KCNN: A 1D Kernel Enhanced CNN Aided Feature Fusion Technique for Multi-level Anxiety Detection from Wearable ECG Sensor Signal

Jyeshtar
Department of ECE
NIT Roukela
Rourkela, India
soujyeshtar3877@gmail.com

Arkapravo Dey
Department of ECE
NIT Rourkela
Rourkela, India
arka2005dey@gmail.com

Saptarshi Chatterjee
Department of ECE
NIT Rourkela
Rourkela, India
chatterjees@nitrkl.ac.in

***Abstract*—This work proposes heart rate variability (HRV) and kernel enhanced 1D-CNN based feature fusion technique for automatic anxiety detection from single channel wearable electrocardiogram (ECG) sensor signals. From the ECG sensor signals R-peaks are detected employing non-linear energy operator to extract efficient HRV features. To extract the inherent temporal features from the ECG signals a set of pre-initialized filters have been incorporated in a kernel-enhanced convolutional neural network model (1D-KCNN). Fused features are used with some standard classifiers to identify normal, light, moderate and severe anxiety levels. The performance of the algorithm for detection of anxiety is evaluated on a publicly available wearable ECG sensor dataset. This work has achieved 97.85% accuracy with 97.49% F1-score for detection of four anxiety classes using cross gradient boosting (XGBoost) model from short duration ECG signals. This work has outperformed recently published works on anxiety detection from ECG and multi-modal physiological signals.**

Keywords—1D-KCNN, Anxiety, Feature fusion, non-linear energy operator, Wearable ECG sensor signal processing

I. INTRODUCTION

Anxiety disorders are among the most common mental health issues, affecting over 301 million people worldwide [1]. Anxiety at excessive levels can result in fear, unease, worry, and a feeling of vulnerability [2]. Thus, early diagnosis is essential to prevent anxiety from developing into a cognitive disorder that can significantly impact an individual's behavior and lifestyle. Physiological signals like electroencephalograms (EEG), electrocardiograms (ECG), electrodermal activity (EDA), photoplethysmograms (PPG), and respiration rates (RSP) provide valuable insights into mental well-being [3].

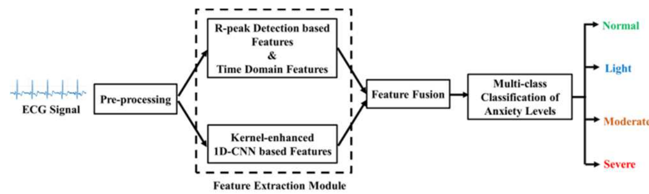
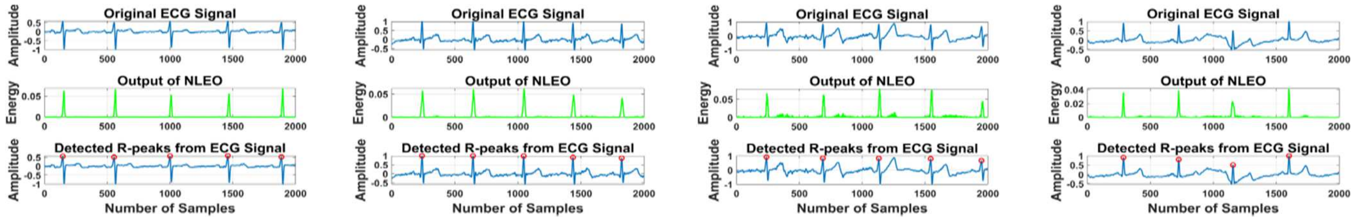


Fig1: Block diagram of feature fusion based multi-level anxiety detection system from ECG sensor signals.

However, specialized expertise and experience are required to interpret these signals, which can be challenging for general practitioners. Therefore, automatic detection of anxiety levels from physiological signals provides a second opinion to the general physicians that may lead to early-stage diagnosis with high accuracy.

Literature suggests that EEG signals are one of the most commonly used bio-signals, reflecting the activity of the nervous system during emotional changes. However, several research groups have investigated that human cognitive variations are also reflected in ECG signals. Determination of heart rate variability (HRV) at different mental conditions provides some insight features to identify cognitive disorders [4]. Hence, the development of efficient R-peak detection algorithm is essential for comprehensive HRV analysis and simultaneous anxiety detection [5-8]. Subsequent researchers have introduced various R-Peak based features for monitoring various cognitive disorders. Selzler et al. in [9] have used ECG along with EDA signals followed by a random forest (RF) classifier to identify three different levels of anxiety. Similarly, Zhao et al. have developed a stress detection system that analyzes ECG and EDA data using a hybrid module of handcrafted with deep learning features, to



(a) (b) (c) (d)

Fig 2: Detected R-peaks from 4s ECG samples of four anxiety levels (a) NA (b) LA (c) MA (d) SA

segregate the subjects into baseline, stressed, or amused states by XGBoost classifier [10]. In [11], Tripathy et al. have employed Fourier-Bessel Domain Adaptive Wavelet Transform (FBDAWT) followed by XGBoost classifier for determining four levels of anxiety with an accuracy of more than 92%. Furthermore, Vulpe-Grigorasi et al. have fused HRV and morphological features extracted from ECG and RSP signals respectively with a 1D-CNN for classifying high and low anxiety levels [12].

In this work, non-linear energy operator-based R-peak detection algorithm is introduced for more precise HRV based feature extraction. A pre-initialized enhanced kernel aided 1-D CNN model is reported here to make a feature fusion based multi-level anxiety detection tool from ECG signals. In this work, we have made the following contributions:

- Introducing a non-linear energy operator-based R-peak detection algorithm for accurate HRV measurement and related feature extraction
- Synergizing signal processing and convolutional neural network through pre-initialized kernel for better explainability and enriched feature extraction.
- Development of feature fusion based lightweight automatic multi-level anxiety detection technique from single channel ECG channels

II. DATASET DESCRIPTION

In this work, a publicly available ECG database on different anxiety levels of 19 participants (14 men and 5 women), recorded using wearable sensors during anxiety-inducing and non-anxiety video clips has been used to evaluate the performance of the reported technique [4]. The ECG signals are sampled at 500 Hz. Based on Hamilton Anxiety Measure (HAM) scores each participant's ECG signals are annotated by experts as Normal (NA), Light anxiety (LA), Moderate (MA) and Severe anxiety (SA) levels.

The ECG signals are decomposed into non-overlapping segments of 4s, 5s, and 6s durations with 2000, 2500, and 3000 number of samples respectively. The segmentation of ECG signals from 19 participants yielded 11881, 9,505, and 7,921 numbers of frames with durations of 4 seconds, 5 seconds, and 6 seconds, respectively.

III. METHODOLOGY

A. Pre-Processing

To eliminate baseline wandering noise from the ECG signals a second-order Butterworth high-pass filter with a cut-off frequency of 0.5 Hz is employed.

B. R-Peak Detection

In ECG signal, difference between consecutive R-peaks in QRS complex helps to determine HRV during anxiety or other stressful events.

Localized high frequency with increased instantaneous energy corresponds to R-peaks in ECG signals. Considering the non-stationarity of R-peaks in ECG signal, non-linear energy operator is employed to weight high frequency components as a function of time with time-varying envelope [13]. In discrete form the non-linear energy operator (NLEO) is defined as:

$$\varphi(x(n)) = x^2(n) - x(n-1)(n+1) \tag{1}$$

Where $x(n)$ is the amplitude of the signal at sample n and $x(n-1)$ and $x(n+1)$ are the previous and next samples of the signal respectively.

Employing this non-linear energy operator, the peak instantaneous energy is determined from the ECG signals. Centering on the location of the peaks in the energy operator output a search window is employed in ECG signals for the identification of the R-peak locations. Non-linear energy operator-based R-peak detection algorithm is given in Algorithm 1. The results of the R-peak detected signals with corresponding energy operator output from the original signals have been shown in Fig.2. This algorithm is less complex and also outperforms the commonly used Pan-Tompkin's algorithm [14] even in some critical ECG patterns during various anxiety stages.

C. HRV and Non-Linear Feature Extraction

From the R-peak detected outcomes, thirteen important features describing HRV have been estimated like heart rate (HR), average and mean average HR, percentage of RR

TABLE1: HRV-based features

Feature	Equation	Feature	Equation
Heart Rate (HR)	$\frac{60}{\mathrm{R-R\ interval(second)}}$	NN50(% of RR interval differences > 50msec)	$\sum_{i=1}^{N-1} 1 \quad \text{if } \lvert RR_{i+1} - RR_i \rvert > 0.05 \text{ seconds}$
Mean average of the RR interval	$\frac{1}{N}\sum_{i=1}^{N} RR_i$	pNN50(% of RR interval differences > 50msec)	$\frac{NN50}{N} \times 100$
Average HR	$\frac{60}{MEAN_{RR}}$ (in bpm)	Root mean square (RMS) of RR interval	$\sum_{i=1}^{N} \frac{(RR_i)^2}{N}$
Root Mean Square of Successive Differences	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(RR_{i+1} - RR_i)^2}$	% of RR intervals with more than one standard deviation above the mean	$\sum \frac{(RR_i > MEAN_{RR} + SD_{RR})}{N} \times 100$
Standard deviation (SD) of the RR interval	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(RR_i - MEAN_{RR})^2}$	% of RR intervals with more than one standard deviation below the mean	$\sum \frac{(RR_i < MEAN_{RR} - SD_{RR})}{N} \times 100$
Standard deviation (SD) of the HR	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(HR_i - AVG_{HR})^2}$	Skewness	$\frac{N}{(N-1)(N-2)}\sum_{i=1}^{N} \frac{(RR_{i+1} - RR_i)^3}{(SD_{RR})^3}$

interval differences, SD of R-R interval, skewness, kurtosis etc., shown in Table 1. Along with these features six non-linear features like entropy, power spectral entropy, Hurst exponent, Katz and Higuchi fractal dimension as well as Lyapunov exponent have been determined.

Algorithm: R-Peak Detection Algorithm

1: **Input:** ECG signal segmented in 4s, 5s, and 6s with number of samples 2000, 2500, and 3000 respectively.
2: **Sampling Frequency (Fs):** 500Hz
3: **Output:** *R-Peaks*
4: ***for*** *e* in *length(data)* ***do***
  $e_n \rightarrow$ *normalize* (e)
  $e_{tkeo} \rightarrow$ ***TKEO*** (e_n)
5: *Initialize: window_length* = ***round*** (0.01 * Fs)
  $e_{smoothed}$ = *movingavg* (e_{tkeo}, window_length)
  $e_{smoothed} \rightarrow$ *normalize* ($e_{smoothed}$)
  threshold = 4 * ***mean***($e_{smoothed}$)
  mean_peak_dist = ***round*** (0.2 * Fs)
  search_window = ***round*** (0.1 * Fs)
6: *Initialize:* pks , locs = [ ] , [ ]
7: ***for*** *i* in $e_{smoothed}$ ***do***
8:   ***if*** *value*[*i*] > *threshold* ***and*** *i* > *mean_peak_dist* ***then***
9:     pks.*append* ($e_{smoothed}$[*i*]) , locs.*append* (*value*[*i*])
  end for
10: *Initialize:* R_peaks , R_peak_locs = [ ] , [ ]
11: ***for*** *i* in ***length***(*pks*) ***do***
  loc = *locs*[*i*]
  lb = ***max***(1, *loc* - *search_window*)
  ub = ***min***(2000, *loc* + *search_window*)
  Max[*i*] = ***max***(e_n[*lb* : *ub*])
  th = 0.4 * ***max***(*Max*)
  for *j* in *length*(*Max*) ***do***
    if *Max*[*j*] > th ***then***
      R_peaks.append(*Max*[*j*])
  end for
  for peak in *R_peaks* ***do***
    R_peak_locs = ***find***(e_n == *peak*)
  end for
end for

D. Kernel-enhanced 1D-CNN(1D-KCNN) Model

This study presents a 1D-CNN based network augmented with residual connections and enhanced by the integration of Daubechies filters [15] to extract some inherent features from the ECG signals. These filters are applied as convolutions with a variable weight matrix, allowing their weights to be dynamically updated during backpropagation. As a result, the filters adapt and incorporate characteristics from other Daubechies kernels making them more versatile and responsive to the signal's underlying patterns. The 1D-CNN architecture consists of convolutional layers followed by a batch normalization layer and an activation function. A skip connection is implemented simultaneously from the convolutional layer using Daubechies filters of 2nd, 4th and 6th order. These filters exhibit fractal characteristics and

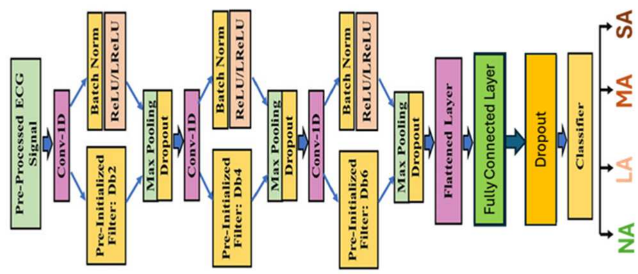


Fig 3: Architecture of 1D-KCNN model

TABLE 2: Specifications of the 1D-KCNN architecture designed for the performance evaluation study

Dataset Information and Parameters	Specifications
ECG signal format	.mat
ECG annotations format	.xlsx
Segment Durations	4s, 5s and 6s
Number of train and test segments	4s: 9505 train and 2376 test segments 5s: 7604 train and 1901 test segments 6s: 6336 train and 1585 test segments
Convolutional layers	CL1: filters = 16, kernel size = 3, stride = 1 CL2 : filters = 32, kernel size = 3, stride = 1 CL3 : filters = 64, kernel size = 3, stride = 1
Fully Connected layers	Input layer: 512 nodes Hidden layer 1: 128 nodes Hidden layer 2: 64 nodes Output layer: 4 nodes corresponding to 4 classes of anxiety
Dropout	p = 0.25 (for convolutional layers) p = 0.5 (for fully connected layers)
Activation Function	ReLU / Leaky ReLU
Learning Rate	0.0001
Decay	0.000001
Loss Function	Binary Cross Entropy
Optimizer	ADAM (Adaptive Moment Estimation)
Epochs	40
Trainable parameters	1,65,15,972(63 MB)
Non-Trainable parameters	480(1.88 KB)
Total parameters	1,65,16,452 (63.01 MB)

waveforms resembling ECG signals. The Daubechies wavelets are defined recursively through their scaling function $\phi(t)$ and wavelet function $\psi(t)$ which are constructed using filter coefficients h_k and g_k and satisfy the following recursive relations as given in equation (2) and (3) respectively:

$$\phi(t) = \sqrt{2}\sum_{k=0}^{N-1} h_k \phi(2t-k) \quad (2)$$

$$\psi(t) = \sqrt{2}\sum_{k=0}^{N-1} g_k \phi(2t-k) \quad (3)$$

where $g_k = (-1)^k h_{N-1-k}$.

The Daubechies filters of orders 2, 4, and 6 are characterized by their unique filter coefficients. For db2, the coefficients are h_0 and h_1. Moving to db4, the set expands to include h_0, h_1, h_2, and h_3. Finally, for db6, the coefficients further extend to h_0, h_1, h_2, h_3, h_4, and h_5. These coefficients are used to recursively define the scaling and wavelet functions, enabling the construction of wavelets with increasing complexity and vanishing moments. This makes them particularly effective for capturing the fractal and non-stationary characteristics of ECG-like signals.

The feature maps from both paths are concatenated and passed to a max-pooling layer for dimensionality reduction, followed by a dropout layer to prevent overfitting. This block is repeated three times and the output is flattened to serve as input for a deep neural network or machine learning algorithms for anxiety classification. The block diagram of this proposed 1-D KCNN network is shown in Fig.3.

The overall specifications of the 1-D KCNN network is summarized in Table 2. The 1-D KCNN network involves three convolutional layers with 16, 32 and 64 number of filters respectively. The dropout layers in the convolution segment have p = 0.25. The outputs of the convolution layers have been normalized as given in equation (4):

$$x_{\text{norm}}[i] = (x[i] - x_{\min}) / (x_{\max} - x_{\min}), \quad (4)$$

where x_{norm} denotes the normalized value of the i^{th} sample in the ECG segment, $x_{\max}$ and $x_{\min}$ denote the maximum and minimum values of the ECG segment x. The normalized segment is directly fed to the activation functions (ReLU / Leaky ReLU). For the fully connected layers, a decreasing number of neurons (512, 128 and 64) have been considered followed by an output layer having 4 neurons for final classification. The dropout layers in the fully connected segment have p = 0.5. For this work, we have used the gradient descent-based ADAM optimizer with a learning rate of 0.0001 and decay of 0.000001. The loss function used is the binary cross entropy loss which can be expressed as in equation (5):

$$Loss = -\frac{1}{4}\sum_{i=0}^{3}[y_i \log(p_i) + (1-y_i)\log(1-p_i)] \quad (5)$$

where y_i is the true label for class i and p_i is the predicted probability for class i. The 1-D KCNN network has a total of approximately 16,516,452 parameters, consuming only around 63.01 MB of memory.

IV. RESULTS AND DISCUSSIONS

This section describes the detailed analysis of the reported techniques for effective feature extraction and performance evaluation for classification of multiple anxiety stages. To evaluate the performance, the complete work flow has been carried out on three different sets of ECG data segmented in 4s, 5s and 6s durations containing all the anxiety levels. Classification performance has been reported by determining the percentage accuracy and F1-score.

TABLE 3: Classification performances on handcrafted features

Classifier	ECG Segments					
	4s		5s		6s	
	Acc(%)	F1(%)	Acc(%)	F1(%)	Acc(%)	F1(%)
NN	84.30	83.21	83.47	82.89	82.20	81.72
DT	84.26	88.39	82.22	79.75	81.61	79.25
ERT	**88.68**	**88.50**	86.27	86.25	**86.81**	**86.75**
XGBoost	86.20	86.92	86.64	86.49	85.67	85.32
RF	88.59	88.25	**86.69**	**85.78**	86.49	85.35

TABLE 4: Classification performances of 1D-KCNN with different activation functions

Activation Function	ECG Segments					
	4s		5s		6s	
	Acc(%)	F1(%)	Acc(%)	F1(%)	Acc(%)	F1(%)
ReLU	94.47	94.51	**95.37**	**95.35**	**95.21**	**95.91**
LReLU	**94.54**	**94.32**	93.53	92.87	93.12	93.13

TABLE 5: Classification Performance for fused feature set using different classifiers

Classifier	ECG Segments					
	4s		5s		6s	
	Acc(%)	F1(%)	Acc(%)	F1(%)	Acc(%)	F1(%)
DT	95.43	95.12	**95.72**	**94.72**	94.55	94.37
ERT	94.45	94.42	94.55	93.61	94.31	94.45
XGBoost	**97.85**	**97.49**	95.21	94.95	**96.59**	**96.23**
RF	95.52	94.83	95.68	95.42	95.22	94.92

TABLE 6: Comparison with SOTA methods for anxiety detection from physiological signals

Method	Physiological Signals	Anxiety Classes	Acc(%)
Giovanni et al.[7]	ECG	Low, Medium, High	78.6
Hao et al. [8]	ECG	Baseline, Stress, and Amusement	90.05
Zhao et al. [10]	ECG , EDA	Anxiety , No Anxiety	77.14
Tripathy et al. [11]	ECG	NA, LA, MA, SA	92.27
Present Work	**ECG**	**NA, LA, MA, SA**	**97.85**

From each of the ECG signal samples, R-peaks are detected employing non-linear energy operator. Combination of HRV based and non-linear features, i.e. 19-D features have been used to classify four anxiety levels using neural network (NN), decision tree (DT), extreme random trees (ERT), XGBoost and random forest (RF) based classifiers. Table-I depicts the performance of four class anxiety classification for three different sets of ECG signals. Here, the entire dataset has been divided in 80% training and 20% testing data, without any overlapping of the samples of same subject. From Table 3 it has been observed that in case of all classifiers, better performance has been obtained for the ECG signal samples with 4s duration with ERT classifier that identified all the four anxiety classes with 88.68% accuracy and 88.50% F1-score. In 5s as well as 6s duration samples, RF and ERT classifiers have reported 86.69% and 86.81% accuracy respectively.

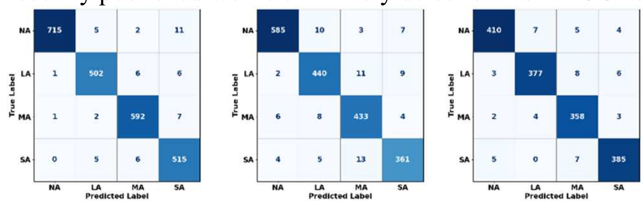
Application of CNN in 1-D signal analysis has come up with a significant potential by extracting complex inherent features with trainable parameters. Updating convolutional kernels during backpropagation employ various operators on the input signal to get various feature representation. However, enhancement of kernel features or pre-initialization of kernels lead to a more explainable network with improved feature extraction ability. In this work, Daubechies filters of three different orders have been incorporated to introduce improved feature planes to the next convolutional layer. The analysis revealed that the Daubechies filters are updated in such a manner that it behaves as a higher order filter in different iterations.

Performances indices tabulated in Table 4 describes that 1-D KCNN model has classified multiple anxiety levels with improved accuracy. To verify the selection of effective activation function for improved classification, ReLU and leaky- ReLU (LReLU) have been used during the experimentation. Table 4 shows that the reported KCNN model has achieved 94.54% accuracy with LReLU from ECG signals of 4s duration. Similarly, this KCNN model identified the anxiety levels with 95.37% and 95.21% accuracy with ReLU activation function from ECG signals of 5s and 6s durations respectively.

Synergism of signal processing techniques and feature learning-based approach provides improved performance with better feature selection and explainability. HRV based features extracted from each of the ECG signals are concatenated with learned features of 1-D KCNN model as in the flattened layer to obtain a fused feature set. This fused feature has been used to classify four anxiety levels using DT, ERT, XGBoost and RF classifiers. Table 5 depicts that from fused feature set XGBoost classifier has identified anxiety levels with 97.85% accuracy from the ECG signals of 4s duration. Table 5 also describes that feature fusion-based anxiety detection technique has consistently improved the performance for all the classifiers mentioned in the table.

It is pertinent to mention that we have achieved acceptable results for short duration ECG signals by extracting inherent temporal features using 1-D KCNN model. Similarly, introduction of this technique has efficiently addressed rapid anxiety level changes and eliminated irrelevant information from long duration samples to provide consistent high identification accuracy for anxiety detection. Fig.4 depicts the corresponding confusion matrices for best classification performance in three different segments of ECG sensor signals for anxiety detection.

As mentioned in earlier section that several researchers have identified anxiety levels from multi-modal physiological signals. This work presents a successful attempt to identify multiple anxiety levels from ECG signals only. In Table 6, performance of the present work has been compared with the recently published works on anxiety detection from ECG as



(a) (b) (c)

Fig 4: Confusion matrices corresponding to the best classification performance in (a) 4s (b) 5s and (c) 6s ECG segments

well as other physiological signals. Table 6 shows that the reported work has identified four different anxiety levels compared to others' two or three levels classification approaches with a higher degree of accuracy of 97.85%. This work has also outperformed the single-channel wearable ECG sensor signal based work in [8] for all three segments (i.e. 4s, 5s and 6s).

V. CONCLUSION

Non-linear energy operator-based tool has been introduced for accurate R-peak detection from wearable ECG sensor signals in anxiety-inducing and non-anxiety conditions and subsequent HRV feature extraction. Introduction of Daubechies filter enhanced kernels in 1-D CNN architecture helps to extract temporal features of the ECG signals more efficiently. This approach has been utilized to identify rapid changes of anxiety levels in short duration ECG signals. In this work, feature fusion technique has been implemented by concatenating HRV and CNN based features and XGBoost classifier to identify multiple anxiety levels with 97.85% accuracy. Development of this tool with lesser number of hyper-parameters makes it a light-weight model for anxiety detection. This reported model can be implemented as an on-board system for automatic identification of anxiety levels from wearable ECG sensor signals in real-time.

REFERENCES

[1] Javaid *et. al* (2023), "Epidemiology of anxiety disorders: Global burden and sociodemographic associations", *Middle East Current Psychiatry.*

[2] R. Kapur (2020), "Effects of Anxiety on Health and Well-being of the Individuals," JOUR, 07.01.

[3] Ancillon *et. al* (2022), "Machine learning for anxiety detection using bio signals: A review," Diagnostics, vol.12, no. 8

[4] Elgendi *et. al* (2022), "Dataset of psychological scales and physiological signals collected for anxiety assessment using a portable device," Data, vol. 7, no. 9, p. 132.

[5] Elgendi *et. al* (2019), "Assessing Anxiety Disorders Using Wearable Devices: Challenges and Future Directions", Brain Sci.

[6] Pascale *et. al* (2020) , "Brain-heart communication: hardware and software strategies through nerves and humoral factors", Brain Heart Dyn. 25‑41.

[7] Giovanni *et. al* (2023), "Adaptive R-Peak Detection on Wearable ECG Sensors for High-Intensity Exercise," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 3, pp. 941-953.

[8] Hao *et. al* (2022), "Adaptive R-Peak Detection Algorithm Based on Brown Exponential Smoothing Model," in *IEEE Access*, vol. 10, pp. 114355-114363.

[9] Selzler *et. al* (2021). Impact of Subject-specific Training Data in Anxiety Level Classification from Physiologic Data.1-6.

[10] Zhao *et. al* (2023), "Stress Detection via Multimodal Multitemporal-Scale Fusion: A Hybrid of Deep Learning and Handcrafted Feature Approach", IEEE Sensors Journal.

[11] Tripathy *et. al* (2023), "Detection of Different Stages of Anxiety from Single-Channel Wearable ECG Sensor Signal Using Fourier–Bessel Domain Adaptive Wavelet Transform",IEEE Sensors Letters,7(5),1–4.

[12] Vulpe-Grigoraşi *et. al* (2021), "A Neural Network Approach for Anxiety Detection Based on ECG," *International Conference on e-Health and Bioengineering (EHB)*.

[13] O'Toole *et. al* (2014), "Assessing instantaneous energy in the EEG: a non-negative, frequency-weighted energy operator" Annu Int Conf IEEE Eng Med Biol Soc. 2014;pp.3288-91. doi: 10.1109/EMBC.2014.6944325.

[14] Pan *et. al* (1985), "A real-time QRS detection algorithm," IEEE Trans. Biomed. Eng., vol. BME-32, no. 3, pp. 230–236.

[15] Aziz *et. al* (2021), "ECG-based machine-learning algorithms for heartbeat classification", *Sci Rep* **11**, 18738.

[16] Mandala *et. al* (2022),"Feature Extraction on Multi-Channel ECG Signals using Daubechies Wavelet Algorithm," *2nd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA).*