ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

REAL-TIME EMOTION RECOGNITION FROM EEG SIGNALS USING ONE ELECTRODE DEVICE

M.Sc. THESIS

Mehmet Ali SARIKAYA

Department of Computer Engineering

Computer Engineering Programme

MAY 2017

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

REAL-TIME EMOTION RECOGNITION FROM EEG SIGNALS USING ONE ELECTRODE DEVICE

M.Sc. THESIS

Mehmet Ali SARIKAYA
(504151517)

Department of Computer Engineering

Computer Engineering Programme

Thesis Advisor: Asst. Prof. Dr. Gökhan İNCE

MAY 2017

**TEK ELEKTROTLU CİHAZ İLE
EEG SİNYALLERİNDEN GERÇEK ZAMANLI DUYGU TANIMA**

YÜKSEK LİSANS TEZİ

**Mehmet Ali SARIKAYA
(504151517)**

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Asst. Prof. Dr. Gökhan İNCE

MAYIS 2017

Mehmet Ali SARIKAYA, a M.Sc. student of ITU Graduate School of Science Engineering and Technology 504151517 successfully defended the thesis entitled "REAL-TIME EMOTION RECOGNITION FROM EEG SIGNALS USING ONE ELECTRODE DEVICE", which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Asst. Prof. Dr. Gökhan İNCE** ....
Istanbul Technical University

Jury Members : **Assoc. Prof. Dr. Mustafa Ersel KAMAŞAK** ....
Istanbul Technical University

Asst. Prof. Dr. Mehmet Amaç GÜVENSAN ....
Yildiz Technical University

....

....

Date of Submission : **05 May 2017**
Date of Defense : **09 June 2017**

To the heroic Turkish Army,

FOREWORD

First, I would like to thank my supervisor Asst. Prof. Dr. Gökhan İnce for his assistance and support. He provided excellent facilities and advices during this project. I would like to thank all the scientists who work to make the country viable despite all the negativities. After all, I would like to thank all my friends and, of course, my family.

x

TABLE OF CONTENTS

ABBREVIATIONS

ANN	**:**	Artificial Neural Network
App	**:**	Appendix
BCI	**:**	Brain-Computer Interfaces
DT	**:**	Decision Tree
EEG	**:**	ElectroEncephaloGraph
HCI	**:**	Human-Computer Interfaces
ITU	**:**	Istanbul Technical University
MSE	**:**	Mean Squared Error
NA	**:**	Not Available
RF	**:**	Random Forest
SVM	**:**	Support Vector Machine

SYMBOLS

P	**:**	Probability
T	**:**	Target
X	**:**	Attributes
c	**:**	Class

LIST OF TABLES

LIST OF FIGURES

REAL-TIME EMOTION RECOGNITION FROM EEG SIGNALS USING ONE ELECTRODE DEVICE

SUMMARY

Since the day the computer was first invented, there has been an interaction between people and computers. This interaction can be performed through software or hardware. The interfaces that provide interaction between people and computers are defined as Brain-Computer Interfaces (BCI).

Electroencephalography (EEG) signals are preferred because of their practical use and portable structure in BCI applications. EEG technology measures the electrical activity of the brain. Many applications developed in recent years benefit from this activity.

In recent years, researchers have concentrated on the development of ElectroEncephaloGraphy (EEG) based Brain-Computer Interfaces (BCI) to increase the quality of life using medical applications. In this context, BCI applications were introduced that could meet the basic needs of people with physical disabilities. In addition, the impact of the content produced in the marketing field on the users was tried to be determined with BCI applications. In recent times it has become an indispensable product for meditation practices and neurotherapy. Many games and entertainment applications were further enriched through the BCI.

BCIs are used for marketing, gaming, and entertainment to provide users with a more personalized experience. Both medical and non-medical applications require the ability to interpret the user's multimedia-induced perception and emotional experience. The changes that both internal and environmental factors create in a person-specific manner are called emotions. Estimating the emotional state of a person is an important step to build effective BCI applications. Therefore, studies are needed to identify the correct emotional state. In the literature, there are few studies on real-time emotion recognition with EEG signals. With this study, it is aimed to close these deficiencies.

It is also possible to foresee that neurotechnology will be the technology of the future. The day when people can upload their thoughts to computers is an unquestionable fact that these studies will prevail with applications that can increase the cognitive capacity of people.

With the developments in EEG technology in recent years, EEG devices can be obtained with more reasonable costs. In addition, more complex advances are being made every day by using less channel EEG devices. The number of studies using a single-electrode device for emotion recognition is low, because single-channel EEG devices have recently been marketed. Single electrode devices are economical and easy to use. It is predicted that it will be included in the literature more widely in the coming years.

In this thesis, a system is designed which detects instantaneous emotional state by using EEG data obtained in real-time with a single-channel commercial BCI device.

While designing this system, it comes from the top of the following challenging problems: The collection of EEG signals, the selection of the most appropriate features, and the creation of the most appropriate detection model using these features.

A software was written to collect the EEG signals. This software can report data coming in real time. After this software has been written, it has been tested whether the software works correctly by looking at the reported data.

For the selection of the right features, firstly visual and auditory content was created to warn emotions correctly. For this reason, visual and auditory items that are both short in duration and as effective as possible have been chosen. The sufficiency of the content found was determined as a result of several pre-evaluation tests. After the content selection was completed, decision trees and random forest algorithms, which are widely used in the literature, were used to correlate EEG data with the emotional state of the content. The order of importance of the features is determined using the random forest algorithm. It has been decided which features should be in the final model by taking the order of importance of the features.

In order to create the most appropriate model, the most suitable parameters were determined by using the deep learning algorithm and the grid search method for different parameters. For the suitability of the specified model, the data were divided into groups as training, validation and test data and the success rate was measured. It is also shown that the response time is fast enough to work in real time. The generated model is tested in real time and the response time is measured. It has been tested to produce accurate results under the required time constraints. The success of the system created on this issue has been revealed.

Both the real-time and non-real-time classification success of the generated model is measured. In both cases, a success rate was achieved that would meet expectations. In addition, the system has been tested for both supervised and unsupervised learning using an incremental learning approach. All the parameters affecting the success of the system are discussed. The created model and method will be a pioneer for such studies. With the experimentation and the increase of collected data, even the ordinary events in human life will reveal the relation with the brain signals.

To sum up, first, while the users watched a video consisting of film fragments and binaural beats that would trigger different emotional states, the EEG data from the users were collected. Then, pre-processing and feature selection were performed on this EEG data and an emotion recognition model with an accuracy of more than 87% was generated using deep learning neural networks. Using this model, a system which detects the instant emotional state on the queuing port structure of a certain length buffer using real time EEG data is designed.

As a result, we present a real-time emotion recognition model using EEG signals. This thesis tackled to figure out the specific link between brain waves and the multimedia-induced emotions. Using an ANN-based classification scheme, the proposed system classifies the multimedia genres such as funny, horror, weepy with an average accuracy of 87% which cause emotional or psychological experiences that are induced in viewers. The system can be used for both medical and non-medical applications. In the future, we plan to develop a real-time game based on the fears of the players which will impact all events in the game. The game will come with a horror story which sets players up to solve a mystery while overcoming their fears.

In the future, with the increase in work in this area, a society is hoped that people will not be sick and will be happy. But it is also necessary to work to ensure that these studies are not used outside ethical and moral boundaries. There is also the possibility that the ideological apparatus of the state may turn into a dystopic cluster by using such studies. For this reason, it is aimed to develop protocols in order to keep such applications in ethical limits in the future.

TEK ELEKTROTLU CİHAZ İLE
EEG SİNYALLERİNDEN GERÇEK ZAMANLI DUYGU TANIMA

ÖZET

Bilgisayarın ilk icat edildiği günden beri insanlar ile bilgisayarlar arasında bir etkileşim söz konusudur. Bu etkileşim yazılım veya donanımlar aracılığıyla icra edilebilmektedir. İnsanlar ile bilgisayarlar arasındaki etkileşimi sağlayan arayüzler Beyin-Bilgisayar Arayüzleri (BBA) olarak tanımlanmaktadır.

BBA uygulamalarında hem pratik kullanımı hem de zamanla daha portatif bir yapıya dönüşmesi nedeniyle ElektroEnsefaloGrafi (EEG) sinyalleri tercih edilmektedir. EEG teknolojisi ile beynin elektriksel aktivitesi ölçülmektedir. Son yıllarda geliştirilen birçok uygulama bu aktiviteden yararlanmaktadır.

Son yıllarda araştırmacılar, tıbbi uygulamalardaki yaşam kalitesini artırmak için ElektroEnsefaloGrafi (EEG) tabanlı Beyin-Bilgisayar Arayüzleri (BBA) geliştirmeye odaklandılar. Bu bağlamda fiziksel engelli insanların BBA uygulamaları aracılığıyla temel ihtiyaçlarını karşılayabilecekleri ürünler ortaya koydular. Ayrıca pazarlama alanında üretilen içeriklerin kullanıcılar üzerindeki etkisi yine BBA uygulamaları ile tespit edilmeye çalışıldı. Son zamanlarda ise meditasyon uygulamaları ve nöroterapi için bir vazgeçilmez ürün haline geldi. Birçok oyun ve eğlence uygulaması BBA aracılığıyla daha da zenginleştirildi.

Hem tıbbi hem de pazarlama, oyun oynama ve eğlence gibi tıbbi olmayan uygulamalar, kullanıcının çoklu ortam kaynaklı algı ve duygusal deneyimini yorumlama yeteneğini gerektirir. Bir bireye ait hem içsel hem de çevresel faktörlerin kişiye özgü olarak yarattığı değişimler duygu olarak adlandırılır. BBA uygulamalarının kişiye özgü ve etkin hale getirilebilmesi için duygusal durumun tespiti önemli bir adımdır. Bu nedenle duygusal durumun doğru tespit edilmesi için çalışmalar yapılmasına ihtiyaç duyulmaktadır. Literatürde EEG sinyalleri ile gerçek zamanlı duygu tanıma alanında yapılmış çalışma yok denilecek kadar azdır. Bu çalışma ile bu eksikliğin kapatılması hedeflenmektedir.

Ayrıca nöroteknolojinin geleceğin teknolojisi olacağını öngörmek mümkün. İnsanların düşüncelerini bilgisayarlara yükleyebileceği günlerin hayali kurulan günümüzde bu çalışmaların insanların bilişsel kapasitesini arttırılabilecek uygulamalara ön ayaklık edeceği su götürmez bir gerçektir.

Son yıllarda EEG teknolojisindeki gelişmeler ile birlikte EEG cihazları daha uygun maliyetler ile elde edilebilmektedir. Ayrıca her geçen gün daha az kanallı EEG cihazları kullanılarak daha karmaşık iler yapılmaya çalışılmaktadır. Duygu tanıma için tek elektrotlu bir cihaz kullanan çalışma sayıca azdır, çünkü tek kanallı EEG cihazları yakın zamanda piyasaya sürülmüştür. Tek elektrotlu cihazlar ekonomiktir ve kullanımı kolaydır. Önümüzdeki yıllarda daha yaygın bir şekilde literatürde yer edineceği öngörülmektedir.

Bu tezde, gerçek zamanlı olarak kullanıcıdan tek kanallı ticari bir BBA cihazıyla elde edilen EEG verileri kullanılarak kullanıcıya ait anlık duygu durumu tespiti yapan bir sistem tasarlanmıştır. Bu sistem tasarlanırken şu zorlu problemlerin üstesinden gelinmiştir: EEG sinyallerinin toplanmasını, en uygun özniteliklerin seçilmesini ve bu öznitelikleri kullanan en uygun tespit modelinin oluşturulması.

EEG sinyallerini toplamak için bir yazılım yazılmıştır. Bu yazılım, gerçek zamanlı olarak gelen verileri raporlayabilmektedir. Bu yazılım yazıldıktan sonra raporlanan verilere bakarak yazılımın doğru çalışıp çalışmadığı test edilmiştir.

Doğru özniteliklerin seçimi için öncelikle duyguları doğru şekilde uyaracak görsel ve işitsel içerik oluşturulmuştur. Bu nedenle hem süre olarak kısa olacak hem de olabildiğince etkili olacak görsel ve işitsel ögeler seçilmiştir. Bulunan içeriklerin yeterliliği birkaç ön değerlendirme testi sonucunda belirlenmiştir. İçerik seçimi tamamlandıktan sonra özniteliklerin seçimi için EEG verilerinin içerikteki duygusal ögelerle korelasyonunu bulmak amacıyla literatürde yaygın kullanılan karar ağaçları ve rasgele orman algoritmaları kullanılmıştır. Rasgele orman algoritması kullanarak özniteliklerin önem sırası belirlenmiştir. Özniteliklerin önem sırası gözetilerek hangi özniteliklerin nihai modelde olması gerektiğine karar verilmiştir.

En uygun modelin oluşturulması için derin öğrenme algoritması ve farklı parametreler için grid arama özelliği kullanılarak en uygun parametreler belirlenmiştir. Belirlenen modelin uygunluğu için veriler eğitim, doğrulama ve test verisi olarak gruplara ayrılmış ve başarı oranı ölçülmüştür. Ayrıca oluşturulan modelin gerçek zamanlı olarak çalışacak kadar yanıt verme süresinin hızlı olduğu ortaya konmuştur. Üretilen model gerçek zamanlı olarak test edilmiş ve yanıt süresi ölçülmüştür. Gerekli zaman kısıtları altında doğru sonuçlar üretebildiği ile test edilmiştir. Bu sayede oluşturulan sistemin başarısı ortaya konmuştur.

Yapılan çalışma sonucu üretilen modelin hem gerçek zamanlı olan hem de gerçek zamanlı olmayan sınıflandırma başarısı ölçülmüştür. Her iki durum için de beklentileri karşılayacak nitelikte bir başarı oranı elde edilmiştir. Ayrıca artımlı öğrenme yaklaşımı kullanılarak hem gözetimli öğrenme hem de gözetimsiz öğrenme için sistemin başarısı test edilmiştir. Sistemin başarısını etkileyen tüm parametreler irdelenmiştir. Oluşturulan model ve yöntem bu tip çalışmalar için bir öncü niteliğinde olacaktır. Yapılan deneylerin ve toplanan verilerin artması ile birlikte insan yaşamındaki sıradan olayların bile beyin sinyalleri ile ilişkisi ortaya çıkarılacaktır.

Özetle, kullanıcılar farklı duygu durumlarını tetikleyecek film parçaları ve binoral vurulardan oluşan bir videoya izlerken kullanıcılardan gelen EEG verileri toplanmıştır. Sonra bu EEG verileri üzerinde ön işleme ve öznitelik seçimi işlemleri yapılmış ve derin öğrenme sinir ağları kullanılarak %87'nin üzerinde bir doğruluğa sahip bir duygu tanıma modeli oluşturulmuştur. Daha sonra ise bu model kullanılarak gerçek zamanlı olarak gelen EEG verilerini tutan belirli bir uzunluktaki kuyruk yapısı üzerinde anlık duygu durumu tespiti yapan bir sistem oluşturulmuştur.

Sonuç olarak, EEG sinyallerini kullanarak gerçek zamanlı duygu tanıma işlemi gerçekleştiren bir model sunulmuştur. Çoklu ortam tarafından uyarılan duygular ile beyin dalgaları arasındaki ilişki ortaya konulmuştur. Yapay sinir ağlarını temel alan bir sınıflandırıcı kullanılarak komedi, korku ve hüzün duygularını barındıran çoklu ortam türleri ortalama %87'lik bir doğruluk ile sınıflandırılmıştır. Gelecekte kullanıcıların korku seviyesine bağlı olarak olayların gelişeceği gerçek zamanlı bir oyun yapmayı

planlıyoruz. Bu oyunda kullanıcılar, çözmek zorunda olduğu gizemlerin yanısıra korkularını da yenmek zorunda oldukları uyarlanabilir bir hikaye ile karşılaşacaklar.

Gelecekte bu tarz çalışmaların zenginleşip artmasıyla birlikte insanların hastalanmayacağı veya üzülmeyeceği daha mutlu yaşayacağı bir toplumun oluşabileceği ümit edilmektedir. Fakat bu çalışmaların ahlaki ve etik sınırların dışında kullanılmamasını sağlamak için de çalışmalar yapılması gerekmektedir. Devletin ideolojik aygıtları tarafından bu tarz çalışmaların kullanılması suretiyle distopik bir topluma dönüşme ihtimali de mevcuttur. Bu nedenle gelecekte bu tarz uygulamaların etik sınırlarda kalması için protokoller geliştirilmesi hedeflenmektedir.

1. INTRODUCTION

In accordance with the developing technology and due to the low-cost Brain-Computer Interfaces (BCI), brain waves can easily be acquired. Originally, BCI was developed for medical purposes; it allowed the physically disabled people to control their limbs via brain waves [1, 2]. In addition, researchers gave efforts to identify Attention Deficit Hyperactivity Disorder (ADHD) using ElectroEncephaloGraphy (EEG) signals [3]. Nowadays, many educational systems have been implemented through the BCI using brain waves [4].

Recent studies have shown the importance of the role of emotions to make decisions [5]. More recently, accessing physiological responses has attracted attention in recognition of the emotional states, unlike conventional methods that use auditory and visual features [6]. All BCI applications might be improved by including emotion sensing and utilizing emotional information in decision making.

EEG is the recording of the bioelectric activity that occurs in the brain during different physical and chemical activities of the brain. The brain produces signals having a wide variety of frequency and amplitude values depending on the state of the human being awake, asleep, or the mood of the person. EEG signals are not periodic, their amplitude and phase are constantly changing. Brain waves of EEG signals are divided into five main frequency bands: delta (1-3Hz), theta (4-7Hz), alpha (8-12Hz), beta (13-30Hz) and gamma (31-50Hz). Low frequency bands are associated with *sleep*, *relaxation* and *meditation*. High frequency bands are associated with *attention*, *information processing* and *pain*. These activities are related to the limbic system. Emotions are linked with the entire nervous system but the limbic system is especially significant. It includes thalamus, hypothalamus, hippocampus, amygdala, and several other nearby areas as shown in Fig. 1.1 [7]. For instance, gamma waves enable to binding mechanism in the thalamus where complex functions are carried out [8]. Just as we can predict the operations performed on the processor by listening to the power
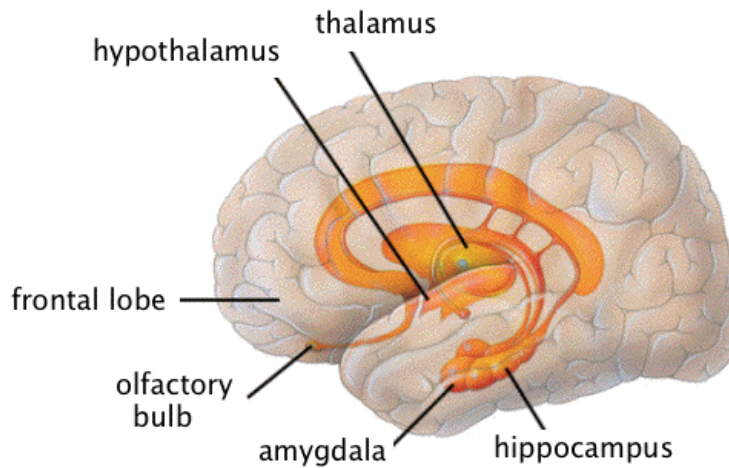
Figure 1.1 : The limbic system [7].

supply of the computer, we can also understand what operations are carried out in the brain with EEG data [9].

1.1 Purpose of Thesis

This thesis aims to provide enrichment of both medical and non-medical applications with real-time detection of the intended emotional state and to prepare the ground for personalized applications. With the estimation of the emotional state, all practices including voice assistants, film and music suggestion systems and even food recommendation systems used in daily life will be specialized and the doors of a specialized world belonging to each individual will be opened. It will also be an important step towards analyzing many complex human behaviors that the scientist have not yet fully grasped, including the decisions we make, as well as the choices we make unintentionally.

Emotion recognition from EEG signals in real-time brings with it the following challenging research problem: The acquiring of EEG signals, the selection of the most appropriate features, and the generating of the most appropriate recognition model using these features.

It is necessary to write a software for acquiring EEG signals from the portable device and to be able to report real-time data coming from this software. By looking at the reported data after this software is written, it will be determined whether the software works correctly.

2

For the selection of the right features, it is first necessary to create visual and auditory stimuli that will alert the emotions correctly. For this reason, it is necessary to find visual and auditory stimuli that will be both short in duration and as effective as possible. The sufficiency of the content of stimuli will be pre-evaluated. Once the content selection is complete, decision trees and random forest algorithms commonly used in the literature will be used to find the correlation between EEG data and emotional state for the selection of features. With these algorithms, it will be tested which features should be selected.

To create the most appropriate model, a deep learning algorithm will be used and the most suitable parameters will be determined using grid search option for different parameters. For the suitability of the determined model, the data will be divided into groups as training, validation and test data and the success rate will be measured. Also, the response time must be fast enough to work in real-time. The generated model will be tested in real-time and the response time will be measured. It will be tested to produce accurate results under the required time constraints. In this way, the success of the real-time emotion recognition system will be measured.

1.2 Literature Review

The literature overview of this research covers a variety of aspects mainly from two domains: 1) General usage of portable EEG-based applications, 2) Emotion recognition applications.

1.2.1 General usage of portable EEG-based applications

The BCI application developed by Rani and Umamakeswari [10] provides safe handling of paralyzed patients using wheelchairs. The application enables to control the wheelchair using alpha and beta waves in the patient's brain. In addition, the patients can turn the wheelchair to the right or to the left using their blink.

Visu et al. [11] proposed a system using brain signals to prevent accidents caused by sleeping drivers. The system, which uses the theta and alpha waves, detects when driver is going to sleep and adjusts the speed of the vehicle by turning control over to the automated system. This prevents the accident that may occur in case of a possible drowsiness.

3

A study by Kan, Lim and Lee [12] developed an application for the detection of distractibility. This study was tested on a group of undergraduate engineering students. As a result of these tests, it was seen that the level of attention of the girls was higher than that of the boys. It has also been found that visual items increase the level of attention more than auditory items and practical exercises.

Bonaci and her colleagues [13] aimed to ensure the safety of people's personal data in BCI applications. Bonaci and his colleagues, exemplifying the ability to restrict access to applications on smartphones, have proposed a model that hides this data to limit access to spyware in personal applications.

Chan et al. [14] have tried to determine brain state (active, rest) with EEG data by using different machine learning algorithms such as Random Forest, Boosting, Naive Bayesian Classifier, k-Nearest Neighbors (KNN) and Support Vector Machine (SVM). As a result of this study, they found that the Random Forest algorithm had the highest performance rate to detect the state of brain.

Ang and his colleagues have developed a BCI application that allows them to use the mouse cursor using a single-channel EEG device [15]. Using this application, it is possible to move the mouse pointer to the desired point by simply winking. Through this application, it is aimed to make it possible for people with physical disabilities to use the computer comfortably.

Kumar et al.'s work is an extraction of features through the 'NeuroSky Mindwave' product and OpenVibe software [16]. In this study, a method has been realized which enables different states of the brain to be detected with the algorithm developed on OpenVibe.

In a study by Morales and colleagues, a low-cost open-source software and navy design has been proposed that enables remote observation of EEG data [17]. EEG data can be transferred to the web server via hardware and made available online via the server.

In the study conducted by Gonzalez et al. [18], three levels of computer programming (code writing, document preparation and debugging) were measured using the low cost EEG device. There is a statistically significant concentration difference between writing code and debugging but this is not the case for document preparation.

Emotional influences (music, video, pictures) to user are compared with the application framework proposed by Chen and colleagues [19]. It is intended to provide emotional adaptation by personalizing the content that users receive through this application which provides feedback to the user.

Singh et al. proposed a new method of communication using BCI technology [20]. According to this method, a 4-bit number consisting of 1 or 0's is generated by the user's eyes, and the predefined text corresponding to this number is converted to voice signals by the method of conversion from text to speech. It is aimed that the speech-impaired people can express themselves by voice.

The study by Perhakaran and colleagues compared the effect of virtual reality and imagery therapy on stress [21]. This comparison was made through the meditation level obtained using the 'NeuroSky Mindwave' product. It has been found that through the use of virtual reality, people become more relaxed.

Vijayaragavan and colleagues [22] used brain waves from the 'NeuroSky Mindwave' to measure the success of music therapy and yoga techniques in their study to reduce stress. In an experiment conducted with over 100 people, 67 were more relaxed with Yoga, while 29 people had lowered their stress level with music therapy. No change was observed in 4 persons.

1.2.2 Emotion recognition applications

EEG-based emotion recognition has been used frequently in recent years in the development of human-computer interface systems. In these studies, researchers have tried to reveal the emotional situations of the persons with a stimulus. These stimuli can be visual, auditory or both. In recent years, some researchers have selected visual stimuli as emotion inducers in their work and used pictures as stimuli [23]. In addition, video clips including visual and auditory stimuli have been widely used in previous studies [24].

In the literature, there are different approaches to emotion recognition. Both low frequency bands [25] and high frequency bands [26] can be used for determining the emotional states. In addition, some researchers claim that the significant information to detect an emotional state is found in the frequency below 30 Hz [27]. Nevertheless,

this research showed the importance of the gamma wave which is above 30 Hz. Also, previous works state that the gamma wave indicates emotional consciousness of a person [28].

In spite of the fact that many researchers have worked for recognition of emotions from the brain waves, the classification results are rather unsatisfactory. An emotion recognition system based on a neural network got a classification accuracy of 64% [29]. Moreover, few research papers have used one electrode device for emotion recognition because single-channel EEG devices are recently introduced. In addition, one electrode devices are affordable and easy to use.

The most similar work to ours has used classification on two classes between relaxation and fear emotions with an average accuracy of 92% using a Support Vector Machine (SVM) based classifier [30]. They recorded EEG data while test subjects were watching a video clip, which includes three emotional states: neutral, relaxation and scary. After that, they classified horror and relaxing movies. Although they have excluded the parts of horror movie that did not carry any fear emotion, our research will keep the movie's integrity.

In the literature, there are few studies on real-time emotion recognition with EEG signals. Liu et al. have tried to realize the real time emotion recognition application with EEG data coming from three electrodes [31]. In our study, a one electrode portable device was used unlike Liu et al.

1.3 Structure of the Thesis

This thesis is organized as follows:

In Chapter 2, first, the general overview of real-time emotion recognition system is portrayed. Then, Sub-parts of the real-time emotion recognition system are described. The relation between each sub-part forming the system and the working principles are explained in detail. In addition, the theoretical and practical information used to create the general design of the system is detailed in this chapter and supported with graphics.

In Chapter 3, the experimental environment established for the creation and testing of the proposed model is described. In addition, hardware and software used in the experimental environment are explained. Information about the subjects involved

in the experiment is given in this section. The theoretical knowledge about the performance criteria applied on the model and the results of the experiments obtained by using these theoretical knowledge are also a part of this chapter. In the last part of this chapter, there is a discussion section where all processes are criticized.

In the fourth chapter, which is the last chapter, a general summary of the study is given. It also emphasizes the importance of future work. In the last part of this section, we mention the future works.

Appendix A.1 provides information about the content of audio and video stimuli used in the experiments. Appendix A.2 includes screenshots from the real-time emotion recognition system.

2. REAL-TIME EMOTION RECOGNITION SYSTEM

The system having real-time EEG signal processing capability basically consists of two modules. The queuing port structure holds the real-time EEG data coming from the subjects. Then using a pre-trained emotion recognition model, emotional state of the user is estimated at certain periods using the incoming data. Fig. 2.1 shows the real-time emotion recognition system.
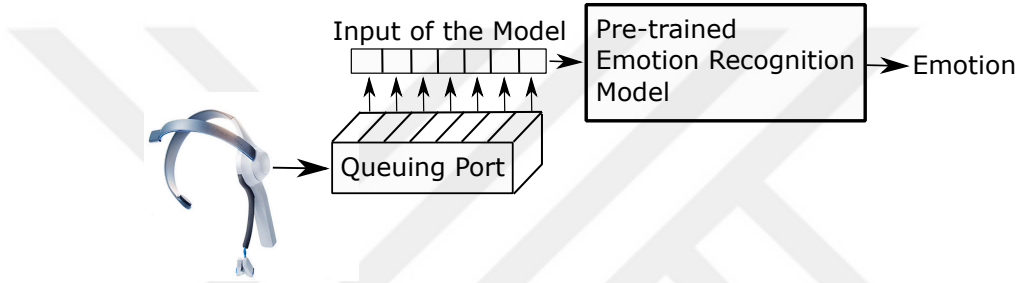
Figure 2.1 : Overview of the real-time emotion recognition system.

2.1 The Queuing Port

The queuing port structure is often used for real-time communication. It can be configured as both FIFO (First In, First Out) and priority-based. Our system uses FIFO-based queuing port structure to buffer EEG data. The last element of the queue is discarded for each incoming data. The pre-trained emotion recognition model checks the queue in fixed periods and uses all data in the queue to determine the emotional state. The FIFO buffer length is an adjustable value, but it must be the same length as the number of the pre-trained emotion recognition model's features in order for the system to work. Fig. 2.2 shows the queuing port structure which is used to manage real-time signal flow.

Our model uses reactive reader to check the queuing port data. The change of emotional state will be observed in the system depending on the period in which it

9
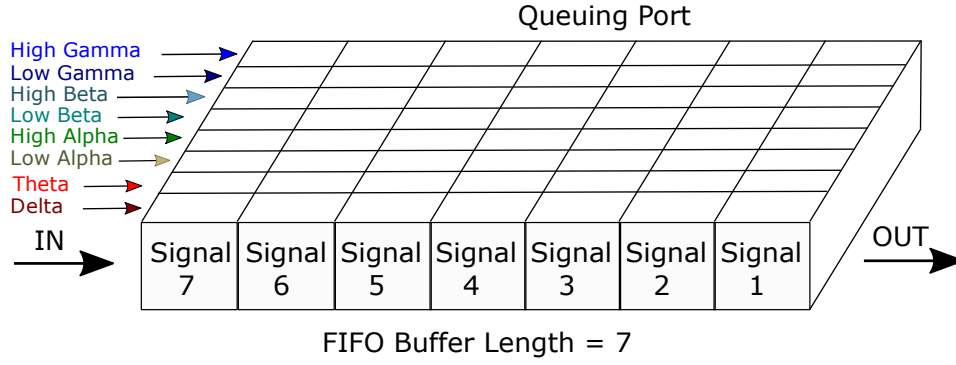
Figure 2.2 : The queuing port structure.

is controlled. Intervals with a length of at least one second and at most three seconds are considered reasonable to make the system real-time compatible.

2.2 Model Generation for Emotion Recognition

In the overview of the system, the sub-steps roughly contained in the model production are shown. In the EEG acquisition section, information related to the part of the brain collected signals and the EEG device used is given. In the preprocessing section, which filters are used for the signals are explained. In the feature selection section, the data obtained from the subjects were examined with decision trees and random forest algorithms. It is determined which features are to be selected depending on the correlation between the features and emotions. Deep learning has been defined in the classification section, which depends on the selected features. This determined model was tested for different parameters and the final model was generated.

2.2.1 Overview of the system

There are four basic steps involved in the proposed methods in Fig. 2.3. EEG acquisition step shows how the electrical signals are captured. Preprocessing and feature selection steps display the transformation action from data to feature. Estimation of the emotion using deep learning is presented in the classification step.

2.2.2 Electroencephalogram acquisition

In this thesis, we have focused our attention to use of a one electrode EEG device. Fig. 2.4(a) shows an internationally recognized 10–20 electrode mounting method to
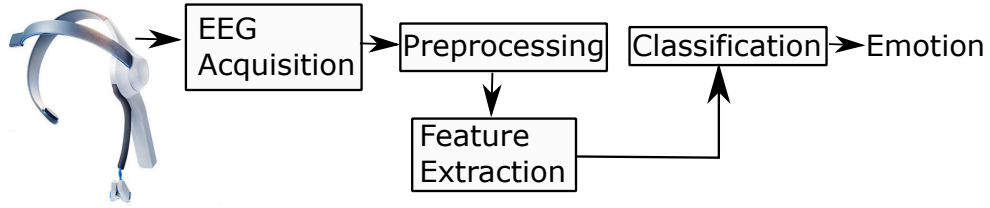
Figure 2.3 : The overview of the model generation system for emotion recognition.

apply the location of EEG electrodes for recording different brain waves [32]. The letters F, T, C, P and O represent Frontal, Temporal, Central, Parietal, and Occipital lobes, respectively. Even numbers (2,4,6,8) represent the positions of the electrodes on the right hemisphere and odd numbers (1,3,5,7) represent the positions of the electrodes on the left hemisphere. The letters A and Fp represent Earlobes and Frontal polar, respectively. As shown in Fig. 2.4(b), using one electrode device we are able to capture the electrical signals in the brain. Sensor arm of this EEG device is located on the frontal lobe (FP1). The ground and reference electrodes are placed on the earlobe (A1). Gel was not used because all electrodes were in particular settings of dry type. This device measures the EEG power spectras (delta, theta, alpha, beta, and gamma waves), eSense meters (attention and meditation) and eye blinks. The device has an output of 12 bit raw-brainwaves (3 - 100Hz) with a sampling rate at 512Hz. EEG analysis involves measuring amplitudes (powers) of activity in certain frequency ranges, so-called bands. ThinkGear that allows the measurement, amplification, filtering, and analysis of EEG signals and brainwaves reports the relative power of each EEG band, typically at 1 second intervals [33].
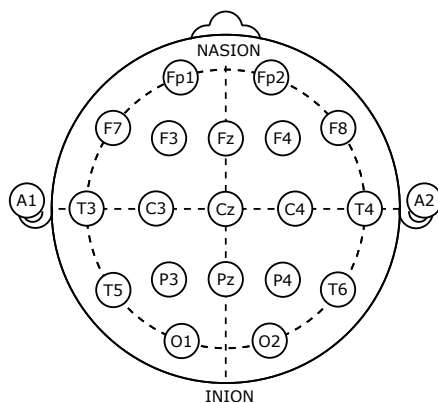
(a) Nomenclature of electrode positions.
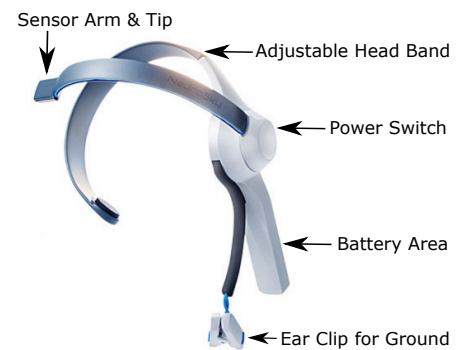
(b) One electrode headset developed by NeuroSky.

Figure 2.4 : EEG acquisition system.

2.2.3 Preprocessing of EEG signals

The data containing Not Available (NA) information are deleted because subjects blink, as shown in Fig. 2.5. High-noise data was discarded. The artifacts and noise

delta	theta	lowAlpha	highAlpha	lowBeta	highBeta	lowGamma	highGamma	blinkStrength
390619	31028	5851	3225	1589	1955	361	889	NA
34488	41203	40866	17793	18092	9429	5162	17119	NA
NA	NA	NA	NA	NA	NA	NA	NA	57
179955	19225	12992	11556	8019	2033	5051	8720	NA
44970	4984	2305	3056	6094	9615	3622	14056	NA
22825	27587	1059	6685	3312	7792	9180	16881	NA
NA	NA	NA	NA	NA	NA	NA	NA	41

Figure 2.5 : A sample data containing NA rows.

produced in the system can be removed by applying data windowing and the binary logarithm to the data sequence. The steps for these operations are shown in Fig. 2.6 step by step. The elimination of weak signals was also carried out at this stage. The weak signal elimination operation was performed based on the data indicating the correctness of the incoming signal provided by the device.

After preprocessing operation, we investigated the EEG power spectrums (delta, theta, alpha, beta, and gamma waves) and eSense meters (attention and meditation) with different classification methods.

2.2.4 Feature selection

Correlation between brain signals and emotions was roughly obtained using decision trees in preliminary evaluation. The order of importance of the features was determined using the random forest algorithm. It was decided which features should be in the final model by considering the importance order of the features. The data used in this section are brain signals that contain emotions such as funny, weepy and fear, which have been obtained as a result of the reaction of the subjects to the visual and auditory stimuli.

2.2.4.1 Preliminary evaluation of features

Decision trees are a visual modeling method which clearly shows the stack of information that the decision maker has about the probing encountered and places the order of the decision options and probabilities in a specific order. It starts from the root in the tree and branches off to the leaves. Best features are selected to divide sample
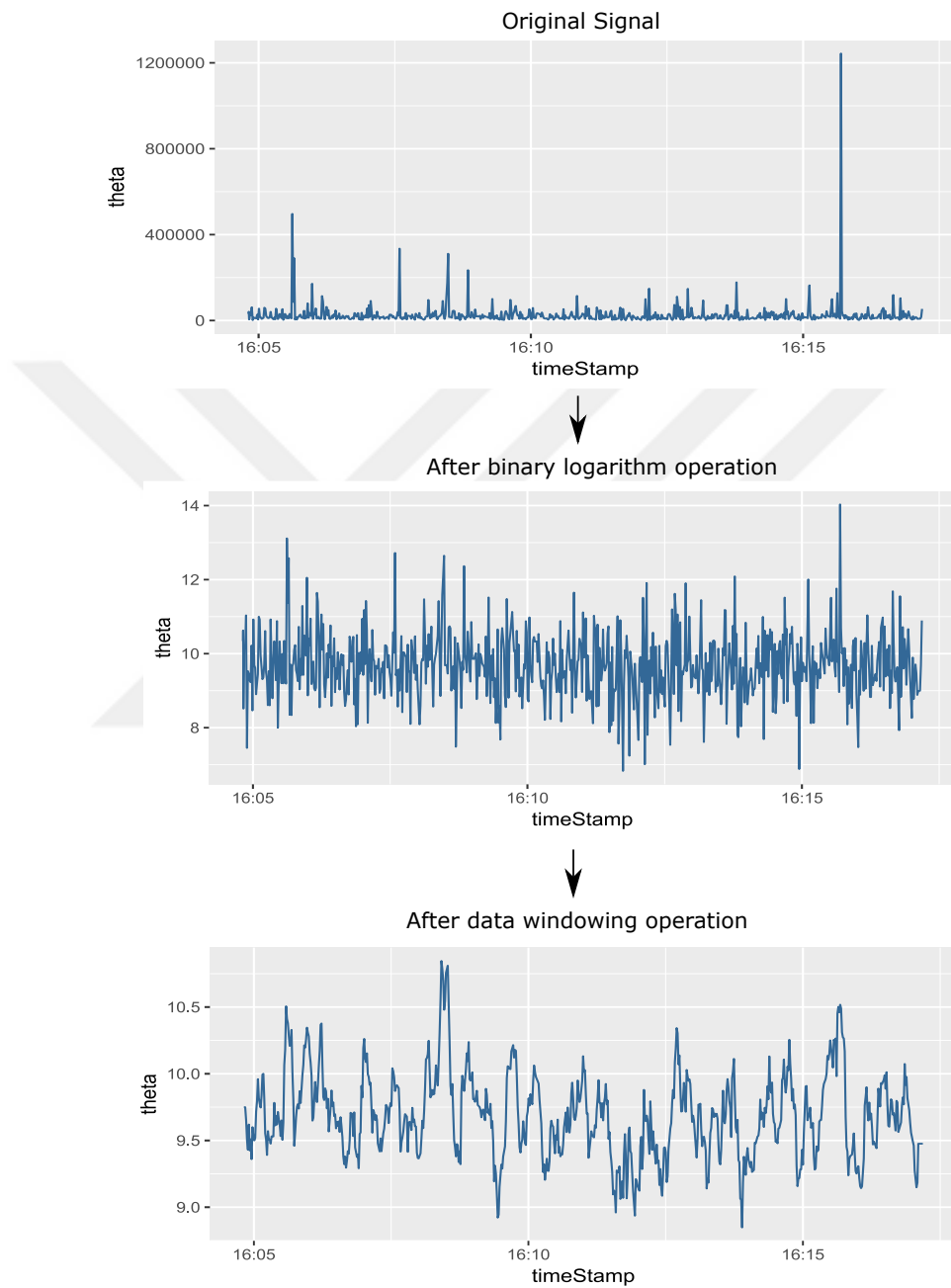
Figure 2.6 : A sample preprocessing operation on theta waves.

nodes to the classes. If all the samples belong to the same class, no new branch is made and the branch ends as a node. Branching continues until all instances of the branch belong to the same class, or until the instances are not qualified to branch [34].

Information gain is an impurity-based criterion that aims to keep impurity at the lowest level when classifying at a node and forms a tree using concepts of information and entropy theory. It assumes that all features that are to be used in the tree receive discrete values. The tree is created by selecting for the partition with the highest information gain from the features [35]. Entropy takes a value between 0 and 1. Entropy becomes zero when all the data in our hand belongs to a single class. When all possibilities are equal, the entropy reaches its maximum value. Mathematically expressed as follows: T represents the target feature, and p_i is the probability of the class in *i-th* node and c denotes the number of classes to separate. Entropy is defined as shown in Equation (2.1).

$$Entropy(T) = -\sum_{i=1}^{c} p_i \log(p_i) \qquad (2.1)$$

To calculate the gain, the entropy of the data and the sum of the weights of the entropies calculated for each feature. X represents attributes.

$$Entropy(T,X) = \sum_{c \in X} P(c) Entropy(c) \qquad (2.2)$$

If this difference is large for subdivision, then branch operation is done correctly. The purpose of decision trees is to find the value that maximizes the value of gain. The formula of the information gain is expressed as follows:

$$Gain(T,X) = Entropy(T) - Entropy(T,X) \qquad (2.3)$$

Decision trees are often used in feature selection. Usage of decision trees is a quick way to find out what changes in features have changed. A binary decision tree was trained with collected EEG data for classification with information gain split method. There are three classes for classification. This decision tree gives a coarse idea about which features might be most important, as shown in Fig. 2.7.

The tree is interpreted as the following: Each value at the bottom of boxes show the percentage after split and values in the middle of boxes indicate distribution of percentage for each class as funny, horror and weepy. Attention and meditation values are device specific measurements which are based on scaled beta and alpha waves,

14
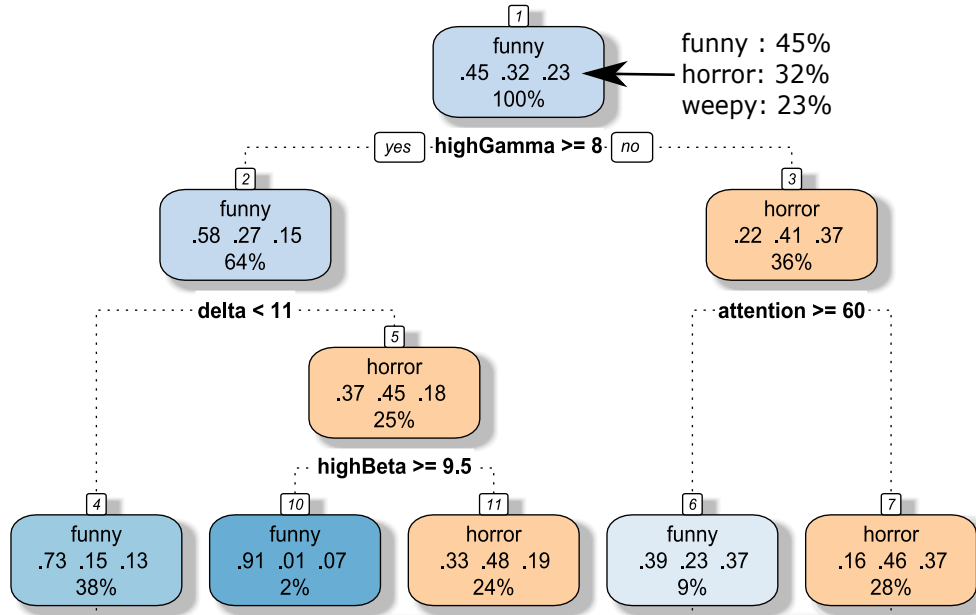
Figure 2.7 : The classification result of decision tree.

respectively. This decision tree shows that funny videos increase high gamma value, and lower high gamma value is related emotions like sadness, grief, regret, fear.

2.2.4.2 Determination of feature importance

Random Forest is a classification algorithm that contains more than one decision tree [36]. Classification results using estimations are produced with more than one classifier, as in other batch classification methods. In the random forest algorithm, the best one is selected from the randomly selected variables in each node. With this algorithm, the tree structure is created by determining the number of trees and samples to be used in each node. The success of the selected features is investigated to ensure the best branching at each node. When the nodes are split in the trees, all of the existing features are processed using the property that will give the best result in a subset of randomly chosen ones, and the classification of the new dataset is completed.

We used Random Forest approach to estimate the importance of the features. The number of user-defined parameters required by random forest classifiers is less and these parameters are easier to define unlike in the support vector machines and neural networks. Fig. 2.8 shows that high gamma is the most important determining factor for predicting emotion, followed by delta, low gamma and high beta.
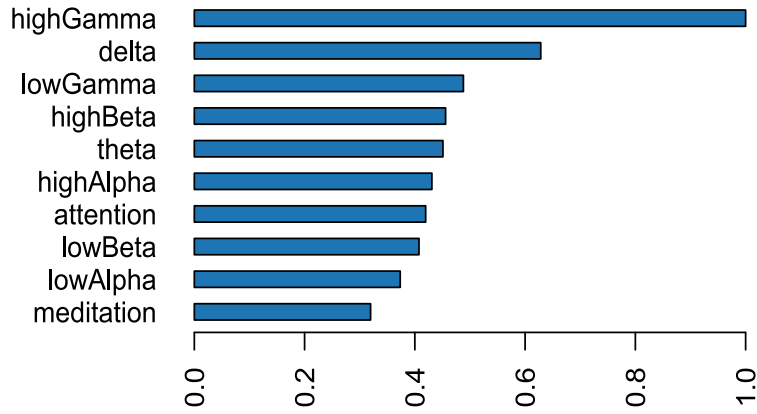
Figure 2.8 : Scaled feature importance of variables.

2.2.4.3 Selected features

Finally, 8 features (delta, theta, low alpha, high alpha, low beta, high beta, low gamma, high gamma) given in Table 2.1 are selected for training a deep learning model. We exclude attention and meditation values because they are device specific measurements. In addition, we can not exclude any of the EEG power spectrums (delta, theta, alpha, beta, and gamma waves) because all of them have a considerable importance acccording to Fig. 2.8.

Table 2.1 : EEG frequency bands used in the proposed system.

Band	Frequency	Activity
Delta	1-3 Hz	In deep sleep, or when unconsious
Theta	4-7 Hz	The region between sleep and wakefulness
Low Alpha	8-9 Hz	Awake,
High Alpha	10-12 Hz	fully conscious
Low Beta	13- 17 Hz	The attention focussed outside,
High Beta	18-30 Hz	solving logical problems
Low Gamma	31- 40 Hz	High-level information processing
High Gamma	41- 50 Hz	

2.2.5 Classification using deep learning

Artificial Neural Networks (ANN) are a class of algorithms that contain many different types of algorithms based on graphs. There are so many types of ANNs. The first type of artificial neural networks were linear, meaning that they could solve only linear

16

decision problems (ie, records that were linearly separable by drawing a line). Over time, this linear neural network model has been known as a perceptron.

Linear neural networks consist of a bipartite graph, the left node, the inputs and the right side nodes are output. Only the weights of the edges between these nodes are learned (the node activation threshold can also be adjusted, but that is only rarely).

A large step was taken when *flat neural networks* were invented: instead of having only one bipartite graph, we use a 3-part diagram: the *input layer*, the *output layer*, and a *hidden layer* between them. Using the *hidden layer*, the network can now make nonlinear decisions and solve problems.

Note that the *flat* term has been retrospectively shaped when a *deep neural network* (also called *n-layers neural networks*) were invented. This is to resist neural networks with only one hidden layer, with *deep neural networks* with *n hidden layers*. With more hidden levels, we can choose more complex data sets as there are more layers to modulate the decision (in other words, to increase the dimensionality of the decision boundary, which can lead to overfitting).

For each feature, we apply a sliding window with a size of 7 (3 values before and after the instance of EEG sample) as shown in Fig. 2.9, because a temporal analysis of subjective feelings is related to before and after that moment. This leads to a total of (7x8) 56 input features for a given value of a wave. The calculation of the change in brain waves to determine the state of emotion significantly increases the success ratio.
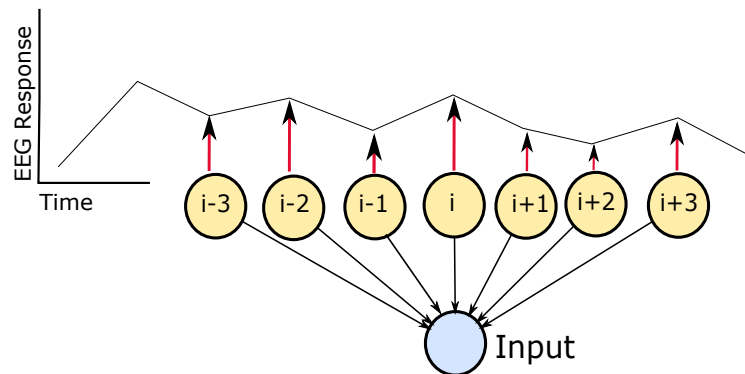
Figure 2.9 : A typical input of the ANN generated using sliding window approach.

The classifier has three output nodes representing funny, horror, and weepy, respectively. The deep Artificial Neural Network (ANN) consists of two hidden layers,

17

each with 150 nodes. In the final model, Maxout activation function which takes the largest input value was used with 60 epochs at each learning rate. Fig. 2.10 shows the proposed deep neural network model for emotion recognition.
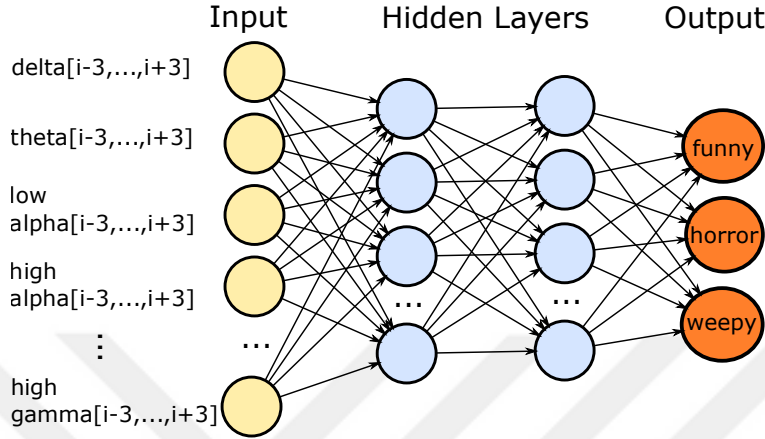
Figure 2.10 : The deep learning neural network model used in the proposed method.

This model can be used in real time once it has been trained. The data read from the queuing port in real time are given as input to this model. This trained model is used to classify the data flowing in real time with low latency.

2.2.6 Incremental learning

Incremental learning which is a machine learning paradigm refers to the situation of continuous model adaptation based on a constantly arriving data stream. The difference of incremental learning from traditional machine learning is that it continues the learning process until it reaches a sufficient training set [37]. The system has been tested for both supervised and unsupervised learning using an incremental learning approach.

In the supervised approach, a model is created through incremental learning using labeled data based on a constantly arriving data stream. Using limited labeled data, the deep learning model has tried to predict remaining data. For testing model, both individual data and group data are used for model generation.

The most widely used non-hierarchical clustering method which is the k-means clustering is used for analyzing data obtained from unsupervised incremental

learning experiments. The K-means clustering method aims to determine relatively homogeneous groups based on predetermined characteristics using an algorithm that can deal with a large number of situations [34]. K-means is an algorithm that groups the data into k clusters. Clustering is achieved by minimizing the sum of distances and squares between the data (growth rates) and the appropriate cluster center (centroid). This method continues until the minimum clustering procedure is stationary, grouping the points according to the minimum distance. Classification is done in such a way that the variability between the clusters is very high and the intra-cluster variability is the least. Using limited data, K-means algorithm has clustered arriving data stream to 3 different clusters which show 3 different emotional states.

3. EXPERIMENTS AND RESULTS

In this chapter, we first describe the software and hardware equipment that we use in our work. Then, the visual and auditory stimuli used in the experiments are explained. In addition, information was given about subjects voluntarily participating in the experiments. The criteria we use to measure the success of our work are also included in this section. In addition, the results of the online and offline experiments are explained. In the last part of this chapter, we discuss the results of the studies in detail.

3.1 Hardware and Software

Mindwave Headset International Rf Version (Single-channel Dry Electrode EEG) was used for brain wave acquisition as shown in Fig. 2.4(b) [38]. A Java program stores EEG data with timestamp. The EEG data of the subjects, which have been collected under visual and auditory stimuli, has been analyzed using different metodologies such as decision trees, random forests and deep learning. We used R Studio as Integrated Development Environment [39]. The random forest and deep learning approach is developed and analysed in R using H2O package [40]. The decision tree is developed in R with rpart package [41]. Clustering operations are developed in R with kmeans function. The proposed method was tested on a personel computer with Intel® Core™ i5-3317U Processor @ 1.70 GHz (4 CPUs) and 8 GB RAM running under Windows 10 operating system. Samsung brand wired earphones with high quality sound transmission were also used.

3.2 Experimental Setup

For the experiments, an experiment room dedicated for user experience evaluation in Istanbul Technical University was used. Everything from room temperature to soundproofing is set for professional experiments in this special environment. This chamber, separated by a glass partition, was observed from the outside during the

experiment and it was noted that the subjects were sufficiently sensitive to the experiment. Participants were instructed to watch a video, and to pay attention to the sounds they hear through the earphones.

A video was prepared, which contains several kinds of movie clips and binaural beats [1]. As stimuli, we chose movies which have high effect to arouse human emotions. We evaluated several different stimuli when selecting movies. When we examined all movies in terms of emotional effects, we realized that they addressed three different emotions. For this reason we focus on producing a content that will include three different emotions, funny, weepy and horror. We watched all the short films on the Internet. As a result, we chose movies that would have the most impact in a short time. Each of the films we choose is one that has been rated as the best stimulus for one of three different emotions. We also chose a suitable sequence for the flow of movies when choosing binaural beat. Thus, the users made a controlled transition from positive emotions to negative emotions. Emotion inducing video parts given in Table 3.1 shows visual and auditory stimuli, their corresponding timing and duration. In addition, pictures from the stimuli are included in the Appendix A.1.

Table 3.1 : Audio and video stimuli used in the experiments.

Stimulus	Start time [min]	End time [min]	Duration [sec]
Binaural beats (8 Hz)	00:00	00:29	29
Binaural beats (alpha)	00:29	01:04	35
Funny animation	**01:04**	**04:19**	**195**
Binaural beats (beta)	04:19	04:51	32
Binaural beats (theta)	06:46	07:16	30
Weepy movie	**07:16**	**08:57**	**101**
Binaural beats (gamma)	08:57	09:26	29
Horror movie	**09:26**	**11:51**	**145**
Binaural beats (delta)	11:51	12:21	30

A binaural beat is an auditory sensation, which appears when two slightly different sounds are received to two ears [42]. We used binaural beats to ensure that users are focused on the actual stimulus. Firstly, we used sounds having a frequency of 8Hz

---

[1]Video and Sound effect on Brain Waves. Retrieved December 31, 2016, from https://www.youtube.com/watch?v=U2Li8hygdm8

because they bring the brain into a relaxed state before the experiment. In additon, alpha, beta, theta, gamma and theta waves as binaural beats enable the human to relax and focus on the stimuli. After the horror movie we played the binaural beat (delta) because the subjects must have been influenced by fear.

3.3 Test Subjects

The group of participants consisted of 10 healthy volunteers (21-28 years old, 8 males, 2 females). All subjects were informed about the aim and scope of the study. The participants did not have any physical or mental illness or condition that would affect the quality of their EEG. Fig. 3.1 shows a photo from the experiment session.

Figure 3.1 : A photo from the experiment.

3.4 Evaluation Criteria

Funny, horror and weepy movies were selected from movies having only one pure emotion which were labeled as funny, horror and weepy, respectively. All test subjects watched the video and we collected all brain waves during video. In summary, the data in total contains 4193 instances (1886 instances for funny movie, 1330 instances for horror movie and 977 instances for weepy movie).

The data is partitioned into 70%, 15%, 15% chunks for training, validation and test, respectively. *Rectifier* is an activation function, which is zero when input value (x) is less than 0 and then linear with slope 1 when x is greater than 0. *Maxout* is an

activation function, which takes the largest value of the inputs. *Dropout* is a type of regularization, which helps to avoid overfitting. *Dropout* remove some neurons randomly to prevent too much co-adaptation of neurons. Using grid search, *Rectifier*, *RectifierWithDropout*, *Maxout* and *MaxoutWithDropout* activation options were tested. Also, the model was tested with different epoch sizes (10 to 70). In addition, we investigated node size of the hidden layers.

The confusion matrix of a classifier which is given in Table 3.2 consists of four outcomes. First outcome is *True Positive (TP)* occuring when both the predicted and

<div align="center">

Table 3.2 : Confusion matrix.

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	TP	FN
	Negative	FP	TN

</div>

actual classes are *positive*. Second outcome is *False Negative (FN)* which occurs when the actual class is *positive* and the predicted class is *negative*. Third outcome is the *False Positive (FP)* occuring when the actual class is *negative* and the predicted class is *positive*. The last outcome, the *True Negative (TN)*, occurs when both the predicted and actual classes are *negative*.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{3.1}$$

The accuracy value shown in Equation (3.1) is the metric based on the success rate of the classifier.

3.5 Results

In this section, the results of experiments performed offline and online are given. It also shows the effect of different activation functions and the changes on the results affected by the ANN parameters.

3.5.1 Results of the offline experiments

Fig. 3.2 shows the effect of epochs. Higher number of training epochs causes overfitting. Therefore, 60 was selected as the best epoch for our system. We find the

time of termination for the training of the model by determining the point of overfitting because the right choice of epoch size allows us to create a learning pattern without recourse to predicted emotional states.
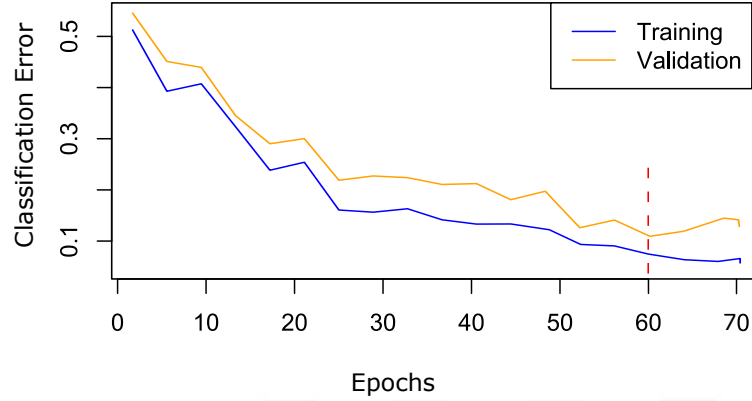
Figure 3.2 : The relation between classification error and epochs.

Confusion matrix of the proposed model given in Table 3.3 shows the performance of the proposed system. According to this table, the model achieved 87.89% accuracy. Also, the performance of predicting horror movie has the highest score. Predicting the weepy movie is less accurate because the feeling of sadness does not have same effect on every people.

Table 3.3 : Confusion matrix of emotion recognition obtained in the offline experiments.

Actual	Predicted			Error
	funny	**horror**	**weepy**	
funny	**249**	22	17	0.1354
horror	5	**180**	4	0.0476
weepy	14	15	**130**	0.1824
Total	268	217	151	**0.1211**

3.5.1.1 The result of person independent tests

To obtain person independent results, each individual's data was tested while other subjects' data were used as training data. As a result, the error rate of deep learning model for each individual's data was found. The average of these error rates is 0.2885.

The proposed model achieved 71.15% accuracy as a person independent. This result shows that the model can also be used as person independent.

3.5.1.2 The effect of pre-processing

Confusion matrix of the proposed model given in Table 3.4 shows the performance of the proposed system without pre-processing. According to this table, the model achieved 49.21% accuracy. The reason why the success rate is low, when there is no pre-processing is that, deep learning model can not distinguish between different emotions. Although binary logarithm and data windowing seem like a simple preprocessing, it is clear how much difference they have made to success. The ANN's success is growing at a serious rate as we deal with the outliers using preprocessing.

Table 3.4 : Confusion matrix of emotion recognition obtained in the offline experiments without pre-processing.

Actual	Predicted			Error
	funny	**horror**	**weepy**	
funny	**195**	48	45	0.3229
horror	76	**73**	40	0.6138
weepy	69	45	**45**	0.7170
Total	340	166	130	**0.5079**

3.5.1.3 The effect of feature reduction

Confusion matrix of the proposed model given in Table 3.5 shows the performance of the proposed system without feature reduction (delta, theta, low alpha, high alpha, low beta, high beta, low gamma, high gamma and attention, meditation). According to this table, the model achieved 80.19% accuracy. The success rate is slightly lower than the model which includes feature selection because the deep learning model confuse emotions when attention and meditation data are included. For this reason, we see that the test results of the model we created with the reduced features (without attention and meditation) are better and feature selection is important.

3.5.1.4 The effect of sliding window approach

Error table for different size of sliding window given in Table 3.6 shows the performance of the proposed system. According to this table, the model achieved

Table 3.5 : Confusion matrix of emotion recognition obtained in the offline experiments without feature selection.

Actual	Predicted			Error
	funny	**horror**	**weepy**	
funny	**227**	21	40	0.2118
horror	4	**159**	26	0.1587
weepy	11	24	**124**	0.2201
Total	242	204	190	**0.1981**

60.67% accuracy without sliding window approach. The success rate is much lower than the model which includes sliding window approach because momentary brain signals do not reflect emotional state correctly. The changes that occur in brain signals are more important to detect feelings. In addition, best accuracy obtained when seven is selected as size of sliding window.

Table 3.6 : Error table for different size of sliding window.

The size of sliding window	Error
1	0.3933
3	0.3002
5	0.2411
7	**0.1211**
9	0.1741

3.5.1.5 The effect of size of hidden layers

Accuracy table for different numbers of hidden layers given in Table 3.7 shows the performance of the proposed system for different sizes of hidden layers. According to this table, the increase in the number of layers causes the deep learning model to memorize rather than to learn the data. Also, although the number of nodes in each layer causes a change in the accuracy, there is no clear reason for this. That is, the design of the layers may change according to the available data.

3.5.2 Results of the online experiments

Screenshots from the real-time emotion recognition system are included in the Appendix A.2. In the online experiment, the response of the real-time system to

Table 3.7 : Accuracy table for different numbers of hidden layers.

Hidden Layers	Accuracy
50, 10	0.6871
10, 50	0.7044
50, 50	0.7736
100, 100	0.8236
150, 150	**0.8789**
200, 200	0.8657
50, 50, 50	0.7662
30, 100, 30	0.7396
150, 30, 150	0.7048
10, 35, 100	0.6799

brain signals produced by three different users excluded from the training set for funny, weepy and horror stimuli was recorded. The results are calculated for each emotional state. Confusion matrix of the real-time model given in Table 3.8 shows the performance of the real-time system. According to this table, the model achieved 73.91% accuracy. Also, the performance of predicting horror movie has the highest score. Predicting the weepy movie is less accurate similar to offline results.

Table 3.8 : Confusion matrix of emotion recognition obtained in the online experiments.

Actual	Predicted			Error
	funny	**horror**	**weepy**	
funny	**48**	10	8	0.2727
horror	10	**61**	9	0.2375
weepy	7	4	**27**	0.2895
Total	65	75	44	**0.2609**

3.5.3 Results of the incremental learning

This section shows two different approaches for incremental learning. The first is a method based on classifying the remaining data that is used for instantaneous training via *supervised classification*. In the second method, the data is used again as instant

data to cluster emotions. In this method, however, the approach is *unsupervised clustering*.

### 3.5.3.1 Results of the supervised incremental learning

Incremental learning trains the model using continuous data that comes in succession compared to normal learning. Instead of statically modeling, a real-time model is created through incremental learning. In this section, firstly the instantaneous data for a selected individual was used as training data. The deep learning model trained using this training data was attempted to predict the data not yet collected. Secondly, the instantaneous data for 10 subjects were combined and used as training data. As in the previous method, the deep learning model trained using this training data was attempted to predict the data not yet collected. Both approaches yielded results for both mean squared error and accuracy.

The mean squared error shows how close a prediction line is to a set of points. It does this by taking the distances which are called as the errors from the points to the prediction line. The squaring is necessary to remove any negative signs. It also gives more weight to larger differences. Finding the average of a set of errors is called as the Mean Squared Error (MSE). Fig. 3.3 shows MSE obtained from supervised incremental learning based classification experiments applied to individual test subjects. An increase in learned brain signals of the individual reduces the error rate in predicting brain waves that have not yet been learned.
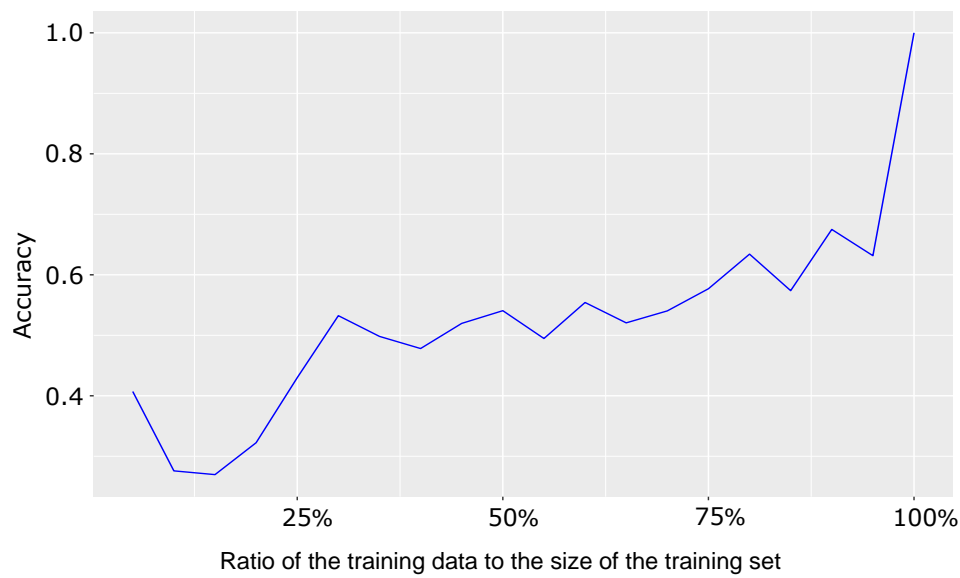
Fig. 3.4 shows accuracy obtained from supervised incremental learning based classification experiments applied to individual test subjects. The accuracy rate is also increasing as opposed to the decrease in the error rate.

Fig. 3.5 shows Mean Squared Error (MSE) obtained from supervised incremental learning based classification experiments applied to all test subjects as a group.

Fig. 3.6 shows accuracy obtained from supervised incremental learning based classification experiments applied to all test subjects. The individual classification seems to be more successful than the collective classification because the brain waves are person-specific. In the incremental learning process it would be more beneficial to treat each individual brain wave separately. Moreover, the success rate of incremental learning seems to be slightly lower than learning offline. In the offline tests, while all
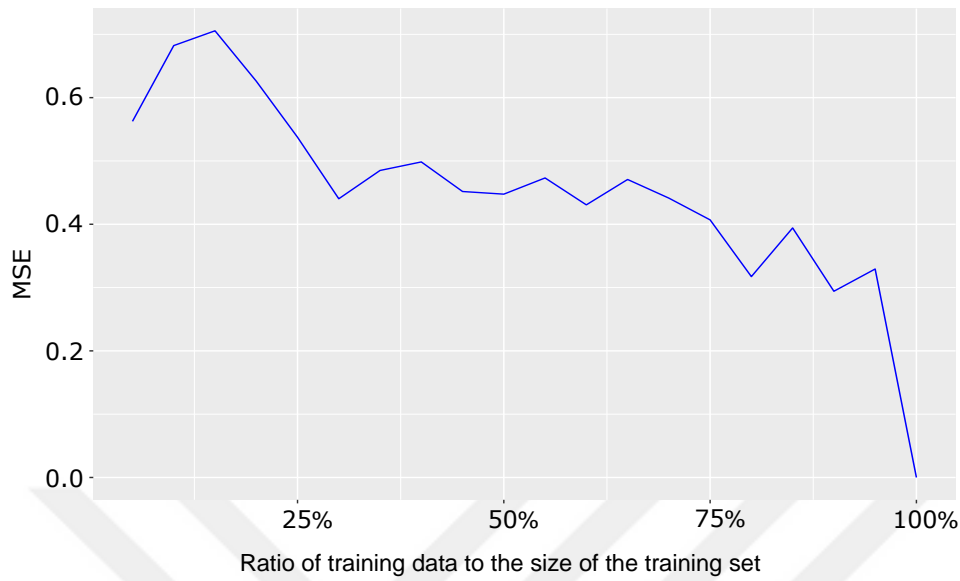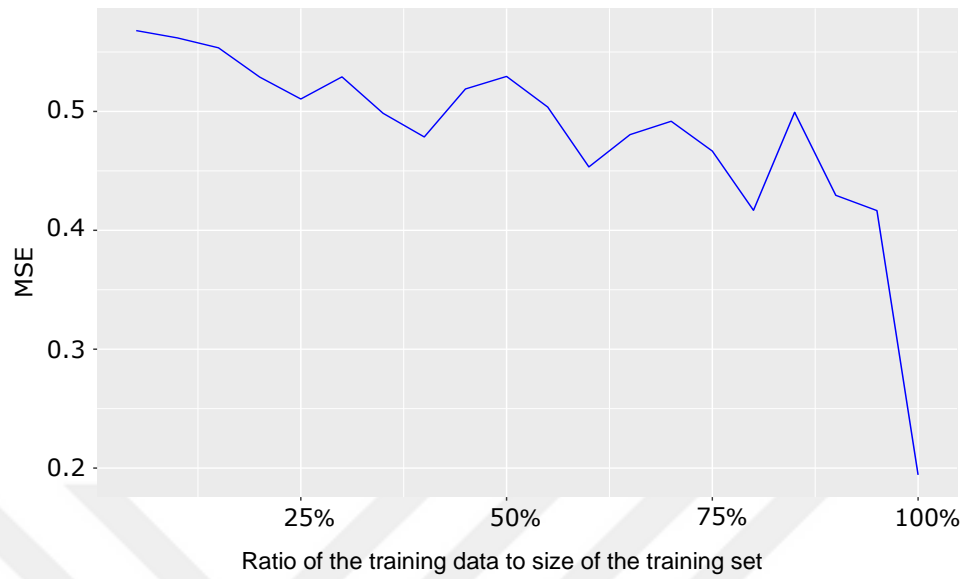
Figure 3.3 : MSE obtained from supervised incremental learning based classification experiments applied to individual test subjects.

Figure 3.4 : Accuracy obtained from supervised incremental learning based classification experiments applied to individual test subjects.

Figure 3.5 : MSE obtained from supervised incremental learning based classification experiments applied to all test subjects.
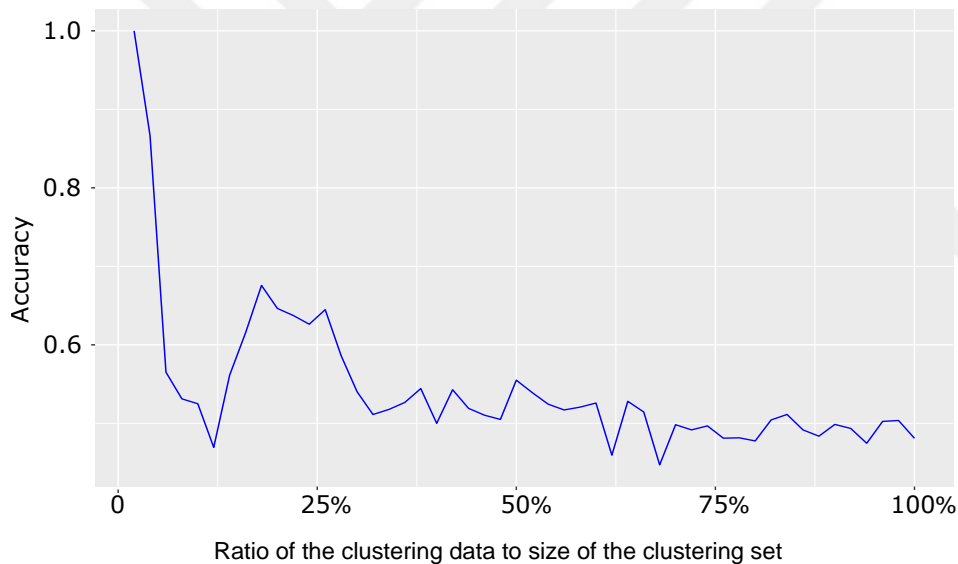

Figure 3.6 : Accuracy obtained from supervised incremental learning based classification experiments applied to all test subjects.

of the parts are selected and tested, new parts are continuously tested in incremental learning. Therefore, the success rate is slightly lower.

### 3.5.3.2 Results of the unsupervised incremental learning

Two approaches: firstly, the instantaneous data for a selected individual is divided into 3 different clusters. In the second approach, the instantaneous data for 10 subjects were combined into 3 different clusters. In both approaches, the accuracy values were calculated. To calculate the accuracy value, 3 different clusters were matched with 3 different emotional states to best overlap. Then the accuracy value of the result of the confusion matrix generated for these 3 emotional states was calculated.

Fig. 3.7 shows accuracy obtained from unsupervised incremental learning based clustering experiments applied to individual test subjects.

Figure 3.7 : Accuracy obtained from unsupervised incremental learning based clustering experiments applied to individual test subjects.

Fig. 3.8 shows obtained from unsupervised incremental learning based clustering experiments applied to all test subjects as a group. In both cases, success is similarly diminishing over time because people's emotional distinctions are different from machines. Mankind thinks that some emotions are different, but mathematically these emotions may be similar because the mathematical facts are different from the ones we have. These results clearly show the difference between human intelligence and machine intelligence. The reason for this distinction is the difference between building
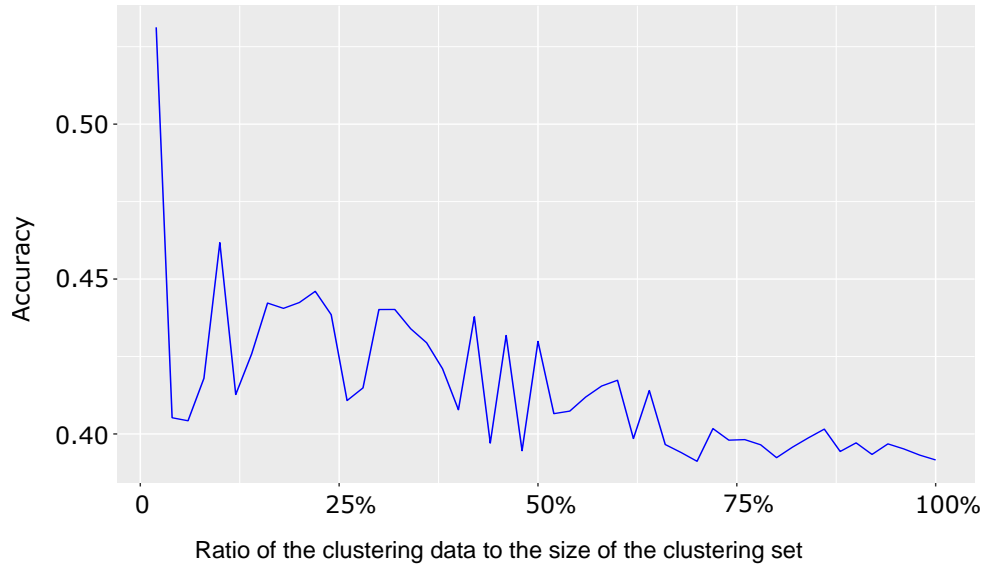
Figure 3.8 : Accuracy obtained from unsupervised incremental learning based clustering experiments applied to all test subjects.

an artificial intelligence on mathematics and building human intelligence on religions and culture.

3.6 Discussion

It is seen how important the preprocessing is for deep learning when the results are examined. For this reason, it is significant to test many different types of filters and tune the system parameters in ANN studies. It is also important that the selection of features and the reduction process are essential to producing a successful model. Because reduced models work faster and produce better results as they are free from noise. Perhaps the most important aspect of this study is that the emotions can be more accurately expressed by the change of brain signals with the sliding window approach. The buffering of the incoming signal by the sliding window approach can be performed in different ways for real-time. It is possible to use future data in the static calculation but in the real-time approach it is possible to go through a process of recognizing the emotions with past and present data.

In both online and offline experiments, the best performance has been achieved in detecting horror movies. Because the effect of fear on the frontal lobe (FP1) of the brain is distinctly different from that of funny and weepy emotions. Also offline results are as much as fourteen percent better than online results. Because it is not certain what

the input value will be in the online data. Any instantaneous brain signal can interfere to change the results. In the offline mode, the performance of the system is higher compared to the online scenerio because it is static and limited.

In these experiments, we labeled the brain signals according to the video tapes. But the reader has to take into account that not every moment of a video does have the same effect on the human brain. For this reason, no matter how well we choose the stimuli, it can also contain e.g., sorrow in fear on yield happiness in sadness. We may be experiencing a few emotions suddenly. For this reason, although our study is important for the recognition of emotions, emotions are so many and varied that they need more research.

4. CONCLUSION

BCIs are used for marketing, gaming, and entertainment to provide users with a more personalized experience. Both medical and non-medical applications require the ability to interpret the user's multimedia-induced perception and emotional experience. In this thesis, a system is designed which detects instantaneous emotional state by using EEG data obtained in real-time with a single-channel commercial BCI device. While the users watched a video consisting of film fragments and binaural beats that would trigger different emotional states, the EEG data from the users were collected. Preprocessing and feature selection were performed on this EEG data. The order of importance of the features is determined using the random forest algorithm. It has been decided which features should be in the final model by taking the order of importance of the features. In order to create the most appropriate model, the most suitable parameters were determined by using the deep learning algorithm and the grid search method for different parameters. Using an ANN-based classification scheme, a model was generated that classifies the multimedia genres such as funny, horror, weepy with an average accuracy of 87% which cause emotional or psychological experiences that are induced in viewers. Using this model, a system which detects the instant emotional state on the queuing port structure of a certain length buffer using real time EEG data is designed. Both the real-time and non-real-time classification success of the generated model is measured. In both cases, a success rate was achieved that would meet expectations. It has been shown that the system can be used for incremental learning with the supervised active learning model. In addition, unsupervised incremental learning based clustering shows the difference between human intelligence based on religions and culture, and machine intelligence based on mathematics. The created model and method will be a pioneer for such studies. With the experimentation and the increase of collected data, even the ordinary events in human life will reveal the relation with the brain signals. The system can be used for both medical and non-medical applications. This paper tackled to figure out the specific link between brain waves and the multimedia-induced emotions. When the results of the experiment are examined, it

can be seen how important each step we have made in this study is. We have identified the basic steps to be taken for deep learning in this work where we have reinstated the basics of a scientific work from the importance of preprocessing to the choice of hidden layers.

In the future, we plan to develop a real-time game based on the fears of the players which will impact all events in the game. The game will come with a horror story which sets players up to solve a mystery while overcoming their fears.

We also aim to work on a protocol to prevent the theft of brain signals via BCI applications. In this context, we are aiming to prevent data theft by spreading astonishing signals instead of restricting certain signals for BCI applications. In this regard, we plan to ensure that technology is used within ethical boundaries.

REFERENCES

[1] **Lusted, H. S., & Knapp, R. B.** (1996). Controlling computers with neural signals. Scientific American, 275(4), 82-87.

[2] **Wolpaw, J. R., & McFarland, D. J.** (2004). Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. Proceedings of the National Academy of Sciences of the United States of America, 101(51), 17849-17854.

[3] **Mann, C. A., Lubar, J. F., Zimmerman, A. W., Miller, C. A., & Muenchen, R. A.** (1992). Quantitative analysis of EEG in boys with attention-deficit-hyperactivity disorder: Controlled study with clinical implications. Pediatric neurology, 8(1), 30-36.

[4] **Szafir, D., & Mutlu, B.** (2013, April). ARTFul: adaptive review technology for flipped learning. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1001-1010).

[5] **Kim, K. H., Bang, S. W., & Kim, S. R.** (2004). Emotion recognition system using short-term monitoring of physiological signals. Medical and biological engineering and computing, 42(3), 419-427.

[6] **Picard, R. W., Vyzas, E., & Healey, J.** (2001). Toward machine emotional intelligence: Analysis of affective physiological state. IEEE transactions on pattern analysis and machine intelligence, 23(10), 1175-1191.

[7] **Boeree, G.** (2003). General psychology. Shippensburg University.

[8] **Minlebaev, M., Colonnese, M., Tsintsadze, T., Sirota, A., & Khazipov, R.** (2011). Early gamma oscillations synchronize developing thalamus and cortex. Science, 334(6053), 226-229.

[9] **Martinovic, I., Davies, D., Frank, M., Perito, D., Ros, T., & Song, D.** (2012, August). On the Feasibility of Side-Channel Attacks with Brain-Computer Interfaces. In USENIX security symposium (pp. 143-158).

[10] **Rani, B. J. A., & Umamakeswari, A.** (2015). Electroencephalogram-based brain controlled robotic wheelchair. Indian Journal of Science and Technology, 8(S9), 188-197.

[11] **Visu, P., Varunkumar, K. A., Srinivasan, R., & Kumar, R. V.** (2016). Brainwave based accident avoidance system for drowsy drivers. Indian Journal of Science and Technology, 9(3).

[12] **Kan, D. P. X., Lim, V. W. W., & Lee, P. F.** (2015). Signal Conversion from Attention Signals to Light Emitting Diodes as an Attention Level Indicator. In 1st Global Conference on Biomedical Engineering & 9th Asian-Pacific Conference on Medical and Biological Engineering (pp. 251-255).

[13] **Bonaci, T., Herron, J., Matlack, C., & Chizeck, H. J.** (2014, June). Securing the exocortex: A twenty-first century cybernetics challenge. In Norbert Wiener in the 21st Century (21CW), 2014 IEEE Conference on (pp. 1-8).

[14] **Chan, A., Early, C. E., Subedi, S., Li, Y., & Lin, H.** (2015, November). Systematic analysis of machine learning algorithms on EEG data for brain state intelligence. In Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on (pp. 793-799).

[15] **Ang, A. M. S., Zhang, Z. G., Hung, Y. S., & Mak, J. N. F.** (2015, April). A user-friendly wearable single-channel EOG-based human-computer interface for cursor control. In Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on (pp. 565-568).

[16] **Kumar, S., Kumar, V., & Gupta, B.** (2015, September). Feature extraction from EEG signal through one electrode device for medical application. In Next Generation Computing Technologies (NGCT), 2015 1st International Conference on (pp. 555-559).

[17] **Morales, J. M., Díaz-Piedra, C., Di Stasi, L. L., & Romero, S.** (2015, June). Low-cost Remote Monitoring of Biomedical Signals. In International Work-Conference on the Interplay Between Natural and Artificial Computation (pp. 288-295).

[18] **Gonzalez, V. M., Robbes, R., Gongora, G., & Medina, S.** (2015, August). Measuring Concentration While Programming with Low-Cost BCI Devices: Differences Between Debugging and Creativity Tasks. In International Conference on Augmented Cognition (pp. 605-615).

[19] **Chen, H. M., Chen, S. Y., Jheng, T. J., & Chang, S. C.** (2015). Design of a Mobile Brain-Computer Interface System with Personalized Emotional Feedback. In Future Information Technology-II (pp. 87-95).

[20] **Singh, A. K., Wang, Y. K., King, J. T., Lin, C. T., & Ko, L. W.** (2015, November). A simple communication system based on Brain Computer Interface. In Technologies and Applications of Artificial Intelligence (TAAI), 2015 Conference on (pp. 363-366).

[21] **Perhakaran, G., Yusof, A. M., Rusli, M. E., Yusoff, M. Z. M., Mahalil, I., & Zainuddin, A. R. R.** (2016). A Study of Meditation Effectiveness for Virtual Reality Based Stress Therapy Using EEG Measurement and Questionnaire Approaches. In Innovation in Medicine and Healthcare 2015 (pp. 365-373).

[22] **Vijayaragavan, G. R., Raghav, R. L., Phani, K. P., & Vaidyanathan, V.** (2015, October). EEG monitored mind de-stressing smart phone application using Yoga and Music Therapy. In Green Computing and Internet of Things (ICGCIoT), 2015 International Conference on (pp. 412-415).

[23] **Schaaff, K., & Schultz, T.** (2009, September). Towards an EEG-based emotion recognizer for humanoid robots. In Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on (pp. 792-796).

[24] **Nie, D., Wang, X. W., Shi, L. C., & Lu, B. L.** (2011, April). EEG-based emotion recognition during watching movies. In Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on (pp. 667-670).

[25] **Lin, Y. P., Wang, C. H., Wu, T. L., Jeng, S. K., & Chen, J. H.** (2007, October). Multilayer perceptron for EEG signal classification during listening to emotional music. In TENCON 2007-2007 IEEE Region 10 Conference (pp. 1-3).

[26] **Kortelainen, J., Väyrynen, E., & Seppänen, T.** (2015). High-frequency electroencephalographic activity in left temporal area is associated with pleasant emotion induced by video clips. Computational intelligence and neuroscience, 2015, 31.

[27] **Min, Y. K., Chung, S. C., & Min, B. C.** (2005). Physiological evaluation on emotional change induced by imagination. Applied psychophysiology and biofeedback, 30(2), 137-150.

[28] **Li, M., & Lu, B. L.** (2009, September). Emotion classification based on gamma-band EEG. In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 1223-1226).

[29] **Choppin, A.** (2000). EEG-based human interface for disabled individuals: Emotion expression with neural networks.

[30] **Jalilifard, A., Pizzolato, E. B., & Islam, M. K.** (2016, October). Emotion classification using single-channel scalp-EEG recording. In Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the (pp. 845-849).

[31] **Liu, Y., Sourina, O., & Nguyen, M. K.** (2010, October). Real-time EEG-based human emotion recognition and visualization. In Cyberworlds (CW), 2010 International Conference on (pp. 262-269).

[32] **Klem, G. H., Lüders, H. O., Jasper, H. H., & Elger, C.** (1999). The ten-twenty electrode system of the International Federation. Electroencephalogr Clin Neurophysiol, 52(3).

[33] **NeuroSky, I. A.** (2010). ThinkGear Socket Protocol (Tech. Rep.).

[34] **Alpaydin, E.** (2014). Introduction to machine learning. MIT press.

[35] **Quinlan, J. R.** (1993). Constructing decision tree. C4, 5, 17-26.

[36] **Breiman, L.** (2001). Random forests. Machine learning, 45(1), 5-32.

[37] **Gepperth, A., & Hammer, B.** (2016). Incremental learning algorithms and applications. In European Symposium on Artificial Neural Networks (ESANN).

[38] **MindWave.** (n.d.). Retrieved December 31, 2016, from http://store.neurosky.com/pages/mindwave

[39] **Studio, R.** (2012). RStudio: integrated development environment for R. RStudio Inc, Boston, Massachusetts.

[40] **Aiello, S., Kraljevic, T., & Maj, P.** (2015). Package 'h2o'. dim, 2, 12.

[41] **Therneau, T. M., Atkinson, B., & Ripley, M. B.** (2010). The rpart package.

[42] **Kasprzak, C.** (2011). Influence of binaural beats on EEG signal. Acta physica polonica A, 119(6A), 986-990.

APPENDICES

APPENDIX A.1 : Pictures from the visual and auditory stimuli
APPENDIX A.2 : Screenshots from the real-time emotion recognition system

APPENDIX A.1

Fig. A.1 shows a picture from experiment while subjects are listening the binaural beats (8 Hz and alpha).

Figure A.1 : A picture for binaural beats (8 Hz and alpha).

Fig. A.2 shows a picture from experiment while subjects are watching the funny animation.

Figure A.2 : A picture from the funny animation.

Fig. A.3 shows a picture from experiment while subjects are listening the binaural beats (beta).

Figure A.3 : A picture for binaural beats (beta).

Fig. A.4 shows a picture from experiment while subjects are listening the binaural beats (theta).

Figure A.4 : A picture for binaural beats (theta).

Fig. A.5 shows a picture from experiment while subjects are watching the weepy movie.

Figure A.5 : A picture from the weepy movie.

Fig. A.6 shows a picture from experiment while subjects are listening the binaural beats (gamma).
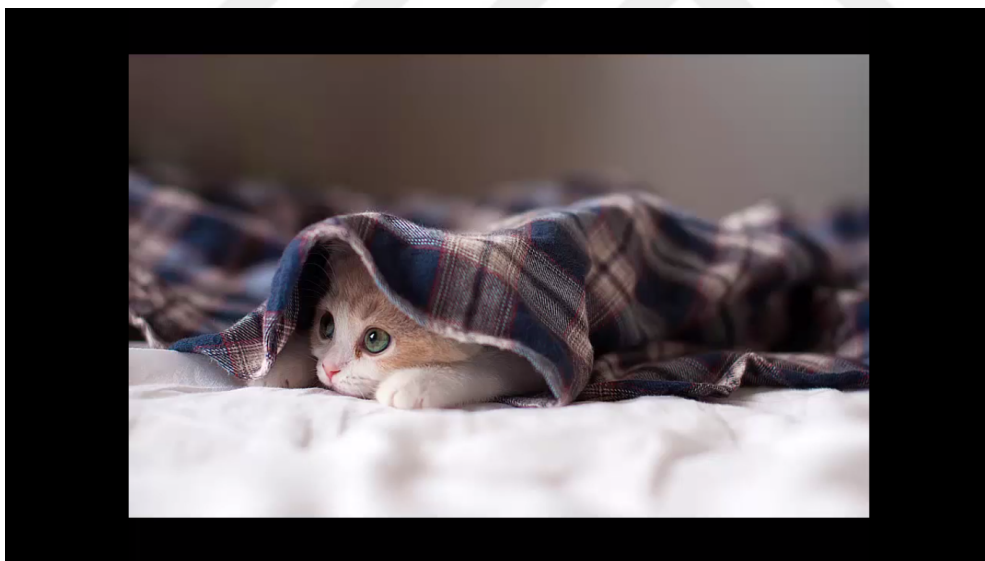

Figure A.6 : A picture for binaural beats (gamma).

Fig. A.7 shows a picture from experiment while subjects are watching the horror movie.

Figure A.7 : A picture from the horror movie.

Fig. A.8 shows a picture from experiment while subjects are listening the binaural beats (delta).
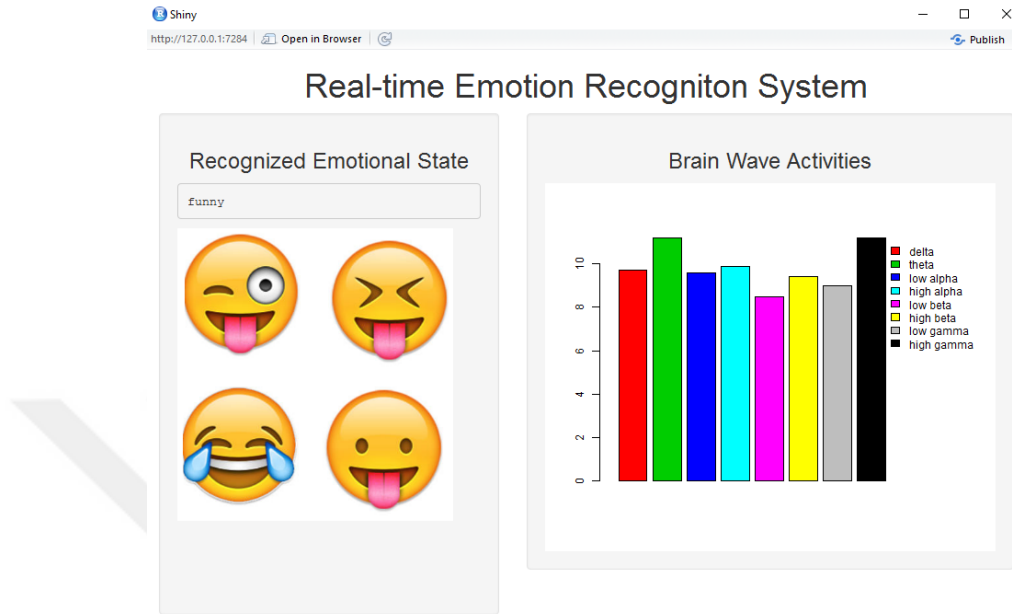

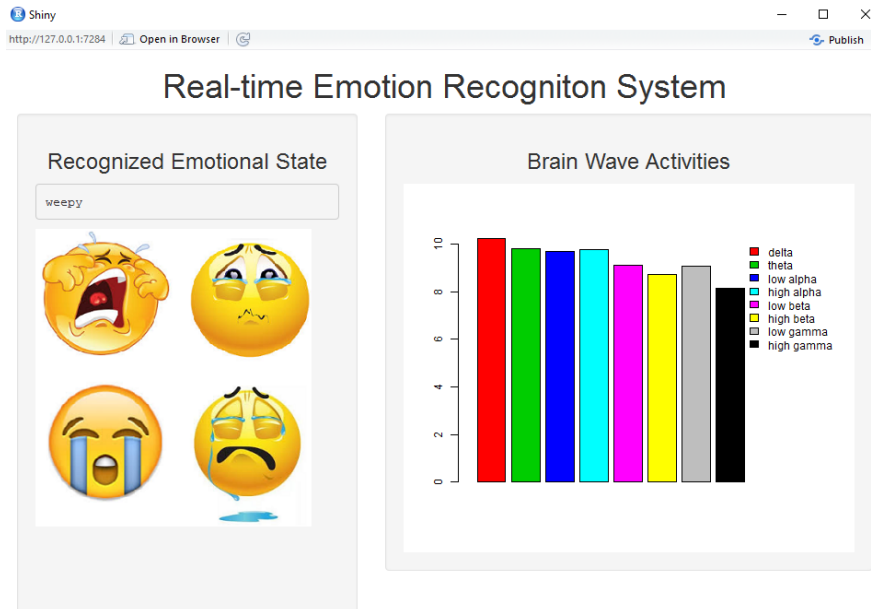**Figure A.8** : A picture for binaural beats (delta).

APPENDIX A.2

Fig. A.9 shows a screenshot for funny state from experiment while subjects are watching the funny animation.
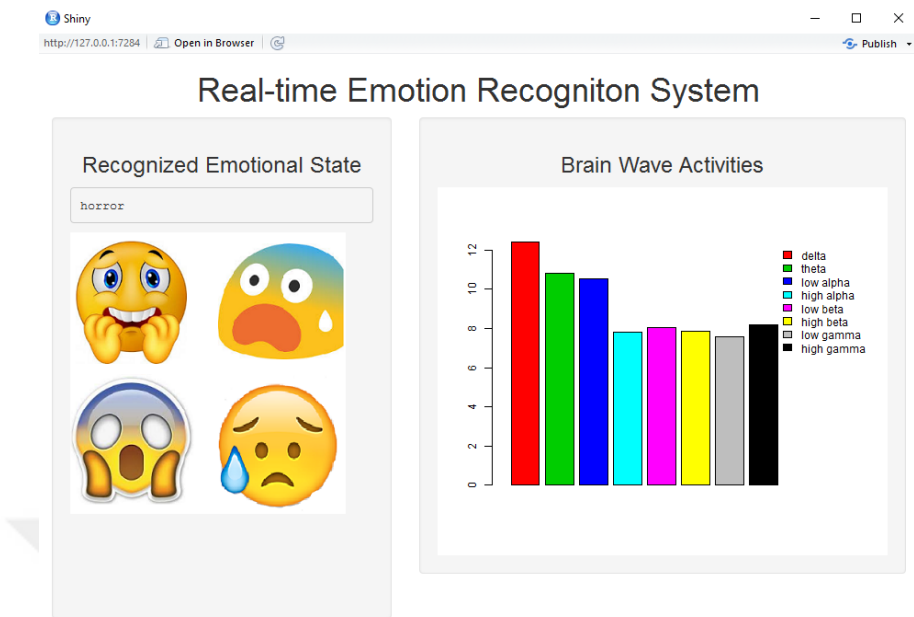

Figure A.9 : A screenshot from the funny state.

Fig. A.10 shows a screenshot for weepy state from experiment while subjects are watching the weepy movie.

Figure A.10 : A screenshot from the weepy state.

Fig. A.11 shows a screenshot for horror state from experiment while subjects are watching the horror movie.

Figure A.11 : A screenshot from the horror state.

CURRICULUM VITAE

Name Surname:
Mehmet Ali SARIKAYA

Place and Date of Birth:
07.01.1991

Adress:
Halıcıoğlu mh. Kalfayamağı sk. No:10 Daire:6 Beyoğlu İSTANBUL

E-Mail:
sarikayameh@itu.edu.tr

HSDIP:
Gaziantep Science High School, 10.2005 - 06.2009

B.Sc.:
Istanbul Technical University, 10.2010 - 02.2014

M.Sc.:
Istanbul Technical University, 09.2015 – 06.2017

Professional Experience and Rewards:
Researcher, Scientific and Technological Research Council of Turkey (TUBITAK), 07.2013 - 09.2016

PUBLICATIONS/PRESENTATIONS ON THE THESIS
M. A. Sarıkaya, G. Ince: Emotion Recognition from EEG Signals through One Electrode Device, Proceedings of the 25th Signal Processing and Communications Applications Conference (SIU 2017), to appear