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

An Instant Depression Screening Method via Valence and Arousal Prediction from Electroencephalogram and Galvanic Skin Response Data via Unsupervised Representation Learning

Sunwoo (Ria) An[1] and Koby Osei-Mensah[#]

[1]Loomis Chaffee School, USA
[#]Advisor

ABSTRACT

Globally, 280 million are suffering from depressive disorders, one of the most common mental disorders referring to long periods of depressed mood that affect life negatively, often leading to suicide. Adolescents are especially vulnerable, with 35% of youth aged 12 to 17 in the United States having major or severe depression. The rate of depression in teenagers has doubled over the past decade, along with the youth suicide rate. Currently, self-report questionnaires such as the Patient Health Questionnaire are utilized to diagnose depression. Self-report questionnaires may be accurate for other population groups that are fully aware of the state of their mental health, and are willing to seek support. However, adolescents do not have the same ability to recognize their symptoms nor answer honestly in school-wide screening tests. In order to solve the aforementioned problem, a method that captures the true emotion without self-reports based on electroencephalogram (EEG) and galvanic skin response (GSR) is proposed. The novel system is composed of representation and transfer learning steps for the extraction of emotion-related features from EEG signals, and a 1D convolutional neural network for extraction in GSR signals. The extracted features are merged and outputted as points on the arousal-valence graph, where emotions can be detected. Through extensive experiments, the proposed model demonstrated exceptional performance. The best MAE was 18.4 when without GSR, 18.8 when without representation learning, and 17.2 when without the proposed equation, but improved to 16.4 in the proposed model.

Introduction

Globally, 280 million people have depressive disorder according to the World Health Organization (WHO, 2023). Depressive disorder, commonly known as depression, is one of the most common mental disorders and can lead to suicide. It refers to a depressed mood and loss of joy for long periods of time, and affects all aspects of life negatively. It results from an intricate interaction of social, psychological, and biological factors. Teenagers, defined as youth between 10 and 19 years old, are vulnerable to mental health problems due to their physical, emotional and social changes. In the United States, 20% of youth aged 12 to 17 have major depression, with 15% experiencing severe depression (Mental Health America, 2024). The rate of depression in teenagers have doubled over the past decade, along with the youth suicide rate which rose 62% from 2007 to 2021 (Wilson & Dumornay, 2022).

Despite the high rate of depression worldwide, the patients often feel misunderstood by society, pushing depressed individuals to hide their mental health conditions (C. Esposito et al.). A substantial amount of teenagers hide their depression and do not receive the appropriate treatment in fear of what others will think. The negative perception towards depression often prevents many depressed individuals from being diagnosed properly, leading to neglection and worsening of their mental health.

Currently, diagnosing depression requires a depression screening test. The Patient Health Questionnaire is a self-administered version of the PRIME-MD tool for common mental health disorders administered by health care professionals. The PHQ-9 and PHQ-2, which are components of the Patient Health Questionnaire, offer psychologists concise, self-administered tools for screening depression. (American Psychology Association) For children and adolescents, the Children's Depression Inventory (CDI) is a modification for the Beck Depression Inventory for adults. It assesses depression severity from children and adolescents 7 to 17 years old. Forms for parents (17 items), teachers (12 items) and a self-report (28 items).

The common feature between all existing depression diagnoses is that it is a self-report questionnaire. This might be accurate for the older population who can fully recognize their symptoms and are willing to accept the diagnosis and receive treatment. However, for adolescents, they are not able to do the same and have a higher chance to be dishonest in their surveys. Therefore, a device that captures the truthful emotion of the adolescents and are not self-reports is necessary.

To gain evidence for the claim above, a group of 50 high school students attending school either in the United States or South Korea was surveyed online about depression screening.

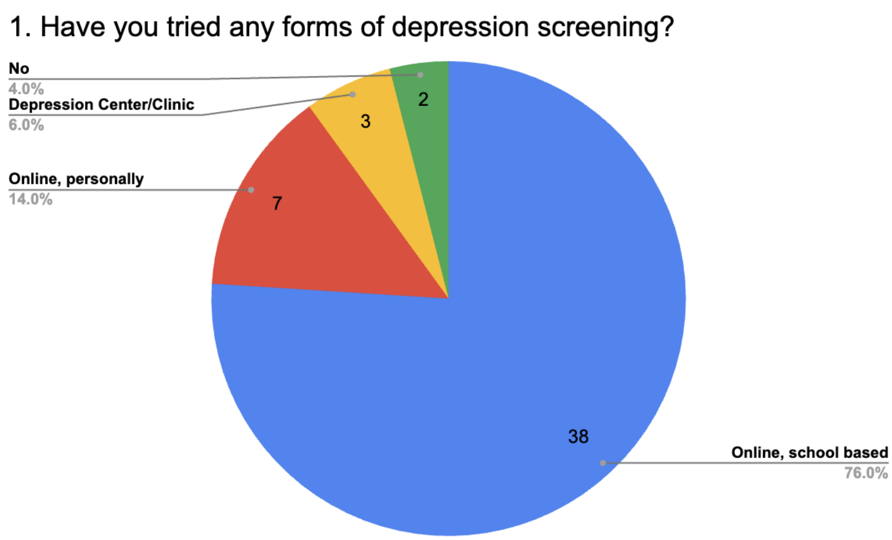

Figure 1. Depression Screening Survey Question on Types of Screening

As seen in Figure 1, 76% of the respondents had completed online, school-based depression screening questionnaires, 14% online, personally, 6% in depression centers or clinics, and 4% never.

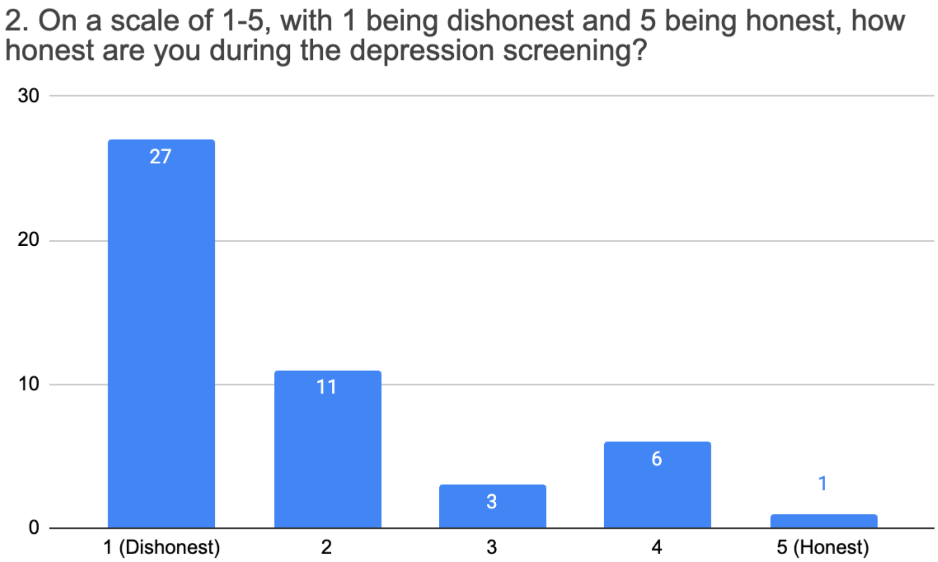

Figure 2. Depression Screening Survey Question on Honesty

The 48 students who had completed any form of depression screening questionnaires mostly answered that they are completely dishonest during the screening. Based on further analysis of the survey data, 36 students out of the 38 students who took the online, school-based tests responded that their level was either 1 or 2 (Dishonest).

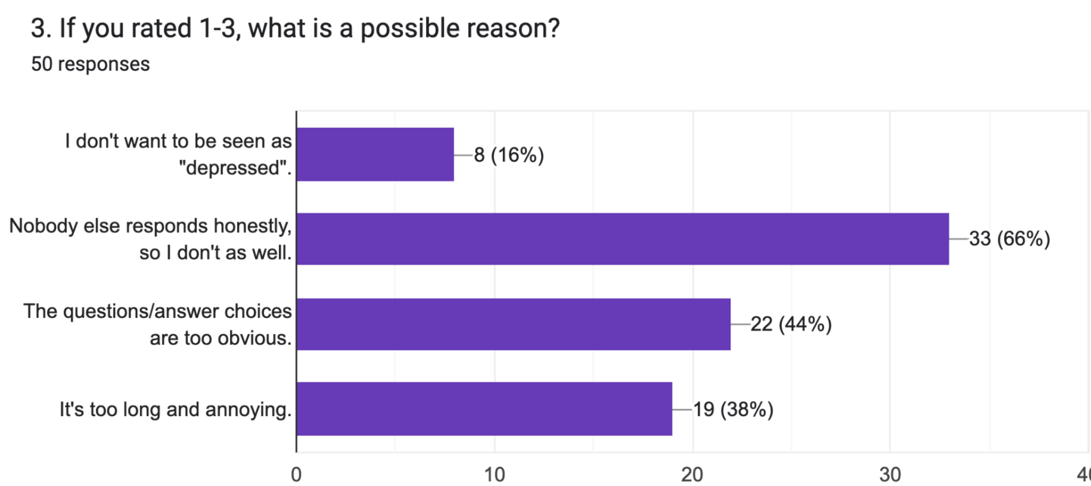

Figure 3. Depression Screening Survey Question on Potential Reasons for Dishonesty

The highest potential reason, with 33 agreeing, for dishonesty during the depression screening questionnaires was the thought that no one else responded honestly, so they did not as well. This suggests that the tests are not taken as seriously as they should be, and a large amount of high school students who take the school-wide tests may be dishonest. The other common reasons were that the questions and answer choices were too obvious, as well as the fact that the test was too long and annoying. These reasons suggest the need for non-survey based, emotionally accurate tests.

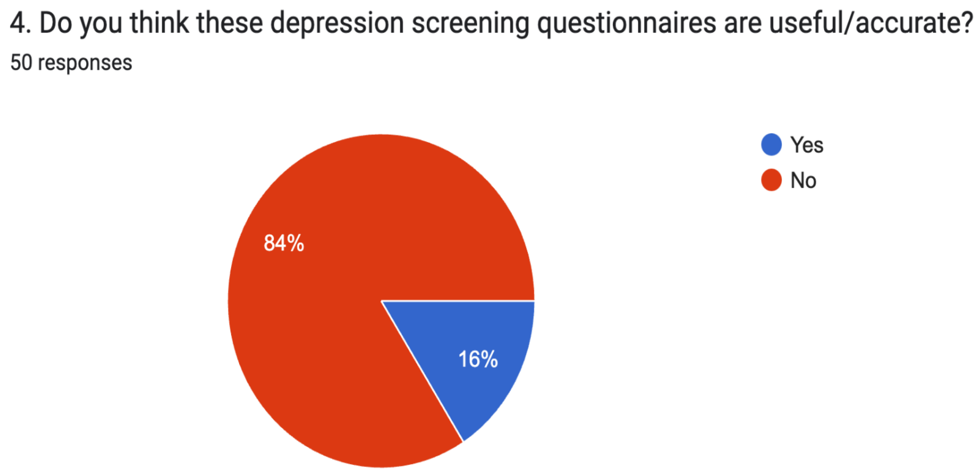

Figure 4. Depression Screening Survey Question on Usefulness of Depression Screening Questionnaires

Lastly, when the respondents were questioned on the usefulness and accuracy of the depression screening questionnaires, 84% selected "no". The high number once again showcases the need for an accurate depression screening method.

In order to address the issue of low credibility based on honesty that arises from the method of self-report questionnaires used to screen depression, this paper proposes an instant depression screening method ia valence and arousal prediction from electroencephalograms (EEGs) and galvanic skin response (GSR) data with unsupervised representation learning.

In brief, there are three main networks in the proposed method. Two inputs, EEG signals and GSR signals are provided, and the final output is the arousal and valence value prediction (Chapter 2.4), named AV-Prob. First of all, EEG signals, converted to topographical maps, are provided as an input for TOP-N, a two-dimensional neural network that extracts emotion-related feature vectors. GSR signals are inputted into GSR-N, a one-dimensional neural network that extracts emotion-related feature vectors. The emotion-related feature vectors from TOP-N and GSR-N are combined and used as an input for the third network, E-RN. It is a regression network that outputs the AV-Prob, the arousal and valence value prediction.

Related Work

Galvanic Skin Response

Galvanic Skin Response (GSR) is a method to measure the electrical conductivity of the skin in response to some stimuli (Science Direct, 2023). When humans experience something emotional, both positive and negative, the sweat glands are triggered unnoticeably, increasing the conductivity of the skin to electricity.

Rapid changes in skin conductance during an emotional trigger is instinctive, and is impossible to consciously control as the sweat glands are part of the autonomic nervous system. This fact makes GSR an effective tool when detecting the strength of emotions without the need of worrying if the participant is intentionally hiding their feelings.

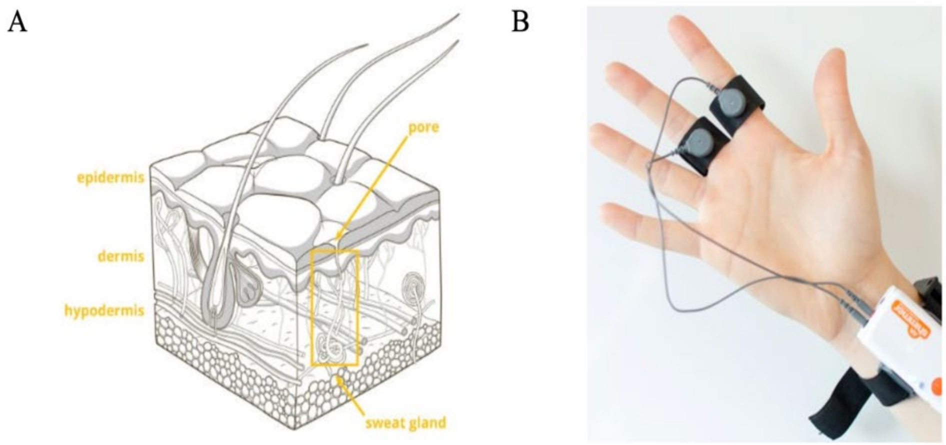

Figure 5. (a) An enlarged diagram of the skin (b) Galvanic Skin Response device (iMotions 2024)

When a person experiences emotional arousal, the sympathetic nervous system activates the sweat glands in the skin, increasing the moisture level on the skin surface as shown in Figure 2(a). The moisture level affects the skin's ability to conduct electricity, with dry skin having lower conductance (higher resistance) and moist skin having higher conductance (lower resistance). The GSR sensor measures the electrical conductance between the electrodes placed on two points of the skin. Usually expressed in microsiemens (μS), higher conductance signifies higher sweat gland activity, as well as higher level or emotional arousal.

To utilize a GSR device, the sensors are attached to the index and middle fingers with Velcro as observable in Figure 2(b). The device is connected to a computer via Bluetooth. The device is remarkably unobtrusive and is easily forgotten by the participant. Participants should try to breathe normally and sit comfortably to reduce their movement to the lowest without engaging in conversation.

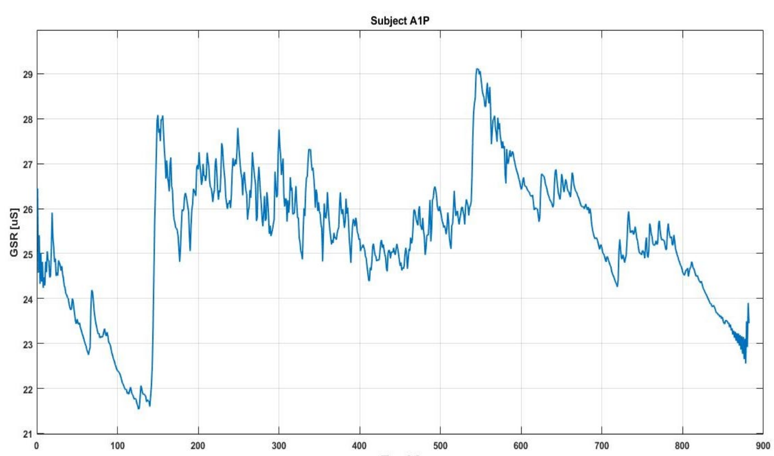

Figure 6. Sample GSR signal (R. Martinez et al.)

There are two components of the GSR signal: the tonic level and phasic response. The tonic level, which varies slowly over a longer amount of time, is the baseline of skin conductance, and is influenced by factors such as room temperature or overall stress levels of the participants. The second component is the phasic response: rapid changes in skin conductance. In the GSR signal, they are observable as peaks or sudden spikes, occurring in response to a rise in arousal or stress levels.

The analysis of GSR is done by measuring the number of phasic responses, or "peaks", over time in the graph of the signals. A peak is defined as a burst in the phasic response approximately 1 to 5 seconds after exposure to emotional stimuli. Higher number of peaks mean the arousal was greater during the experience (Science Direct, 2023).

Electroencephalography

Electroencephalograms (EEG) is a non-invasive neurophysiological technique that detects abnormalities in the brain waves or in the electrical activity of the brain using small, metal discs (electrodes) attached to the scalp. Neurons, which make up the brain's electrical charge.

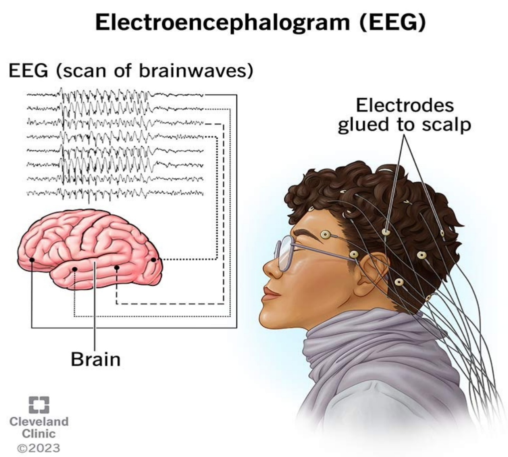

Figure 7. EEG device and signal sample diagram (Cleveland Clinic 2024)

When the wave reaches the electrodes on the scalp, they can push or pull the electrons located in the metal of the electrodes. Since metal conducts the interaction of electrons easily, the difference is electrically charged through membrane transport proteins that pump ions across their membranes. They incessantly exchange ions with the extra-cellular milieu, pushing out countless ions simultaneously. This causes volume conduction, which is where ions released push their nearest ions, and those ions the next bordering ones, and as such in a wave as oppositely charged ions repel each in push or pull voltages between any two electrode are measured by voltmeters; making up the EEG signals as time passes (Tatum Wo et al., 2008). The charges are then amplified and graphed for the healthcare provider to interpret the readings. When epilepsy is present, seizure activity will appear as rapid spiking waves on the EEG recordings (Johns Hopkins University, 2023).

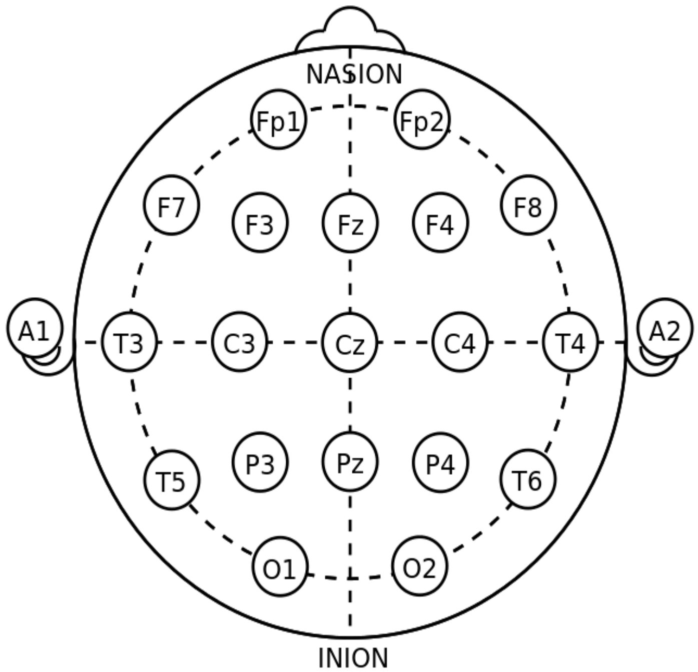

Figure 8. Electrode locations of International 10-20 system for EEG (electroencephalography) recording

The International 10-20 system is used to place the nodes on the scalp. The 10-20 system is based on the relationship between the location of an electrode and the underlying area of the cerebral cortex. As shown in Figure 1, each point on the scalp indicates a possible electrode position. Each site has a letter that identifies the lobe and a number or letter to identify the hemisphere location. The letters representing the lobes stands for pre-frontal (Fp), frontal (F), temporal (T), parietal (P), occipital (O), central (C), and mastoid process (A). The electrodes with even numbers (2,4,6,8) refer to the placement on the right side of the head, whereas odd numbers refer to those on the left. The letter 'z' represents the center vertical line of the scalp (Journal of Clinical Neurophysiology, 1991).

Previous Depression Screening Methods

The most commonly used instruments used for depression screening are PHQ-9 surveys. It normally takes two to five minutes to complete, and has demonstrated 61% sensitivity and 94% specificity for mood disorders in adults, and 89.5% sensitivity and 77.5% specificity in adolescents. (American Family Physician, 2023) One has to take into consideration that the accuracy is solely based on those who were both willing to take a depression test and were honest on the questionnaire. The total score indicates depression severity of the patient.

PATIENT HEALTH QUESTIONNAIRE-9 (PHQ-9)

Over the last 2 weeks, how often have you been bothered by any of the following problems? *(Use "✔" to indicate your answer)*	**Not at all**	**Several days**	**More than half the days**	**Nearly every day**
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself — or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3

FOR OFFICE CODING ___0___ + ______ + ______ + ______

=Total Score: ______

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult
☐	☐	☐	☐

Figure 9. PHQ-9 in patient health questionnaire (PHQ) screeners (Pfizer 2024).

Arousal-Valence Model

The arousal-valence model is a framework used to describe and categorize emotions. It plots emotions on a two-dimensional graph with arousal and valence axes. Arousal refers to the level of physiological and psychological activation or intensity of the emotion, and the valence refers to the positivity and negativity of the emotion (Neurodivergent Insights, 2024).

Arousal ranges from low arousal, with examples being emotional states such as calm and relaxed, to high arousal, such as the emotion of excitement or agitation. Valence ranges from negative valence, common emotions such as sadness or anger included, to positive valence like happiness and delight.

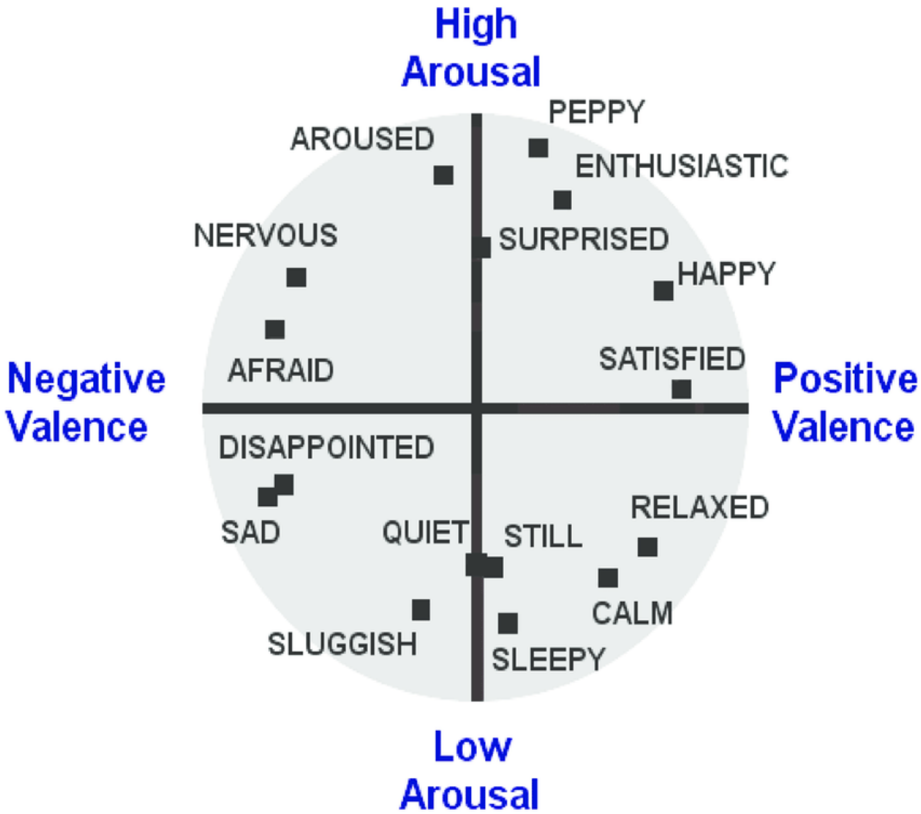

Figure 10. Arousal-Valence Model Examples of Emotions (J. Foucher et al.)

As shown in Figure 10, the four quadrants represent a diverse range of emotions. The first quadrant with high arousal and positive valence includes emotions like excited and surprised, the second quadrant with high arousal and negative valence includes emotions such as anger, fear, and anxiety, the third quadrant with low arousal and positive valence includes emotions like calmness and relaxation, and finally, the fourth quadrant with low arousal and negative valence includes emotions like sadness, depression, and boredom.

Methods

The proposed system utilizes topographical maps from EEG (Electroencephalogram) signals and GSR (galvanic skin response) signals. Each signal is processed through TOP-N (Topographical Map Network) and GSR-N (Galvanic Skin Response Network) respectively, which extract emotion-related feature vectors. In the overall architecture, these emotion-related feature vectors are merged and used as an input for the E-RN (Emotion Regression Network), with the output being the arousal and valence value as its prediction.

TOP-N: EEG Representation Learning

The TOP-N (Topographical Map Network) has two phases of learning: representation learning and transfer learning. The goal of representation learning is to create a CNN (convolutional neural network) that best projects emotions from the raw EEG signal by training its weight and bias.

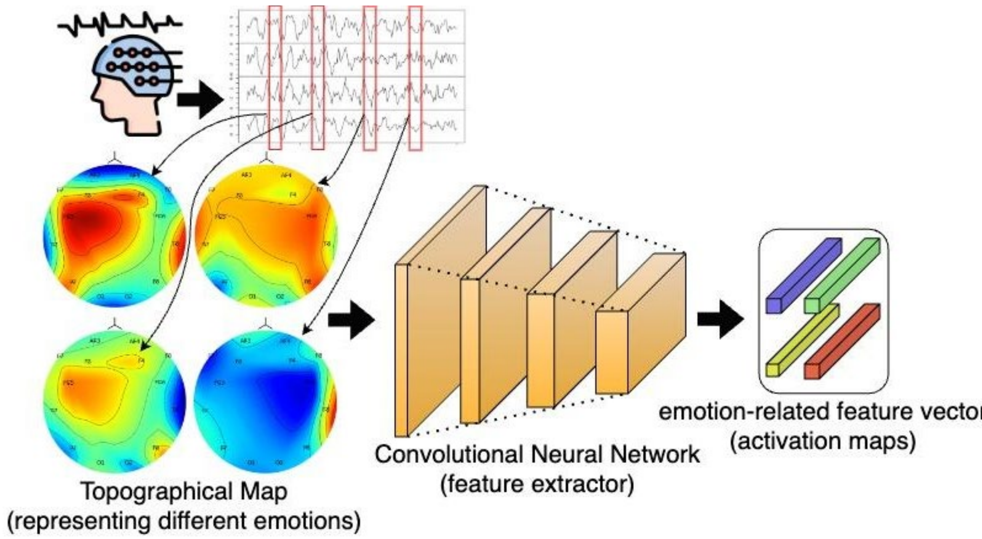

Figure 11. Architecture of TOP-N: EEG Topography Representation Learning

As seen in Figure 11, two-dimensional signals from the EEG device are converted into topographical maps, each map representing one epoch from the original signal. The maps are inputted into the unsupervised CNN, a type of neural network that is designed to learn patterns and features from data without labeled outputs paired with inputs. The output of the network is emotion-related feature vectors.

The underlying assumption is that feature vectors created from similar emotions should be projected by the CNN as mathematically similar. To achieve this, a new loss function was designed.

The first loss function calculates the cosine value between two feature vectors as similarity can be defined by the angular difference. As shown in Equation 1, the cosine value is calculated using the cosine similarity equation.

Equation 1. Cosine Similarity

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

Here, *a* and *b* denotes the two feature vectors representing different emotions being compared. The cosine similarity equation, which measures the cosine of the angle between two vectors, is used. It can be used to determine if the two vectors are pointing in a roughly similar direction. The value of *similarity(a,b)* will range from -1 to 1. -1 can be interpreted as the angle between the two vectors compared being 180°, or “least similar”, while 1 can be interpreted as the two vectors having the same direction with the angle between them being 0°.

In order to evaluate the accuracy of the classification, a cross-entropy loss function is used as shown in Equation 2. It measures the difference between predicted probabilities and the true labels of a given dataset.

Equation 2. Cross Entropy Loss

$$L_{(i,j)} = -log_e(\frac{e^{similarity(z_i,z_j)}}{\sum_k e^{similarity(z_i,z_k)}})$$

Here, z_i and z_j represent the feature vectors of the i-th and j-th data points, respectively. The value from *similarity*(z_i, z_j) is exponentiated, and divided by the sum of the exponentiated similarity of all possible pairs to normalize the exponentiated similarity into a value between 0 and 1. After the normalization, the negative logarithm transforms the resulting value into a comprehensible loss value. It is used to maximize the similarity between z_i and

z_j while minimizing the similarity between z_i and all other feature vectors z_k. Thus, the function effectively distinguishes the target pair (z_i, z_j) from other pairs.

TOP-N & E-RN: EEG Transfer Learning

The second phase of TOP-N (Topographical Map Network) is transfer learning, as shown in Figure 12. Transfer learning is a machine learning technique where knowledge gained from solving one problem is applied to a different but related problem. The ultimate goal is to build a CNN network based on the pre-trained network from the representation learning phase.

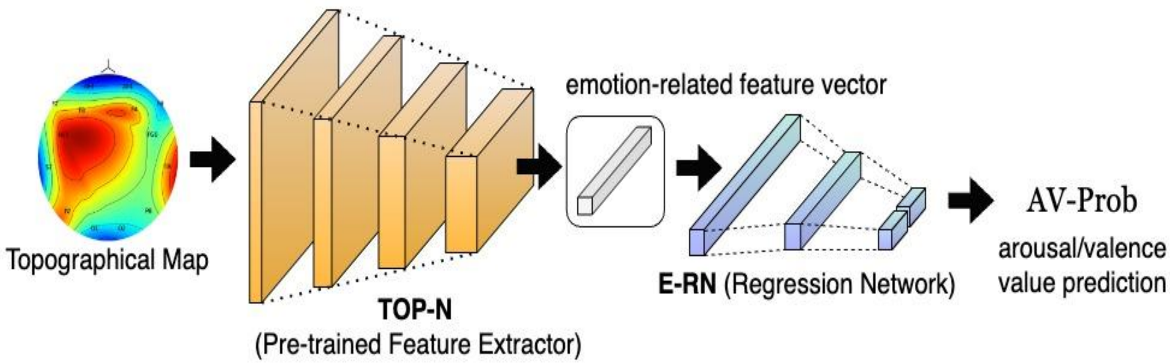

Figure 12. Architecture of TOP-N & E-RN: EEG Transfer Learning

In this step, the regression task, used to predict a continuous numerical value based on input data, predicts the arousal and valence values based on the emotion-related feature vector from the EEG topography map. The topographical map of the EEG signal is inputted into TOP-N, the pre-trained feature extractor, outputting the extracted emotion-related feature vector. These vectors are inputted into the E-RN (Emotion Regression Network), which outputs predicted arousal and valence values. Emotional valence and arousal refers to the extent to the positivity and negativity of an emotion and its intensity, respectively as mentioned in Chapter 2.4.

In order to evaluate the performance of E-RN, the square difference between the actual arousal-valence value in a two-dimensional graph and the predicted value is calculated. The lower the difference, the more similar the actual and predicted arousal and valence values are. As shown in Equation 3, Mean Square Error (MSE) function is used as a loss function.

Equation 3. Mean Squared Error Function

$$L_{av-prob} = \frac{1}{N}\sum\nolimits_i (arousal_i - \widehat{arousal_i})^2 + \frac{1}{N}\sum\nolimits_i (valence_i - \widehat{valence_i})^2$$

Here, N represents the number of feature vectors. The difference between each arousal and valence feature vector is squared, added and finally divided by the number of total feature vectors. In other words, the equation measures the average squared difference between the predicted and the actual target feature vector value.

TOP-N & GSR-N & E-RN: EEG Transfer Learning with GSR

The overall architecture of the proposed model is shown below in Figure 13. Two inputs, topographical maps of the EEG signals and the GSR signals are inputted in the TOP-N and GSR-N respectively, where the emotion-related feature vectors of each are extracted. The extracted emotion-related feature vectors are then inputted in the E-RN, finally outputting the AV-Prob, the arousal and valence value prediction.

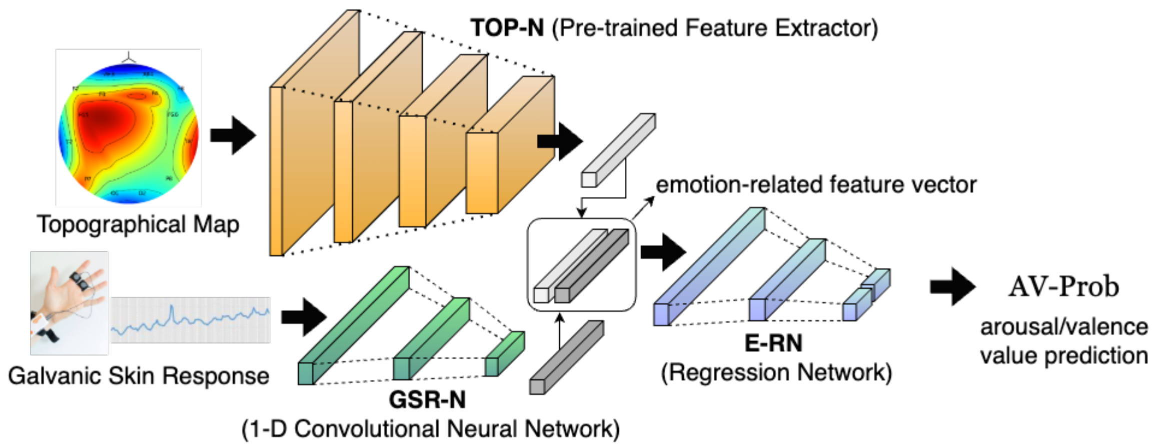

Figure 13. Overall Architecture of Proposed Model

As explained in Chapter 2.1, Galvanic Skin Response (GSR) is a method to measure the electrical conductivity of the skin in response to some stimuli (Science Direct, 2023). The GSR was chosen to be utilized in the proposed method to compensate for the high sensitivity to noise of EEG devices, as it can accurately display the strength of emotions the participant is feeling unconsciously.

Experimental Results

Datasets

Two datasets were used to train and test the proposed model.

The first dataset used is the SEED-IV dataset from Brain-like Computing & Machine Intelligence (BCMI). 72 film clips were chosen and categorized by a preliminary study as having a tendency to induce happiness, sadness, fear or neutral emotions. 15 subjects participated in the experiment as shown in Figure 14, with 3 sessions performed on different days for each, and one session containing 24 trials. In one trial, the participant watched one of the film clips, while his or her EEG signals and eye movements were collected with the 62-channel ESI NeuroScan System. Each session is sliced into 4-second nonoverlapping segments, with one segment regarded as one data sample during the model training.

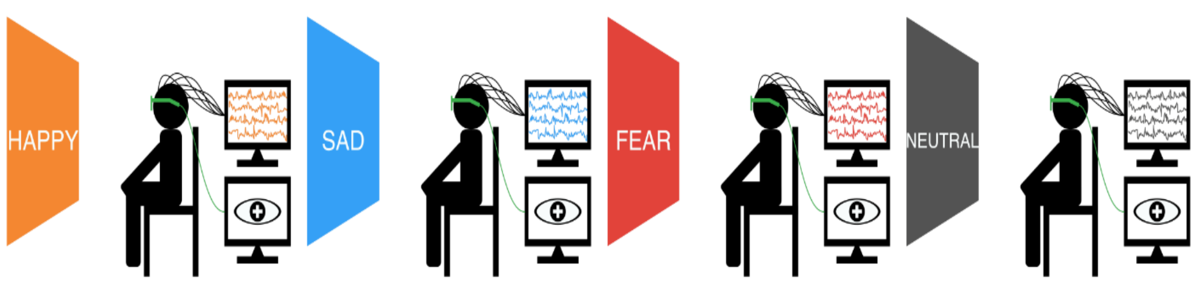

Figure 14. SEED IV Dataset Diagram (Zheng et al. 2018)

The second dataset used is DEAP (A Database for Emotion Analysis using Physiological Signals), a multimodal dataset for the analysis of human affective states. The electroencephalogram (EEG) and peripheral physiological signals of 32 participants were recorded as each watched 40 one-minute long excerpts of music videos. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance and familiarity.

Evaluation Metric & Comparison

To evaluate the performance of the E-RN, Equation 4. Mean Absolute Error (MAE) is utilized. It is a commonly used metric for the performance evaluation of regression models, and it measures the average magnitude of the errors between the predicted and actual values without considering the vectors' directions (positive or negative). A lower MAE value indicates high accuracy, implying that the predicted values are close to the actual values, and that the model is performing well. On the contrary, a higher value indicates low accuracy.

Equation 4. MAE (Mean Absolute Error)

$$MAE = \frac{1}{N}\sum_{i}|t_i - \hat{t}_i|$$

Here, y_i and x_i represent the actual and predicted values respectively. n represents the number of outputs, and — is the absolute error for the i-th observation. The absolute error for each output is added and divided by the number of outputs, measuring the average magnitude of errors without considering their directions. Overall, two evaluations based on model architectures were conducted with and without Galvanic Skin Response (GSR), and one more evaluation was conducted to compare both of them together.

Table 2 and Figure 15 below presents the MAE depending on four different neural network architectures without GSR usage. Firstly, VGG19 (Simonyan et al. 2014) includes 16 convolutional layers, 3 fully connected layers, and 5 max-pooling layers, with the total depth being 19 weight layers. Secondly, HRNet-W30-C consists of 82 layers. Thirdly, ResNet-101 consists of 101 layers, including the initial convolutional layer, followed by 33 bottleneck blocks, each having 3 layers (99 layers), as well as the final fully connected layer. Lastly, DenseNet-264 (Huang et al., 2017) contains a total of 264 layers. The evaluation consists of the MAE of arousal and valence predictions and the total value, the average of the previous two values.

As shown in Figure 15, the performance of DenseNet-264 is the highest compared to the other models. The performance is lowest when VGG19 is used, and increases as the depth of the neural networks increases. Considering that this evaluation is a baseline assessment, the results are evident.

Table 2. Mean Absolute Error (MAE) based on Architecture (without Galvanic Skin Response)

without Galvanic Skin Response	Mean Absolute Error (Arousal)	Mean Absolute Error (Valence)	Mean Absolute Error (Total)
VGG19 (Simonyan et al. 2014)	30.9	22.5	26.7
HRNet-W30-C	27.4	19.0	23.2

ResNet-101	22.3	17.2	19.8
DenseNet-264 (Huang et al. (2017)	21.3	15.5	18.4

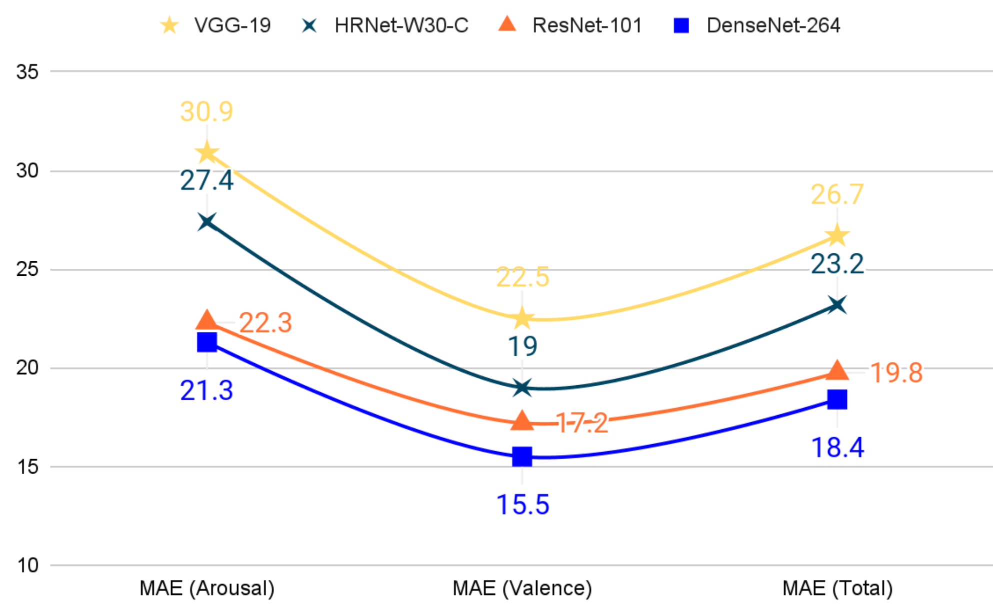

Figure 15. Graph of Table 1. Mean Absolute Error (MAE) based on Architecture (without Galvanic Skin Response)

Table 2 and Figure 15 below presents the MAE depending on four different neural network architectures with GSR usage. As shown in Figure 15, the performance of DenseNet-264 is the highest compared to the other models.

Table 3. Mean Absolute Error (MAE) based on Architecture (with Galvanic Skin Response)

	Mean Absolute Error (Arousal)	Mean Absolute Error (Valence)	Mean Absolute Error (Total)
VGG-19 (Simonyan et al. 2014)	27.5	20.7	24.1
HRNet-W30-C	23.7	17.5	20.6
ResNet-101	19.7	15.1	17.4

DenseNet-264 (Huang et al. (2017)	18.9	13.9	16.4

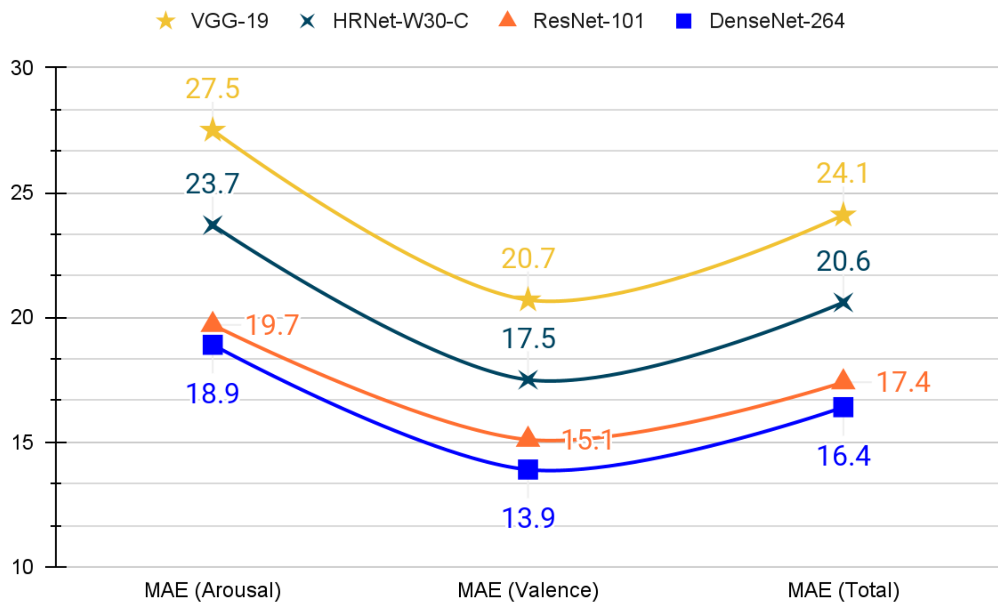

Figure 16. Graph of Table 3. Mean Absolute Error (MAE) based on Architecture (with Galvanic Skin Response)

The MAE value for each arousal, valence, and total prediction with and without the usage of GSR is compared in Figure 17. Overall, the values of the MAE are lower when GSR is included in the model, indicating that utilizing GSR enhances the proposed model's performance. Using mathematical logic and multimodal data, the proposed method improves the consistency of the model's training, thus improving the accuracy.

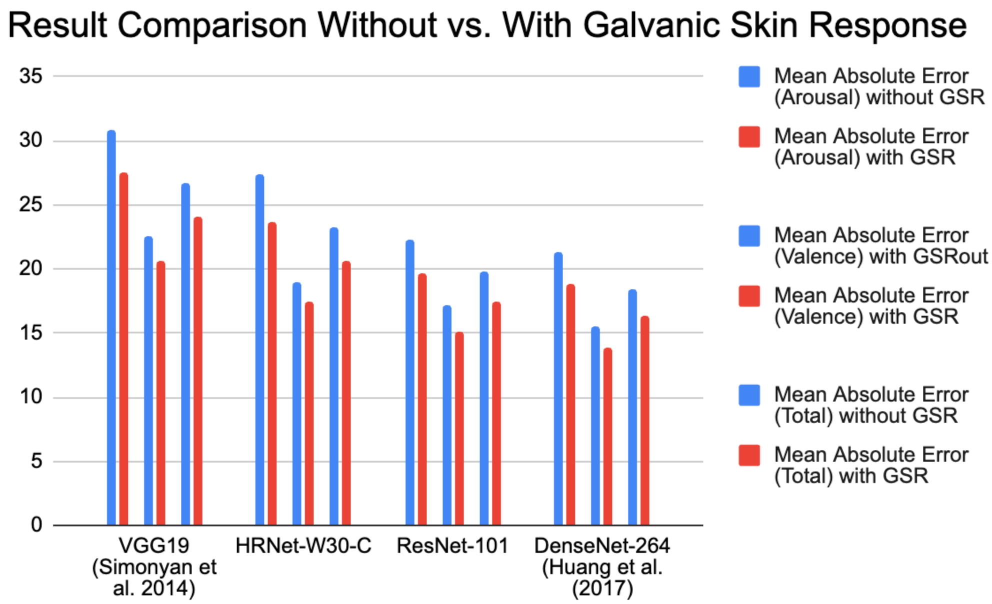

Figure 17. Comparison of Mean Absolute Error (MAE) based on Architecture with vs. without Galvanic Skin Response

Ablation Study

An ablation study was conducted to evaluate the contribution and significance of various components in the proposed model. It is done by systematically removing parts of the model and observing the impact on performance.

The first part of the ablation study, with its results shown in the first column of values, is when the representation learning (Figure 11) is removed, with only the basic networks in place. When compared with the second column, the proposed model that includes representation learning has a lower MAE value overall, showing that representation learning improves the model's performance and is thus pivotal.

The second part of the ablation study, with its results shown in the third column of values, is when the L1 squared function is used to compare the feature vectors of the emotions to train the model. When compared with the fourth column, the proposed model that utilizes cosine similarity function (Equation 1) has a lower MAE value overall, showing that cosine similarity function improves the model's performance and is thus significant.

Table 4. Ablation study

	Mean Absolute Error (Total) Evaluation on Representation Learning		Mean Absolute Error (Total) Evaluation on Equation	
	Baseline (without representation learning)	Proposed Model (with representation learning)	L1 squared function	Cosine Similarity Function (proposed model)

VGG19 (Simonyan et al. 2014)	26.9	24.1	25.3	24.1
HRNet-W30-C	23.4	20.6	22.1	20.6
ResNet-101	20.1	17.4	17.9	17.4
DenseNet-264 (Huang et al. (2017)	18.8	16.4	17.2	16.4

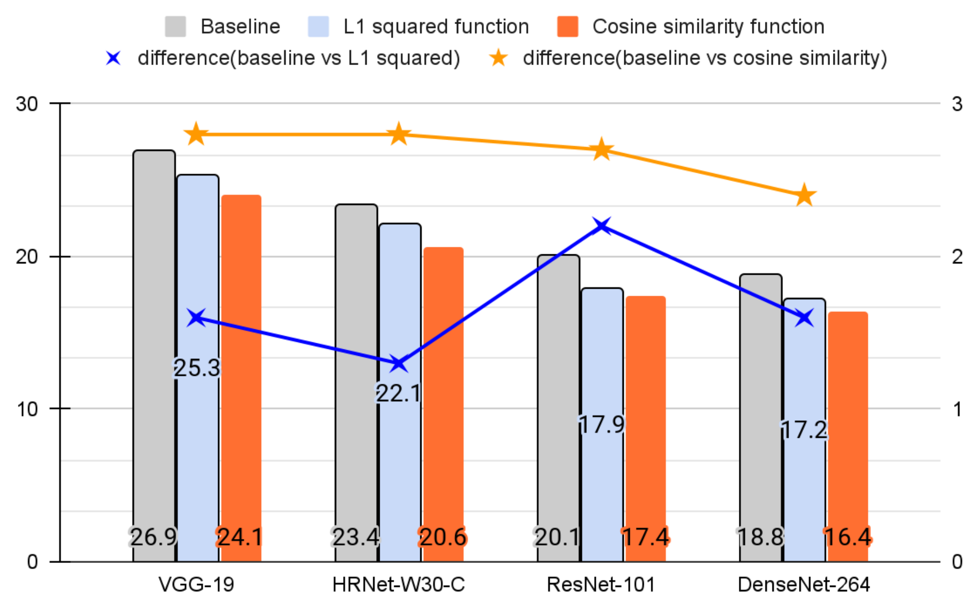

Figure 18. Ablation Study

t-SNE Evaluation

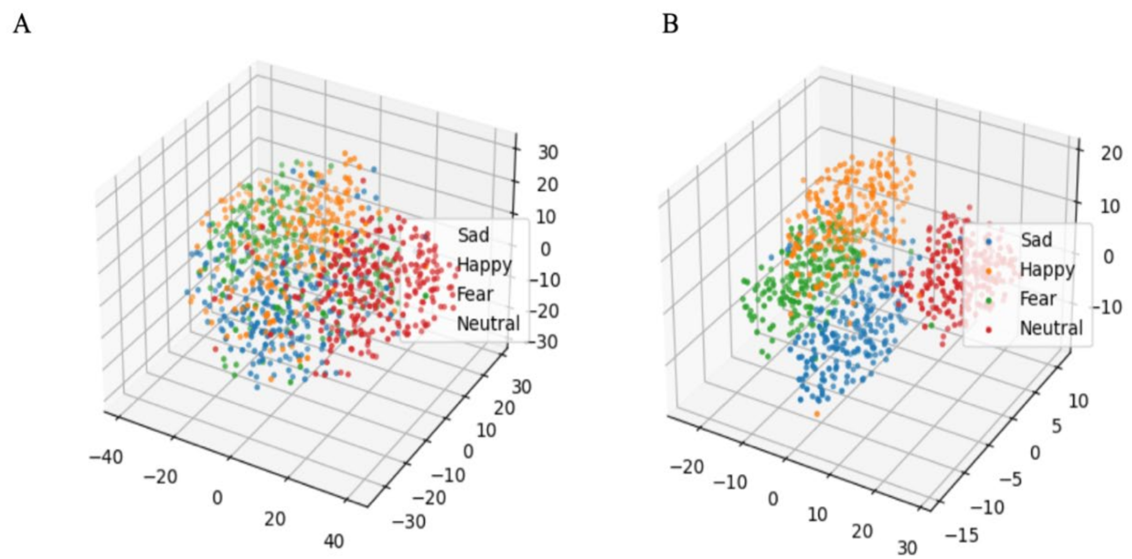

Figure 19. t-SNE Evaluation Results, (a) Baseline t-SNE evaluation (b) Proposed method t-SNE evaluation

Additionally, I conducted a t-SNE evaluation to visually assess how the proposed approach disentangles features related to each emotion category. I trained two models: one with the proposed approach and another without it, referred to as the baseline. Figure X illustrates the evaluation results. Each dot in the figure represents a feature vector extracted from an input sample, with dots of the same color representing the same categories. The feature vectors extracted from the model trained with the proposed approach show more distinct patterns which indicates better disentanglement, while the feature vectors from the baseline model appear more entangled.

Conclusion

In this paper, a novel depression screening method that is distinguished from the conventional methods of self-report questionnaires was developed. The proposed method utilizes several machine learning networks including representation and transfer learning, outputting the current emotional state of the participant using electroencephalogram (EEG) and galvanic skin response (GSR) signals. When evaluated, the proposed model demonstrated exceptional performance. The best MAE was 18.4 without GSR; 18.8 without representation learning; 17.2 without the proposed equation, but improved to 16.4 in the proposed model.

The proposed method has three key contributions. First of all, based on mathematical logic, representation and transfer learning were applied. Their effectiveness was shown through the MAE value ranging from 18.8 to 26.7 when baseline was used and the MAE being improved to 16.4 with the proposed model. Secondly, GSR signals were utilized on top of EEG signals for enhanced accuracy. This was shown by the best MAE value being 18.4 and the worst 30.9 when the model excluded GSR signals, and the value being lowered to 16.4 when the model utilized GSR. Thirdly, various CNN networks were tested. Four networks, VGG19, HRNet-W30-C, ResNet-101 and DenseNet-264 was used.

Future applications include depression screening devices, especially for adolescents. As observed early in the first chapter of this research, many adolescents have a high chance of responding dishonestly to standard depression screening self-report questionnaires, and may not realize their mental health and emotional states precisely. Thus, the proposed method will help screen depression in adolescents, helping to decrease the high number of depression and suicide. When implemented in schools, each adolescent will be shown different scenes that they are commonly exposed to, such as their peers, teachers, classrooms, family, social media, and more. Individuals exhibiting symptoms of depression will be differentiated as they will have different, often more negative, emotional responses to the scenes they were exposed to. This will help adolescents who are unaware of their psychological condition receive support and restore their mental health to an optimal state.

Acknowledgments

I would like to thank Mr. Koby Osei-Mensah (Loomis Chaffee School) for all his advice in this research.

References

Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., Adeli, H., & Subha, D. P. (2018). Automated EEG-based screening of depression using deep convolutional neural network. Computer methods and programs in biomedicine, 161, 103-113.

Arnrich, Bert. “Galvanic Skin Response - an Overview | ScienceDirect Topics.” ScienceDirect.Com | Science, Health and Medical Journals, Full Text Articles and Books., Journal of Biomedical Informatics, 2019, https://www.sciencedirect.com/topics/computer-science/galvanic-skin-response.

Cleveland Clinic. (2024, Mar 8). “Electroencephalogram (EEG)”: Cleveland Clinic https://my.clevelandclinic.org/health/diagnostics/9656-electroencephalogram-eeg.

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.

iMotions. (2024, Apr 17). “Galvanic Skin Response (GSR): The Complete Pocket Guide”: iMotions https://imotions.com/blog/learning/research-fundamentals/galvanic-skin-response/.

Lee, Chris. “Roughly 1 in 5 Adolescents Report Experiencing Symptoms of Anxiety or Depression | KFF.” KFF, 6 Feb. 2024, https://www.kff.org/coronavirus-covid-19/press-release/roughly-1-in-5-adolescents-report-experiencing-symptoms-of-anxiety-or-depression/.

“Mental Health of Adolescents.” World Health Organization (WHO), World Health Organization (WHO), 2024, https://www.who.int//news-room/fact-sheets/detail/adolescent-mental-health/?gad_source=1&gclid=CjwKCAjwvIWzBhAlEiwAHHWgvTmF1owwHxALm4223l3lQLOPPHHCtjcJTaKONRQIQxrPbVDZjfEaRxoCRZYQAvD_BwE.

Mental Health America. (2023, May 23). “Quick Facts and Statistics About Mental Health”: Mental Health America https://mhanational.org/mentalhealthfacts.

Neff, Megan Anna. “How to Use Arousal-Valence Model.” Insights of a Neurodivergent Clinician, Insights of a Neurodivergent Clinician, 26 Feb. 2023, https://neurodivergentinsights.com/blog/arousal-valence-model.

Pfizer. (2024, Sep 11). “Patient Health Questionnaire (PHQ) Screeners”: Pfizer. https://www.phqscreeners.com/.

Plass-Oude Bos, Danny. (2006). EEG-based Emotion Recognition. The Influence of Visual and Auditory Stimuli.

"Depression - National Institute of Mental Health (NIMH)." National Institute of Mental Health (NIMH), 2024, https://www.nimh.nih.gov/health/publications/depression.

"Depression Assessment Instruments." APA.Org, Jan. 2023, https://www.apa.org/depression-guideline/assessment#:~:text=The%20Patient%20Health%20Questionnaire%20.

Wilson, S., & Dumornay, N. M. (2022). Rising rates of adolescent depression in the United States: Challenges and opportunities in the 2020s. Journal of Adolescent Health, 70(3), 354-355.

WHO. (2023, Mar 31). "Depressive disorder (depression)": WHO https://www.who.int/news-room/fact-sheets/detail/depression/?gad_source=1&gclid=CjwKCAjwvIWzBhAlEiwAHHWgvVf4dbrXX3cCxmrPtpwE8Q3H5ev64adn_KF72gjGZG_MllXE4T4LgRoCVXMQAvD_BwE.

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.