Human Emotion Detection with ECG and EEG Signals Using ML Techniques

Denuka Jayaweera
University of Peradeniya
Sri Lanka
e17144@eng.pdn.ac.lk

Lahiru Wijesinghe
University of Peradeniya
Sri Lanka
e17405@eng.pdn.ac.lk

Adithya Gallage
University of Peradeniya
Sri Lanka
e17091@eng.pdn.ac.lk

Roshan Ragel
University of Peradeniya
Sri Lanka
roshanr@eng.pdn.ac.lk

Isuru Nawinne
University of Peradeniya
Sri Lanka
isururnawinne@eng.pdn.ac.lk

Mahanama Wickrmasinghe
University of Peradeniya
Sri Lanka
mahanamaw@eng.pdn.ac.lk

Vajira Thambawita
Simula Metropolitan for Digital Engineering
Norway
vlbthambawita@gmail.com

I. Abstract

In this research, the integration of Electroencephalography (EEG) and Electrocardiography (ECG) data for the classification of human emotions through Convolutional Neural Networks (CNN) is explored to enhance the accuracy of human emotion classification. The data collection process involved eliciting four distinct emotions and neutral emotion in subjects via video and audio stimuli. However, due to the limited sample size of the collected data, the model was trained using the DREAMER dataset, a multi-modal database capturing signals recorded during affect elicitation via audio-visual stimuli from 23 subjects. The self-assessment of subjects in terms of valence, arousal, and dominance, provides a valuable benchmark for the approach. The DREAMER dataset was explored using CNN-based models, with the aim of improving emotion classification accuracy. Emotion classification methodologies are tried to be identified in this research which extends the applications of CNN in real-world scenarios utilizing time series data such as Electroencephalography and Electrocardiography. Finally, it is hoped that the findings of this research will contribute to further studies in this field among the research community.

II. Introduction

One of the most important phenomena that revolve around human day-to-day life is emotions, and emotions are encountered in every activity that is performed by humans. Therefore, the understanding and accurate identification of human emotions is widely important in various fields of work such as healthcare [1], Human-Computer Interaction (HCI) [2], and psychology [3]. Traditionally, the identification of emotions has been done via self-reporting, which has several limitations. The process of identifying emotions via self-reporting can be biased and subjective depending on the perspective of the person who is giving out the self-report. Now, the era where physiological signals generated by the human body are being used as an emotion-measuring technique has been there, which can be hectic but will give good outputs.

Over the years, various methods have been used to tackle emotion recognition through physiological signals. In this research, ECG and EEG signals, signal processing, Machine learning techniques, and other state-of-the-art techniques are used to enhance the signal that was acquired and processed. Different methods have been developed and tried to provide emotional intelligence to computers, including facial expression-based emotion recognition, speech analysis, and biosignals such as ECG and EEG. Among them, facial recognition was widely used, while bio signals provided highly accurate results due to their unmaskable nature.

In the area of Human-Computer Interaction (HCI) research, the main focus is to make computer interactions more productive, seamless, and interactive. A promising method to achieve this task is to provide emotional intelligence to computing systems. These systems can adapt based on the emotional state of the user, which can be useful in areas such as entertainment systems, healthcare systems, adaptive learning systems, and computer games.

The final aim of the project was to create a robust framework for detecting and classifying human emotions into different categories which can be used in the future for the above-mentioned applications.

III. Related Work

A. What is an Emotion

Emotion is a complicated state expressed by human consciousness in response to environmental stimuli. In general, reactions to ideas, memories, or events that occur in our environment are expressed by them. It involves a variety of individual thoughts, feelings, behaviors, and psychological experiences, which play a vital role in human decision-making and mutual interaction [4].

B. Models of Emotions

Throughout, researchers adhered to using the basic discrete emotion model and the dimensional emotion model [5] as the two main approaches to studying and categorizing emotions. When considering the basic discrete emotion model, categorizes a variety of core emotions such as joy, sadness, fear, anger, disgust, etc., while the dimensional model categorizes emotions in various matrices such as valence, arousal, dominance, etc. These models provide the basis for researchers to understand, organize, and study various emotions to explore more complex aspects of emotions in different fields.

One of the well-known discrete emotion models is the Tomkins emotion classification model [6] that covers nine core emotions: interest (excited), surprise (startled), amusement (joy), distress (angry), disgust (disgust), scared (terrified), anger (anger), contempt (disgust), and shame (humiliation). These nine emotions are thought to be the building blocks for complex emotions. The common case of discrete emotion models is that all the other emotions that were not in the base model were considered to be variations or combinations of the core emotional states.

The discrete emotion model has drawbacks when representing a range of different emotions because it categorizes only core, basic emotions in a distinct way. To mitigate the above-mentioned drawback, the dimensional emotion model has been introduced. In this model, emotions are organized in a multidimensional space, with each dimension representing a different emotional aspect or measure. Russell‘s circumplex 2D model [5] uses arousal and valence as the dimensions and shelters up to 150 different effective emotions. Schloberg introduced a three-dimensional emotion model [7] that adds an attention-rejection dimension to the two-dimensional model. Out of these models, Russell’s 2D emotion model is the most frequently used, and this research was aided by it as well. In Russell’s 2D emotion model, the vertical axis represents the arousal dimension, indicating the intensity of the emotional experience from low to high, while the horizontal axis represents the valence dimension, representing the degree of cheerfulness from negative to positive. Based on this coordinate system, emotions can be categorized into four main groups as shown in Figure 1.

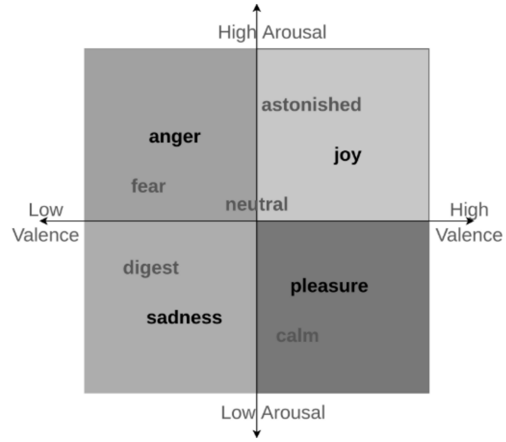

Fig. 1: The 2D emotion model

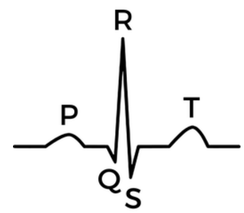

Fig. 2: Illustration of the PQRST wave-form

C. Electrocardiography in the Field of Emotion Recognition

ECG provides valuable information on heart activity based on the electrical pulses of the heart, which helps with research on complex aspects of heart activity in various fields. ECG signals have been widely used in developing machine learning and deep learning solutions, such as for detecting improper heart rhythms, identifying emotions, and even for biometric identification purposes. Scientific studies conducted in the domain of emotion detection using psychological signals have shown a strong relationship between cardiovascular activity and emotions, as emotions can affect the autonomic nervous system that governs heart activity.

When considering a detailed ECG, generally consists of three main waves, as shown in Figure 2. These are P, QRS, and T waves. The P wave represents atrial depolarization, the QRS wave marks the beginning of ventricular contractions, and the T wave appears when the ventricles repolarize. Each of these wave forms contains valuable information that helps understand an individual’s cardiac condition [8]. So these physiological aspects of the ECG can be used as a parameter that coincides with the change of emotions. therefore these correlating features of the ECG can be used in experiments to identify the emotional state of a person.

D. Electroencephalography in the Field of Emotion Recognition

EEG is used to monitor brain activity or brain waves [8] which can be recognized as one of the most efficient ways to measure brain waves. The process of measuring these signals includes placing electrodes on the scalp to detect tiny electrical signals generated by the brain’s neurons and it will be amplified for further analysis.

By analyzing those brain waves, emotional states, and emotional changes can be identified. When an electrode is placed on the scalp it will measure the electrical activities of a group of neurons rather than a single neuron. Those neuronal signals include functional and physiological changes in the central nervous system(CNS). Therefore, the EEG signals contain useful and meaningful psychological and physiological information.

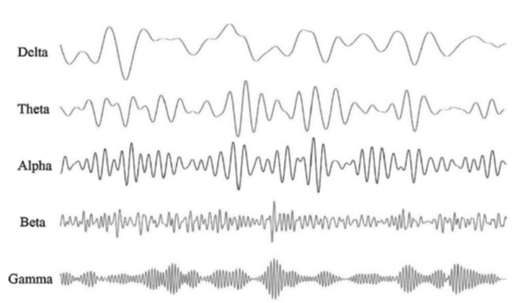

Fig. 3: The wave-forms of EEG bands

The EEG signals can be classified into five categories based on the variation in frequency bands: Delta, Theta, Alpha, Beta, and Gamma. There can be a strong correlation among these EEG signals concerning frequency range and different brain activities.

EEG provides valuable information about different emotional states. By understanding them it can be seen that low-frequency brain waves are highly related to emotions than high-frequency waves, and also negative emotions are usually stronger than positive emotions. By analyzing specific brain waves such as Beta, Alpha, and Theta waves, we can identify emotions like happiness, sadness, and fear. Also, it can be identified that certain brain areas are connected to specific emotions. For example, enjoyment activities are associated with the left front of the brain, while fear will reduce the activity in the left front of the brain. So the possibility of understanding the relationship between the EEG data and the emotion elicitation can be used to detect emotions. Which will be great from the perspective of using physiological signals to detect the emotions of humans.

E. How Different Experimental Processes were Done in Related Work

The field of study of physiological signals in human emotion detection is not novel but it is so vast and there is a lot of research done. In this section, we talk about the work that is relative to the research that is presented in this paper.

For emotion recognition, there have been many methods used, such as speech analysis, facial expression analysis, and physiological signal processing [9]. All these methods come with pros and cons, speech-based emotion recognition needs continuous speech [10], [11] while in the facial feature analysis method, some people might and can hide and mask their facial features so the emotion recognition will be harder [12], [13]. Therefore using physiological signals might be a better approach because they are continuously produced and extremely hard to mask and using physiological signals has provided some great results as well [9], [13], [14].

The DREAMER model [14] has followed a multi-model approach including ECG and EEG signals to classify human emotions. Data was collected using signal elicitation by audio-visual stimuli which was similar to the pattern that was followed in this research. Here the research was based on the 3D emotion classification model on valence, arousal, and dominance.

MPED a multi-model physiological emotion database for discrete emotion recognition was also a similar research that was done, through physiological signals [15]. This model has four types of signals. Electroencephalography, galvanic skin response, electrocardiogram, and respiration. In this research, the researchers talk about the limitations of culture-dependent stimuli. Here the experiment was done for seven distinct emotions including the neutral emotion. The MPED research introduces a method to use the attention-long short-term memory model that increases the effectiveness of the sequence analysis and the feedback is taken to compare the predicted and self-reported outputs.

The paper ”A Survey on Psycho-Physiological Analysis and Measurement Methods in Multi-modal Systems” [16] provides an analysis of the multi-model systems of the subject’s emotional state during the experiment. Here the important feature is that they reveal many measurements that can be used to identify the emotional state by a physiological signal. This model has an extremely high accuracy rate and, therefore, can be used for emotion classification using EEGs and the accuracy was measured using the self-reported data and predicted output.

The paper by Morteza et al. [17] uses EEG signals for emotion recognition, this research provides a good step-by-step procedure. The problem in this research was it is not multi-model therefore there is no comparison between different physiological signals and the accuracy that it provides can be improved as well. But an excellent paper for understanding the challenges and progress of emotion recognition using EEG.

The paper ECG Pattern Analysis for Emotion Detection [18] is uni model research that was conducted using ECG signals, this research understands the complexity of emotion recognition using physiological signals and how to tackle some of the issues that every researcher will face while conducting similar research. This model has a relatively high accuracy of the predictions on the valence-based emotion recognition so it emphasizes the importance of comparing only one dimension of the emotional model as well. However, the research itself says that there might be an accuracy increment in using multi-model approaches.

The review paper by Essam H. Houssein et al. [19] discusses the usage of brain-computer interfaces (BCI) along with machine learning to detect human emotions. The paper reviews various approaches to recognizing emotions through EEG-based BCI, along with an overview of the datasets and techniques used in emotion stimuli. Moreover, the paper focused on the general architecture of research based on BCI and ML techniques. The complete architecture involves distinct steps, such as EEG signal acquisition, data preprocessing, feature extraction, feature selection, emotion classification, and performance evaluation. According to the review paper, machine learning algorithms such as the k-nearest neighbor, support vector machine, decision tree, artificial neural network, random forest, and naive Bayes, as well as deep learning algorithms such as convolutional and recurrent neural networks with long short-term memory, were used in previous studies conducted by the researchers. The paper reviews more details of the usage of the above algorithms in the field of study. Finally, it suggests several challenges and directions for future research in the recognition and classification of human emotions using EEG. One suggestion is to adhere to deep learning techniques such as convolutional (CNN) and recurrent neural networks (RNN) with long short-term memory for significantly improving emotion recognition performance. Another important point emphasized by the paper is that improving feature extraction and selection methods is crucial to enhancing the performance of emotion detection systems. So the paper suggests a more novel research direction based on using multi-model data (combining EEG data with other psychological data such as ENG and ECG) to improve the accuracy of emotion recognition.

The research, "Research on Emotion Recognition Based on ECG Signal" [20] focuses on emotion recognition using cardiographic waves and explores various aspects of how ECG signals control human emotions. This study is based on using existing, well-formed machine learning algorithms for human emotion detection using correlation features and time-frequency domain statistical features of ECG signals. It compares the performance of commonly used classification algorithms such as Support Vector Machines(SVM), Classification and Regression Trees(CART), and K Nearest Neighbor (KNN) in emotion detection. The main output of the study indicates that using correlation features of ECG has a greater impact on the performance of the above-mentioned classification algorithms than using time-frequency domain features, which give accuracy of 19.7% and 16.7%, respectively. When considering algorithmic performance, K Nearest Neighbor, also known as KNN, gives the highest accuracy over other algorithms. Furthermore, this research paper introduces a possible optimization of the KNN algorithm to improve performance in emotion recognition based on combining the Max-Min Ant system with the KNN algorithm. It improves the emotion recognition performance of the KNN algorithm by up to 92% compared to direct use of the KNN algorithm, which gives 16.9%. The data collection procedure used for this study is based on video stimuli. For exploring algorithmic performance over different emotion dimensions, one emotion from each quadrant is selected from the two-dimensional emotion model. Further, this paper suggests some future research areas in the emotion recognition domain. These suggestions include developing multi-model emotion recognition algorithms that have higher accuracy by integrating Electrocardiography signals with other physiological signals such as Electroencephalography, Electromyography, and Skin Conductance. Moreover, it suggests a newer research area on emotion detection using deep learning techniques, which holds the possibility of improving the accuracy of emotion recognition.

Taking the above-related works as the basis for the research "Human Emotion Detection with ECG and EEG Signals using Machine Learning Techniques", it is aimed to explore novel knowledge in the field of emotion recognition involving both EEG and ECG signals. The data set is planned to be collected through video stimuli while targeting four distinct emotions (Joy, Fear, Sadness, Relaxation) and the neutral emotion that cover a broader range of affective emotion dimensions in the 2D emotion model. This simplification allows for the capture of a more accurate set of data and allows for easier supervision of the data during the training phase.

In summary, existing literature and previous studies done in the domain of human emotion recognition manifest the potential of using ECG and EEG signals for emotion recognition. Furthermore, the combination of various physiological signals, such as ECG and EEG, along with deep learning techniques offers improved classification accuracy compared to pure machine learning techniques. Integrating those models into computerized systems leads to more effective and personalized human-computer interaction systems.

IV. Methodology

As mentioned previously, the primary aim of the research was to develop and publish a dataset that classifies human emotional (affective) states based on ECG and EEG signals, along with an ML model for classification. Following the literature survey, a deep learning-based Convolutional Neural Network(CNN) model was used for emotion classification. The initial part of the methodology outlines how the ECG and EEG data were collected. Moving forward, the next section discusses challenges encountered during the data collection process while the subsequent section addresses the steps taken to resolve these issues, specifically highlighting the use of the DREAMER dataset [14] for developing the ML model. Lastly, the final part of the methodology illustrates the machine learning procedure employed.

A. Experiment for ECG and EEG Data collection

This section outlines the procedure for collecting subject data in order to create the machine learning model. The first section of this subsection offers guidance on choosing appropriate ECG and EEG devices for data collection. The second section explains the challenges and ethical considerations involved in collecting such physiological signals. Moving forward, the third section illustrates the data acquisition setup that was developed to collect both ECG and EEG data from the subject while evoking emotions through video stimulation. Subsequently, the fourth section illustrates the adapted discrete emotional model and how it employs selecting videos for emotion elicitation. Finally, the last section explains the ECG and EEG data collection experiment in detail.

1) ECG and EEG Wearable Devices: The selection of the devices was based on a few considerations because there are many different kinds of hardware available that record ECG and EEG data. First and foremost, the subject's comfort is directly affected by the gear. Therefore, they shouldn't feel constrained by the hardware setup while wearing it. Secondly, recorded signals should not be too noisy, which would complicate preprocessing. And lastly, the gear ought to be reasonably priced.

After careful consideration, the Spiker Shield Heart and Brain Bundle was chosen for ECG data acquisition due to its low cost and integrated noise-canceling sensors. Additionally, it is not uncomfortable for the subject because what was required to record the ECG signals are a few adhesives on

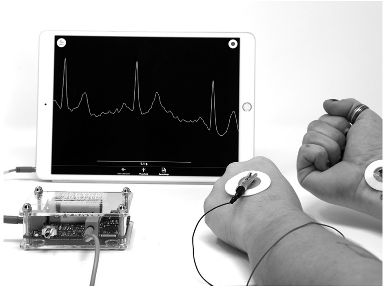

Fig. 4: The Spiker Shield Heart and Brain Bundle

Fig. 5: The OpenBCI Cyton Biosensing Board with Active Electrode Kit

their hands. The 'OpenBCI Cyton Biosensing Board with Active Electrode Kit' was chosen for EEG signal acquisition due to its ability to deliver high-quality research-grade signals while offering subjects comfort and a user-friendly experience similar to consumer-grade products. Figure 4 and Figure 5 show the images of the wearable devices used for ECG and EEG signal acquisition, respectively.

2) Ethical Considerations in the Data Collection Process: In the experiment of data collection, a video stimuli-based mechanism was used to induce emotions in the subjects. This section addresses the ethical considerations in the process of ECG and EEG data collection that were labeled, along with the human emotion state that was elicited by the video stimuli mechanism.

Prior to participation, informed consent was obtained from all subjects, describing the study's purpose, possible risks, data collection mechanism, and setup, as well as their right to leave the study at any time. Furthermore, approval was obtained from the university's Ethical Review Committee (ERC) for the study, confirming its compliance with ethical guidelines and guaranteeing the privacy and safety of its subjects.

To recognize the psychological effects of emotion-eliciting stimuli, subjects underwent an interrogation process following exposure to each and every one of the stimuli, and in the instance that distress showed up, support options, including university counseling, were made available. Also, through the previously discussed point, which is to protect subjects from psychological effects during the emotion-eliciting stimuli, voluntary participation was assured and allowed individuals to leave whenever they wished without any pressure or obligation to stay. Another important aspect that was taken into account was maintaining clear and open communication throughout the experiment with the subject to convey research objectives, risks, and benefits. Finally, to protect privacy, the data collected was anonymized and securely stored, with limited access granted only to authorized parties adhering to stringent confidentiality protocols.

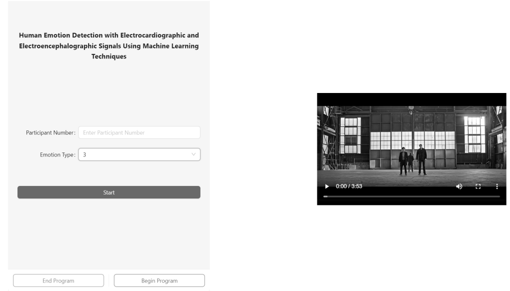

Fig. 6: The UI of Data Acquisition Software Application

3) Data Acquisition Setup: This section describes the entire setup for the collection of data, including the synchronization of data obtained from ECG and EEG devices, the interactive engagement of subjects with the video stimuli, and the accurate placement of ECG and EEG sensors on the subjects.

Here, a software application was developed to synchronize data obtained from the ECG and the EEG devices and to allow subjects to interact with video stimuli. This software application consists of a frontend developed using ReactJs and a backend developed using Python Flask framework. The backend functionality is mainly compromised with three processes. They are the main process, the ECG process, and the EEG process. The ECG process directly connects to the ECG signal-capturing device via a USB port. The EEG process works in the same nature, and it also connects to the EEG signal-capturing device via a USB port. Both of these processes have the same capability, which is recording observed data. The main process is in charge of handling these two processes while communicating directly with the frontend as shown in Figure 6.

Regarding the frontend, it consists of an interactive User Interface (UI) that allows the researcher to input the subject identification number and choose required video stimuli through an accordion-style input element. The selected video will appear within the right section of the User Interface as in Figure 1. Even though the researcher selected the specific emotion-related video, the subjects were not aware of the intended emotion type of each video stimulus. Since the accordion-type input was encoded and the encoding algorithm was only known by the researcher. When the subject is prepared for the process of collecting data using ECG and EEG wearables, the researcher presses the play button to start the clip, and the video instantly takes the full-screen mode. Simultaneously, an automatic signal is triggered for the backend main process to start both ECG and EEG processes to record data. This synchronizes the data acquisition process from ECG and EEG wearable devices. Upon video completion, the frontend triggers an automatic signal to the backend main process to halt the ECG and EEG processes, thereby concluding the data collection phase.

As for the hardware setup the ECG adhesive sensors are placed on the hands, as shown in Figure 7 utilizing three sensors, two of which are placed nearby to the wrist on the

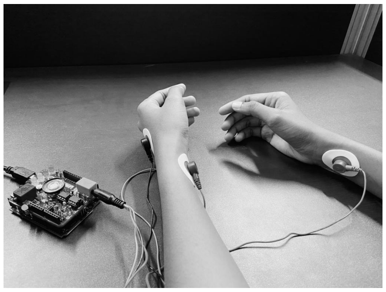

Fig. 7: The ECG Sensor Placement

Fig. 8: The EEG Sensor Placement

inner side of the arm, while the remaining one is placed on the opposite side of the palm of the left hand. The EEG active electrode placement was placed on the subject's scalp. Here, 8 active electrodes were used and were fixed within a wearable cap, as shown in Figure 8. Accurate EEG data acquisition can be done by wearing the electrode-embedded cap correctly.

4) Video selection for stimuli: In the study, one of the main objectives was to collect ECG and EEG data labeled based on the subject's emotional state. As mentioned, five emotions were focused on. They are joy, fear, sadness, relaxation, and neutral. The chosen emotions covered all four quadrants in the discrete 2D emotion classification model [5] shown in Figure 1. The first quadrant of the model contains highly valenced and aroused (HVHA) emotions like joy, excitement, astonishment, and delight. The second quadrant of the model represents low-valenced and high-aroused (LVHA) emotions, including emotions like fear, anger, and annoyance. The third quadrant of the model represents low-valenced and aroused (LVLA) emotions such as sadness and disgust. Finally, the fourth quadrant of the emotion model includes relaxation, calm and pleasure, and annoyed emotions, which represent the high-valenced and low-aroused (HVLA) scenarios.

The decision to limit the study to the five key emotions that are drawn from the four quadrants of the discrete 2D emotional model, along with neutrality, was based on enhancing the precision of labeling the collected ECG and EEG data. Expanding the study to select multiple emotions from each quadrant would have an impact on labeling accuracy due to variations in subjects' ability to distinguish and feel between these emotions. Moving forward, a set of videos was collected by consulting with domain experts and the university Ethical Review Committee in order to invoke specified target

Fig. 9: The Experiment Environment - I

emotions. These videos lasted between three and five minutes each. Table I illustrates the information regarding the selected video clips and duration.

TABLE I: Details regarding videos

Name	Target Emotion	Duration(min)
The World's Worst Translator	Joy	3.9
Lights out	Fear	2.7
My Dad is a Liar	Sad	3.45
Relaxing Music	Relaxation	2
A black screen	Neutral	2

5) Experiment: Subjects were recruited through the University of Peradeniya and the surrounding environment. Subjects of any age, gender, and background who are capable of watching videos and providing informed consent were welcome to participate in the study. There were no specific restrictions or requirements for participation, allowing for a diverse range of individuals to contribute to the research. Prior to participation, each subject was provided with a consent form outlining the purpose, procedures, risks, benefits, and confidentiality of the study, adhering to the ethical guidelines mentioned above. Subjects were given adequate time to review the consent form, ask questions, and provide written informed consent before proceeding.

Afterward, subjects were seated comfortably in a controlled environment, free from distractions.ECG and EEG electrodes were placed in specific areas of the subjects's body according to established guidelines as discussed in the previous section. Subsequently, the whole data acquisition setup, including the software application, was made operational. Then, each and every video was shown to the subjects, with around one to two minutes of breaks in between. At the end of each video, subjects were asked to write down their emotional experience throughout the video on a pre-designed feedback form. Finally, data were labeled based on those feedbacks. Figure 9 and Figure 10 shows the experimental setup and the environment while a subject was going through the designed experiment. Furthermore, data was successfully acquired from a total of six subjects who participated in the experiment.

B. Failures and Challenges in the Data Collection Process

A total of seventeen subjects participated in the experiment. Out of these, the ECG signals of eleven subjects had to be

Fig. 10: The Experiment Environment - II

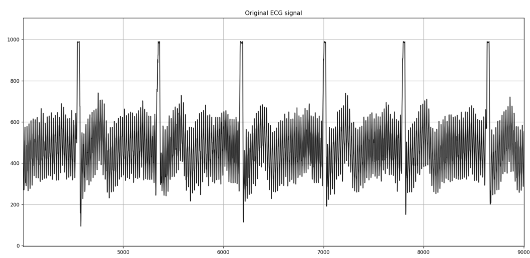

Fig. 11: The Distorted ECG Signals

removed due to noise issues and signal anomalies, as shown in Figure 11. The primary cause of these signal anomalies and noise issues might be the quality of the ECG adhesive electrodes used in the experiment. The high-quality ECG electrodes that were included in the 'Heart and Brain SpikerShield Bundle' had expired, and unfortunately, due to the current situation in the country, sourcing high-quality adhesive electrodes in a short period of time was challenging. Since the limitation of data to train a CNN model, the focus has changed to using public datasets published in this domain to continue the research.

C. Available Data Sets

Due to ECG signal anomalies and limitations in data capture, a publicly available dataset called the DREAMER dataset, was used to develop and evaluate the proposed CNN-based solution in-depth on a wide range of different subjects.

1) DREAMER Data set: The DREAMER dataset [14] was collected from 23 subjects. In order to collect the dataset, video clips containing both video and audio stimuli were shown to the subjects, who provided their consent for the data collection. Nine distinct feelings were elicited by the 18 video clips that were screened. They are amusement, excitement, happiness, calmness, anger, disgust, fear, sadness, and surprise. Every clip lasted for sixty seconds. Before every video, a neutral video was also played to assist the subject in regaining emotional stability. Both the ECG and EEG signals of the subject were captured at a 256 Hz sampling rate. At the end of the session, subjective experience of arousal and valence scores were collected using self-assessment methods in the range of 1 to 5 for labeling the data.

2) Data Pre-Processing: The pre-processed version of the DREAMER data set was used in the study. According to the literature, since the ECG data acquisition device that was used was integrated with built-in noise filters, any data preprocessing mechanism was not applied. However, the captured EEG signals were pre-processed since they contained various types of noise and artifacts, such as those caused by cardiac, ocular, and muscular activities, and power line interference. To enhance the EEG data quality, band-pass filters were applied to extract relevant frequency ranges. While 50 and 60 Hz interferences were managed by the built-in filters of the EEG data acquisition device, additional artifact removal was performed using the Artifact Subspace Reconstruction (ASR) method, followed by the Common Average Reference (CAR) technique after eliminating bad channels in order to avoid introducing noise to other channels.

After acquiring the pre-processed version of the DREAMER dataset, several additional pre-processing steps were undertaken to prepare the dataset for further research work. Initially, data labeling was carried out for the DREAMER dataset to find to which quarter of the discrete 2D emotional model a given collected data belongs using the score that was given by the subjects according to whether valence and arousal have a higher value. Then the data sample is classified as HVHA, which stands for high valence and high arousal, and similarly, HVLA, LVHA, and LVLA are the other three categories, standing for high valence and low arousal, low valence and high arousal, and low valence and low arousal, respectively. Afterward, the last 60 seconds from both ECG and EEG data samples within the DREAMER dataset were extracted and subsequently segmented into fixed window sizes of 1 second, 10 seconds, 20 seconds, and 30 seconds, resulting in the formation of five distinct versions of datasets. To prevent any possible data leakage between training and testing data, no overlap has been put between segments. Table II above illustrates more information regarding those three datasets.

TABLE II: Information on datasets extracted from DREAMER

Name	Window size(s)	Sample size	Data Points per Sample
Set 1	60	414	15,360
Set 2	30	828	7,640
Set 3	20	1,242	5,120
Set 4	10	2,484	2,560
Set 5	1	24,840	256

D. ML Model

In addressing the research gap, a CNN-based approach was taken during the development of the classification algorithm. Two types of CNN models were developed for the experimental process. One was designed to classify emotions based on either ECG or EEG data, while the other was designed to classify emotions based on a combination of both ECG and EEG data.

1) Individual Mode Approach: The ML model was designed to classify emotions based on either ECG or EEG data. Here, the baseline model was developed with two parts, where the first part was convolution, which is used for feature extraction, and the second part was a fully connected deep

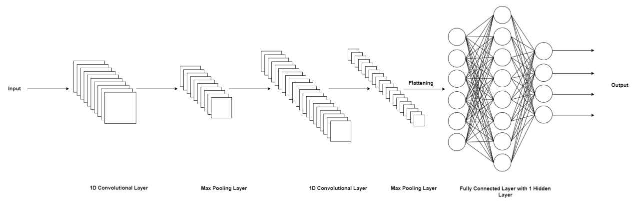

Fig. 12: Architecture of the Individual CNN Model

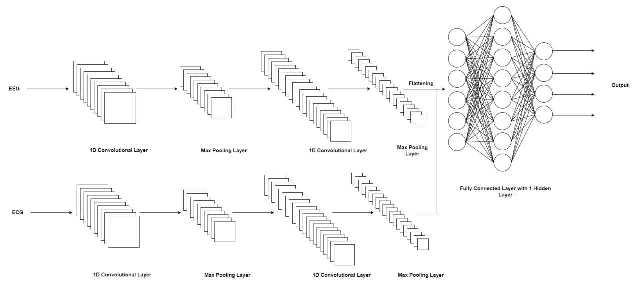

Fig. 13: Architecture of the Combined CNN Model

neural network for the classification. The feature extraction part of the model consisted of two sets of 1D convolutional layers with 64 filters and ReLU as the activation function, followed by a Max pooling layer for downsampling. And at the end, they were converted into a 1D vector using the flattened layer. A dropout layer with a dropout rate of 0.5 was used to prevent overfitting. Then, for the prediction part, there is a fully connected deep neural network that consists of an input layer and an output layer, with a hidden layer in between. The high-level architecture of the model is shown in Figure 12.

2) Combined Mode Approach: The ML model was designed to classify emotions based on both ECG and EEG data as input streams. Here, the baseline model was developed with two parts, where the first part was convolution, which is used for feature extraction, and the second part was a fully connected deep neural network for the classification as similar to the previous model. The feature extraction part consists of two pipelines, and each pipeline holds the responsibility of extracting features of ECG and EEG data respectively. Each pipeline consisted of two sets of 1D convolutional layers with 64 filters and ReLU as the activation function, followed by a Max pooling layer for downsampling. And at the end of the pipeline, they were converted into a 1D vector using the flattened layer. Finally, both flattened layers from pipelines were merged and pushed through a dropout layer with a dropout rate of 0.5 to prevent overfitting. Then, for the prediction part, there is a fully connected deep neural network that consists of an input layer and an output layer, with a hidden layer in between. The high-level architecture of the model is shown in Figure 13.

V. Results and Discussion

A. Results

After selecting the DREAMER data set for the experiment, the previously explained models were trained and tested. The use of the DREAMER model produced varying results for different models. This section will discuss the obtained results and their comparisons.

For the comparison, the researchers have employed the data labeling approach of prioritizing emotion detection. The model was designed to predict the emotion the subject felt during the experiment by converting the score of valence and arousal.

1) ECG Model: The Baseline accuracy received by the model discussed in the methodology section yielded an accuracy of 43 percent which seems improvable. Then the hyperparameter optimization for the model was carried out in different ways to improve the accuracy which eventually led to an accuracy of 59.1 percent.

One important improvement to the accuracy came with the data division. the sixty-second data sample was divided into smaller portions of thirty seconds (7680 data samples in a data), twenty seconds (5120 data samples in a data), ten seconds(2560 data samples in a data), and one Second(256 data samples in a data) as shown in the Table II. the accuracy mentioned above 59.1 percent came with the thirty-second data. and the accuracy eventually reduced when the division got smaller.

2) EEG Model: Similarly for the EEG model the initial baseline accuracy of 46.3 percent without dimension reduction. then it was found that the accuracy tends to increase with the use of PCA channel reduction, after the reduction of channels to two the accuracy eventually increased. To reach the accuracy of 50 percent similar to the ECG model the EEG model data sample size was also changed during the experiment process. The data division took the same approach as the ECG model with 60-second data samples 30-second samples 20-second samples 10-second samples and 1-second samples were taken and accuracy was compared. The highest accuracy came with the 30-second data sample.

3) Combined Model: The Combined model is the model that uses both ECG data and EEG data. To maintain consistency EEG data are reduced to two channels while the ECG retains its two-channel nature. The reduction was done using PCA similar to the above PCA-based EEG Model. In this model, the results were yielded initially without the PCA reduction of EEGs, which was further improved after the EEG channel reduction using the PCA. Accuracy at the end of all the hyper-parameter optimizations has improved to 52.4 percent.

One important point of hyperparameter optimization is data division, similar to the two models above the data divisions of 60 seconds to 1 Second were carried out with EEG and ECG before training the model for divided data samples. the 30-second data division of the EEG and ECG yielded the highest accuracy of 52.4 percent while the accuracy reduced after the 30-second interval. These results are summarized in the Table III.

4) ML Model Accuracy comparisons: According to Table IV the ECG model has yielded the highest accuracy out of the three models that were developed during the experiment and as the authors try to embed the EEG data into the model. The accuracy reduces gradually where pure EEG model has the lowest accuracy and it can seen in the table III.

TABLE III: Accuracy comparison data division

Model	60 Seconds	30 Seconds	20 Seconds	10 Seconds	1 Second
ECG Model	50	59.1	54.7	54.7	50
EEG Model	47.6	50	50	47.6	47.6
Combined Model	47.6	52.4	50	50	47.6

Training Set					
TARGET / OUTPUT	HVHA	HVLA	LVLA	LVHA	SUM
HVHA	5 11.9%	3 7.1%	2 4.8%	1 2.4%	11 45.5% 54.5%
HVLA	2 4.8%	6 14.3%	1 2.4%	2 4.8%	11 54.5% 45.5%
LVLA	1 2.4%	2 4.8%	4 9.5%	3 7.1%	10 40.0% 60.0%
LVHA	0 0.0%	1 2.4%	2 4.8%	7 16.7%	10 70.0% 30.0%
SUM	8 62.5% 37.5%	12 50.0% 50.0%	9 44.4% 55.6%	13 53.8% 46.2%	22 / 42 52.4% 47.6%

Fig. 14: The Confusion Matrix of the Combined Model

TABLE IV: ML Model Accuracy Comparison with the DREAMER Research

Model	Own	Valence	Arousal	Dominance
Combined Model	52.4	61.8	62.3	61.8
ECG Model	59.1	62.3	62.3	61.5
EEG Model	50	62.4	62.1	61.8

B. Discussion

1) ML Model Accuracy comparison with the original DREAMER Research: The DREAMER Research has a different accuracy for all the valence arousal and dominance models where this research reports significantly lower accuracy values compared to the DREAMER accuracy values and that comparison is shown in the table IV.

2) Improvements for the DREAMER Classifications: The DREAMER Research[14] has considered following the results on top of the valence, arousal, and dominance dimensions but it had not considered the Emotion Detection as a final outcome. therefore the results are in how accurately it has predicted the valence, arousal, and dominance level of the subjects that faced the experiment.

C. Solutions and Future improvements

The data labeling was done using how the subject is assumed to have felt, but that contradicts the question that the research tries to answer. Taking what the subject felt was what the research was trying to omit altogether. A good approach would be to label the video beforehand and try to figure out whether the subject has felt the emotion the video is intended to.

As mentioned above emotional elicitation has intentions to predict emotions. But ideally, a video chosen should elicit the same emotion from every subject. Therefore it is easier for the model training process and also for comparison purposes, so during the data acquisition part of the experiment the subject should be asked which emotion that the emotional elicitation procedure has given out, and if it is the intended emotion keep the data collected as a sample, if not that data sample should be removed from the data set by naming it as a outlier. The data scarcity and hardness in collecting such data in this sort of experiment is understandable therefore the trade-off value should be considered and carried out accordingly.

When it comes to the extraction of data samples as the training and testing data the last sixty seconds of the data sample was taken as recommended by the DREAMER Research[14]. Even though the best selection was done it still doesn't correlate to the part of the data sample that the emotion was elicited. Therefore as an improvement during the data collection procedure, the part of the video that is used to elicit the emotion should be marked beforehand and then that part of the data sample that was collected should be taken as train/test data.

VI. Conclusion

In the work that has been carried out, the procedure carried out to detect emotions using Electroencephalography and Electrocardiography was presented. The initial purpose was to present a data set consisting of ECG signals and EEG signals of subjects who are subjected to emotion elicitation via video stimuli. then after a few altercations, the paper includes a comparison of using ANN with Support Vector Machines and Deep Neural Network techniques of Convolutional Neural Networks for emotion recognition using EEG and ECG. The paper presents how the pre-experimental procedure of doing a survey to collect basic required data and the subject education on the experiment. Then the development of the Hardware setup that was carried out by the researchers to collect ECG data and EEG data simultaneously. Furthermore, the procedure of data collection was mentioned as how the two setups are used in collecting ECG data and EEG data. The Methodology should be very specific so as not to gain external interference and not to disturb the subject which could induce unwanted EEG and ECG signals.

In the paper it has mentioned an investigation to select a data set to develop a machine learning model by diving into the literature and why the DREAMER data set[14] is chosen to carry out the experiment forward. Development of a baseline machine learning model using the DREAMER Data set[14] and after that how it can be improved was also carried out by the authors.

last but not least the authors have presented the results of the experiments carried out and how the models and the data sets can be improved. This eventually reaches the problem of whether can emotions be predicted through ECG signals and EEG signals. the answer would be yes and this paper has contributed in a way to improve the pathway for researchers to carry out more experiments in the future.

VII. Acknowledgment

The authors would like to thank the DREAMER database owners for providing the data set to carry out the research, the Association for Computer Engineering Students(ACES) for providing the financial support, and the subjects of the experiment from the University of Peradeniya who made the research feasible.

References

[1] T. J. Mayne. Emotions and health. In T. J. Mayne and G. A. Bonanno, editors, *Emotions: Current issues and future directions*, pages 361–397. The Guilford Press, 2001.

[2] Nikolas Fragopanagos and J. G. Taylor. Emotion recognition in human–computer interaction. *Neural Networks*, 18(4):389–405, 2005.

[3] Carroll E Izard. *The psychology of emotions*. Springer Science & Business Media, 1991.

[4] Iris B Mauss and Michael D Robinson. Measures of emotion: A review. *Cognition & Emotion*, 23(2):209–237, Feb 2009.

[5] Jonathan Posner, James A Russell, and Bradley S Peterson. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(3):715–734, 2005.

[6] Tomkins S. S. Affect imagery consciousness. *Springer*, 1, 1962.

[7] Harold Schlosberg. Three dimensions of emotion. *Psychological Review*, 61(2):81–88, 3 1954.

[8] E.H. Houssein, A. Hammad, and A.A. Ali. Human emotion recognition from eeg-based brain–computer interface using machine learning: a comprehensive review. *Neural Computing and Applications*, 34:12527–12557, 2022.

[9] K. Rattanyu and M. Mizukawa. Emotion recognition based on ecg signals for service robots in the intelligent space during daily life. *J. Adv. Comput. Intell. Intell. Inf.*, 15:582–591, 2011.

[10] R. Cowie, E. Douglas-Cowie, J. Taylor, S. Ioannou, M. Wallace, and S. Kollias. An intelligent system for facial emotion recognition. In *Proceedings of the 2005 IEEE International Conference on Multimedia and Expo*, pages 904–907, Amsterdam, Netherlands, July 2005.

[11] C.-T. Tu and J.J.J. Lien. Automatic location of facial feature points and synthesis of facial sketches using direct combined model. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40:1158–1169, 2010.

[12] S. Jerritta, M. Murugappan, K. Wan, and S. Yaacob. Emotion recognition from electrocardiogram signals using hilbert huang transform. In *Proceedings of the 2012 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology, STUDENT 2012 - Conference Booklet*, pages 6–9, Kuala Lumpur, Malaysia, October 2012.

[13] J. Kim and E. André. Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30:2067–2083, 2008.

[14] Stamos Katsigiannis and Naeem Ramzan. Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices. *IEEE Journal of Biomedical and Health Informatics*, 22(1):98–107, 2018.

[15] Tengfei Song, Wenming Zheng, Cheng Lu, Yuan Zong, Xilei Zhang, and Zhen Cui. Mped: A multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access*, 7:12177–12191, 2019.

[16] Muhammad Zubair Baig and Manolya Kavakli. A survey on psycho-physiological analysis & measurement methods in multimodal systems. *Multimodal Technologies and Interaction*, 3:37, 2019.

[17] Morteza Zangeneh Soroush, Keivan Maghooli, Seyed Kamaledin Setarehdan, and Ali Motie Nasrabadi. A review on eeg signals based emotion recognition. *International Clinical Neuroscience Journal*, 4(4):118–129, 2017.

[18] F. Agrafioti, D. Hatzinakos, and A. K. Anderson. Ecg pattern analysis for emotion detection. *IEEE Transactions on Affective Computing*, 3(1):102–115, 2012.

[19] Essam Houssein, Asmaa Hamad, and Abdelmgeid Ali. Human emotion recognition from eeg-based brain–computer interface using machine learning: a comprehensive review. *Neural Computing and Applications*, 34, 05 2022.

[20] Zhongze Zhang, Xiaofeng Wang, Pengfei Li, Xi Chen, and Liwei Shao. Research on emotion recognition based on ecg signal. *Journal of Physics: Conference Series*, 1678(1):012091, nov 2020.