmentation before feature extraction. The extracted features are then labelled and used as inputs to the classification model. Finally, the system evaluates model performance using standard metrics such as accuracy, precision, recall, and F1-score.

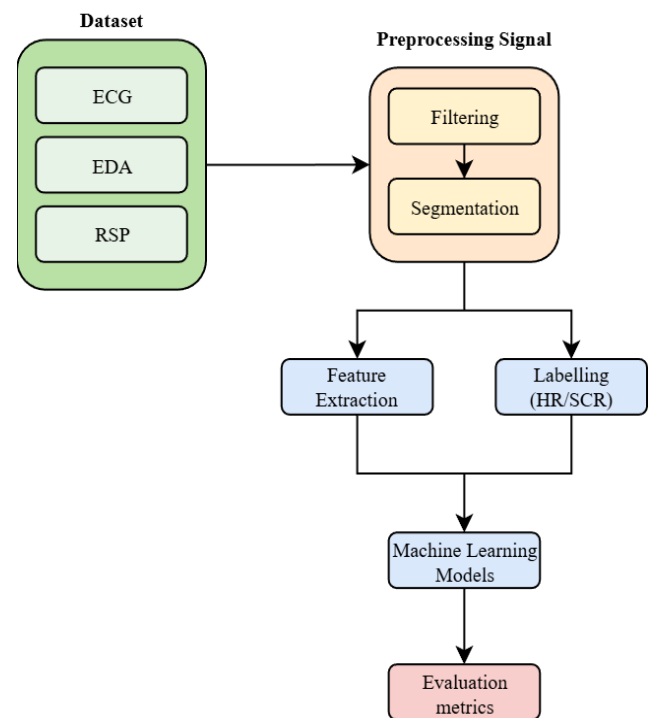


Figure 1. Anxiety detection flowchart

2.1 Dataset

This study employed the Spider ECG-EDA-RSP dataset, a publicly available database from PhysioNet that provides multimodal physiological recordings from 57 participants exposed to spider-related fear stimuli [18, 27, 28]. The dataset included three synchronized biosignals—ECG, EDA, and RSP,—recorded using BITalino (r)evolution wearable devices at a sampling frequency of 100 Hz [26, 29-31]. Each participant underwent a session divided into three distinct phases: baseline (resting state), stimulus (fear exposure), and recovery, which enabled the observation of autonomic responses linked to anxiety [18, 32, 33]. This dataset has been widely used in studies involving anxiety and stress detection, owing to its rich annotations and clear event markers [18, 27, 34].

2.2 Labelling process

The labeling process in this study defines physiological arousal levels associated with anxiety-eliciting stimuli, rather than clinical or subjective anxiety diagnoses. Each signal segment is assigned to one of three arousal categories (Low, Medium, High), which serve as the target labels for supervised classification.

Two physiological markers were used to generate these arousal labels: Heart Rate (HR) derived from ECG and SCR derived from EDA. These signals reflect complementary aspects of autonomic nervous system activation. ECG captures rapid cardiovascular responses, while EDA reflects sustained sympathetic activation through sweat gland activity. Using both allows the evaluation of anxiety-related arousal from two physiological perspectives.

Accordingly, two labeling schemes were considered: HR-derived arousal labels and SCR-derived arousal labels. In both schemes, segments belong to one of the experimental phases (Baseline, Exposure, Recovery) provided by the dataset. Within each phase, predefined HR or SCR thresholds were applied to assign Low, Medium, and High arousal levels, following the well-established relationship between autonomic activation and anxiety.

These HR- and SCR-derived labels do not represent independent psychological or clinical ground truth. Instead, they define objective physiological arousal categories associated with anxiety-provoking stimuli. The HR- and SCR-based thresholds were applied globally across subjects within the baseline, exposure, and recovery phases, rather than being individually calibrated. The classification task therefore evaluates how well multimodal biosignal features (ECG, EDA, and RSP) reproduce these physiological arousal definitions. These labels are generated prior to feature extraction and are independent of the ML input features. Both labeling approaches have been widely used in anxiety detection studies, ensuring consistency and comparability with previous works [11, 18, 24]. Therefore, the reported classification accuracy should be interpreted as the ability of multimodal biosignals to model physiological arousal patterns under anxiety-eliciting stimuli, rather than as a measure of diagnostic accuracy for anxiety disorders.

2.3 Preprocessing

Pre-processing steps were performed to ensure signal quality and consistency prior to analysis. The ECG signal was filtered using a 4th-order Butterworth bandpass filter (0.5–45 Hz) to remove baseline drift and high-frequency noise. Z-score normalization was applied to standardize signal amplitudes [7, 9, 15]. R-peaks were detected using the NeuroKit2 ECG processing pipeline, which applies robust QRS detection and signal-quality assessment. RR intervals were computed from successive R-peaks, and NN intervals were obtained after automatic detection and removal of ectopic beats and motion artifacts using RR-interval outlier correction and interpolation. HRV features were then computed from these physiologically valid NN intervals.

The EDA signal was first low-pass filtered with a cutoff frequency of 1.5 Hz to remove high-frequency noise and to serve as an anti-aliasing filter prior to downsampling to 10 Hz. The signal was then baseline-corrected to separate slow-varying tonic components from fast phasic SCR activity, following standard EDA preprocessing practice [6, 15, 17].

The RSP signal was filtered using a bandpass filter (0.1–1 Hz) to emphasize respiratory cycles and suppress slow drift and high-frequency noise. A baseline correction was applied to remove slow trends and inter-subject offsets in the respiration signal [7, 15]. All signals were segmented into 30-second windows based on the annotated event markers corresponding to the baseline, stimulus, and recovery phases [2, 6]. Although 30-s windows provide limited frequency resolution for low-frequency HRV components, this window length is commonly used in short-term stress and anxiety analysis to capture relative autonomic changes under emotionally salient stimuli [35, 36].

2.4 Feature extraction

Feature extraction focused on deriving meaningful descriptors from ECG, EDA, and RSP signals across the time, frequency, and non-linear domains. For the ECG, HRV metrics were computed. HRV represents the physiological variation in the time intervals between consecutive heartbeats (NN intervals), reflecting the balance between the sympathetic and parasympathetic branches of the autonomic nervous system (ANS). NN intervals refer to the time difference between two normal sinus beats, free from artifacts, noise, and arrhythmias. Although NN intervals originate from normal heart rhythm, they do not necessarily indicate a relaxed physiological state, as individuals experiencing anxiety may still exhibit sinus rhythm [37, 38]. The features extracted from the ECG signal are as follows:

A. Mean_HR_BPM

Mean_HR_BPM represents the average HR in beats per minute within a segment. It is computed from the average duration of NN intervals (MeanNN), where NN refers to the time difference between two consecutive normal sinus beats. Physiologically, an increase in BPM reflects stronger sympathetic activation, which commonly occurs during anxiety. The feature is calculated in Eq. (1), where 60,000 ms corresponds to one minute.

$$Mean_{HR}BPM = \frac{60000}{MeanNN} \quad (1)$$

B. HRV_MeanNN

HRV_MeanNN (Eq. (2)) denotes the average interval between two consecutive normal R-peaks. If NN_i represents the i -th NN interval and N is the total number of NN intervals, this value reflects the overall rhythm of the heart. Shorter MeanNN indicates higher HR and increased sympathetic activity.

$$HRV_{MeanNN} = \frac{1}{N} \sum_{i=1}^N NN_i \quad (2)$$

C. HRV_SDNN

HRV_SDNN measures the standard deviation of all NN intervals. Here, NN_i is the i -th NN interval, and MeanNN is the mean NN interval computed previously. A lower SDNN reflects reduced long-term variability, commonly associated with anxiety. The feature is defined in Eq. (3).

$$SDNN = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (NN_i - MeanNN)^2} \quad (3)$$

D. HRV_RMSSD

HRV_RMSSD (Eq. (4)) represents short-term variability, computed from the differences between adjacent NN intervals. If NN_i and NN_{i+1} are two successive intervals and N is the total number of NN intervals, RMSSD captures vagal (parasympathetic) modulation. Lower values typically appear in individuals experiencing anxiety.

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (NN_{i+1} - NN_i)^2} \quad (4)$$

E. HRV_pNN50

HRV_pNN50 (Eq. (5)) gives the percentage of successive NN interval pairs that differ by more than 50 ms. Here, N_{50} is the number of NN pairs with a difference > 50 ms, and N is the total number of NN intervals. A low pNN50 indicates reduced adaptability and decreased parasympathetic modulation.

$$pNN50 = \frac{N_{50}}{N-1} \times 100\% \quad (5)$$

F. HRV_LF (0.04-0.15 Hz)

The LF component represents spectral power in the low-frequency band, influenced by both sympathetic and parasympathetic systems. LF commonly increases during anxiety due to heightened sympathetic activation.

G. HRV_HF (0.15-0.4 Hz)

The HF component reflects parasympathetic (vagal) activity. HF tends to decrease when individuals experience anxiety, indicating reduced vagal tone.

H. HRV_LFHF (LF/HF Ratio)

Ratio between LF and HF power, often used as an indicator of ANS balance. Higher LF/HF suggests sympathetic dominance commonly associated with anxiety. The feature is defined in Eq. (6).

$$LF / HF = \frac{LF}{HF} \quad (6)$$

I. HRV_SD1

HRV_SD1 (Eq. (7)) represents short-term variability derived from the Poincaré plot. It is calculated from RMSSD as:

$$HRV_{SD1} = \sqrt{\frac{1}{2}} \cdot RMSSD \quad (7)$$

Lower SD1 indicates decreased parasympathetic modulation typical in anxious individuals.

J. HRV_SD2

HRV_SD2 (Eq. (8)) reflects long-term variability and global autonomic regulation. Using SDNN and RMSSD, it is computed as:

$$HRV_{SD2} = \sqrt{2 \cdot SDNN^2 - \frac{1}{2} \cdot RMSSD^2} \quad (8)$$

Decreased SD2 may signal autonomic dysregulation related to anxiety.

K. HRV_HFD (Higuchi Fractal Dimension)

HRV_HFD quantifies the fractal complexity of HRV using

Higuchi's algorithm. Higher HFD values indicate better physiological adaptability, whereas lower values reflect reduced complexity often seen in anxiety.

For the EDA, features such as the number of SCR peaks, mean SCR amplitude, skin conductance level (SCL), and statistical measures (mean and standard deviation) were calculated [6, 12]. Feature extraction from EDA signal is also known as Galvanic Skin Response (GSR), it aims to quantify sympathetic nervous system activation through skin conductance, which is highly sensitive to emotional states such as anxiety and stress.

In this study, the normalized and segmented EDA signals were processed to detect SCR peaks with a minimum threshold of $0.03 \mu\text{S}$ to avoid noise-induced detections, after which the derived SCR characteristics and tonic phasic conductance measures were computed to represent the physiological reactivity associated with anxiety. After detecting the SCR peaks, the following parameters were subsequently calculated:

A. Mean SCL

Mean SCL represents the average tonic level of EDA within a segment. If EDA_i denotes the EDA value at the i -th sample and N is the total number of samples, Mean SCL is formulated in Eq. (9). This feature reflects the baseline conductance level associated with sustained arousal.

$$Mean\ SCL = \frac{1}{N} \sum_{i=1}^N EDA_i \quad (9)$$

B. N_SCR

N_SCR (Eq. (10)) represents the total number of SCR peaks detected in the EDA signal. A peak is counted when the EDA value EDA_i exceeds a predefined threshold.

$$N_{SCR} = \text{count}(EDA_i > \text{Threshold}) \quad (10)$$

C. SCR Amplitude

SCR amplitude represents the average amplitude of all detected SCR peaks. If A_j denotes the amplitude of the j -th SCR peak and N_{SCR} is the total number of SCRs, the feature is calculated using Eq. (11).

$$SCR\ Amplitude = \frac{1}{N_{SCR}} \sum_{j=1}^{N_{SCR}} A_j \quad (11)$$

D. SCR Duration

SCR duration estimates the average temporal spacing between consecutive SCR peaks. If t_j and t_{j+1} are the sample indices of two successive SCR peaks, f_s is the sampling frequency (Hz), and N_{SCR} is the total number of peaks, the duration is defined in Eq. (12).

$$SCR\ Duration = \frac{1}{N_{SCR} - 1} \sum_{j=1}^{N_{SCR}-1} \frac{t_{j+1} - t_j}{f_s} \quad (12)$$

E. SCR Frequency

SCR frequency measures how often SCRs occur within the total duration of the signal. If T is the total duration (in seconds), the feature is defined in Eq. (13).

$$SCR\ Frequency = \frac{N_{SCR}}{T} \quad (13)$$

F. Standard Deviation of EDA

Standard Deviation quantifies the overall variability of the EDA signal within a segment. If EDA_i is the EDA value at the i -th sample and Mean SCL is the average EDA value, the standard deviation is defined in Eq. (14).

$$std_{EDA} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (EDA_i - MeanSCL)^2} \quad (14)$$

For RSP, the average respiratory rate, amplitude, respiratory volume per time (RVT), and cycle irregularity were extracted [7, 9]. These features have been validated in prior studies as reliable markers of sympathetic nervous system activation associated with anxiety [1, 5, 13]. Feature extraction from the RSP signal was performed to capture breathing patterns during exposure to emotional stimuli, as psychological states such as stress and anxiety often alter both the frequency and depth of RSP. The RSP signal used in this study was recorded using a piezoelectric sensor that measures thoracic expansion and was preprocessed through filtering, normalization, and segmentation. The primary features derived from the RSP signal include: Respiratory peaks and troughs were detected from the filtered RSP signal after smoothing using a moving-average filter. A minimum peak-to-peak distance constraint was applied to avoid false detections caused by noise. Breathing rate was computed from successive respiratory cycles identified from these peaks.

A. Breathing Rate (Breaths per Minute)

Breathing Rate represents the number of respiratory cycles occurring within one minute. If N_{peaks} denotes the number of detected RSP peaks and T is the total signal duration in seconds, the breathing rate is computed by using Eq. (15).

$$Breathing\ Rate = 60 \times \frac{N_{peaks}}{T} \quad (15)$$

B. RSP Amplitude

RSP Amplitude measures the average amplitude of each respiratory cycle. For each cycle, P_i is the inhalation peak value and V_i is the exhalation trough value, while N represents the total number of cycles in the segment. The amplitude is defined in Eq. (16).

$$RSP\ Amplitude = \frac{1}{N} \sum_{i=1}^N (P_i - V_i) \quad (16)$$

C. Respiratory Volume per Time

RVT quantifies the combined effect of breathing amplitude and frequency. If $(P_i - V_i)$ is the amplitude of the i -th breath and $(t_{i+1} - t_i)$ is the duration between successive breaths, then with N respiratory cycles, RVT is defined in Eq. (17), where t_i denotes the time interval between two successive respiratory peaks (peak-to-peak).

$$RVT = \frac{1}{N} \sum_{i=1}^N \frac{P_i - V_i}{t_{i+1} - t_i} \quad (17)$$

D. Cycle Symmetry

Cycle Symmetry measures the ratio between the inhalation duration and exhalation duration within one respiratory cycle. If D_{inh} is the inhalation duration and D_{exh} is the exhalation duration, the feature is expressed in Eq. (18).

$$Cycle\ Symmetry = \frac{D_{inh}}{D_{exh}} \quad (18)$$

E. Standard Deviation of RSP

std_RSP represents the overall variability of the RSP signal. If RSP_i denotes the RSP value at the i -th sample and \bar{RSP} is the mean RSP value across the segment, the standard deviation is calculated using Eq. (19).

$$std_{RSP} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (RSP_i - \bar{RSP})^2} \quad (19)$$

2.5 Machine learning classifiers

Five machine-learning classifiers were employed to evaluate the performance of anxiety detection: RF, SVM, KNN, Logistic Regression, and XGBoost [8, 15, 39]. RF was configured with 100 estimators, SVM with an RBF kernel and optimized gamma, KNN with $k = 5$, Euclidean distance, and Logistic Regression with L2 regularization. XGBoost, an ensemble gradient boosting model, was tuned with a learning rate of 0.1 and a maximum depth of 5 [16, 18]. The dataset was split into 80% training and 20% testing subsets. Model training used 5-fold cross-validation, and hyperparameter tuning was conducted through a grid search [32, 40].

Hyperparameter tuning was performed using 5-fold cross-validation within the training set only, while the held-out test set was used exclusively for final performance evaluation. A fully nested cross-validation scheme was not employed in this study. This choice may introduce optimistic bias; however, all classifiers were evaluated under the same protocol to ensure fair comparative analysis.

2.6 Evaluation metrics

The classification performance in this study was evaluated using overall classification accuracy. Accuracy was selected as the primary evaluation metric to enable direct comparison with previous anxiety and fear detection studies, which predominantly report accuracy as the main performance indicator. Using a consistent metric allows a fair and interpretable comparison of different classifiers and signal configurations under similar experimental conditions. Accuracy is defined as the proportion of correct predictions to the total number of predictions, as expressed in Eq. (20):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

where True Positive (TP) refers to samples correctly classified into a given class, False Positive (FP) refers to samples incorrectly assigned to a class, False Negative (FN) represents samples belonging to a class that were misclassified, and True Negative (TN) represents samples correctly classified as not belonging to that class. This evaluation strategy provides a concise and consistent assessment of model performance across all experiments conducted in this study.

2.7 Experimental protocol and validation strategy

The dataset was divided into training and testing sets using a record-wise splitting strategy, in which individual signal segments were randomly assigned to each subset. As a result,

segments from the same subject may appear in both the training and testing sets. Therefore, the reported performance reflects within-subject physiological arousal classification, rather than subject-independent generalization across unseen individuals. The distributions of the three arousal classes (Low, Medium, High) were approximately balanced, and macro-averaged precision, recall, and F1-score were used to avoid bias toward any dominant class.

Hyperparameter optimization was performed using 5-fold cross-validation within the training set. A fully nested cross-validation scheme was not employed; therefore, the reported performance values may be slightly optimistic. However, all classifiers were evaluated using the same training–testing protocol, ensuring fair and consistent comparison across models.

3. RESULT

This study evaluated the performance of a biosignal-based anxiety detection system using four datasets: HR BPM (ECG+EDA), HR BPM (ECG+EDA+RSP), SCR Peaks (ECG+EDA), and SCR Peaks (ECG+EDA+RSP). Each dataset was trained and tested using five classifiers: Logistic

Regression, KNN, SVM, RF, and XGBoost.

The experimental results as shown in Table 1 demonstrate that including respiratory signals consistently improves model performance for both feature types. For HR BPM, the XGBoost model with three signals (ECG+EDA+RSP) achieved the highest accuracy of approximately 92%, whereas without RSP (ECG+EDA) the accuracy was slightly lower at approximately 90%. F1-scores also improved across all classes when RSP was added, particularly achieving around 0.99 for the Low class.

A similar trend was observed for SCR Peaks, where the addition of RSP also improved the accuracy. XGBoost on SCR Peaks (ECG+EDA+RSP) achieved an accuracy of approximately 69%, compared to approximately 67% without RSP. Although SCR-based labelling yielded lower accuracy compared to HR BPM, these findings highlight that integrating RSP has a positive contribution.

Other models such as RF and Logistic Regression also showed good performance with RSP (around 90% for HR BPM and 69–71% for SCR Peaks). Meanwhile, SVM achieved moderate accuracy (80–86% for HR BPM, ~69% for SCR Peaks), and KNN yielded the lowest classification accuracy (59–69%).

Table 1. Classifier accuracy for three anxiety levels using combined HR, SCR, and RSP features

Model	HR BPM		EDA SCR	
	EDA + ECG	+RSP	EDA + ECG	+RSP
Logistic Regression	85	90	65	72
K-Nearest Neighbors (KNN)	69	69	59	64
Support Vector Machine (SVM)	80	86	71	69
Random Forest (RF)	89	90	71	69
XG Boost	91	92	69	67

Note: HR = heart rate; ECG = electrocardiogram; EDA = electrodermal activity; RSP = respiration; SCR = skin conductance response.

4. DISCUSSION

The findings indicate that combining HR BPM and SCR peaks features, especially when including respiratory signals, yields strong physiological arousal classification performance

under anxiety-eliciting stimuli. Simultaneous changes in HR, skin conductance, and RSP reflect ANS activity during anxiety. Gradient boosting models such as XGBoost can capture these complex non-linear patterns, achieving high accuracy even with individual variability across participants.

Table 2. Comparison with previous study

Study	Signals	Method	Binary Accuracy	Three-Class Accuracy
Vulpe-Grigorași et al. [24]	ECG	Neural Network	79.7%	-
Petrescu et al. [11]	ECG, EDA	SVM (best)	84.5%	-
Ihmig et al. [18]	ECG, EDA, RSP	Bagged Trees	89.8%	74.4%
Proposed Method	ECG, EDA, RSP	XGBoost	-	92%

Note: ECG = electrocardiogram; EDA = electrodermal activity; RSP = respiration; SVM = support vector machine.

Compared to other studies as shown in Table 2, the results are competitive. Vulpe-Grigorași et al. [24] employed ECG with neural networks and reported an accuracy of approximately 85–90% for binary anxiety classification. Petrescu et al. [11] integrated biosignal measurements in virtual reality settings and achieved 70–88% accuracy depending on the scenario. Ihmig et al. [18] reported 89.8% accuracy for binary classification 74.4% accuracy for three-class classification using ECG, EDA, and RSP. The present findings are competitive with previous studies, particularly when considering that this work focuses on three-class physiological arousal classification, whereas several earlier studies primarily reported binary classification performance. This highlights the benefits of including respiratory signals and selecting suitable machine-learning methods.

Despite these promising results, some limitations of this study remain. The relatively small sample size limited the generalizability of our results. Class imbalances and controlled laboratory conditions may also restrict applicability in real-world scenarios. Demographic and clinical factors were not extensively analyzed. Accordingly, the reported results should be interpreted as physiological arousal classification under anxiety-eliciting stimuli rather than clinical anxiety diagnosis.

These findings provide an important foundation for the development of non-invasive biosignal-based anxiety detection systems. The inclusion of respiratory signals has proven beneficial and should be considered in the design of wearable devices and physiological monitoring systems. Moreover, this approach has potential for adaptation in various contexts, including exposure-based therapy (as in Ihmig et al.

[18]) and integration with virtual-reality technologies (as in Petrescu et al. [11]).

5. CONCLUSIONS

This study designed and evaluated a three-level physiological arousal classification system based on multimodal biosignals, including ECG, EDA, and RSP. The experimental results demonstrate that integrating respiratory signals consistently improves classification performance across different labeling strategies. The XGBoost model achieved the highest accuracy of approximately 92% for HR-based three-class arousal classification and around 69% for SCR-based three-class classification, highlighting the effectiveness of multimodal feature fusion.

These findings indicate that respiratory signals provide complementary information to cardiovascular and electrodermal measures in modeling anxiety-related autonomic activation. Nevertheless, this study is limited by its relatively small sample size and controlled laboratory conditions. Accordingly, the reported results should be interpreted as physiological arousal classification under anxiety-eliciting stimuli rather than clinical anxiety diagnosis. Future work should focus on subject-independent evaluation, external ground-truth validation, and deployment in real-world wearable settings.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Biomedical Engineering Study Program, Telkom University, Bandung, Indonesia.

REFERENCES

- [1] Stein, D.J., Lim, C.C., Roest, A.M., De Jonge, P., et al. (2017). The cross-national epidemiology of social anxiety disorder: Data from the world mental health survey initiative. *BMC Medicine*, 15(1): 143. <https://doi.org/10.1186/s12916-017-0889-2>
- [2] Daniel-Watanabe, L., Fletcher, P.C. (2022). Are fear and anxiety truly distinct? *Biological Psychiatry Global Open Science*, 2(4): 341-349. <https://doi.org/10.1016/j.bpsgos.2021.09.006>
- [3] Spielberger, C.D., Gonzalez-Reigosa, F.E.R.N.A.N.D.O., Martinez-Urrutia, A., Natalicio, L., Natalicio, D.S. (1971). Development of the Spanish edition of the state-trait anxiety inventory. *Interamerican Journal of Psychology*, 5(3-4): 145-158.
- [4] Grabowska, A., Sondej, F., Senderecka, M. (2024). A machine learning study of anxiety-related symptoms and error-related brain activity. *Journal of Cognitive Neuroscience*, 36(5): 936-961. https://doi.org/10.1162/jocn_a_02126
- [5] Pandit, M., Azwaan, M., Wani, S., Ibrahim, A.A., Abdulghafor, R.A.A., Gulzar, Y. (2023). Examining factors for anxiety and depression prediction. *International Journal on Perceptive and Cognitive Computing*, 9(1): 70-79. <https://doi.org/10.31436/ijpcc.v9i1.368>
- [6] Zhou, E., Soleymani, M., Matarić, M.J. (2023). Investigating the generalizability of physiological characteristics of anxiety. In 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Istanbul, Turkiye, pp. 4848-4855. <https://doi.org/10.1109/BIBM58861.2023.10385292>
- [7] Rizal, A., Siregar, F.D.A.A., Fauzi, H.T. (2022). Obstructive sleep apnea (OSA) classification based on heart rate variability (HRV) on electrocardiogram (ECG) signal using support vector machine (SVM). *Traitement du Signal*, 39(2): 469-474. <https://doi.org/10.18280/ts.390208>
- [8] Arif, M., Basri, A., Melibari, G., Sindi, T., Alghamdi, N., Altalhi, N., Arif, M. (2020). Classification of anxiety disorders using machine learning methods: A literature review. *Insights Biomed Res*, 4(1): 95-110. <https://doi.org/10.36959/584/455>
- [9] Tzevelekakis, K., Stefanidi, Z., Margetis, G. (2021). Real-time stress level feedback from raw ECG signals for personalised, context-aware applications using lightweight convolutional neural network architectures. *Sensors*, 21(23): 7802. <https://doi.org/10.3390/s21237802>
- [10] Alkurdi, A., He, M., Cerna, J., Clore, J., Sowers, R., Hsiao-Wecksler, E.T., Hernandez, M.E. (2025). Extending anxiety detection from multimodal wearables in controlled conditions to real-world environments. *Sensors*, 25(4): 1241. <https://doi.org/10.3390/s25041241>
- [11] Petrescu, L., Petrescu, C., Mitruț, O., Moise, G., Moldoveanu, A., Moldoveanu, F., Leordeanu, M. (2020). Integrating biosignals measurement in virtual reality environments for anxiety detection. *Sensors*, 20(24): 7088. <https://doi.org/10.3390/s20247088>
- [12] Wickramasuriya, D.S., Faghih, R.T. (2020). A mixed filter algorithm for sympathetic arousal tracking from skin conductance and heart rate measurements in Pavlovian fear conditioning. *Plos One*, 15(4): e0231659. <https://doi.org/10.1371/journal.pone.0231659>
- [13] Ain, K., Nur Rahma, O., Purwanti, E., Varyan, R., et al. (2025). Measuring anxiety level on phobia using electrodermal activity, electrocardiogram and respiratory signals. *International Journal of Electrical & Computer Engineering*, 15(1): 337-348. <https://doi.org/10.11591/ijece.v15i1.pp337-348>
- [14] Klimek, A., Mannheim, I., Schouten, G., Wouters, E.J., Peeters, M.W. (2023). Wearables measuring electrodermal activity to assess perceived stress in care: A scoping review. *Acta Neuropsychiatrica*, 37: e19. <https://doi.org/10.1017/neu.2023.19>
- [15] Ancillon, L., Elgendi, M., Menon, C. (2022). Machine learning for anxiety detection using biosignals: a review. *Diagnostics*, 12(8): 1794. <https://doi.org/10.3390/diagnostics12081794>
- [16] Henry, J., Lloyd, H., Turner, M., Kendrick, C. (2023). On the robustness of machine learning models for stress and anxiety recognition from heart activity signals. *IEEE Sensors Journal*, 23(13): 14428-14436. <https://doi.org/10.1109/JSEN.2023.3276413>
- [17] Bolpagni, M., Pardini, S., Dianti, M., Gabrielli, S. (2024). Personalized stress detection using biosignals from wearables: A scoping review. *Sensors*, 24(10): 3221. <https://doi.org/10.3390/s24103221>
- [18] Ihmig, F.R., Neurohr-Parakenings, F., Schäfer, S.K., Lass-Hennemann, J., Michael, T. (2020). On-line anxiety level detection from biosignals: Machine learning based

- on a randomized controlled trial with spider-fearful individuals. *Plos One*, 15(6): e0231517. <https://doi.org/10.1371/journal.pone.0231517>
- [19] Petrescu, L., Petrescu, C., Mitruț, O., Moise, G., Moldoveanu, A., Moldoveanu, F., Leordeanu, M. (2020). Integrating biosignals measurement in virtual reality environments for anxiety detection. *Sensors*, 20(24): 7088. <https://doi.org/10.3390/s20247088>
- [20] Ham, S.M., Lee, H.M., Lim, J.H., Seo, J. (2023). A negative emotion recognition system with internet of things-based multimodal biosignal data. *Electronics*, 12(20): 4321. <https://doi.org/10.3390/electronics12204321>
- [21] Panda, R., Kumar, R., Biradar, O. (2025). Anxiety detection on ECG signal using fuzzy deep learning. *Procedia Computer Science*, 258: 1823-1832. <https://doi.org/10.1016/j.procs.2025.04.434>
- [22] Fauzi, H., Rizal, A., Oktarianto, A., Said, Z. (2023). Classification of normal and abnormal heart sounds using empirical mode decomposition and first order statistic. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 5(2): 82-88. <https://doi.org/10.35882/jeeemi.v5i2.287>
- [23] Abd-Alrazaq, A., AlSaad, R., Aziz, S., Ahmed, A., et al. (2023). Wearable artificial intelligence for anxiety and depression: Scoping review. *Journal of Medical Internet Research*, 25: e42672. <https://doi.org/10.2196/42672>
- [24] Vulpe-Grigorași, A., Grigore, O. (2021). A neural network approach for anxiety detection based on ECG. In *2021 International Conference on e-Health and Bioengineering (EHB)*, Iasi, Romania, pp. 1-4. <https://doi.org/10.1109/EHB52898.2021.9657544>
- [25] Alvarez Espezua, C.B., Cruz de la Cruz, J.E., Apaza Davila, F.A., Cruz de la Cruz, T.D., Huaquipaco Encinas, S., Mamani Machaca, W.A. (2024). Classification of depression and anxiety with machine learning applying random forest models. In *Proceedings of the 2024 5th International Conference on Intelligent Medicine and Health*, pp. 128-132. <https://doi.org/10.1145/3715931.3715955>
- [26] Staib, M., Castegnetti, G., Bach, D.R. (2015). Optimising a model-based approach to inferring fear learning from skin conductance responses. *Journal of Neuroscience Methods*, 255: 131-138. <https://doi.org/10.1016/j.jneumeth.2015.08.009>
- [27] Schäfer, S.K., Ihmig, F.R., Lara H.K.A., Neurohr, F., et al. (2018). Effects of heart rate variability biofeedback during exposure to fear-provoking stimuli within spider-fearful individuals: Study protocol for a randomized controlled trial. *Trials*, 19(1): 184. <https://doi.org/10.1186/s13063-018-2554-2>
- [28] Zhang, M., Karner, A., Kistorz, K., Shea, S., et al. (2025). SpiDa-MRI: Behavioral and (f) MRI data of adults with fear of spiders. *Scientific Data*, 12(1): 284. <https://doi.org/10.1101/2024.02.07.578564>
- [29] Rodríguez-Arce, J., Lara-Flores, L., Portillo-Rodríguez, O., Martínez-Méndez, R. (2020). Towards an anxiety and stress recognition system for academic environments based on physiological features. *Computer Methods and Programs in Biomedicine*, 190: 105408. <https://doi.org/10.1016/j.cmpb.2020.105408>
- [30] Elgendi, M., Galli, V., Ahmadzadeh, C., Menon, C. (2022). Dataset of psychological scales and physiological signals collected for anxiety assessment using a portable device. *Data*, 7(9): 132. <https://doi.org/10.3390/data7090132>
- [31] Erfianto, B., Rizal, A. (2022). IMU-Based respiratory signal processing using cascade complementary filter method. *Journal of Sensors*, 2022(1): 7987159. <https://doi.org/10.1155/2022/7987159>
- [32] Alkurdi, A., Clore, J., Sowers, R., Hsiao-Wecksler, E.T., Hernandez, M.E. (2024). Resilience of machine learning models in anxiety detection: Assessing the impact of gaussian noise on wearable sensors. *Applied Sciences*, 15(1): 88. <https://doi.org/10.3390/app15010088>
- [33] Lor, C.S., Steyrl, D., Karner, A., Götzendorfer, S.J., et al. (2025). SpiderPhy dataset: A multimodal dataset of physiological, psychometric and behavioral responses to fear stimuli. *Scientific Data*, 12(1): 599. <https://doi.org/10.1038/s41597-025-04908-x>
- [34] Aristizabal, S., Byun, K., Wood, N., Mullan, A.F., et al. (2021). The feasibility of wearable and self-report stress detection measures in a semi-controlled lab environment. *IEEE Access*, 9: 102053-102068. <https://doi.org/10.1109/ACCESS.2021.3097038>
- [35] Bernardes, A., Couceiro, R., Medeiros, J., Henriques, J., et al. (2022). How reliable are ultra-short-term HRV measurements during cognitively demanding tasks? *Sensors*, 22(17): 6528. <https://doi.org/10.3390/s22176528>
- [36] Castaldo, R., Montesinos, L., Melillo, P., James, C., Pecchia, L. (2019). Ultra-short term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life. *BMC Medical Informatics and Decision Making*, 19(1): 12. <https://doi.org/10.1186/s12911-019-0742-y>
- [37] Huang, Y., Wang, Y., Xu, B., Zeng, Y., Chen, P., Huang, Y., Liu, X. (2025). The association between constipation and anxiety: A cross-sectional study and Mendelian randomization analysis. *Frontiers in Psychiatry*, 16: 1543692. <https://doi.org/10.3389/fpsy.2025.1543692>
- [38] Rosenbaum, D., Leehr, E.J., Kroczeck, A., Rubel, J.A., et al. (2020). Neuronal correlates of spider phobia in a combined fNIRS-EEG study. *Scientific Reports*, 10(1): 12597. <https://doi.org/10.1038/s41598-020-69127-3>
- [39] Tabares, M.T., Álvarez, C.V., Salcedo, J.B., Rendón, S.M. (2024). Anxiety in young people: Analysis from a machine learning model. *Acta Psychologica*, 248: 104410. <https://doi.org/10.1016/j.actpsy.2024.104410>
- [40] Park, J.H., Shin, Y.B., Jung, D., Hur, J.W., et al. (2025). Machine learning prediction of anxiety symptoms in social anxiety disorder: Utilizing multimodal data from virtual reality sessions. *Frontiers in Psychiatry*, 15: 1504190. <https://doi.org/10.3389/fpsy.2024.1504190>