HIGH SCHOOL EDITION
Journal of Student Research
Volume 14 Issue 1 (2025)

Enhancing the Reliability of Electrocardiogram Signals for Stress Management: Learning Transferable Models from High-Resolution Electroencephalogram Supervision

Yerin Cho[1] and Esther Dettmar[#]

[1]Emma Willard School, USA
[#]Advisor

ABSTRACT

Stress, though often overlooked, is a significant driver behind the high suicide rates in South Korea, which leads among OECD countries with 24.1 suicides per 100,000 people. This issue is not unique to South Korea; in the United States, adolescent suicide rates have surged by 35% since 1999, now affecting 12 per 100,000 individuals. To address these alarming trends and prevent stress-related health issues, schools frequently rely on traditional checklist tests. However, these tests are often inaccurate, are time-consuming, and lack scientific rigor. Recent studies have started to explore the use of electroencephalogram (EEG) technology, which provides a scientific and quantitative measure of stress by closely monitoring brain activity. Despite its promise, electrocardiogram (ECG) signal, which also correlates with stress, has not been as thoroughly investigated. To address this problem, I propose an integrated approach that combines EEG and ECG assessments to develop a more reliable and cost-effective method for stress detection using machine learning. During the training phase, the proposed system takes both EEG and ECG signals as input and learns to map these signals into an emotion-related feature space. After training, the pre-trained network is then used to predict arousal and valence from ECG signals. Extensive experiments demonstrated that the proposed approach significantly improved performance, reducing RMSE by 8.24.

Introduction

One of the OECD (Organization for Economic Cooperation and Development) countries, South Korea has a tremendously high suicide rate that is skyrocketing and marks the first out of those countries. There are numerous factors that contribute to these statistics, but the underlying one is stress.

Stress management and quantification have become critical areas of focus in educational settings particularly as the pressures of academic life continue to rise. Traditionally, schools have relied on self-checklist assessments to gauge students' stress levels. While these methods provide some insight, they are often time-consuming, are ineffective, and lack scientific rigor. The subjective nature of self-reported checklists can lead to inaccuracies, making it very challenging for educators and mental health professionals to obtain an understanding of students' stress levels.

The primary technique to identify stress related conditions is EEG, a clinical instrument to monitor for abnormal brain activities. Furthermore, since the cardiovascular system has a high correlation with stress exposure, the ECG is also commonly utilized to detect the cause of abnormal heart conditions. The ECG monitors for cardiac rhythms using electrical charges. While EEG holds benefits of generating detailed data, it is not cost-efficient and not user-friendly in wearing the EEG instruments. On the other hand, the ECG is widely utilized in clinical settings and low cost. Both data are heavily impacted by stress, thereby allowing space to generate more efficient methods to track stress by combining two techniques.

Inspired by this, I propose a machine learning-based stress management system that utilizes both EEG and ECG data. The proposed method employs weakly supervised learning to detect ECG signals with EEG, leveraging its high temporal resolution in relation to stress. I also propose an innovative scenario using only ECG signals to detect abnormalities in individuals who may need mental health assistance, utilizing transfer learning.

The following chapters of this research paper are organized as follows: Chapter 2 provides background knowledge on ECG and EEG. Chapter 3 details the procedure of the proposed approach, while Chapter 4 demonstrates its effectiveness. Finally, Chapter 5 summarizes the paper.

Background Knowledge

Electroencephalogram

A diagnostic test, EEG monitors brain activities by using electrical charges. In figure 1, the patches attached to the patient's head are called the electrodes that facilitate data collecting in the EEG machine. The electrodes detect the brain's electrical signals and depict them through brainwave patterns.

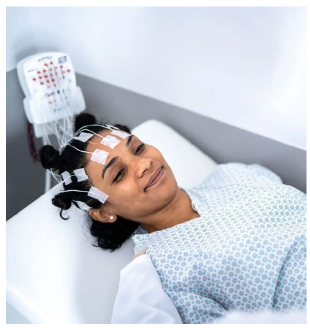

Figure 1. A patient being tested using EEG (Epilepsy Action 2024)

EEG is particularly effective at capturing stress levels in individuals, as it directly affects brain activity. By detecting brain activities with dense temporal resolution, numerous brain-related symptoms and disorders can be identified, such as epilepsy seizures, sleep disorders, brain tumors, and mood disorders. Taking advantage of this characteristic, I developed a system that utilizes EEG signals to determine whether individuals are experiencing stress.

Electrocardiogram

ECG uses a similar mechanism as the EEG but applies it into the cardiovascular system. Shown in figure 2, ECG obtains data through detecting electrical charges through the electrodes that are attached on the upper body and generate electrical waveforms: it provides information of the heart rate, heart rhythm problems, and other heart conditions, including heart attack, arrhythmias, chest pain, etc.

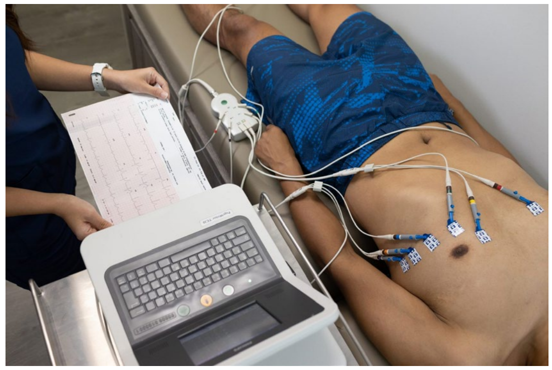

Figure 2. A patient being tested using ECG (Capital Heart Centre 2023)

Both ECG and EEG signals are used as inputs for the proposed stress management method. Even though the ECG is more widely accessible and familiar, ECG has indirect correlations with stress, making the data less evident compared to EEG. Combining these techniques will help enhance the availability of stress detection and management to the general public. A detailed explanation of the proposed system will be provided in Chapter 3.

Proposed Approach

The proposed stress management system consists of two main modules: an EEG and ECG representation learning network and an emotion regression network. The representation learning network takes both EEG and ECG signals as input, learning to map these signals into a shared emotion-related feature space. This network is designed to project features that are mathematically similar when the input EEG and ECG signals correspond to the same emotional state. The regression network, built on the pre-trained model from the representation learning phase, is then used to predict arousal and valence values. Chapters 3.1 and 3.2 provide detailed explanations of the procedures for each step.

EEG and ECG Representation Learning Network

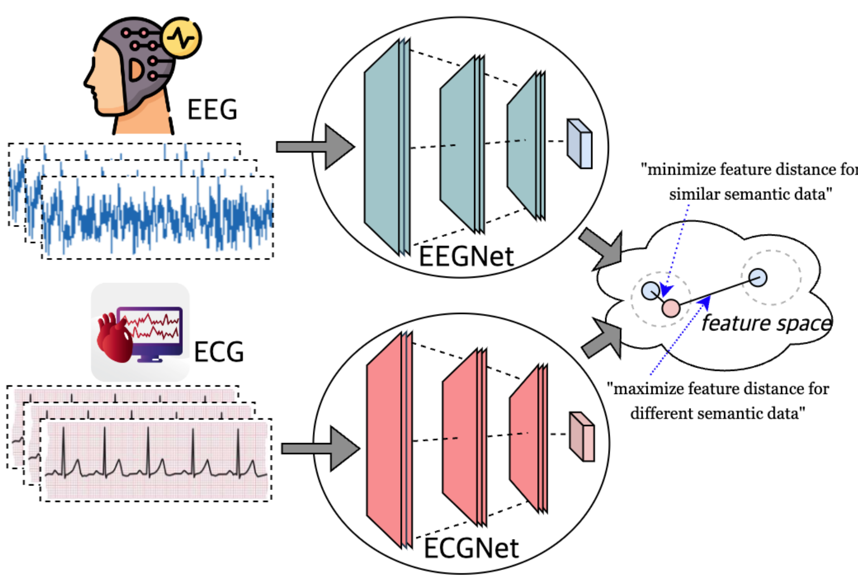

Figure 3. Structure of the proposed eeg and ecg representation learning network

Figure 3 illustrates the representation learning network. The representation learning network is divided into ECGNet and EEGNet, which are 1-D convolutional neural networks designed to extract features. ECG and EEG signals, categorized into four emotional states (happy, sad, fear, and neutral), are fed into ECGNet and EEGNet, respectively. The generated features should be mathematically similar if the signals belong to the same category and dissimilar if they belong to different categories. To apply this hypothesis into the loss function, the difference between the two features is first calculated. I used the cosine-based similarity measure, as described in Equation 1.

Equation 1: Similarity measure function

$$Sim(a,b) = \frac{a * b}{|a| \times |b|}$$

$$Sim(a,b) = \frac{a_1b_1 + a_2b_2 + \cdots + a_kb_k}{\sqrt{a_1^2 + a_2^2 + \cdots + a_k^2} \times \sqrt{b_1^2 + b_2^2 + \cdots + b_k^2}}$$

To explain the mechanism behind the equation, if there are two points of A and B on a plain, the addition of the two points' product is denoted by *; the root-squared addition with squared data is the norm denoted by |.|.

Equation 2: Softmax function

$$Softmax(S_k) = P_k = \frac{e^{S_k}}{\sum_j e^{S_j}}$$

Applying Equation 1 to EEG and ECG, the evaluation occurs to find correlation between numerous data sets. The obtained score, subsequently, converts to the scale of probability and substitutes into the softmax function in Equation 2. In general, the softmax function converts data into probability on a scale of 0 to 1. The softmax function's K represents real numbers; the numerator has the exponential form; and the denominator has the sum of data of all real numbers'. By making the data exponential, the logarithmic probability is generated.

Equation 3: Cross entropy loss function

$$L_{ce} = -log_e P$$

Using the softmax function's output of probability, the cross entropy loss function in Equation 3 evaluates how good the network model is in doing tasks and makes regression and classification according to the outputs. Making the graph as the negative ln (log e) function allows it to be easily read on a scale of 0 to 1, from high error to near to 0 error.

Emotion Regression Network

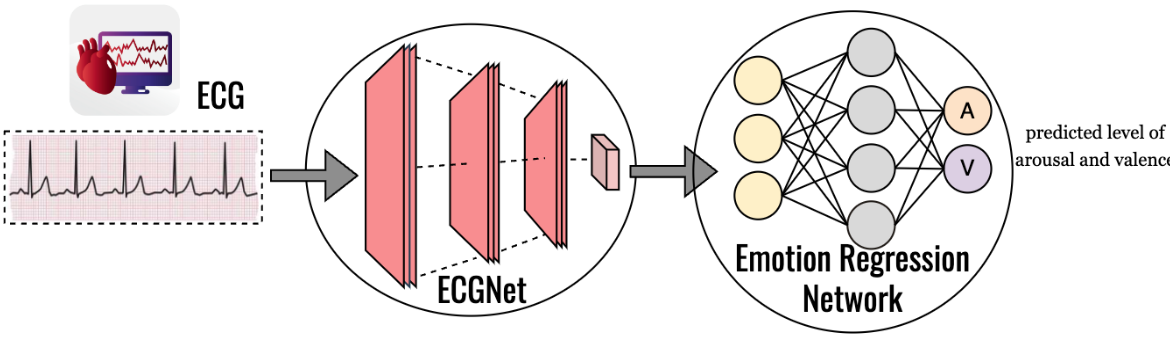

Figure 4. Structure of the proposed emotion regression network

Figure 4 illustrates the proposed Emotion Regression Network, where the pre-trained ECGNet is used as a feature extractor. I developed a two-layer neural network to predict arousal and valence values. Arousal indicates the intensity of an emotion, while valence reflects its positive or negative nature. Since the ECGNet is pre-trained with a representation learning network, it can extract robust and rich features from the input ECG signal. The effectiveness of the proposed approach is evaluated in Chapter 4. The Emotion Regression Network is trained using the mean squared error function, as explained in Equation 4.

Equation 4: Mean square error function

$$L_{mse} = \frac{1}{N}\sum_{i}(A_i - \widehat{A}_i)^2 + (V_i - \widehat{V}_i)^2$$

Based on the prediction through the Emotion Regression Network, the evaluation on whether its prediction is accurate takes place. Here, Equation 4 portrays a new type of loss function that measures the degree of error in the model. Gathering data from the regression network prediction, A represents arousal, and $\hat{A}$ shows the predicted value of arousal. Similarly, V represents valence, and $\hat{V}$ shows the predicted value of valence. After subtracting $\hat{A}$ from A and squaring the difference, adding the same of valence generates one predicted error value. The sigma adds all the error values, and the 1/N helps produce the mean of the squared errors.

The effectiveness of the proposed eeg and ecg representation learning will be demonstrated through extensive experiments, such as the series of machine learning techniques and the corresponding loss functions, and explained in Chapter 4.

Experimental Results

In order to extract data of electrocardigam's effectiveness in detecting emotion, data adapted from the Young Adult's Affective Data (YAAD), 25 subjects' emotions of happiness, sadness, fear, surprise, anger, disgust, and neutrality were tracked through ECG and Galvanic Skin Response (GSR) while watching emotion-inducing films by using ECG (Dar et al. 2022). Among the researchers' emotions, four emotions of happiness, sadness, fear, and neutrality and their arousal and valence (A&V) are chosen and tracked, respectively. Similarly, the EEG's efficiency of extracting emotional detection was adapted from the SEED-IV dataset and is portrayed in Figure 6, where 20 subjects' four emotions and A&V were detected (Zheng et al. 2019) while watching various films as displayed in Figure 5.

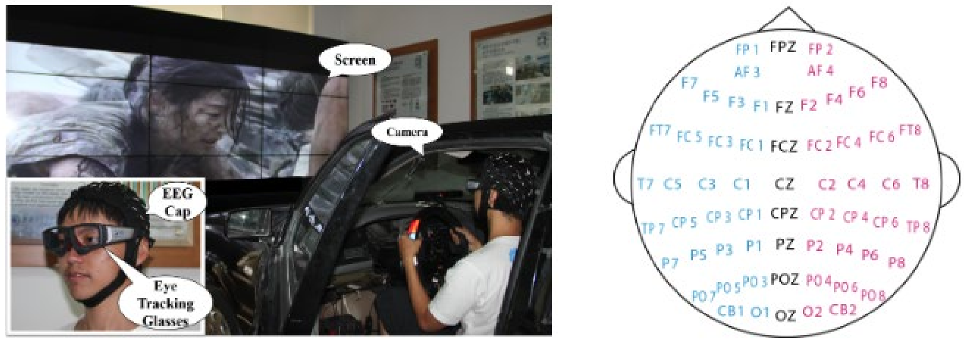

Figure 5. EEG electrode layout from SEED-IV (Zheng et al. 2019)

After obtaining both ECG and EEG data, EEG input and ECG input are generated. In order to develop a novel cost-efficient and more accurate method, both inputs are required. Combining the intricacy of EEG input and ECG input, an emotion regression network is established. As shown in Table 1, five emotion classification convolutional neural networks of VGG-17, ResNet-18, ConvNext, Xception, and ResNet-152 measure four categories of inference metric for the newly developed emotion regression network: accuracy, recall value, precision, and F1 score. The accuracy shows the correctness of the proposed model; recall represents the model's frequency of predicting positive data from all data; precision measures the model's ability to detect true positives from all positive data; and F1-Score is the mean of precision and recall values that evaluates the predictive ability of the model. In Table 1, ResNet-152 demonstrates the highest in all four inference metric sectors, hinting its efficiency and overall prime performance.

Table 1. Each CNN's Inference Metric Value in Table

Emotion Classification	**Accuracy**	**Recall**	**Precision**	**F1-Score**
VGG-16 (Simonyan et al. 2014)	76.57	76.44	76.60	76.52
ResNet-18 (He et al. 2016)	76.72	76.68	76.01	76.34
ConvNext (Liu et al. 2022)	78.66	78.27	78.53	78.40
Xception (Fran et al. 2017)	78.69	78.62	78.68	78.65
ResNet-152 (He et al. 2016)	81.91	81.84	81.96	81.90

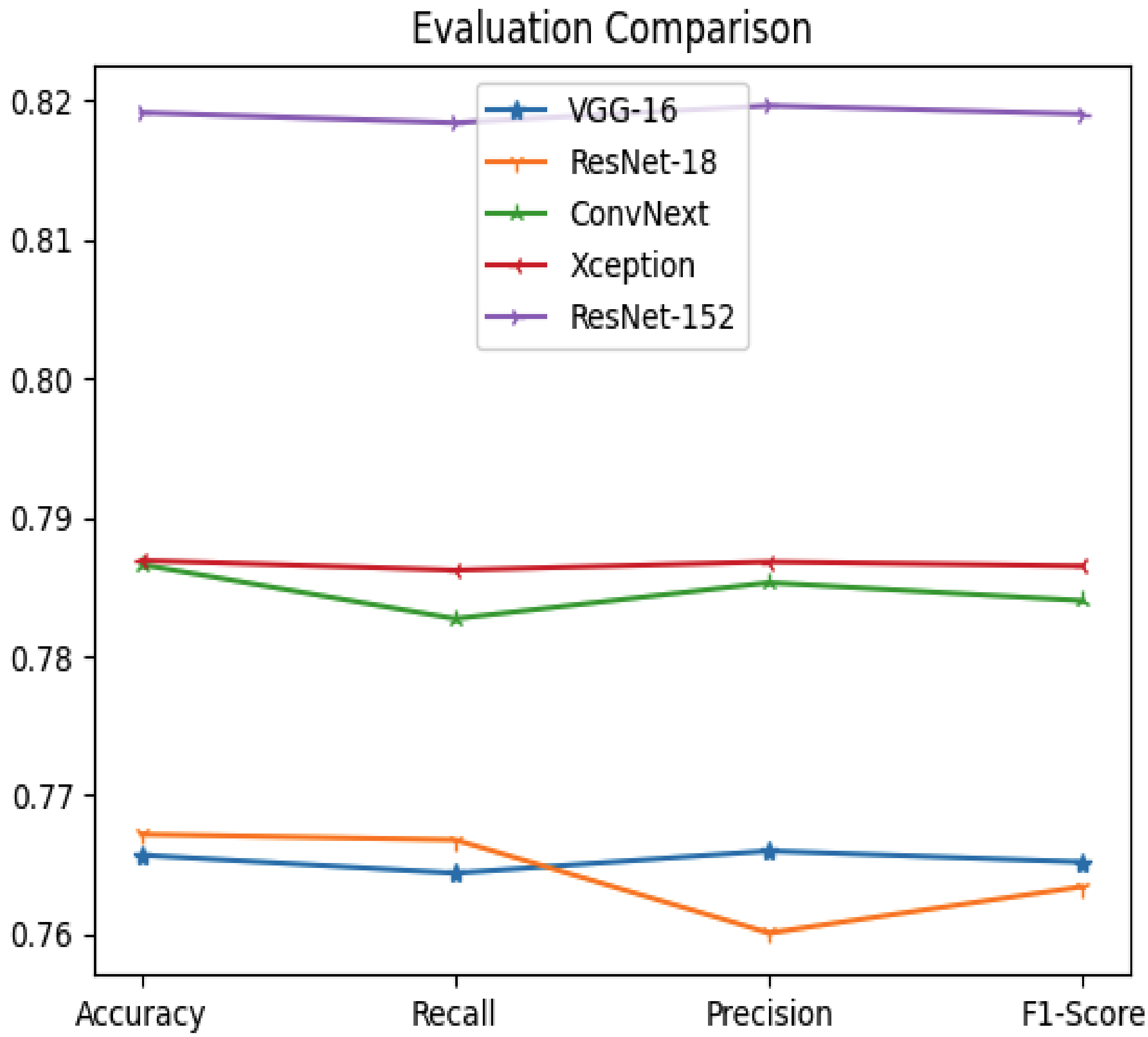

Figure 6. Evaluation Comparison Across Five CNNs in Graph

Along with creating a confusion matrix to evaluate the efficiency of a method, evaluation comparison is held to compare the different inference metric values across the CNNs. The visualization of Table 1 is Figure 6. Among the five CNNs, ResNet-152 demonstrates the highest in all sectors.

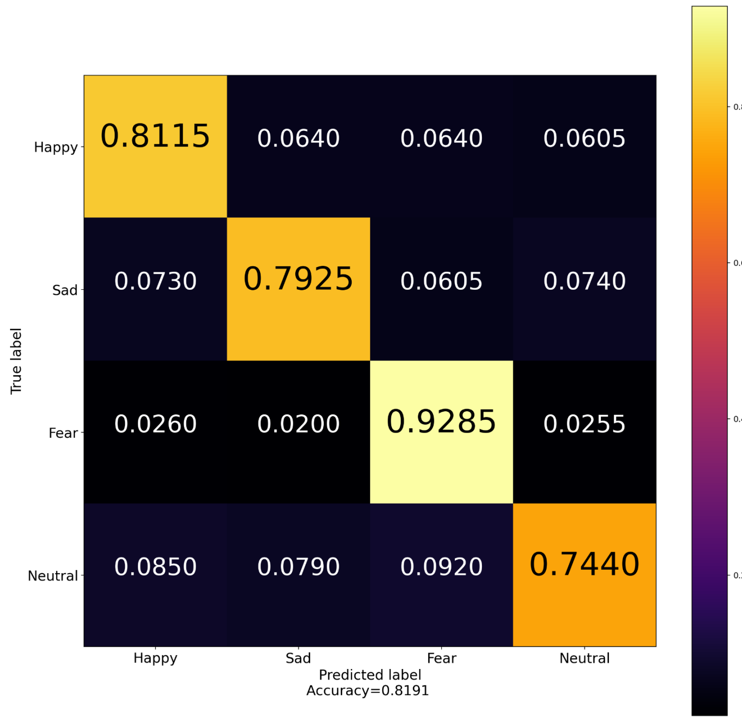

Figure 7. Confusion Matrix (ResNet-152's each category's true positive ratio analysis)

Based on the four emotion classification data, a confusion matrix in Figure 7 is generated in order to evaluate the performance of the conventional method of just ECGNet and combined EEGNet, showing 81% for happiness, 79% for sadness, 93% for fear, and 74% for neutrality. By analyzing data from the confusion matrix, other parameters of accuracy, precision, and recall are yielded.

Equation 5: Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_i (y_i - \hat{y}_i)^2}$$

Equation 5. demonstrates the root mean squared error that shows the deviation in the actual data and the predicted data of the proposed model. The y is the predicted value, and the $\hat{y}$ is the actual value. N represents the number of data points. Adding the squared difference and dividing the value into the number of data points lead to the RMSE. Shown across Table 2, the VGG has 2.72 RMSE; ResNet-18 has 2.54 RMSE; ConvNext has 2.04 RMSE; Xception has 1.95 RSME; and ResNet-152 has the lowest RMSE of 1.74. According to Figure 6 and Figure 7, ResNet-152 overall has the highest accuracy and lowest RMSE, indicating that the model generated by the CNN is the most adequate.

Table 2. RMSE for Each CNN in Table

Arousal / Valence Prediction	RMSE
VGG-16	2.72

ResNet-18	2.54
ConvNext	2.04
Xception	1.95
ResNet-152	1.74

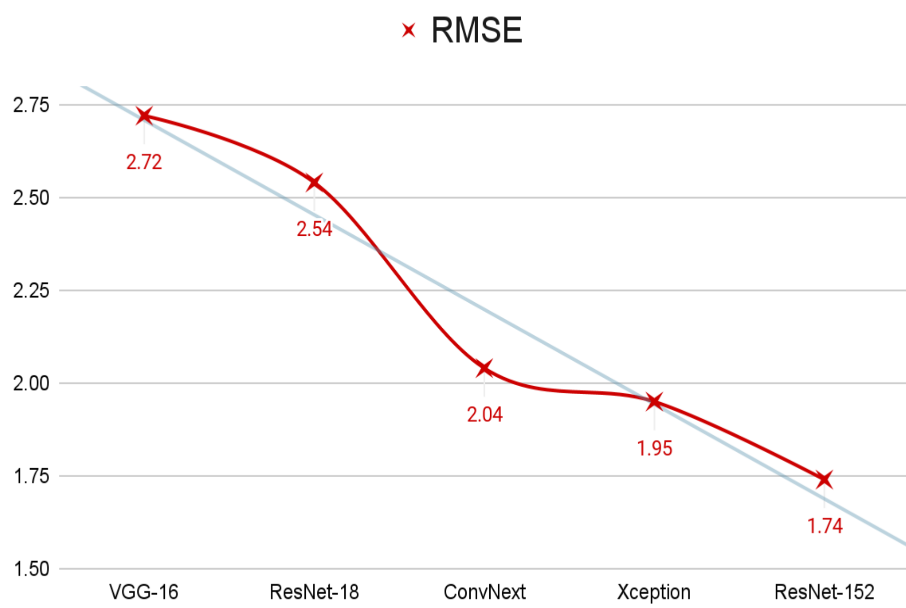

Figure 8. RMSE for Each CNN in Graph. Converting Table 2 to Figure 8 successfully indicates the vast difference of each CNN's operative accuracy.

Table 3. Accuracy of Proposed Method and ECGNet only Method in Table

Emotion Classification	**Accuracy (proposed)**	**Accuracy (ECG only)**
VGG-16	76.57	68.52
ResNet-18	76.72	69.04
ConvNext	78.66	71.25
Xception	78.69	72.54
ResNet-152	81.91	73.67

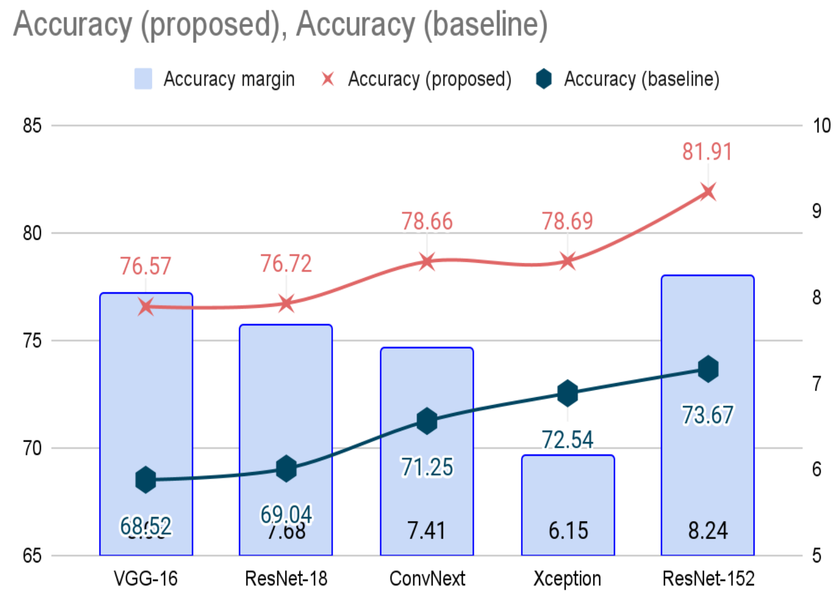

Figure 9. Accuracy of Proposed Method and ECGNet only Method in Graph

After AI Model Training in all five CNNs, the accuracy of the ECGNet and that of combined Net are compared in Figure 9. While the proposed method of combining EEGNet and ECGNet shows the highest of 81.91% and lowest of 76.57%, the ECGNet only method has the highest accuracy of 73.67% and the lowest of 68.52%. The accuracy margin highlighted in the blue sector ranges from 6.15% to 8.24%, designating the considerable difference in accuracy of the proposed method and that of just ECGNet.

Table 4. Difference in RMSE in Proposed Method and ECG only Method in Table

Arousal / Valence Prediction	RMSE (proposed)	RMSE (ECG only)
VGG-16	2.72	5.76
ResNet-18	2.54	5.18
ConvNext	2.04	4.56
Xception	1.95	4.26
ResNet-152	1.74	4.15

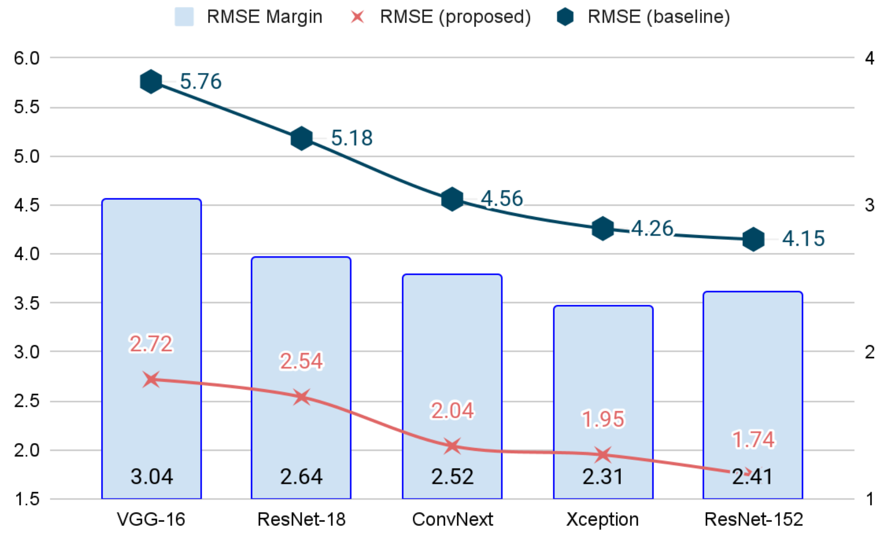

Figure 10. Difference in RMSE in Proposed Method and ECG only Method in Graph

Table 4 represents the RMSE in each CNN for the proposed method and ECG only method. Having the table more readable, Figure 10 demonstrates the difference in RMSE in the two methods. While the range of RMSE in the proposed method denoted by X is from 2.72 to 1.74, that of ECG only method denoted by dark blue ranges from 5.76 to 4.15. The RMSE Margin in light blue visualizes the substantial difference between the RMSE. Taken altogether, the ResNet-152 shows the most accuracy and least RMSE.

Conclusion

In this research, I proposed an innovative method to combine ECG and EEG data to successfully detect stress in a more accessible and cost efficient way by utilizing machine learning. Throughout the experiment, the best output was generated by the ResNet-152, yielding about 81.91% accuracy in detecting emotions and having about 1.74 of RMSE. These data indicate the reliability of the method, and the ablation study conducted demonstrates that the combination of ECG and EEG indeed has potential to help future research into stress detection. As such, developing a model that has this proposed feature but is more wearable will contribute vastly to the medical field, and I will be conducting experiments using this technique that connects to the real world scenario.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Capital Heart Centre. (2023, Mar 13). "*What Does An Abnormal ECG Mean?*": Capital Heart Centre https://www.capitalheart.sg/what-does-an-abnormal-ecg-mean/

Dar, M. N., Rahim, A., Akram, M. U., Khawaja, S. G., & Rahim, A. (2022, May). YAAD: young adult's affective data using wearable ECG and GSR sensors. In 2022 2nd International Conference on Digital Futures and Transformative Technologies (ICoDT2) (pp. 1-7). IEEE.

Epilepsy Action. (2024, May 1). "*EEG tests for epilepsy*": Epilepsy Action. https://www.epilepsy.org.uk/info/diagnosis/eeg-electroencephalogram

Fran, C. (2017). Deep learning with depth wise separable convolutions. In IEEE conference on computer vision and pattern recognition (CVPR). https://doi.org/10.48550/arXiv.1610.02357

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://doi.org/10.48550/arXiv.1512.03385

Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A convnet for the 2020s. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 11976-11986).

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

Zheng, W. L., Liu, W., Lu, Y., Lu, B. L., & Cichocki, A. (2018). Emotionmeter: A multimodal framework for recognizing human emotions. IEEE transactions on cybernetics, 49(3), 1110-1122.