AN EMOTION RECOGNITION METHOD WITH LOW SPEC ECG AND EEG DEVICES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

EMRE KUTLU KÖSE

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MEDICAL INFORMATICS

MAY 2022

Approval of the thesis:

AN EMOTION RECOGNITION METHOD WITH LOW SPEC ECG AND EEG DEVICES

submitted by **EMRE KUTLU KÖSE** in partial fulfillment of the requirements for the degree of **Master of Science in Medical Informatics Department, Middle East Technical University** by,

Prof. Dr. Deniz Zeyrek Bozşahin
Dean, Graduate School of **Informatics** ________________

Assoc. Prof. Dr. Yeşim Aydın Son
Head of Department, **Medical Informatics** ________________

Prof. Dr. Ünal Erkan Mumcuoğlu
Supervisor, **Department of Health Informatics, METU** ________________

Examining Committee Members:

Assoc. Prof. Dr. Tolga Esat Özkurt
Department of Health Informatics, METU ________________

Prof. Dr. Ünal Erkan Mumcuoğlu
Department of Health Informatics, METU ________________

Prof. Dr. Osman Abul
Department of Computer Engineering, TOBB ETU ________________

Date: 09.05.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Emre Kutlu Köse

Signature :

ABSTRACT

AN EMOTION RECOGNITION METHOD WITH LOW SPEC ECG AND EEG DEVICES

Köse, Emre Kutlu
M.S., Department of Medical Informatics
Supervisor: Prof. Dr. Ünal Erkan Mumcuoğlu

May 2022, 67 pages

In recent years, developments on sensory technologies make electroencephalography and electrocardiography devices portable and more accessible. This provides new opportunities in emotion recognition area for researches and the industry. Interest of industry grows on the subject. Moreover, popularity of emotion recognition research is increased. To increase further possibilities of emotion recognition systems, this study provides an emotion recognition method that uses portable electroencephalography (EEG) and electrocardiography (ECG) devices. This study proposes an EEG feature called zero-crossing variance that detects small frequency chances on data and tests effectiveness of three different classifiers with it. Different from similar works, the study solves three class classification problem for valence and arousal and achieves accuracy of 92% for valence and 95% for arousal.

Keywords: EEG, ECG, HRV, Emotion, Classification

ÖZ

DÜŞÜK ÖZELLİKLİ EEG VE EKG CİHAZLARIYLA DUYGU TAHMİNİ

Köse, Emre Kutlu
Yüksek Lisans, Sağlık Bilişimi Bölümü
Tez Yöneticisi: Prof. Dr. Ünal Erkan Mumcuoğlu

Mayıs 2022 , 67 sayfa

Sensör teknolojilerinde oluşan son yıllardaki gelişmeler, elektroensefalografi (EEG) ve elektrokardiyo-grafi (EKG) cihazlarının daha taşınabilir ve daha erişilebilir hale getirmiştir. Bu durum araştırmacılar ve endüstri için duygu tanımada yeni fırsatlar sunmaktadır. Duygu tanıma sistemlerini daha kullanılabilir hale getirebilemek için bu çalışmada, taşınabilir elektroensefalografi ve elektrokardiyogram cihazlarını kullanan bir duygu tanıma yöntemi sunulmaktadır. Çalışma yeni bir EEG değişkeni önermiştir. Bu değişken için üç farklı sınıflandırma yönteminin performanslarını karşılaştırmıştır. Benzer çalışmalardan farklı olarak, bu çalışmada üç gruplu sınıflandırma problemi çözülmüştür. Çalışmada değerlik için $\%92$, uyarılma için $\%95$ doğruluk değeleri elde edilmiştir.

Anahtar Kelimeler: EEG, EKG, HRV, Duygu, Sınıflandırma

ACKNOWLEDGMENTS

Words cannot express my gratitude to my professor Prof. Dr. Ünal Erkan Mumcuoğlu invaluable patience and feedback. I also could not have undertaken this journey without Assoc. Prof. Dr. Tolga Esat Özkurt and Prof. Dr. Osman Abul who generously provided knowledge and expertise.

I am also grateful to İsmail Çağdaş Yılmaz for his editing help, feedbacks , and moral support. Thanks should also go to the İlker Sönmez and Özkan Yılmaz, who impacted and inspired me.

Lastly, I would be remiss in not mentioning my family, especially my parents, Timur Köse and Gülay Köse, and brother Oğuz Alp Köse. Their belief in me has kept my spirits and motivation high during this process.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ANS	Autonomic Nervous System
CNN	Convolutional Neural Network
DVC	Dorsal Vagal Complex
ECG	Electrocardiography
ECOC	Error-correcting output codes
EEG	Electroencephalography
EVI	Emotional Valence Index
GSA	Galvanic Skin Response
HRV	Heart Rate Variability
KNN	K Nearest Neighbor
PNS	Parasympathetic Nervous System
RMSSD	Root Mean Squared Successive Differences
RRA	Respiration Rate Analysis
PSD	Power Spectral Density
SKT	Skin Temperature Measurements
SNS	Sympathetic Nervous System
SVM	Support Vector Machine
VVC	Ventral Vagal Complex
ZCV	Zero Crossing Variance Feature

CHAPTER 1

INTRODUCTION

This chapter provides brief introduction to study. Motivation of work, background research and similar studies are explained in this chapter.

1.1 Motivation and Problem Definition

Emotion recognition is becoming an important subject in robotics, marketing, education, and the entertainment industries. In marketing and entertainment industries it is used for targeting purposes, in robotics it is used to improve human machine interactions, in education it is used to investigate learning process and to improve learning methodo-logies [7].

Need of industry also increases popularity of emotion recognition studies and development of related sensory and devices [8]. Currently, electroencephalography (EEG) and electrocardiography (ECG) are among the most popular human body parameters that are used in emotion recognition [7]. However, there are not definite standards for emotion recognition through electroencephalography signals [9].

Even though, more affordable and acceptable performing EEG devices such as Open BCI and Emotiv EPOC headset exist, according to literature review of Torres et al., new devices will become even more useful [9]. Moreover, They state that current models of EEG devices gives discomfort after long usage time [9].

Since there is no standard for emotion recognition using EEG signals, this work aims to provide a new method that performs on par with existing researches by using a few number of EEG scalp locations to reduce cost and discomfort of a such system, with the help of electrocardiography recordings.

1.2 EEG

EEG is used to measure the electrical activity of the cerebral cortex. EEG can only detect pyramidal neurons on the higher levels of the cerebral cortex. Pyramidal neurons create dipoles that can be detectable via EEG thanks to their unique shape and positioning on the cerebral cortex. Moreover, EEG can sense the voltage change on the scalp due to excitatory and inhibitory postsynaptic potentials of multiple neurons. Even though; the major cause of potential change is action potential, these can not be detected by EEG since they occur too rapidly to sum on each other to create enough potential on the scalp [1]. Action potentials occur around 1 ms, whereas postsynaptic potentials last up to a couple of 10 ms.

Summed neuron potentials create voltage on EEG scalp. If these neurons are fired simultaneously, resulted EEG wave will have high amplitude and low frequency. On the other hand, asynchronously fired neurons cause high frequency and low amplitude EEG signals. These processes are called synchronization and de- synchronization. Figure1.2 shows how 3 Hz firing rate results in different EEG waves for synchronization and de-synchronization.

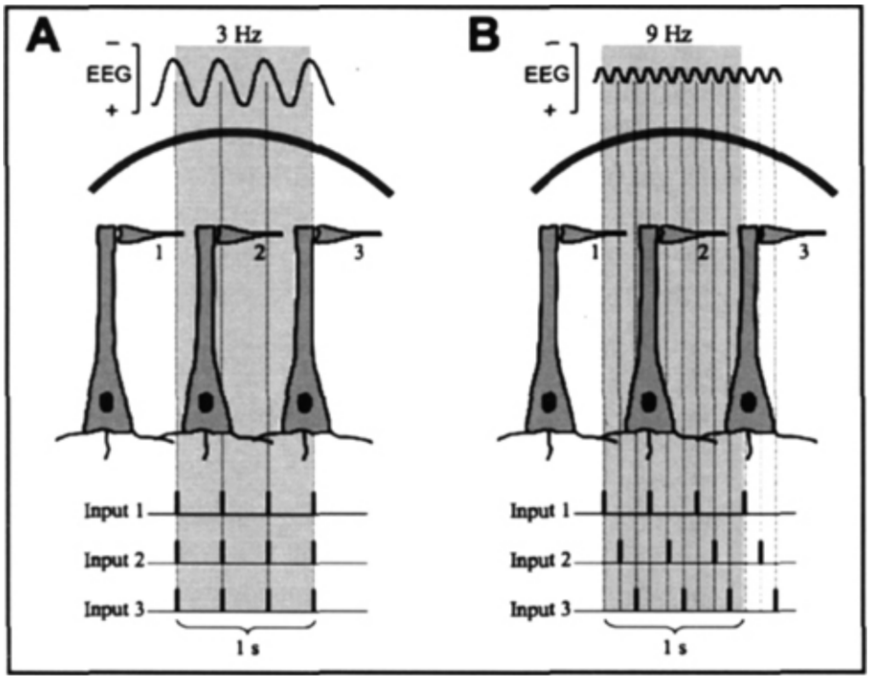

Figure 1.1: Synchronization and desynchronization [1]

EEG signals are divided into frequency bands which are the delta (0.5-4Hz), theta (4-8Hz), alpha (8–13Hz), beta (13–30Hz),and gamma (>30Hz). According to Wang's review on literature, alpha, beta and gamma bands are better to use on emotion recognition studies [8]. Which also complies with Suhaimi's work [10]. Moreover to frequency bands, different brain regions relate different activities. Zhang's findings on EEG scalp locations and emotion recognition task shows that prefrontal cortex activity correlates with emotion in humans [11]. Suhaimi's review also shows that frontal lope has an important role in emotion-related studies [10].

Emotional valence index(EVI) is a EEG feature that relates emotional valence [12]. According to Tomarken's work that shows the relation between frontal asymmetry and emotional valence, frontal lope alpha power on the left sphere decreases with positive emotions while right frontal lope alpha power decreases with negative ones [13]. Some existing works use other frequency bands related to frontal asymmetry metrics [14] [12].

According to Wang's review one of the most popular information theory related EEG feature is sample entropy [8]. Approximate sample entropy (ApEn) was proposed to quantify the complexity of short and noisy time series. Aproximate sample entrophy is calculated as follows. Firstly, probability of vectors with in range r of $X_m(i)$ calculated as $C_i^m(r)$. Then,logaritmic average of all posible vector's probabilities gives $\Phi^m(r)$. Finally, approximate sample entrophy becomes $ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r)$.

$$C_i^m(r) = 1/(N - m + 1) \sum_{j=1}^{N-m+1} \Theta[r- \parallel X_m(i) - X_m(j) \parallel] \tag{1.1}$$

$$\Phi^m(r) = 1/(N - m + 1) \sum_{j=1}^{N-m+1} ln[C_i^m(r)] \tag{1.2}$$

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r) \tag{1.3}$$

1.3 Artifact Subspace Reconstruction

Artifact Subspace Reconstruction is an automatic pre-processing method for EEG signals. ASR learns statistical properties of clean calibration data and compares these statistics with statistics of new data during the processing. ASR consists of two parts; calibration and processing. Calibration includes recording approximately 1 min EEG data in ideal condition (which gives clean data). The algorithm learns the statistical properties of this data and rejects noisy parts that have different statistical properties than the calibration data, in the processing part.

ASR works as follows. It takes low artifact, clean signal $X_c = R^Q xM$ (Q: number of channels, M: sample length). Then decompose it to its principle components $W = [w_1 w_Q] \epsilon R^Q xQ$. Component activations are calculated as $Y = X_c W^T$. For each component's RMS value, mean m_k and standard deviation s_k are calculated in overlapping short time windows (0.5 seconds). Using per-component mean $m = [m_1 m_k]$ and standard deviation $s = [s_1 s_k]$, the vector of per-component thresholds $t_{ij} = m_{ij} + cs_{ij}$ is calculated. "c" is a tunable cut-off parameter that determines how much of signal variance will be removed. Typical it is taken as $c = 5 - 7$. However, according to Chang et al., that is too aggressive and removes parts containing meaningful brain signals, and they suggest that "c" should be between 10 and 100 [15]. In the last step, $U \epsilon R^Q xQ$ is chosen as $U_i j = 1$,if $\sigma_{ij} < t_{ij}$; $U_{ij} = 0$ otherwise. Then X_c becomes $X_{clean} = VM(M \otimes U)^+ V^T X$, where $M = V^T \bar{M}$ and $\bar{M}\bar{M}^T = Cor(X)$.

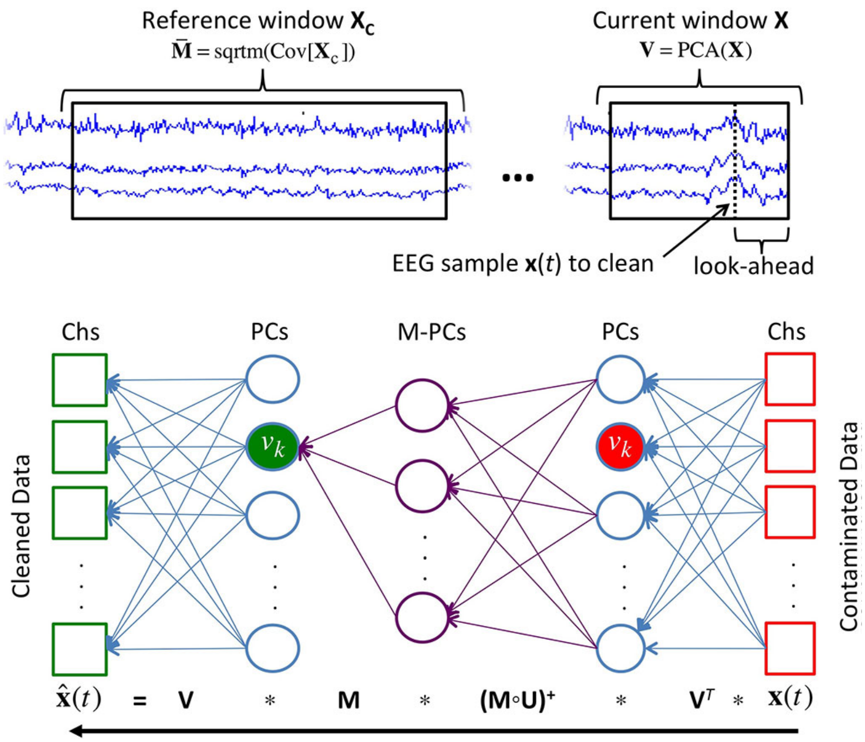

Figure 1.2: ASR method [2]

1.4 ECG

ECG senses electrical signals that occurs during heart beat. Heart consist of four regions which are left atrium, right atrium, left ventricle and right ventricle. Electrical signals that occurs during heart beat, stars from sinoatrial node which resides in right atrium. Then, after small delay, it spreads through ventricle. These signals are measured from skin and ECG signals emerge from them. There are 4 electrodes in a full ECG setup. They can be seen from Figure1.3.

HRV is calculated through an ECG signal. It is the measure of the difference in time between consecutive R intervals of ECG signal which can be seen in idealized form in Figure 1.4.

The autonomic nervous system (ANS) has a critical role in physiological arousal [4]. There are two parts of ANS, which are the excitatory sympathetic nervous system (SNS) and the inhibitory parasympathetic nervous system (PNS). SNS becomes domi-nant and increases heart rate in case of environmental stress to make the body able to adapt to changing environmental factors. PNS, on the other hand, reduces heart rate in stable conditions. ANS causes varying heartbeats in case of stress or different emotions. This property makes HRV useful in emotion recognition studies. Quintana's work provides direct evidence of a relationship between emotion and HRV [16].

There are two different theories regarding the relation between ANS and emotions.

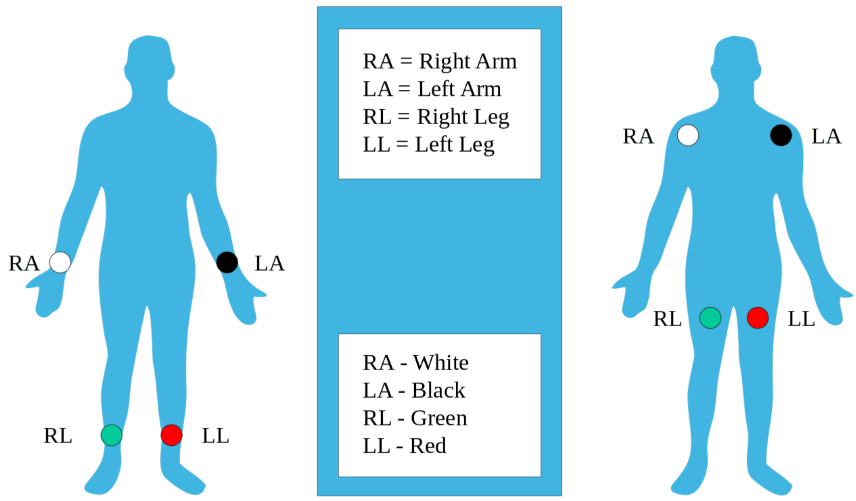

Figure 1.3: ECG Electrode Locations [3]

One of these theories is the polyvagal theory proposed by Porges [17]. The polyvagal theory explains emotional responses (also environmental stressors) through 3 subsystems that become active in sequence depending on their evolutionary stage. In the case of external stimuli, the ventral vagal complex (VVC) which is the most recent one in the evolutionary line among 3 subsystems, becomes active first and causes a rapid increase in heart rate. Then, the sympathetic nervous system is activated to increase heart beat further, due to insufficiency of rapid heart rate increase coming from VVC. Finally, the dorsal vagal complex (DVC) which is the most primitive one among 3 subsystems, becomes active due to excessive usage of oxygen sources in the body caused by increased heart rate and reduced heart rate. Another theory proposed by Thayer and Lane is the neurovisceral model [18]. Their model explains that the emotional process is controlled by a system that is a part of a even bigger dynamic system. This complex system is also connected to ANS which affects HRV.

There are several metrics related to HRV from the literature review of Kim [19]. Some of these metrics are calculated in the time domain such as the root mean square of the successive differences (RMSSD), the number of interval differences of successive NN intervals (consecutive R-R interbeat intervals) greater than 50 ms (NN50). There are also metrics calculated in the frequency domain. As explained in Bradley's work, the high-frequency (HF)component is a power spectral density of 0.15–0.40 Hz band and primarily reflects PNS activity [4]. The low-frequency (LF) component is a power spectral density of 0.04 –0.15 Hz band and mostly relates to SNS activity. The ratio of LF to HF (LH) is also used by researchers to investigate the balance between PSN and SNS.

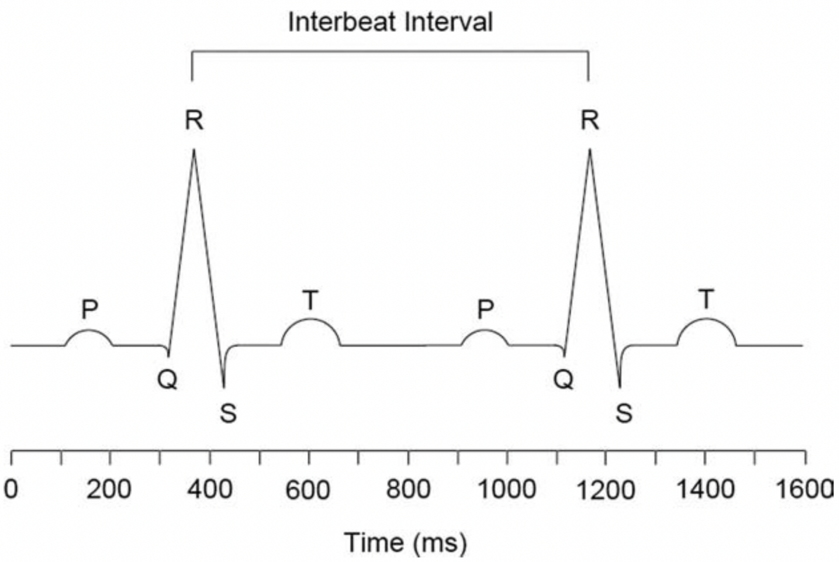

Figure 1.4: Idealized ECG segment [4]

1.5 Emotional Models

The emotional models can be divided into two categories. One of these categories is categorical models. Categorical models put emotions in distinct emotion classes. The other category is dimensional models. Dimensional models represent emotions in a dimensional space that uses a set of common dimensions.

Ekman's basic emotions model is a well-known categorical model. Ekman's model contains six distinct emotional states which are happiness, anger, fear, enjoyment, sadness, and disgust [20]. Ekman's theory comes from the research on facial express-ions. According to his findings, there are six distinct universal facial expressions that repre-sent six different emotions.

Circumplex model of emotion is one of the most widely used emotional models [21]. Model is proposed by Russel, after Schlosberg's proposal which states that emotions can be represented by two bipolar dimensions (valence and arousal) [5] [22]. Dimensions in the circumplex model are valence and arousal. Figure 1.5 shows representation of eight emotions in valence-arousal space. The valence scale represents the direction of the feeling or emotion, such as how positive the emotion is. The arousal scale represents the intensity of the emotion experienced in response to the stimuli. Arousal is mostly about the physical response of the body.

Categorical models may cause inaccuracy. When the exact names of emotion felt by the subject doesn't appear in categories, the subject may choose different emotions due to personal differences. Therefore, bipolar dimensions provide more accurate representation especially in self-reporting studies [21]. Schlosberg's model comes from the examination of errors that occur during the classification of emotional states of facial expressions [22]. Russell also favors bipolar dimensions in self-reporting studies [5]. Moreover, Thayer states that the dynamic system nature of the neurovis-

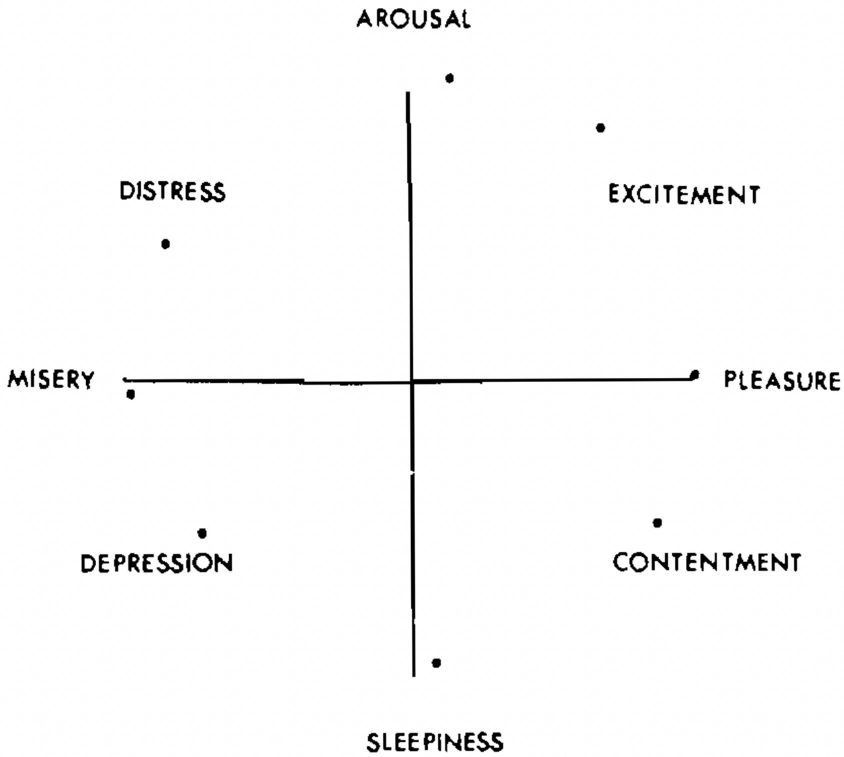

Figure 1.5: Circumplex Model of Affect as Eight Emotions Represented in Valence-Arousal Space [5]

ceral model makes sense with dimensional models [18].

1.6 Classification Methods

Literature reviews show that KNN and SVM are the two most popular classification algorithms in emotion recognition studies [23]. Since this work uses 3 classes that should be classified, a binary classifier such as SVM is not applicable. Therefore, This subsection provides a brief explanation of KNN, ECOC (which is a machine learning method such as SVM and can be used on classification problem that involves more than 2 classes), and random forest method which is the main classification method that used in this work.

Random forest classifier is made of multiple number of tree predictors that cast a vote on prediction result according to their prediction [24]. Prediction is made according to popular result among predictions of each individual tree.

In the training procedure of random forest, each tree is trained with random samples from the training set. These random samples may or may not return the training set depending on the random forest model. If the selected samples return to the training set, they may be chosen to train another tree in the model.

Trees are grown depending on features' Gini scores. Gini score represents impurity of classification related to the feature (so, it is the probability of the wrong classification in case of a random selection related to a specific feature). Equation 1.4 shows calculation of Gini score. Each node is divided according to features with the lowest Gini score.

Hyperparameters of the random forest method are the number of trees, the maximum number of split, mtry, and replacement [25]. Number of trees is a self-explanatory parameter that represents the number of trees in a random forest model. Number of the split is the parameter that limits how deep can a tree split. Mtry is the number of variables that are randomly sampled in each split and for classification problems, optimum usage is $sqrtp$ (where p is a variable number). The replacement parameter chooses whether or not randomly chosen samples will be available for other trees.

$$G = \sum_{i=1}^{C} p(i) * p(1 - i) \tag{1.4}$$

K nearest neighbor algorithm works as follows. Let's assume there are sample pairs as X_s for variables and Y_s for class. While trying to estimate point y from variables x, K number of points from X_s that have the lowest distance to x is determined and they become X_k and related classes become Y_k. The class that has the most number of members in Y_k, is predicted as y.

Error-correcting output codes models can be used for classification problems that involve more than two classes.Support vector machine models can not be used, since they are binary classifiers. In ECOC models, multi-class classification problems are solved as several binary class problem by using multiple SVM classifiers.

There are two different methods to reduce multiple class problems to several binary classification problems. The first method is one against one. In this method, from all exiting classes, two pair is chosen an d they are treated as a binary class problem. Therefore, for each pair of classes one SVM classifier is trained. That means for the K class, the $K(K-1)/2$ model is trained. The second method is one against all. In this method, each exiting class and all remaining classes becomes a binary classification problem. Therefore, for each class one SVM classifier is trained. That means for the K class, the K model is trained.

1.7 Classification Performance Metrics

This sub section explains calculations and meanings of classification performance metrics used in this work. Since the work classifies 3 class problem, this chapter focuses on only performance metrics related 3 class problem.

From Table1.1, for i^{th} class, T_i shows the number of times it predicted correctly and F_{ji} shows the number of times it predicted as j instead of i.

Table 1.1: Example Confusion Matrix

		True/Actual		
		Class1	Class2	Class3
Predicted	Class1	T_1	F_{12}	F_{13}
	Class2	F_{21}	T_2	F_{23}
	Class3	F_{31}	F_{32}	T_3

Precision is measurement of how many times prediction made as i, actually i. Equation 1.5 shows how precision can be calculated for "Class1". It is the ratio between positive and false positives for that specific class.

$$Precision_1 = \frac{T_1}{T_1 + F_{12} + F_{13}} \tag{1.5}$$

Recall is measurement of how many times class i is correctly classified. Equation 1.6 shows how precision can be calculated for "Class1". It is the ratio between positive and false negatives for that specific class.

$$Recall_1 = \frac{T_1}{T_1 + F_{21} + F_{31}} \tag{1.6}$$

F1 score provides insight about how successful classifier is. For a specific class, it is the harmonic mean of the class's precision and recall score. Equation 1.7 shows the calculation of F1 score for Class1.

$$F1score_1 = \frac{2 * Precision_1 * Recall_1}{Precision_1 + Recall_1} \tag{1.7}$$

Accuracy is ratio between number of correct predictions and total predictions. Equation 1.8 shows how it is calculated.

$$Accuracy = \frac{T_1 + T_2 + T_3}{T_1 + F_{21} + F_{31} + T_2 + F_{12} + F_{13} + T_3 + F_{13} + F_{23}} \tag{1.8}$$

F1 scores are calculated each class separately,. However, in order to make clear comparison between classifiers, overall F1 score should be calculated. There are two different methods for that purpose. First one is Macro-F1 score and it is calculated by taking arithmetic mean of F1 scores of classes. Equation 1.9 shows calculation of Macro-F1 score. Second one is weighted F1 score which is calculated by taking weighted average of F1 scores. In weighted average case, class distribution of true data is used as weights. Example of calculated weight for Class1 is shown in equation **??** and Weighted-F1 score calculation is shown in 1.11.In same way both macro and weighted average of recall and precision can be calculated.

$$Macro - F1 = \frac{F1score_1 + F1score_2 + F1score_3}{3} \tag{1.9}$$

$$Weight_1 = T_1 + F_{21} + F_{31} \tag{1.10}$$

$$Weighted{-}F1 = \frac{weight_1 * F1score_1 + weight_2 * F1score_2 + weight_3 * F1score_3}{weight_1 + weight_2 + weight_3} \tag{1.11}$$

Finally, After calculation for other classes, performance metrics becomes as table1.2.

Table 1.2: Arousal Classification Results for Fixed Channel Case

Class Name	Prec	Recall	F1	Sample Size
Class1	$Precision_1$	$Recall_1$	$F1score_1$	$weight_1$
Class2	$Precision_2$	$Recall_2$	$F1score_2$	$weight_2$
Class3	$Precision_3$	$Recall_3$	$F1score_3$	$weight_3$
macro avg	$Macro - Precision$	$Macro - Recall$	$Macro - F1$	
weighted avg	$Weighted - Precision$	$Weighted - Recall$	$Weighted - F1$	
		Accuracy		

1.8 Chi-Squared Test

Chi-squared test is a non-parametric statistical test that investigates the relationship between two categorical variable. Relationship is investigated through comparison between frequencies observed in the categories and expected frequencies which is the value if the frequency is obtained by pure chance. Equation 1.12 shows calculation of chi-squared test.

$$X^2 = \sum \frac{(observed_{ij} - expected_{ij})^2}{expected_{ij}} \tag{1.12}$$

Table 1.3: Example Frequency Distribution Table For 2 Classes

Class Name	B1	B2	B3
A1	V_{11}	V_{12}	V_{13}
A2	V_{21}	V_{22}	V_{23}
A3	V_{31}	V_{32}	V_{33}

From Table1.3, A and B are two variables with each having 3 classes. V_{11} to V_{33} are observed values.

For each observation, expected value is calculated through ratio between multiplication of total values in the row and the column that observed value fills and the total number of observation. Example calculation for observed value V_{11} is shown in Equation 1.13.

$$Expected_{11} = \frac{(V_{11} + V_{12} + V_{13}) * (V_{11} + V_{21} + V_{31})}{V_{11} + V_{12} + V_{13} + V_{21} + V_{22} + V_{23} + V_{13} + V_{31} + V_{32} + V_{33}} \quad (1.13)$$

As X^2 value is increased, distribution between variables are less likely to happen by pure chance. Therefore, two variables are more likely to be related to each other. This method also can be used to compare features in a classification problem. Importance of each feature can be understood by comparing chi-squared test results between each feature and the class.

In this study matlab function fscchi2() is used to get importance scores of features. However, as mentioned before chi-squared test can be used between categorical variables and features in this study are continuous. Therefore, function quantizes continuous data to bins then applies chi-squared test as if it is categorical data.

1.9 Similar Studies

This study proposes an emotion recognition method that uses ECG and EEG signals. In this section, methods used by similar emotion recognition studies are investigated. EEG metrics in current studies are discussed here since they are different than the most popular metrics previously mentioned in EEG section. However, HRV metrics are the same as previously mentioned features, therefore they are not discussed in this section. Moreover to EEG metrics other available human body parameters are briefly mentioned.

There are different human body parameters used in similar studies. Two of the available parameters are discussed in previous sections since they are the topic of this study. Other parameters are galvanic skin response (GSA), respiration rate analysis (RRA), skin temperature measurements (SKT), and electromyogram. GSA is the measurement of the electrical conductance of the skin. It is usually used with ECG features to stress assessment. RRA is the measurement of pace and depth of breathing. It is possible to predict it from ECG and is usually used with other metrics to predict emotions [7]. SKT is the measurement of the surface temperature of the skin. It is related to parasympathetic system activity like HRV features. Generally, there is a large delay between emotion occurrence and change in SKT measurement. EMG is used to detect muscle activity in humans through electrodes placed on the skin.

The most commonly used EEG features are power spectral density, statistical, wavelet transform, differential entropy, sample entropy, wavelet empirical mode decomposition according to Wang's review [8]. Kim's review also mentions power spectral density and adds brain asymmetry-related features [26]. Among studies that compared to this study, Cui's work extracts features related to brain asymmetry via extracting information through several CNN [27]. In Cheng's work pre-processed EEG data directly feed into a cascaded random forest model which output of each level of random forest model becomes the next level's features, as a 2D vector [28]. In Liu's work again, EEG data is processed through series of specialized deep learning models to acquire features then to be classified [29]. In Wang's work symmetric and positive definite (SPD) matrix which is a convolutional network, method is used to extract features from EEG data [30]. Maheshwari's work power of different frequency bands of EEG signals are used as features [31]. In Song's work, the graph that consists of nodes made by differential entropy feature of each EEG channel is used as a feature [32].

According to Wang's review on literature, Support Vector Machine(SVM), KNN, Random Forest, Naive Bayes(NB) are the most popular machine learning methods in the literature [8]. Moreover, Kim's 2013 review on literature mentions SVM and KNN as popular methods [26]. However, neural network methods are gaining popularity in recent years [8]. This also complies with recent studies that are compa-red to the results of this study in Table 3.7.

In this study, the Dreamer data set is used. There are also other available data sets that exist. DEAP data set contains EEG data of 32 subjects collected during music video watching sessions. For 22 subjects, frontal face video is also provided. After sessions participants rate music pieces according to valence, arousal, and dominance. For the

SEED data set, emotional video clips are shown to 15 participants. EEG recordings and frontal face videos are collected. Emotional states are collected through a question-naire.

CHAPTER 2

METHOD

In this chapter, methods and design choices are explained. In this work, the Dreamer data set is used [33]. Dreamer data set contains arousal and valence scores which are given by 23 participants to 18 videos. In video sessions, ECG and EEG data of participants are collected. Due to the difference in the length of videos and to provide the necessary time to emotions which are related to video become dominant, the last 60 seconds of each video is used, as suggested by Katsigian-nis [33].

Valence and arousal scores are predicted through ECG and EEG recordings, rather than the actual emotions assessed to videos. Since assessment has been done in different study beforehand, the response of Dreamer study participants may differ from the assessed emotions. Moreover, There are works favoring the valence and arousal model in self-reporting studies [21] [18] [22].

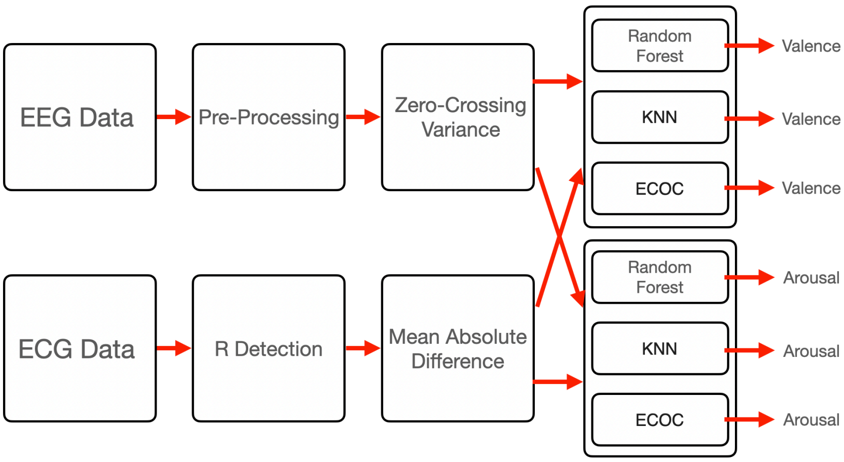

Figure 2.1: Valence-Arousal Prediction System

There are 3 main blocks to the proposed prediction system, which are pre-processing, feature extraction and classification. The following sub-chapters explain those steps. Figure 2.1 shows overall system blocks.

2.1 Dreamer Data Set

Dreamer is a data set that contains EEG and ECG recordings of 23 participants [33]. Recordings are collected during participants are exposed to audio and visual stimuli. 18 video clips from Gabert-Quillen's work are used as stimuli [34]. Video clips represent 9 different emotions which are amusement, excitement, happiness, anger, disgust, fear, sadness, surprise, calmness.

ECG recordings in the data set are collected with SHIMMER wireless sensor platform. The platform uses Class 2 Bluetooth for communication. Moreover, the platform provides LA -> LL and RA -> LL readings with 256 Hz sampling frequency. Burns et al. show that ECG readings taken from SHIMMER achieved a sensitivity of 99.6%, a mean positive predictivity of 99.8%, and a mean detection error rate of 0.5% with QRS detection algorithm [35].

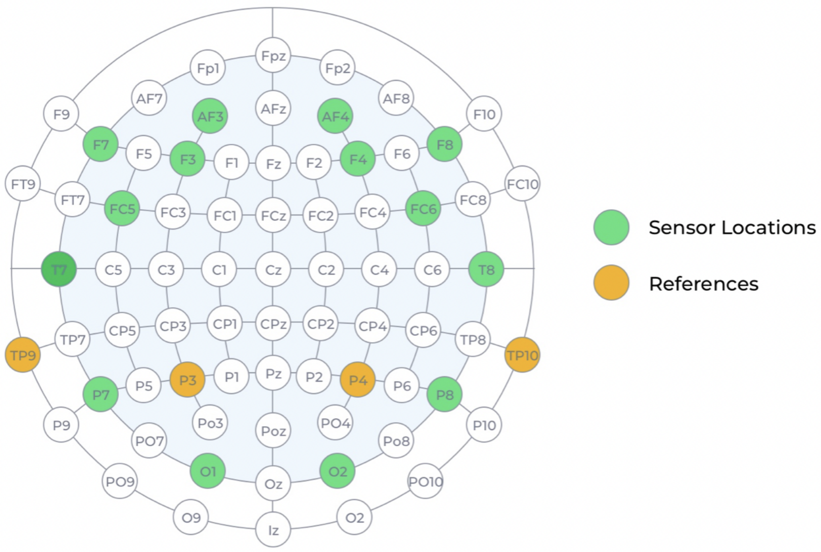

Figure 2.2: Emotiv EPOC wireless EEG headset Scalp Locations [6]

EEG recordings in the data set are collected with an Emotiv EPOC wireless EEG headset. Headset provides EEG signal on locations international 10-20 system AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, M1, and M2. Locations also provided in Figure 2.2. Reference points M1 and M2 are shown as P3 and P4 in Figure 2.2. M1 is used as a ground reference for all other sensors and M2 is used as a reference for reducing external electrical interference. Emotiv EPOC wireless EEG headset works with a 128 Hz sampling frequency.

Low grade Emotive Epoc Headset is expected to provide EEG signal with poor signal to noise ratio. However, Badcock's work that compares Emotive Epoc Headset to research-grade Neuroscan Synamps shows that Emotiv EPOC headset can be an alternative to research-grade devices [36]. Through investigating correlations between signals in same channels of research grade device and EPOC Emotive Headset, the study shows that resulted wave forms are significantly similar. Moreover, study compares peak amplitude and latency measures of resulted signals through t-test, and finds no significant difference between measurements. Therefore, Emotive EPOC headset could be a valid alternative to research grade devices.

2.2 Pre-processing

EEG signals are contaminated by artifacts from different sources. These sources could be external or internal. External artifacts may occur due to faulty electrodes, line noise and high electrode impedance. Internal artifacts are caused by physiological sources such as muscle activity, cardiac activity or eye movement [37]. Those artifacts should be eliminated to increase the quality of data for classification. ECG signals, on the other hand, carry little noise which does not require pre-processing. Therefore, this section only covers pre-processing on the EEG signals.

According to Stamos Katsigiannis and Naeem Ramzan, in Dreamer Database, EEG data is corrupted by ocular artifacts below 4 Hz, muscle movement related artifacts above 30 Hz and power line noise between 50 and 60 Hz [33]. Therefore, there is a relatively clean part between 4 Hz and 30 Hz. Moreover, this part carries information related action potentials an it is divided into different frequency bands. These bands are alpha, beta, and gamma.

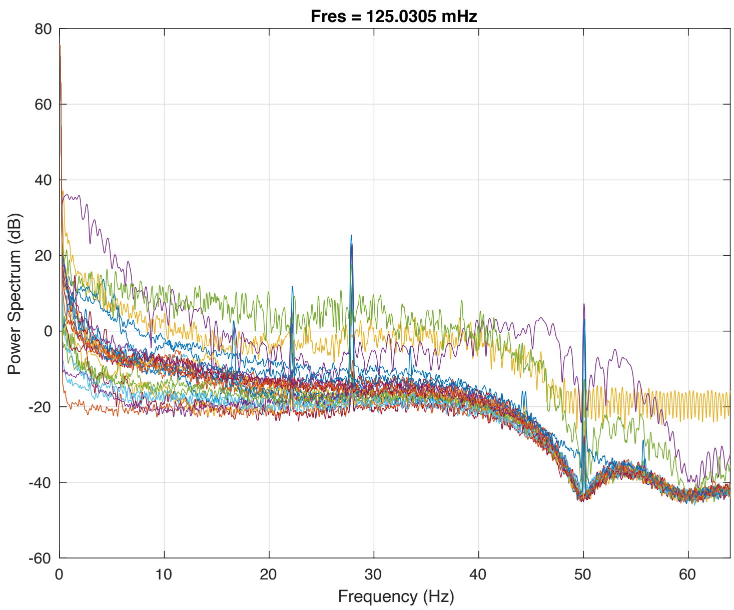

Figure 2.3: PSD of EEG signal on P7 for 23 participants before filters applied

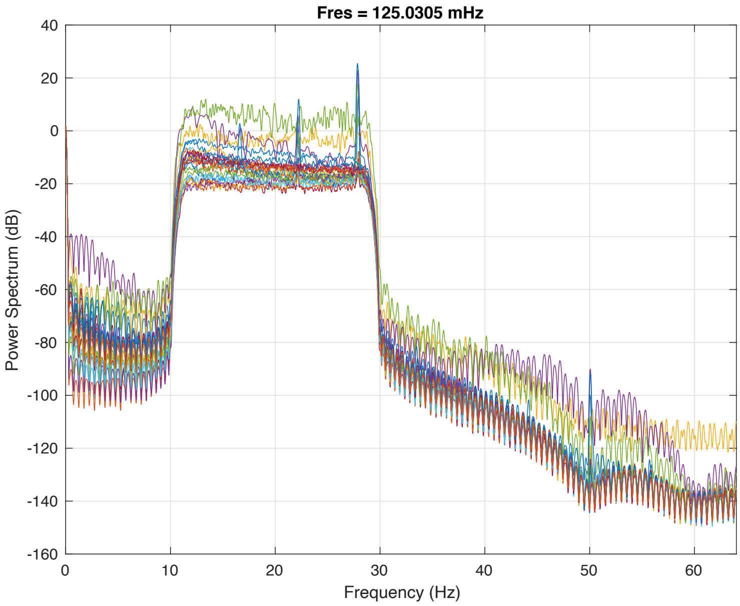

Figure 2.4: PSD of EEG signal on P7 for 23 participants after filters applied

In this study, alpha and beta bands are utilized. To get this band, firstly, a bandpass filter with a passband between 12 Hz and 30 Hz is applied to the signal. Figure 2.3 shows power spectral density of raw signal and Figure 2.4 shows PSD of the filtered signal for 23 participants. As Figures show, higher and lower frequency component of signals are sufficiently reduced. Matlab's higpass() function is used to build filter.

2.3 EEG Features

The frequency of EEG signals varies over time due to the nature of fired neurons on the cerebral cortex. As mentioned before, de-synchronization and synchronization of fired neurons result in low and high-frequency signals. These frequency changes are related to brain activity, so the relation between the frequency of EEG signals and the emotion of a person is investigated by many different works. According to Wang's review on literature, different frequency bands reflect the different state of mind, and also positive and negative emotions relate to different frequencies of EEG signal [8]. EEG feature proposed in this work uses these frequency changes in the signal.

The feature represents the amount of frequency changes over a specified time window. First step of the feature calculation is the determination of the time difference between consecutive zero-crossing points of EEG signal. The second step is calculation of variance of the time differences in a 4.5 second time window.

The emotive epoc headset is a low spec device with poor sampling frequency. To reduce the effects of the sampling frequency, an extra step is introduced into the determination of zero-crossing points. As mentioned in work of Albert et al. , to determine zero-crossing points more accurately values of samples around sign change occurs are taken into account [38]. Where ZC is zero-crossing point. v_1 is the sample before and v_2 is the sample after than the zero-crossing point. t_1 and t_2 are their respective times; ZC is calculated as equation 2.1. Time and values are also visible in figure2.5.

$$ZC = t_1 + ((t_2 - t_1) * v_1/(v_1 - v_2)) \tag{2.1}$$

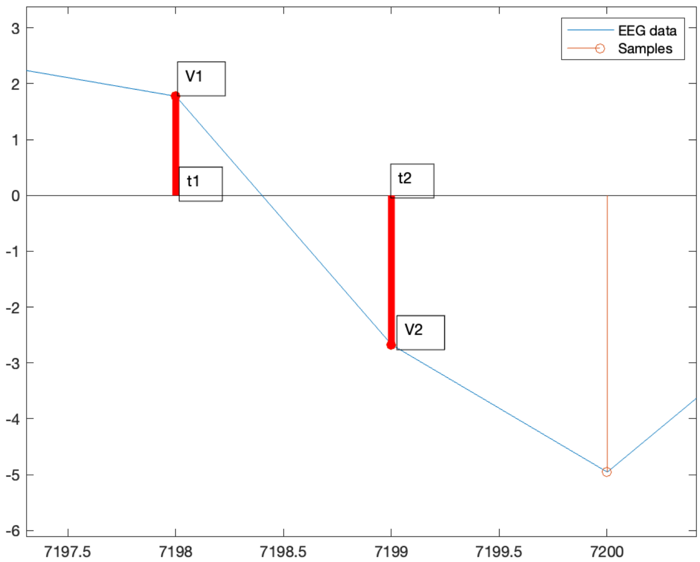

Figure 2.5: Example Signal and Samples Around Zero Crossing Point

2.4 ECG Features

ECG feature in this study is based on heart rate variability. HRV is the measurement of time between consecutive R curves in ECG signals. Detection of R peaks is done as follows. Firstly, a mean filter with a window length of four samples is applied to the ECG signal to reduce P and Q peaks. This prevents miss identification of P and Q peaks as R peaks. Then, the Matlab function "findpeaks()" is applied to the filtered signal to detect R points. "findpeaks()" function detects samples that have a higher value than both of its neighbors and also higher than a given threshold. The threshold is chosen as 0.9, since R peaks occur higher than 0.9 in available data.

HRV feature that is used in this work, is the mean absolute difference. The mean absolute difference feature is calculated by averaging absolute values of difference in consecutive HRV values. Where W is window length, HRV and MAD are hearth rate variability and mean absolute difference respectively; Equation 2.2 shows how the mean absolute feature is calculated.

$$MAD(r) = 1/(W) \sum_{i=rW}^{W*(r+1)-1} |HRV(i+1) - HRV(i)| \tag{2.2}$$

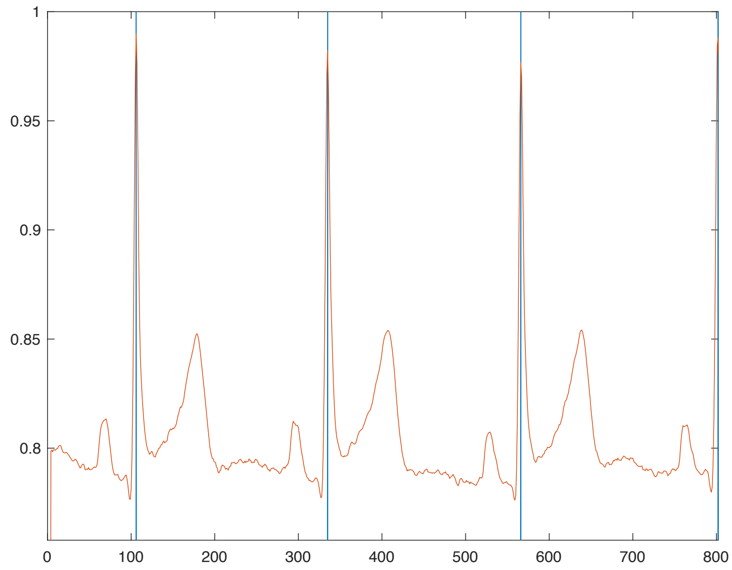

Figure 2.6: Filtered and Raw ECG with Detected R Peaks

2.5 Classification

The classification method chosen in this work is the random forest method due to its performance compared to KNN and SVM methods. Both arousal and valence prediction models trained with 2 EEG and 1 ECG feature which are chosen according to chi-square test results(optimum locations for each participant given in Table A.1 - Table A.4).

In the random forest model, Matlab's fitcensemble() method with 'bag' option is used. The method provides good classification accuracy even with default parameters, however, training and prediction take a longer time. Therefore, without losing too much accuracy, the number of trees and maximum split of trees are reduced. In Figure 2.7 and Figure 2.8, average resubstitution loss for 23 participants related to maximum number of split and number of trees is provided. Resubstitution loss is the weighted fraction of misclassified observations and calculated through resubLoss() function of Matlab. Where n is number of samples, w_j is weight of related observation, $\hat{y_j}$ is predicted class and y_j is actual class; resubstitution loss becomes as Equation 2.3.

$$L = \sum_{j=1}^{n} w_j I(y_j \neq \hat{y_j}) \tag{2.3}$$

Random Forest models are tested with 2,4,16,32,64,128 and 256 maximum number of split and 1 to 150 number of trees. Optimization is done according to the average resubstitution loss of 23 models (one model for each participant). Substitution loss is calculated by using out of bag samples for optimization. However, overall classifica-tion performances given in "Results" section, are calculated with separate samples(vali-dation set). Figure 2.7 and Figure 2.8 shows the optimized parameters. For valence, the maximum number of the split is 64 and the number of trees is 9. For arousal, the maximum number of the split is 128 and the number of trees is 11.

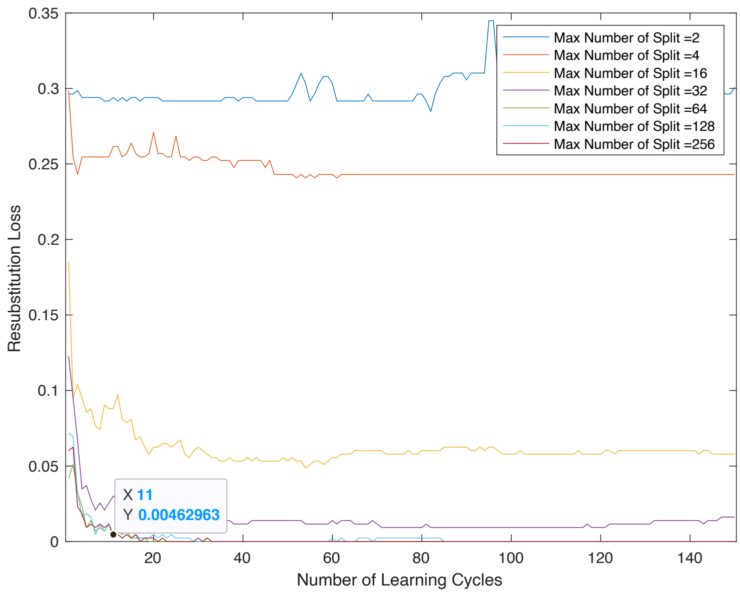

Figure 2.7: Arousal Resubstitution Loss

In KNN models, Matlab function fitcknn() is used. As mentioned before, the only parameter to control in such model is the K which controls the number of neighbors that is used to predict the class of a data point. For different overall accuracy of 23 participants that is calculated by averaging 5 trials for different "K" values is provided in Figure 2.9 and 2.10.

KNN models have tested 2 to 50 neighbors by increasing the number of neighbors by 2 in each iteration. Optimization is done by considering the average accuracy of models for 23 participants. The number of neighbors is 27 for valence classification. The number of neighbors is 26 for valence classification. Those can be seen through Figure 2.9 and 2.10 respectively.

For ECOC models, Matlab function fitcecoc() is used with one vs all option. In one vs all method, ECOC models train SVM models for each class to treat the multi-class problem as several binary classification problems. Learner type choice that is used in SVM is investigated in this section.

In ECOC models, 3 different SVM learners are tested. These learners are linear, Gaussian, and polynomial. Optimization is done by considering the average accuracy of models for 23 participants. Linear learners performed poorly, while polynomial and Gaussian learners show similar results. Since polynomial learners show more robust classification results, polynomial learners are chosen. Accuracy related valence and arousal cases can be seen through Figure 2.11 and 2.12 respectively.

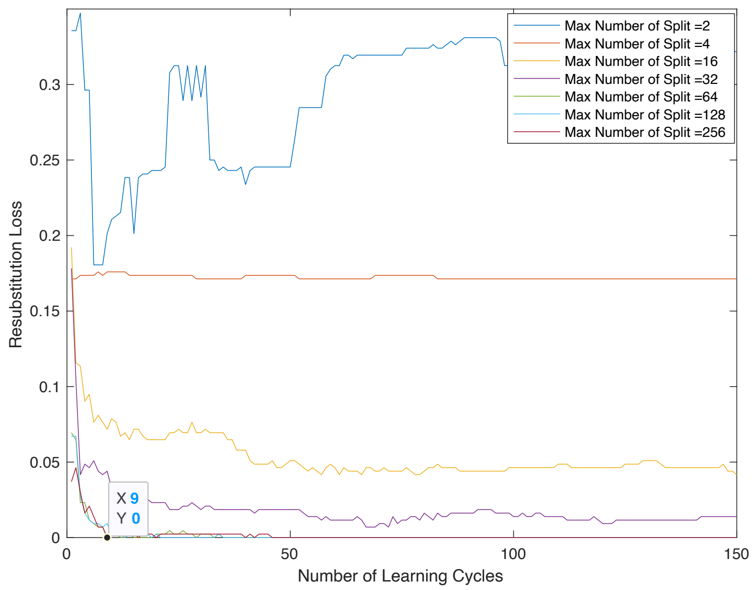

Figure 2.8: Valence Resubstitution Loss

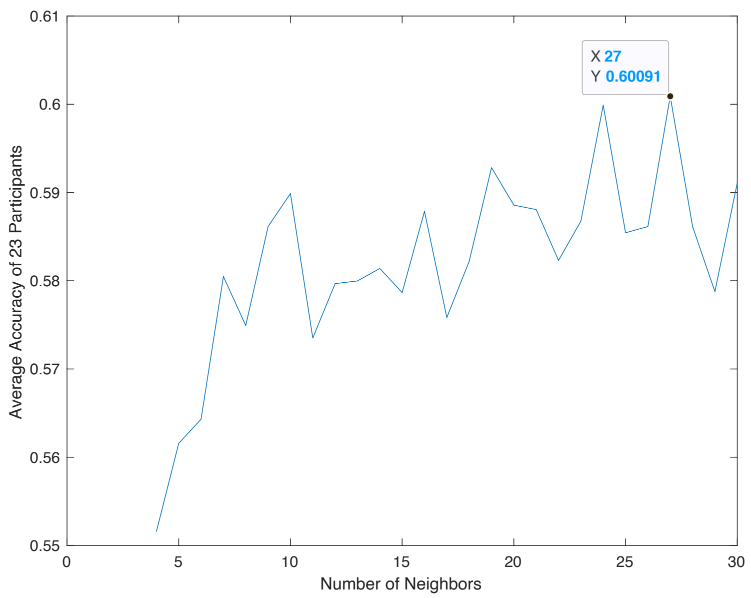

Figure 2.9: Valence Accuracy vs Number of Neighbors for KNN Model

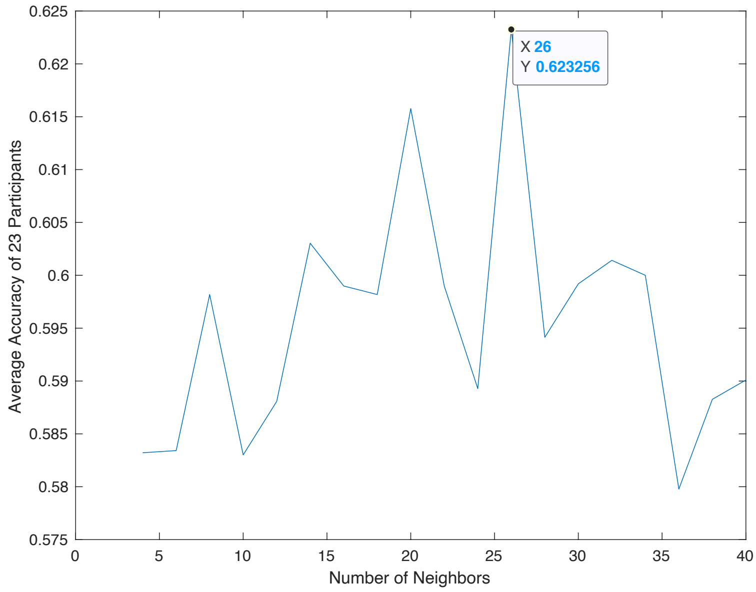

Figure 2.10: Arousal Accuracy vs Number of Neighbors for KNN Model

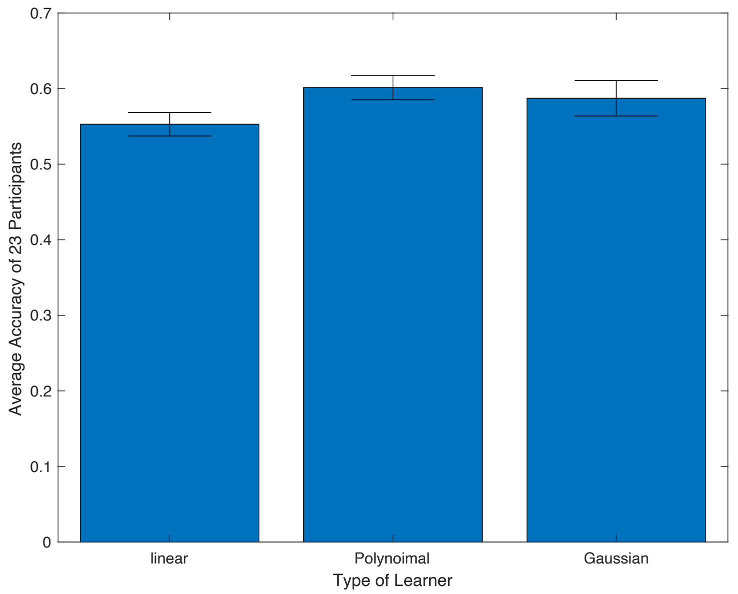

Figure 2.11: Valence Accuracy vs Learners for ECOC Model

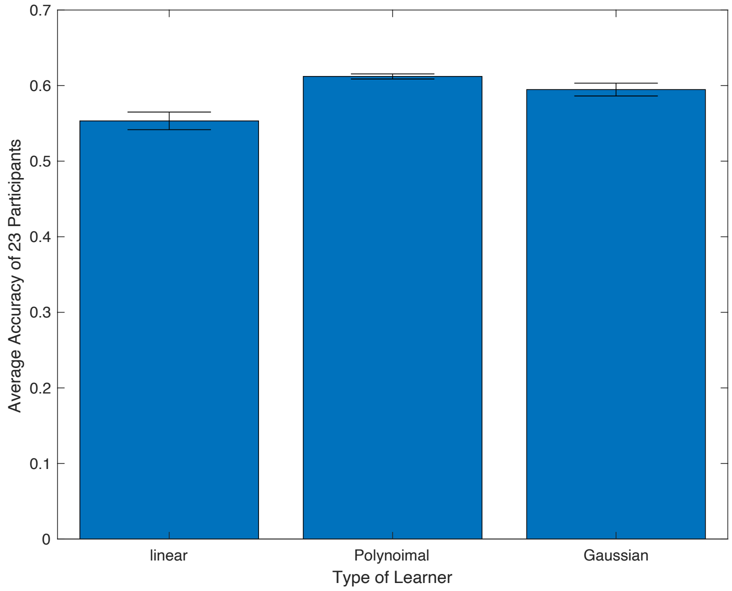

Figure 2.12: Arousal Accuracy vs Learners for ECOC Model

CHAPTER 3

RESULTS

In this chapter, classification performance metrics related to different classifiers and features are presented. Classifier performances that are presented in this chapter are calculated as follows. For each participant, 5 trials are performed. For each trial, new random samples are selected to train and validate the new model. Moreover, a confusion matrix is calculated for that model. A total of 115 (for 23 participants and 5 trial) confusion matrices are added to each other. Then, the performance metrics are calculated from resulted confusion matrix.

Valence and arousal data is given as 5 scale data in the Dreamer data set. However, in this work, the scale is reduced to 3 (high, neutral, low). Values 4 and 5 in the Dreamer data set are mapped into the "High" class, 3 is mapped into the "Neutral" class, 1 and 2 are mapped into "Low". Resulted class distributions are given in Fig. 3.1 and Fig. 3.2. High class dominates class distribution in arousal data, while neutral and low classes are sharing remaining data points equally. In valence data, low and high classes have a similar number of data points and neutral class has lower data points than them.

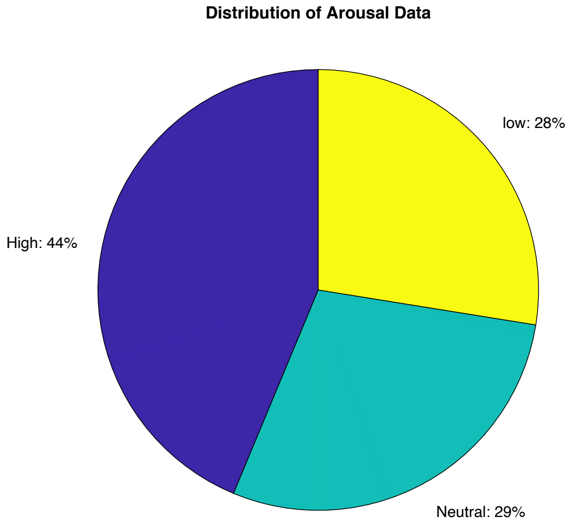

Figure 3.1: Distribution of Arousal Classes

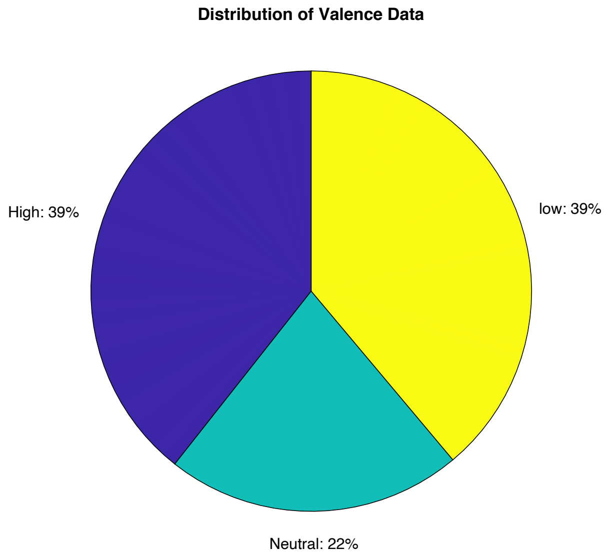

Figure 3.2: Distribution of Valence Classes

3.1 Evaluation of EEG Features

In this section, performance analysis of EEG features are provided for both valence and arousal estimation. Predictor importance scores are calculated through the chi-squared test and test results are presented as bar graphs. Each bar in the bar graphs in this section represents the average importance score of 23 participants for the related feature.

In Figure 3.3 and Figure 3.4, power, brain asymmetry, sample entropy and zero-crossing variance features are compared to each other. Each power, sample entropy and zero-crossing variance feature is calculated through one channel. This channel is different from participant to participant. "ZCV Ch1" , "Pow Ch1" and "SampEnt Ch1" are zero-crossing variance, power and sample entropy features calculated through best channels respectively. For example, "ZCV Ch1" for arousal case is calculated by using channels "EEG_1" from Table A.3 and Table A.4. These channels achieve highest chi-squared test scores for the related participant with "ZCV Ch1". Similarly, for "Pow Ch1" and "SamoEnt Ch1" best channels with related feature is used. In "Ch2" features, second best channels are used. Brain asymmetry feature is calculated by using all channels.

According to Figure 3.3 and Figure 3.4, "ZCV Ch1" and "ZCV Ch2" are the most important features for both arousal and valence estimation. Brain asymmetry is the third most important feature for arousal and valence.

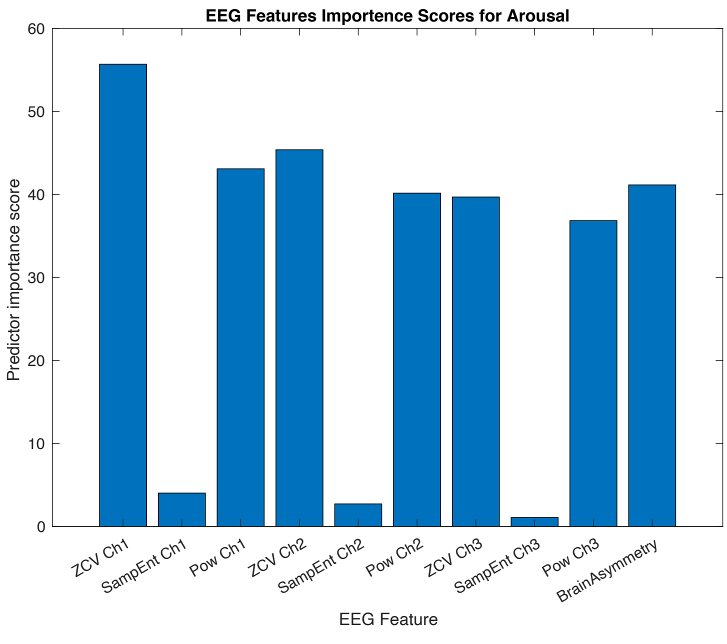

Figure 3.3: EEG Feature Importance Scores for Arousal Estimation by using chi-square test

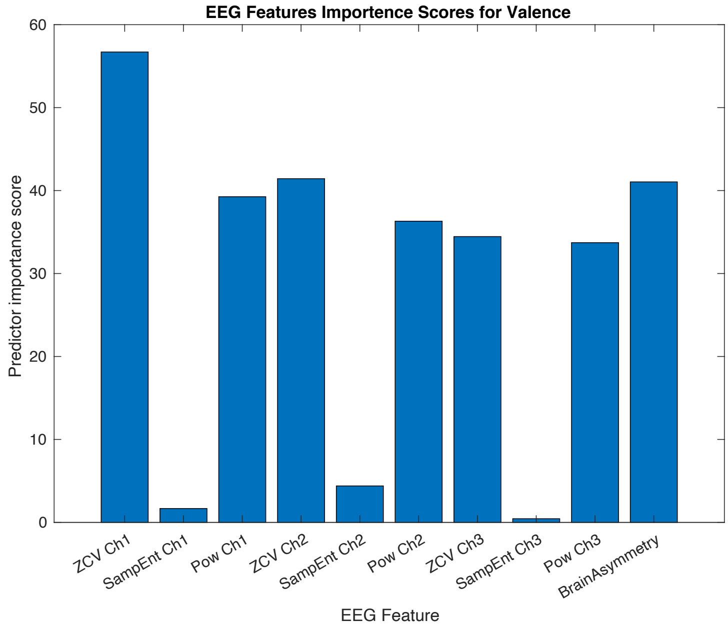

Figure 3.4: EEG Feature Importance Scores for Valence Estimation by using chi-square test

3.2 Evaluation of Features

In this section, performance analysis of features are provided for both valence and arousal estimation. Predictor importance scores are calculated through the chi-squared test and test results are presented as bar graphs. Each bar in the bar graphs in this section represents the average importance score of 23 participants for the related feature.

RMSSD, NN50, MAD, LH have previously mentioned HRV features. EVI and EEG_1-EEG_6 features are EEG features. EGG_i feature is a zero-crossing variance metric obtained from the EEG channel that results in i^{th} highest importance score. For example; EEG_1 is a zero-cross variance metric calculated from the EEG channel that provides the best importance score.

In Figure 3.5; importance score of different features for arousal estimation is given. The most important feature is MAD which is the HRV feature. Since other HRV features have low importance scores, MAD is the only HRV feature that is used in classification. Zero crossing variance metrics have relatively higher importance than the EVI. EVI has higher importance than only after EEG_4 and less important zero-crossing variance metrics. Therefore, EVI is not used in arousal estimation.

In Figure 3.6; the importance score of different features for valence estimation is given. The importance of HRV features in valence estimation is similar to arousal estimation. While MAD is the most important feature, other HRV features are not important. Scores of EEG features in valence estimation is also similar to arousal estimation case. However, in this case, EVI is slightly more important. Even though,

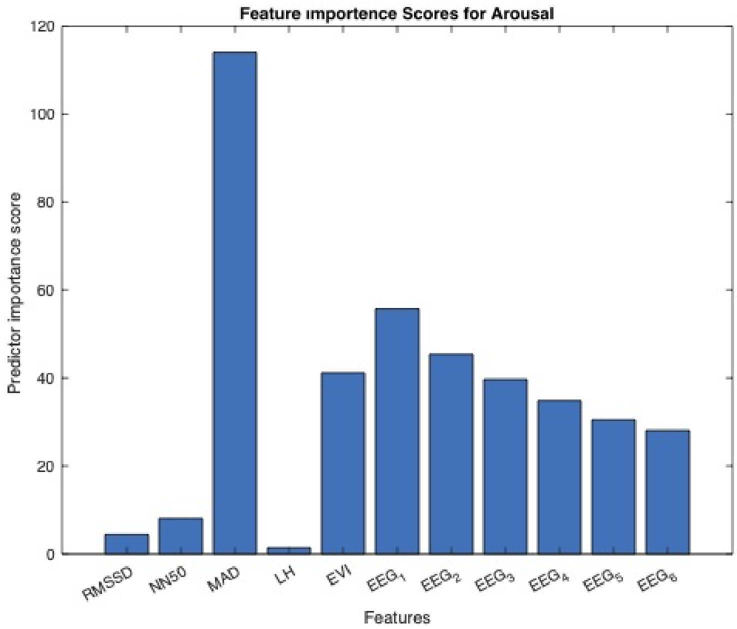

Figure 3.5: Feature Importance Scores for Arousal Estimation by using chi-square test

EVI feature is more important than arousal case, it is not used, since EEG_1 and EEG_2 are sufficient for classification.

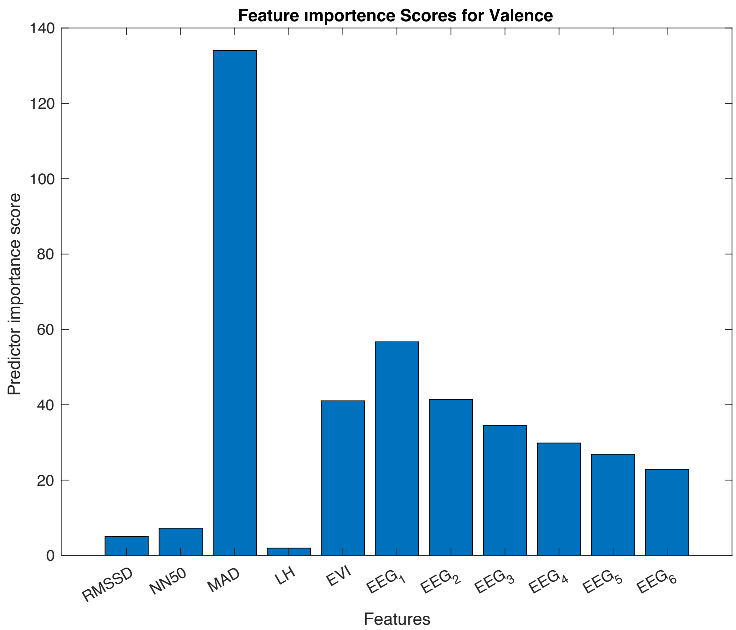

Figure 3.6: Feature Importance Scores for Valence Estimation by using chi-square test

3.3 Evaluation of ZCV Feature in Terms of EEG Scalp Locations

In this section, the zero-crossing variance of different EEG locations is evaluated in terms of importance to arousal and valence estimation. Importance is calculated through the chi-squared test. Scores and rankings of each EEG location are presented in the bar graphs. Each bar in the bar graphs that relates importance scores (Fig.3.7, Fig. 3.9), represents average importance score of 23 participants for related EEG scalp location. Bars in ranking graphs (Fig.3.8, Fig. 3.10) represents total points for 23 participant. For each participant, the most important EEG location gets 13 points, while the least important one gets 0. Each remaining location gets one less point as its rank gets lower. Therefore, a higher bar represents a better EEG location in ranking graphs.

Figure 3.8 and Figure 3.7 shows that P7, T7, FC5, F3, F7, T8 scalp locations provide better overall valence estimation performance with zero-crossing variance feature. Figure 3.10 and Figure 3.9 shows that P7, T7, FC5, F3, F7 scalp locations are important in arousal classification like valence case. However, for arousal case, F3 is the most important location.

In this chapter both rank and score graphs are provided in case of outlying chi-squared test results that might occur in some participants. However, rank and score graphs show similar results regarding the importance of EEG scalps locations. Therefore, it is safe to say that the locations provided in this chapter are important.

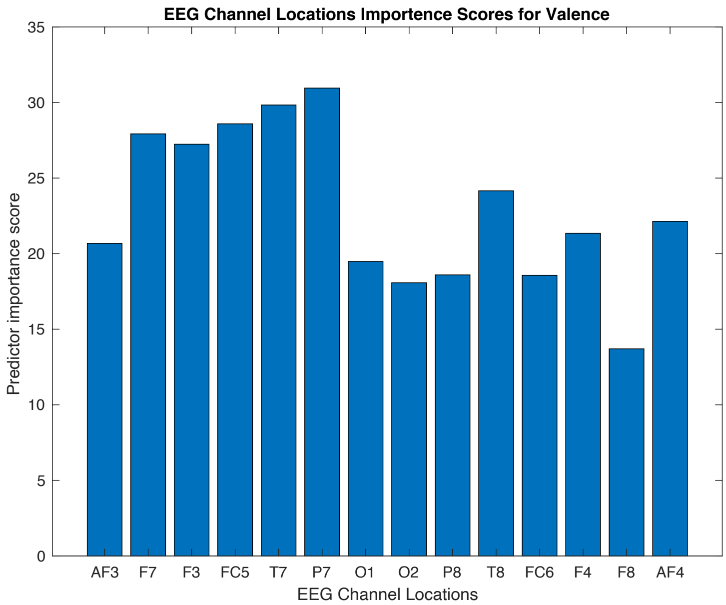

Figure 3.7: EEG Channel Locations' Importance Scores for Valence Estimation by using chi-square test

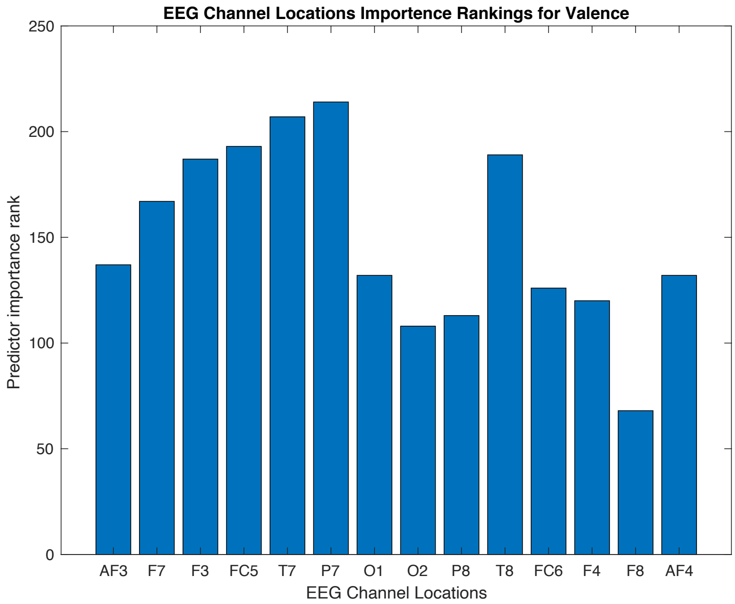

Figure 3.8: EEG Channel Locations' Importance Ranks for Valence Estimation by using chi-square test

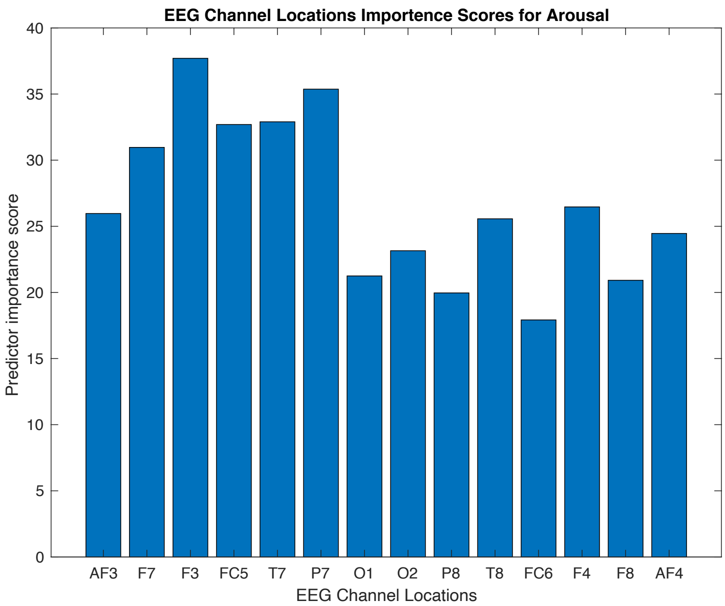

Figure 3.9: EEG Channel Locations' Importance Scores for Arousal Estimation by using chi-square test

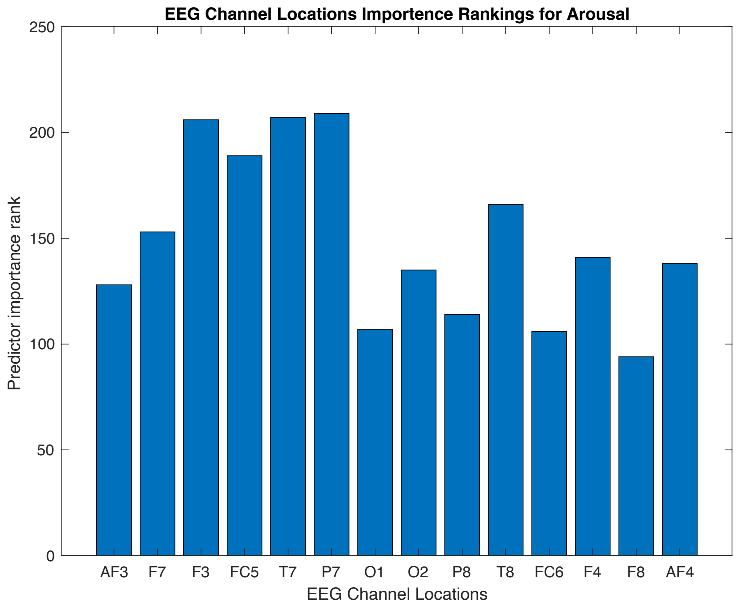

Figure 3.10: EEG Channel Locations' Importance Ranks for Arousal Estimation by using chi-square test

3.4 Evaluation of Different Classifier

In this section, the performance of different classifiers is evaluated. ECOC, KNN and random forest methods are concerned in this study. Classifiers are evaluated only using multi-modal features (ECG and EEG features together), individual modalities are investigated in a different chapter.

Table 3.1: Arousal Classification Results for Different Classifiers

	KNN			ECOC			Random Forest			
Class Name	Prec	Recall	F1	Prec	Recall	F1	Prec	Recall	F1	Sample Size
Low	0.63	0.69	0.66	0.64	0.67	0.66	0.93	0.95	0.94	1429
Neutral	0.63	0.55	0.59	0.61	0.58	0.59	0.95	0.95	0.95	1353
High	0.71	0.72	0.71	0.71	0.72	0.71	0.95	0.95	0.95	2163
macro avg	0.66	0.65	0.65	0.66	0.65	0.65	0.95	0.95	0.95	
weighted avg	0.66	0.67	0.66	0.66	0.66	0.66	0.95	0.95	0.95	
	Accuracy = 0,66			Accuracy = 0,67			Accuracy = 0,95			

Arousal classification performance for different classifiers are given in Table 3.1. The random forest method provides better precision, recall, and F1 score for all three classes. Moreover, It has the highest overall accuracy among the methods used.

Table 3.2: Valence Classification Results for Different Classifiers

	KNN			ECOC			Random Forest			
Class Name	Prec	Recall	F1	Prec	Recall	F1	Prec	Recall	F1	Sample Size
Low	0.65	0.69	0.67	0.66	0.71	0.68	0.93	0.93	0.93	1913
Neutral	0.60	0.50	0.55	0.63	0.51	0.57	0.91	0.89	0.90	1074
High	0.64	0.66	0.65	0.64	0.66	0.65	0.92	0.92	0.92	1958
macro avg	0.63	0.62	0.62	0.64	0.63	0.63	0.92	0.92	0.92	
weighted avg	0.64	0.64	0.64	0.65	0.64	0.64	0.92	0.92	0.92	
	Accuracy = 0,64			Accuracy = 0,65			Accuracy = 0,92			

Valence classification performance for different classifiers are given in Table 3.2. The random forest method provides better precision, recall, and F1 score for all three classes. Moreover, It has the highest overall accuracy among the methods used.

Since highest performing method is random forest, accuracy scores for each subject are also provided in Figure3.11 and Figure3.12 for random forest method. Results are provided as bar graph as each bar represents different participants.

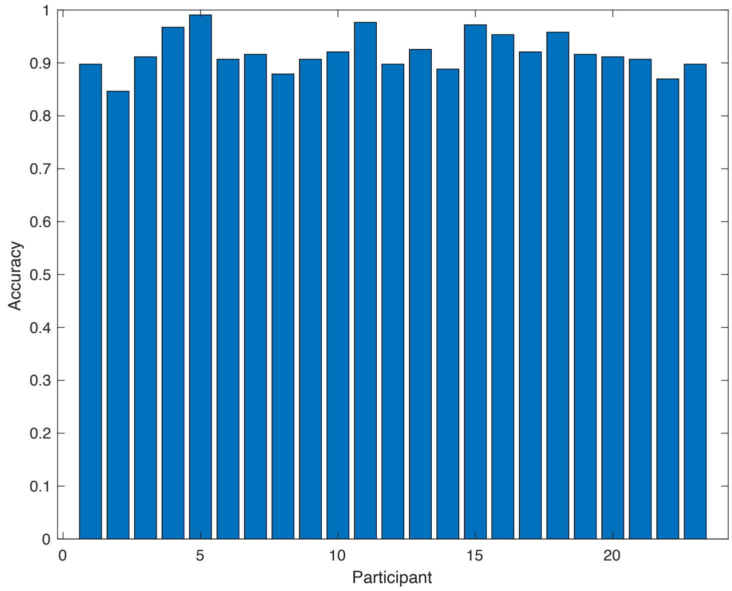

Figure 3.11: Accuracy of Valence Classification for Different Participants

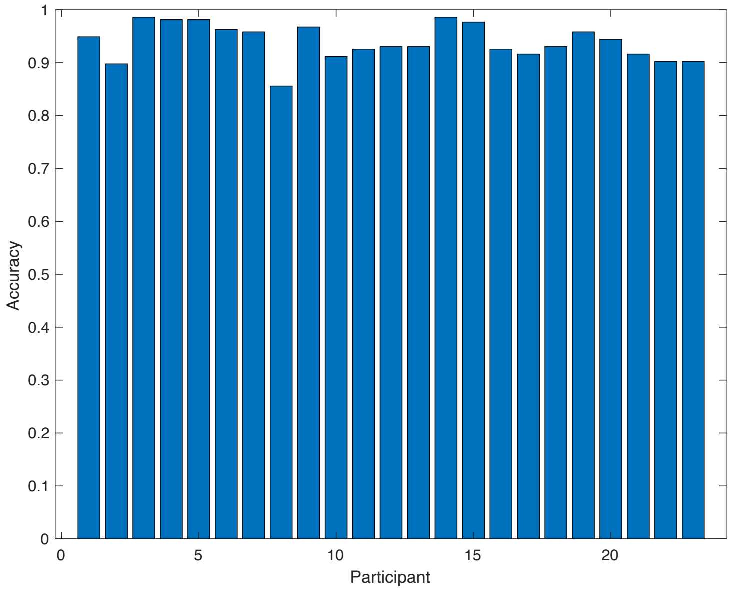

Figure 3.12: Accuracy of Arousal Classification for Different Participants

3.5 Evaluation of Modality Performances

In this section, different modalities used in both arousal and valence classification are compared. HRV and EEG features are used separately in individual modalitie cases. However, in multi-modal case, combination of HRV and EEG features are used for classification. Only random forest method is used for classification in this section since previous section shows that it outperforms other methods.

In EEG-only models, EVI and EEG_1-EEG_3 features are used in both arousal and valence classification. In HRV only models, RMSSD, NN50, MAD, LH features are used in both arousal and valence classification. Moreover, in the Multi-modal case, MAD, EEG_1 and EEG_2 features are used in the arousal and valence classification.

Figure 3.4 and Figure 3.3 show that combining EEG and HRV features outperforms individual modality features in both arousal and valence classification.

Table 3.3: Arousal Classification Results for Different Modalities

	EEG			HRV			EEG+HRV			
Class Name	Prec	Recall	F1	Prec	Recall	F1	Prec	Recall	F1	Sample Size
Low	0,69	0,72	0,70	0,87	0,90	0,89	0,91	0,94	0,92	1320
Neutral	0,68	0,64	0,66	0,93	0,88	0,90	0,95	0,93	0,94	1387
High	0,77	0,77	0,77	0,91	0,92	0,92	0,95	0,95	0,95	2238
macro avg	0,71	0,71	0,71	0,90	0,90	0,90	0,94	0,94	0,94	
weighted avg	0,72	0,72	0,72	0,90	0,90	0,90	0,94	0,94	0,94	
	Accuracy = 0,72			Accuracy = 0,90			Accuracy = 0,94			

Table 3.4: Valance Classification Results for Different Modalities

	EEG			HRV			EEG+HRV			
Class Name	Prec	Recall	F1	Prec	Recall	F1	Prec	Recall	F1	Sample Size
Low	0,73	0,77	0,75	0,90	0,92	0,91	0,93	0,94	0,94	1930
Neutral	0,68	0,62	0,65	0,89	0,84	0,86	0,89	0,90	0,90	1056
High	0,73	0,72	0,73	0,88	0,89	0,89	0,93	0,93	0,93	1959
macro avg	0,71	0,71	0,71	0,89	0,88	0,89	0,92	0,92	0,92	
weighted avg	0,72	0,72	0,72	0,89	0,89	0,89	0,92	0,92	0,92	
	Accuracy = 0,72			Accuracy = 0,89			Accuracy = 0,92			

3.6 Evaluation of Fixed Channel Location Performance

Up to this point, presented classification results are achieved by using optimum EEG scalp locations for each participant (EEG_1 and EEG_2). This chapter investigates the case which channel locations did not change between participants.

First 3 locations that are chosen as optimum points for 23 person according to chi-squared test, can be seen in detail from Table A.1-A.4. Popular choices for EEG_1 and EEG_2 locations for valence and arousal shown in Figure 3.13 and Figure 3.14. From most popular locations, T7 and F3 are chosen as fixed locations.

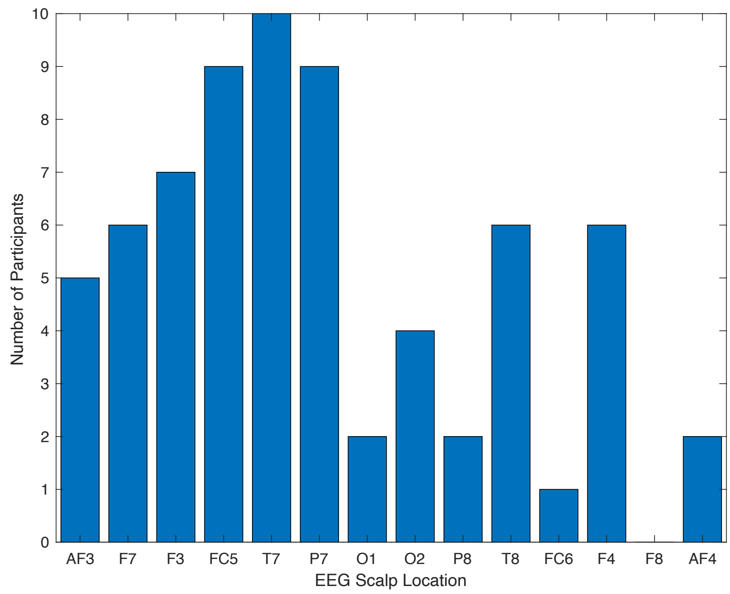

Figure 3.13: Number of Times EEG Channel Locations Chosen for Valence Estimation

In Table 3.5 and Table 3.6, classification performance results are presented for valence and arousal respectively. 91% accuracy, F1, precision and recall scores are achieved for valence classification. For arousal classification, 92% accuracy, F1, precision and recall scores are achieved.

Accuracy for each participant given in Figure3.15 and Figure3.16 for valence and arousal classification. Participants 2 and 8 show low accuracy scores in valence classification. Participant 8 shows relatively poor results compared to other partici-pants, also in arousal classification. Participants 4 and 5 shows overall best results.

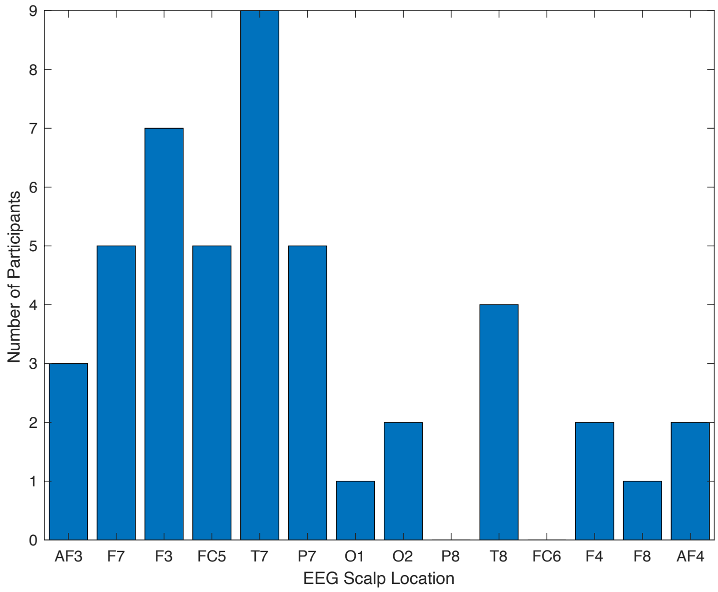

Figure 3.14: Number of Times EEG Channel Locations Chosen for Arousal Estimation

Table 3.5: Valance Classification Results for Fixed Channel Case

Class Name	Prec	Recall	F1	Sample Size
Low	0,94	0,94	0,94	1910
Neutral	0,89	0,90	0,89	1073
High	0,93	0,92	0,93	1962
macro avg	0,92	0,92	0,92	
weighted avg	0,93	0,93	0,93	
	Accuracy = 0,93			

Table 3.6: Arousal Classification Results for Fixed Channel Case

Class Name	Prec	Recall	F1	Sample Size
Low	0,90	0,92	0,91	1345
Neutral	0,92	0,93	0,93	1449
High	0,94	0,92	0,93	2151
macro avg	0,92	0,92	0,92	
weighted avg	0,92	0,92	0,92	
	Accuracy = 0.92			

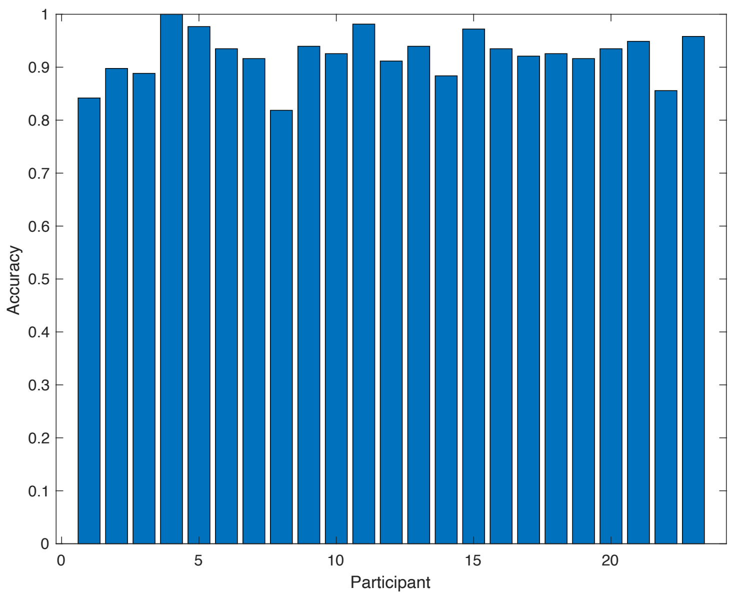

Figure 3.15: Accuracy of Valence Classification with Fixed Channel for Different Participants

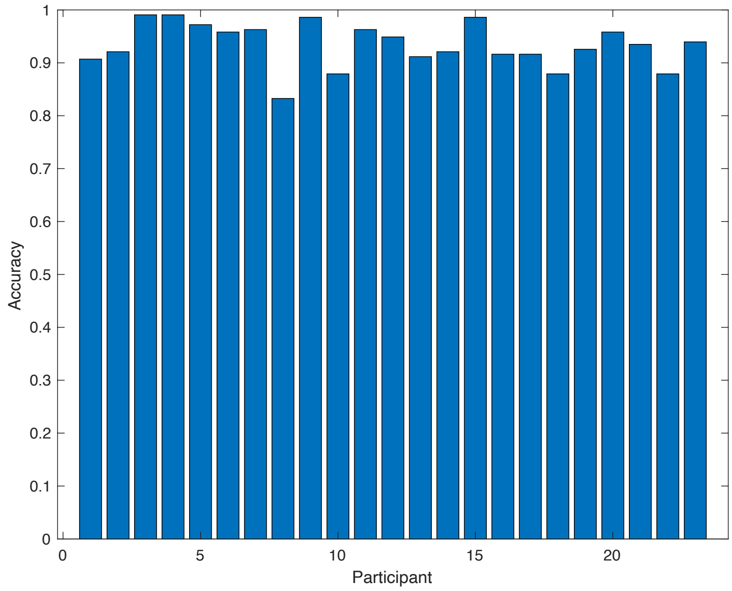

Figure 3.16: Accuracy of Arousal Classification with Fixed Channel for Different Participants

3.7 Comparison of Classification Results for Different Channel Pairs

In this section, classification results of the method that uses fixed locations T7 and F3 and other possible two channel combinations are compared. For every model, MAD feature is used with two zero-crossing variance features. 5x2 cross-validation with t-test applied to compare classification results achieved by using T7-F3 channel pair to other 90 possible combinations of channel pairs for 23 participants.

In 5x2 cross-validation test, for two trials data is randomly but evenly split into 5 subsets. Four of these subsets are used to train two models and one subset is used to get accuracy scores for two models. This process repeated for five subsets. Then, for second trial new random five subsets are acquired and the process is repeated for these five subsets. Finally, ten paired accuracy values are obtained for two models. After that, paired t-test is applied to compare the difference between accuracy values. This process is implemented via Matlab function "testckfold()". The function provides information about which model performs better, in addition to the significance of differences between accuracy values. Table A.10, Table A.11 and Table A.9 shows the locations which T7-F3 pair fails to shows significantly better results ($p>0.05$).

From Table A.9, for participants 3, 4, 7, 9, 10, 11, 14, 17 and 23 T7-F3 pair shows significantly better performance than any other location for arousal classification ($p<0.05$). For other participants, there are better or similar performing channel pairs. For some of the participants that T7-F3 shows significantly better results, even though, T7-F3 pair is not optimum location according to chi-squared test.

From Table A.10 and Table A.11, for participants 2,3, 5, 9, 11, 15,19 and 23 T7-F3 pair shows significantly better performance than any other location for valence classification ($p<0.05$). For some of the participants that T7-F3 shows significantly better results, even though, T7-F3 pair is not optimum location according to chi-squared test.

To sum up, Accuracy results show significant differences depending on the channel choice. For some participants, T7-F3 shows significantly better performance than the other possible locations, even though some of these locations are expected to perform better according to chi-squared test results.

3.8 Evaluation of Results

In this sections results presented in this work is discussed and compared to other studies.

When the results are examined in terms of EEG features, even though optimum locations depend on the participant, overall importance scores show some differences with existing works. In Zhang and Peng's work which investigates optimum EEG electrodes for emotion recognition, most of the locations presented in their work are similar to this work. However, P8, F8, and O2 are reported as important locations in their work. They are not in this work [11]. Moreover, P7 and T7 locations which are among the most important locations in this work, are not considered important in

Zhang's work [11]. According to Sarno's [39] work AF3, F7, FC5, T7, T8, and AF4 scalp locations have the highest importance in arousal. T8 and T7 locations have the highest importance in valence according to Pearson-correlation test results. Although T7 is an important location both in this and Sarno's study, P7 is also not considered in Sarno's work [39].

HRV features which are investigated by previous studies such as LH ratio, RMSSD, NN50 do not show promising results in this study [4] [19]. However, the mean of average difference in consecutive RR intervals(MAD) feature shows high importance score.

Classification accuracy achieved through the random forest are 92% for valence and 95% for arousal. The best classification result was achieved via using EEG and ECG metrics together. When the result is compared to similar works that use Dreamer data set from Table 3.7, the proposed work achieves comparable results with them.

Methods proposed in Table 3.7 rely on more than two EEG locations. This work uses ECG signals to reduce the number of EEG scalps. Neural networks are popular among recent studies for both feature extraction and classification. This work, on the other hand, relies on more conventional methods for classification.

Table 3.7: Results of Similar Works

References	Method	Accuracy	
		Valence	Arousal
Cui et al. [27]	Regional-Asymmetric Convolution Neural Network, Asymmetric Differential Layer	0.95	0.97
Cheng et al. [28]	spatial position relationship, deep forest	0.89	0.90
Liu Y. et al. [29]	Multi-level features guided capsule network	0.94	0.95
Wang Y. et al. [30]	SPD matrix network	0.67	0.76
Maheshwari et al. [31]	Multi-channel, rhythm selection, CNN	0.97	0.96
Song et al. [32]	Dynamic graph convolution neural network	0.86	0.85

CHAPTER 4

CONCLUSIONS

4.1 Discussions

In this work, an emotion recognition method that uses EEG and ECG signals is presented. This chapter revolves around discussions of the results.

The proposed work detects arousal and valence scores rather than the distinct emotion classes. One of the reasons for this, Dreamer data set does not provide actual emotions of subjects. It only provides arousal and valence scores provided by participants for each video. The second reason is, as mentioned before, dimensional representation of emotions reduces the error caused by participants while scoring emotions in self-assessment studies.

Presented pre-processing method and EEG features provide high accuracy together. Therefore, more complex pre-processing methods are not deployed. EEG signals carry many physiological artifacts and some of these are dealt in data collection state by asking participants to stand still as much as possible. However, in real life applications more complex pre-processing methods may be required. For this purpose ASR method might be useful.

Proposed zero-crossing variance feature aims to provide a measure of small frequency changes on EEG signal. In similar works, P7 does not show importance in general. However, it becomes an important location with the zero-crossing feature in this study.

In ECOC method, linear learners provide poor results compared to Gaussian and polynomial learners. KNN also performs similar to polynomial learners and random forest performs even better. Both KNN and random forest method can handle non linear distribution. These results indicate, features shows non linear relationships with classes.

In this work, intrasubject classification method is utilized due to personalized nature of HRV and EEG data. Moreover, to calculate HRV from ECG signal, relatively large time window is used, and also only the last 60 seconds of each video is used as suggested by Katsigiannis [33]. Therefore, sample size for each classifier is limited. This requires a classification method that can be used with limited data (not too small sample size) such as random forest. Neural network methods were avoided due to number of available data points don't exceed (even though, it is close due to low number of features used) Alwosheel's suggestions [40]. However, there are recent works that use CNN and achieve good accuracy. They solve 2 class problem for valence and arousal [27] [32]. Since, this work focuses on 3 class problem, required neural network would be more complex for this work. Therefore, such method still may not be applicable. Features proposed in this work may be tried in a system that uses neural network methods with a larger data set.

When the accuracy achieved for each subject is examined from Figure3.11, Figure3.12, Figure3.15 and Figure3.16, similar accuracy values are observed. Accuracy for partici-pant 8 is poor all cases. Accuracy values related to subject, show similar results with [29] and [28] works. This implies, there might be personal errors (especial for participant 8) in arousal and valence scores that cause inaccuracy independent from channel choice and classification method.

Accuracy achieved by optimum scalp locations are 92% for valence and 95% for arousal and by fixed locations 92% for valence and 93% for arousal. Accuracy values are similar for both case, because fixed locations are the actual best locations for almost half of the participants and they are at least good locations for remaining participants, since, there is few better location pairs that provides significantly better performance among 90 possible pairs(Table A.9 - Table A.11).Fixed location case might be advantageous, since it reduces total number of required EEG scalp to two. This can reduce cost and also may provide better comfort to user of a such system.

Even though, system solves 3 class classification problem, system manages to achieve high accuracy values. Moreover, it proposes an option that only requires two scalps. These are the advantages of this work over the similar ones. However, still this work does not provide definitive standard for emotion recognition, it just provides an alternative to existing works.

4.2 Future Work

This chapter tries to provide new research and investigation ideas related to proposed method.

Proposed zero-crossing variance feature helps system to recognize emotion of a person. Even though, it is not investigated in this work, this feature may show correlations with other brain functions. Therefore, further investigation may be useful and provide an alternative investigation method for EEG signals.

In similar works, P7 doesn't show importance in general. Significance of the location may be investigated with further studies. Even though, the importance is presented, reasons behind it is not enlightened in this work. Therefore, activity of channel might be further investigated.

In this work, the Dreamer data set is used. Data collection step in the Dreamer data set is performed in ideal conditions where participants stand still as much as possible. However, in real life cases, EEG signals carry many physiological artifacts due to muscle movement. A database prepared in such environment may help researchers to investigate relation between EEG signals and emotion in real life scenario.

This study and similar ones, does not provide definitive standards for emotion recognition. Therefore, further research may focus on identifying reasons behind the high accuracy of proposed systems and provides definitive standard.

4.3 Conclusion

In this work, an emotion recognition method that uses of EEG and ECG signals, is presented.

Simple band pass filtering was applied to EEG signals to reduce noise and also extract alpha and beta bands. After pre-processing, proposed zero-crossing variance feature was collected from EEG channels. For two different channel selection methods system is tested. These methods are fixed and optimum. In the optimum method, different locations are chosen for each individual by comparing their chi-squared test results. In the fixed case, most used locations from optimum case are used for all participants.

To acquire HRV features from ECG, firstly a mean filter applied to ECG signal. Then, R points are detected. Mean absolute difference of consecutive R-R time intervals are calculated and the MAD feature is extracted.

Performance for different modalities and classification methods are investigated. KNN, ECOC and random forest methods are compared and random forest achieves the best result. For modalitiy differences, this work shows that EEG and ECG features provide better results together.

To sum up, in this work an emotion recognition system presented which uses off the shelf, low spec devices. System performs on a par with similar works, while it solves 3 class classification problem. That might provide better insight about persons emotional state since, neutral state for both valence and arousal is provided by system. Moreover, this system can achieve high accuracy scores by only using two electrodes. Therefore, this work successfully achieves desired results.

REFERENCES

[1] T. Kirschstein and R. Kohling, "What is the source of the eeg ?," vol. 40, pp. 146–149, 2009.

[2] T. R. Mullen, C. A. E. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T. ping Jung, and G. Cauwenberghs, "Real-time neuroimaging and cognitive monitoring using wearable dry eeg," vol. 62, pp. 2553–2567, 2015.

[3] F. Bryn, "What is ecg and how does it work?," *https://imotions.com/blog/what-is-ecg/*, 1 2022.

[4] B. M. Appelhans and L. J. Luecken, "Heart rate variability as an index of regulated emotional responding," *Review of General Psychology*, vol. 10, pp. 229–240, 9 2006.

[5] J. A. Russel, "A circumplex model of affect," *Journal of Personality and Social Psychology 39(6):1161-1178*, vol. 39, pp. 1161–1178, 1980.

[6] EMOTIV, "Epoc x user manual," *https://emotiv.gitbook.io/epoc-x-user-manual/introduction/introduction-to-epoc-x/coverage*, 1 2022.

[7] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, "Human emotion recognition: Review of sensors and methods," *Sensors*, vol. 20, p. 592, 1 2020.

[8] J. Wang and M. Wang, "Review of the emotional feature extraction and classification using eeg signals," *Cognitive Robotics*, vol. 1, pp. 29–40, 2021.

[9] E. P. Torres, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "Eeg-based bci emotion recognition: A survey," *Sensors*, vol. 20, p. 5083, 9 2020.

[10] N. S. Suhaimi, J. Mountstephens, and J. Teo, "Eeg-based emotion recognition: A state-of-the-art review of current trends and opportunities," *Computational Intelligence and Neuroscience*, vol. 2020, 2020.

[11] J. Zhang and P. Chen, "Selection of optimal eeg electrodes for human emotion recognition," vol. 53, pp. 10229–10235, Elsevier B.V., 2020.

[12] A. Tandle, N. Jog, A. Dharmadhikari, and S. Jaiswal, "Estimation of valence of emotion from musically stimulated eeg using frontal theta asymmetry," pp. 63–68, Institute of Electrical and Electronics Engineers Inc., 10 2016.

[13] A. J. Tomarken, R. J. Davidson, and J. B. Henriques, "Resting frontal brain asymmetry predicts affective responses to films," *Psychological Association, Inc*, vol. 59, pp. 791–801, 1990.

[14] S. Wu, X. Xu, L. Shu, and B. Hu, "Estimation of valence of emotion using twofrontal eeg channels," 2017.

[15] C. yuan Chang, S. Member, S. hsiou Hsu, and S. Member, "Evaluation of artifact subspace reconstruction for automatic eeg artifact removal," pp. 1242–1245, 2018.

[16] D. S. Quintana, A. J. Guastella, T. Outhred, I. B. Hickie, and A. H. Kemp, "Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition," *International Journal of Psychophysiology*, vol. 86, pp. 168–172, 11 2012.

[17] S. W. Porges, "The polyvagal theory: phylogenetic substrates of a social nervous system," *International Journal of Psychophysiology*, vol. 42, p. 123146, 2001.

[18] J. F. Thayer and R. D. Lane, "A model of neurovisceral integration in emotion regulation and dysregulation," *Journal of Affective Disorders*, vol. 61, pp. 201–216, 2000.

[19] H. G. Kim, E. J. Cheon, D. S. Bai, Y. H. Lee, and B. H. Koo, "Stress and heart rate variability: A meta-analysis and review of the literature," *Psychiatry Investigation*, vol. 15, pp. 235–245, 3 2018.

[20] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, pp. 169–200, 5 1992.

[21] G. Mahalakshmi and P. SREEJA, "Emotion models: A review," *International Journal of Control Theory and Applications*, vol. 10, pp. 651–657, 2017.

[22] H. Schlosberg, "The description of facial expressions terms of two dimensions," *Journal of Experimental Psychology*, vol. 44, 1952.

[23] A. Al-Nafjan, M. Hosny, Y. Al-Ohali, and A. Al-Wabil, "Review and classification of emotion recognition based on eeg brain-computer interface system research: A systematic review," *Applied Sciences (Switzerland)*, vol. 7, 12 2017.

[24] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.

[25] P. Probst, M. Wright, and A.-L. Boulesteix, "Hyperparameters and tuning strategies for random forest," 4 2018.

[26] M. K. Kim, M. Kim, E. Oh, and S. P. Kim, "A review on the computational methods for emotional state estimation from the human eeg," *Computational and Mathematical Methods in Medicine*, vol. 2013, 2013.

[27] H. Cui, A. Liu, X. Zhang, X. Chen, K. Wang, and X. Chen, "Eeg-based emotion recognition using an end-to-end regional-asymmetric convolutional neural network," *Knowledge-Based Systems*, vol. 205, 10 2020.

[28] J. Cheng, M. Chen, C. Li, Y. Liu, R. Song, A. Liu, and X. Chen, "Emotion recognition from multi-channel eeg via deep forest," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, pp. 453–464, 2 2021.

[29] Y. Liu, Y. Ding, C. Li, J. Cheng, R. Song, F. Wan, and X. Chen, "Multi-channel eeg-based emotion recognition via a multi-level features guided capsule network," *Computers in Biology and Medicine*, vol. 123, 8 2020.

[30] Y. Wang, S. Qiu, X. Ma, and H. He, "A prototype-based spd matrix network for domain adaptation eeg emotion recognition," *Pattern Recognition*, vol. 110, 2 2021.

[31] D. Maheshwari, S. K. Ghosh, R. K. Tripathy, M. Sharma, and U. R. Acharya, "Automated accurate emotion recognition system using rhythm-specific deep convolutional neural network technique with multi-channel eeg signals," *Computers in Biology and Medicine*, vol. 134, 7 2021.

[32] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, pp. 532–541, 7 2020.

[33] S. Katsigiannis and N. 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*, vol. 22, pp. 98–107, 1 2018.

[34] C. A. Gabert-Quillen, E. E. Bartolini, B. T. Abravanel, and C. A. Sanislow, "Ratings for emotion film clips," *Behavior Research Methods*, vol. 47, pp. 773–787, 9 2015.

[35] A. Burns, B. R. Greene, M. J. McGrath, T. J. O'Shea, B. Kuris, S. M. Ayer, F. Stroiescu, and V. Cionca, "Shimmer™ – a wireless sensor platform for noninvasive biomedical research," *IEEE Sensors Journal*, vol. 10, pp. 1527–1534, 9 2010.

[36] N. A. Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and G. McArthur, "Validation of the emotiv epoc eeg gaming system for measuring research quality auditory erps," *PeerJ*, vol. 1, p. e38, 2 2013.

[37] X. Jiang, G. B. Bian, and Z. Tian, "Removal of artifacts from eeg signals: A review," *Sensors (Switzerland)*, vol. 19, 3 2019.

[38] D. Albert, B. R. Satchwell, and K. N. Barnett, "Wireless,ultrasonicpersonal health monitoring system us 8,301.232 b2," 10 2012.

[39] R. Sarno, M. N. Munawar, and B. T. Nugraha, "Real-time electroencephalography-based emotion recognition system," *International Review on Computers and Software*, vol. 11, pp. 456–465, 5 2016.

[40] A. Alwosheel, S. van Cranenburgh, and C. G. Chorus, "Is your dataset big enough? sample size requirements when using artificial neural networks for discrete choice analysis," *Journal of Choice Modelling*, vol. 28, pp. 167–182, 9 2018.

APPENDIX A

DETAILED RESULTS FOR EACH PARTICIPANT

A.1 Best Channel Locations For Participants

EEG channel locations that provides zero-crossing variance feature with the highest chi-squared test score for each participant given in tables in this section.

Table A.1: EEG_1-EEG_3 Locations for Valence (Participant 1-16)

Participant Number	EEG_1	EEG_2	EEG_3
1	P7	P8	F3
2	T7	F4	P7
3	T7	T8	P7
4	P8	P7	T7
5	P7	O1	O2
6	P7	T7	FC5
7	T7	FC6	T8
8	T8	T7	FC5
9	T8	P7	T7
10	F4	AF3	O2
11	T7	T8	F7
12	F4	F7	F3
13	F7	AF3	F4
14	FC5	AF3	F7
15	AF4	FC5	F4
16	AF4	FC5	F3

Table A.2: EEG_1-EEG_3 Locations for Valence (Participant 17-23)

Participant Number	EEG_1	EEG_2	EEG_3
17	O2	F3	FC5
18	F7	P7	AF3
19	FC5	F4	F3
20	O1	O2	T7
21	T7	FC5	F3
22	F7	FC5	F3
23	AF3	P7	T8

Table A.3: EEG_1-EEG_3 Locations for Arousal (Participant 1-16)

Participant Number	EEG_1	EEG_2	EEG_3
1	P7	T8	F8
2	P7	T7	FC6
3	F3	T7	O2
4	F3	F8	O2
5	O1	O2	FC5
6	T7	FC5	T8
7	T7	T8	FC5
8	O2	T7	T8
9	T8	FC5	P7
10	F3	F4	AF4
11	T7	AF4	FC6
12	F4	F3	P7
13	F7	T7	F3
14	F7	AF3	F3
15	F3	F7	P7
16	AF4	FC5	F7

Table A.4: EEG_1-EEG_3 Locations for Arousal (Participant 17-23)

Participant Number	EEG_1	EEG_2	EEG_3
17	FC5	T7	AF3
18	P7	F3	F7
19	AF3	P7	F3
20	F3	T7	T8
21	P7	F7	F4
22	F7	FC5	F3
23	T8	AF3	F3

A.2 Random Forest Accuracy Results For Participants

Tables in this section covers accuracy of classification results in 5 trails for each participant. Tables for valence, arousal classification results for optimum and fixed channel case are provided.

Table A.5: Valence Classification Accuracy Results for Random Forest(Optimum Location Case)

Participant Number	trail-1	trail-2	trail-3	trail-4	trail-5
1	0.95	0.95	0.93	0.81	0.84
2	0.84	0.79	0.98	0.93	0.91
3	0.91	0.91	0.91	0.88	0.88
4	1	1	1	0.98	0.95
5	0.98	1	0.95	0.98	1
6	0.91	0.93	0.93	0.93	0.91
7	0.91	0.91	0.93	0.91	0.88
8	0.84	0.84	0.91	0.91	0.88
9	0.91	0.95	0.93	0.86	0.95
10	0.93	0.91	1	0.88	0.88
11	1	0.98	0.95	0.98	1
12	0.86	0.93	0.86	0.88	0.95
13	0.86	0.95	0.91	0.91	0.93
14	0.91	0.86	0.91	0.95	0.88
15	0.98	0.98	0.98	1	0.98
16	0.98	0.98	0.95	0.93	0.95
17	0.95	0.91	0.86	0.95	0.84
18	0.95	0.95	0.91	0.93	0.98
19	0.93	0.93	0.88	0.93	0.81
20	0.86	0.95	0.95	0.93	0.91
21	0.95	0.93	0.95	0.98	0.95
22	0.84	0.84	0.86	0.98	0.88
23	0.91	0.88	0.86	0.77	0.91

Table A.6: Arousal Classification Accuracy Results for Random Forest(Optimum Location Case)

Participant Number	trail-1	trail-2	trail-3	trail-4	trail-5
1	0.91	1	0.93	0.98	0.93
2	0.88	0.95	0.95	0.95	0.98
3	0.98	1	0.98	0.98	1
4	1	0.98	1	1	1
5	0.98	0.93	0.98	0.98	0.98
6	0.95	0.93	0.98	0.98	1
7	0.98	1	0.98	0.95	0.95
8	0.77	0.86	0.79	0.77	0.84
9	0.98	1	0.98	1	0.98
10	0.95	0.86	0.98	0.98	0.95
11	0.98	0.95	0.88	1	0.98
12	0.98	0.98	0.98	0.98	0.93
13	1	0.98	0.98	0.98	0.86
14	0.91	1	1	0.95	0.93
15	1	0.98	1	0.98	1
16	0.98	0.93	0.86	0.98	0.93
17	0.95	0.91	0.95	0.93	0.93
18	0.95	0.98	0.98	0.91	0.95
19	0.95	0.98	0.91	0.95	0.98
20	0.98	0.93	0.93	0.95	0.98
21	0.98	0.86	0.98	0.95	1
22	0.91	0.88	0.77	0.91	0.86
23	0.93	0.86	0.88	0.93	0.95

Table A.7: Valence Classification Accuracy Results for Random Forest(Fixed Location Case)

Participant Number	trail-1	trail-2	trail-3	trail-4	trail-5
1	0.91	0.98	0.98	0.88	0.86
2	0.86	0.91	0.74	0.86	0.79
3	0.93	0.88	0.93	0.95	0.88
4	1	1	1	1	1
5	1	0.98	0.95	1	0.95
6	0.91	0.93	0.98	0.88	0.93
7	0.98	0.79	0.81	0.78	0.78
8	0.77	0.84	0.84	0.72	0.84
9	0.95	0.93	0.95	0.95	0.95
10	0.91	0.93	0.93	0.86	0.91
11	0.98	1	1	0.98	1
12	0.91	0.88	0.93	0.86	0.93
13	0.91	0.95	0.91	0.93	0.93
14	0.79	0.88	0.86	0.77	0.84
15	0.98	1	0.98	1	0.98
16	0.93	0.88	0.91	0.93	0.88
17	0.88	0.84	0.95	0.86	0.84
18	0.98	0.98	0.93	0.95	0.93
19	0.91	0.81	0.98	0.91	0.93
20	1	0.95	0.91	0.93	0.98
21	0.95	0.79	0.95	0.95	1
22	0.81	0.86	0.86	0.86	0.86
23	0.86	0.95	0.95	0.91	0.93

Table A.8: Arousal Classification Accuracy Results for Random Forest(Fixed Location Case)

Participant Number	trail-1	trail-2	trail-3	trail-4	trail-5
1	0.91	0.86	0.95	0.91	0.88
2	0.86	0.88	0.88	0.91	0.93
3	1	1	1	1	1
4	1	0.98	0.95	1	1
5	1	1	1	1	0.98
6	0.98	0.98	1	1	0.93
7	0.86	0.95	0.88	0.98	0.98
8	0.81	0.88	0.84	0.81	0.79
9	0.95	0.98	0.98	0.98	0.98
10	0.88	0.91	0.93	0.86	0.88
11	1	1	0.98	0.98	1
12	1	0.91	0.91	0.86	0.93
11	0.95	0.72	0.93	0.98	1
14	0.91	0.95	0.86	0.91	0.98
15	0.98	0.95	0.95	1	1
16	0.93	0.81	0.95	0.91	0.86
17	0.88	0.95	0.91	0.84	0.91
18	0.91	0.88	0.84	0.95	0.81
19	0.88	0.93	0.86	0.86	0.91
20	0.93	0.98	0.93	0.95	0.91
21	0.95	0.98	0.95	0.93	0.91
22	0.91	0.86	0.86	0.86	0.79
23	0.84	0.88	0.84	0.91	0.91

A.3 Cross Validation of Classification Results Between T7-F3 and Other Channels

Channel and participants pairs that does not provide significantly better accuracy than T7-F3 pair are provided with relevant p value.

Table A.9: Cross Validation of Classification Results Between T7-F3 and Other Channels for Arousal Classification

Participant Number	Channels	P-Values	Participant Number	Channels	P-Values
1	'P7-O1'	0.094	13	'O1-FC6'	0.056
2	'T7-P7'	0.109	13	'P8-F4'	0.099
2	'T7-O1'	0.051	15	'P7-FC6'	0.072
2	'P7-AF4'	0.073	16	'T7-F4'	0.052
5	'F7-FC5'	0.076	18	'AF3-F3'	0.132
6	'F7-AF4'	0.054	18	'AF3-P7'	0.062
8	'F7-O1'	0.084	18	'F7-F4'	0.052
8	'O2-P8'	0.088	18	'FC5-FC6'	0.100
12	'F7-FC5'	0.119	19	'AF3-O2'	0.076
12	'FC5-AF4'	0.073	19	'AF3-F4'	0.077
13	'AF3-O2'	0.074	19	'FC5-T7'	0.206
13	'AF3-P8'	0.079	19	'P7-O2'	0.078
13	'AF3-F4'	0.060	19	'P7-F8'	0.070
13	'F7-F3'	0.226	20	'AF3-P8'	0.068
13	'F7-FC5'	0.052	20	'F7-P7'	0.104
13	'F7-O2'	0.056	21	'P7-O2'	0.158
13	'F3-F4'	0.052	22	'FC5-T7'	0.058
13	'FC5-T8'	0.057			
13	'FC5-F4'	0.057			
13	'P7-F8'	0.091			

Table A.10: Cross Validation of Classification Results Between T7-F3 and Other Channels for Valence Classification (Part-1)

Participant Number	Channels	P-Values	Participant Number	Channels	P-Values
1	'AF3-P7'	0.085	8	'T7-F8'	0.055
1	'F3-P7'	0.255	8	'O2-T8'	0.373
1	'P7-AF4'	0.081	10	'F7-FC6'	0.170
1	'T8-F4'	0.057	10	'FC5-F4'	0.054
4	'T7-P7'	0.203	10	'P8-F4'	0.177
4	'P7-F4'	0.063	10	'T8-F4'	0.142
4	'O1-O2'	0.067	12	'F7-F4'	0.117
6	'FC5-P7'	0.122	13	'AF3-F3'	0.148
6	'P7-T8'	0.096	13	'AF3-FC5'	0.070
7	'FC5-O1'	0.160	13	'AF3-FC6'	0.063
7	'FC5-F8'	0.050	13	'F7-O1'	0.067
7	'FC5-AF4'	0.084	13	'F7-F4'	0.178
7	'T7-T8'	0.060	13	'F7-AF4'	0.055
7	'P7-FC6'	0.114	13	'P7-FC6'	0.051
7	'P7-AF4'	0.080	13	'O1-F8'	0.063
7	'O1-F4'	0.082	14	'AF3-P7'	0.086
7	'O2-FC6'	0.354	14	'AF3-FC6'	0.204
8	'F7-T8'	0.118	14	'F7-F3'	0.072
8	'T7-F4'	0.053	14	'F7-T7'	0.067

Table A.11: Cross Validation of Classification Results Between T7-F3 and Other Channels for Valence Classification (Part-2)

Participant Number	Channels	P-Values	Participant Number	Channels	P-Values
14	'F3-P7'	0.051	16	'P7-P8'	0.108
14	'F3-F8'	0.089	16	'O1-P8'	0.131
14	'FC5-P7'	0.173	16	'P8-T8'	0.071
14	'T7-P7'	0.076	16	'FC6-F4'	0.085
14	'P7-O1'	0.090	16	'FC6-AF4'	0.061
14	'P7-O2'	0.134	17	'F7-O2'	0.057
14	'P7-T8'	0.051	17	'F3-O2'	0.118
14	'O2-F4'	0.053	18	'F7-FC5'	0.052
16	'AF3-F3'	0.097	18	'F7-P7'	0.135
16	'AF3-AF4'	0.384	20	'AF3-P8'	0.061
16	'F7-P8'	0.074	20	'FC5-O2'	0.088
16	'F7-AF4'	0.078	21	'AF3-FC6'	0.165
16	'F3-FC5'	0.155	21	'O2-FC6'	0.061
16	'F3-P8'	0.178	22	'AF3-FC5'	0.052
16	'FC5-P7'	0.300	22	'F7-T7'	0.054
16	'FC5-P8'	0.175	22	'FC5-O2'	0.120
16	'FC5-T8'	0.208	22	'P7-T8'	0.057
16	'FC5-AF4'	0.678	22	'P7-F8'	0.066
16	'T7-FC6'	0.080			