Non-oddball ERP Paradigms with Joint Temporal-Frequency Learning in Convolutional Neural Network

by

Madina Saparbayeva

Submitted to the Department of Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Computer Science

at the

NAZARBAYEV UNIVERSITY

April 2022

Author . . . .
Department of Computer Science
April 27, 2022

Certified by. . . . .
Min-Ho Lee
Assistant Professor of Computer Science
Thesis Supervisor

Certified by. . . . .
Berdakh Abibullaev
Assistant Professor of Robotics
Thesis Co-supervisor

Accepted by . . . .
Vassilios D. Tourassis
Dean, School of Engineering and Digital Sciences

Non-oddball ERP Paradigms with Joint Temporal-Frequency Learning in Convolutional Neural Network

by

Madina Saparbayeva

Submitted to the Department of Computer Science
on Apr 27, 2022, in partial fulfillment of the
requirements for the degree of
Master of Science in Computer Science

Abstract

Brain-Computer Interface (BCI) helps people who have a severe disease to interact with external devices. Event-Related Potential (ERP) based BCI systems send stimuli, then detect brain signals responding to stimuli. Stimuli play loud or flashes with high intensity to distinguish the target brain signals in most researches. Thus, typical oddball-paradigm causes psychological and psychical discomforts. This thesis proposed non-oddball BCI paradigms which send stimuli which has almost zero volume or intensity. Users tend to produce voluntary mental task during experiment to cover typical stimuli's brain signal, thus it compensates for the loss of accuracy of a result of the reduced stimulus. As an outcome, a specific mental task was used to investigate task-relevant endogenous components, and system performance was significantly enhanced. The proposed Convolutional Neural Network (CNN) approach has a decoding accuracy more than 90% for both the non-oddball visual and auditory paradigms, respectively, outperforming the linear classifier model noticeably. These discoveries offer new opportunities for pragmatic ERP systems, potentially improving the usability of current brain-computer interfaces remarkably.

Thesis Supervisor: Min-Ho Lee
Title: Assistant Professor of Computer Science

Thesis Co-supervisor: Berdakh Abibullaev
Title: Assistant Professor of Robotics

Acknowledgments

I would first like to thank my thesis advisor **Assistant Professor Min-Ho Lee** of the Computer Science, School of Engineering and Digital Sciences at Nazarbayev University, for patience, guidance, and support. I am grateful that he took me on as a student and research assistant.

Thank to my colleague and best friend, **Madina Kudaibergenova**, for her support and for being with me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A brain-computer interface (BCI) detects the brain signals of subject by special equipment to control the computer which helps people with various diseases. The program of BCI shows the subject intention while his brain produces strong potentials which can be detected by electroencephalography (EEG). The most common used type is EEG because of its invasive and low-cost quality. Additionally, EEG is user-friendly for usage for users and developers [1].

The brain produce the potentials after given stimulus and the highest positive point called P300 in range of 300 and 400 milliseconds. The P300 is used in the oddball paradigm which classifies target and non-target stimuli. The target and non-target stimulus flashes at different frequencies where first stimulus appears much less than second stimulus. The conventional ratio of stimulus's ration is 20 - 80 [2]. The more convenient signals elicit by showing the target stimuli repeatedly and randomly to improve signal-to-noise ratio (SNR) [3].

The signal feature is distinguished characteristics of the brain signals which produces with user's intention. Human and computer classify the target or non-target stimuli by feature extraction. The process of feature extraction is meaningful and differentiates signal from noises. Event-Related Potential (ERP) or P300 of ERP depends on external stimulus which showed higher responses in the oddball paradigm [4, 5, 6]. These study focused on external stimuli to increase the brain responses by changing the characteristics of stimulus. For instance, authors applied volume, spatial arrangement, modulated frequency in the auditory BCIs, and shape, color [6, 7], intensity, or highly recognizable images [8, 9] in the visual BCIs. Therefore, ERP-based BCI system can be improved by the changing the characteristics of target stimuli [6, 7, 8]. On other words, the authors suggest to increase auditory or visual stimuli to obtain higher responses of subject's brain. The limits of this study is subjects was feeling mentally and physical unwell condition i.e stress, fatigue.

We hypothesized that user intention can generate strong potentials in certain BCI tasks. The main idea is proposed non-oddball cues can be detected in the endogenous features. Thus, user intent to produce high the brain responses which help program to detect the endogenous potentials which is almost similar the conventional oddball paradigm. Proposed non-oddball paradigm carried out visual and auditory binary system. Before the experiment, subjects were instructed to memorize the type of sound which played in high-pitched and high-volume type condition. During the training experiment, subjects were listening the sound cue, but with lower volume, and perform mental task in the first experiment and gaze non-impact visual stimuli in second experiment. Then in the testing experiment, subjects tried intent mental task with memorized sound ten times in one trial. In each experiment, there were three mental tasks: non-intention (NI), passive concentration (PC), and active concentration (AC). These mental tasks used for investigate the neuro-physiological differences on ERP responses in non-oddball paradigm.

This investigation advances a high-performance Convsolutional Neural Network (CNN) model for visual and auditory non-oddball paradigm. In online experiment, CNN provide the output of mental task after one or more mental task's attempts. Averaging received raw brain signal minimizes the noises and get the pattern of P300 signals. Alongside with CNN, Linear Discriminant Analysis (LDA) was tested, which is mostly used linear classifier in BCI papers, to examine the differences of classifying accuracy and precise output of mental task. Our results investigated a late positive potential (LPP) by the active mental task and sufficiently decodes the user's intention in the non-oddball paradigms. The maximum decoding accuracy of the non-oddball visual and auditory systems were more than 90% using EEGNet that significantly outperformed the linear classifier model.

Chapter 2

Literature Review

2.1 Main Focus Group in BCI

The BCI users can have an individual approaches. F.Nijbour have conducted a surveys where the professional specialists were participated and summarized their answers. The potential user groups have concurrent physical, sensory and cognitive problems beside physical disability. The problem is some BCI technologies require the steady position to acquire signals, but users who have ticks or another problem with nerves could not complete experiments. The author stated this approach will be fitted for people in locked-in-state (ALS, MS or SMA) and people with high SCI. The another group is people with Duchenne disease, CP or Rett syndrome. They have condition that could have problem with visual BCI techniques that can be lead to epileptic seizures or spasms. The overall groups that author resulted were: late-stage amyotrophic lateral sclerosis (ALS), late-stage multiple sclerosis, spinal muscular atrophy type II (SMA type II), Duchenne disease, Rett syndrome, cerebral palsy (CP), high spinal cord injury (SCI) and the locked-in sybdrome (LIS).

Amyotrophic lateral sclerosis or ALS is neurological disease that affects on nerve cells in the brain and spinal cord. The disease can lead to the loss of the muscle control. The main damages occurs on motor neurons which is **responded** to moving, communication skills, feeding and breathing functions.

Multiple sclerosis or MS is chronic disease with no medical treatment. The finalstage MS users can be faced with trouble with balance, coordination, posture, limited mobility, muscle spasms, breathing problem, speech problem, vision problem, cognitive difficulties etc.

Spinal muscular atrophy type II or SMA2 is a genetic neural disease that affects on the nerve cells of motor muscles. The provided list is the part of group who needs BCI, but it can be used by any people in daily life. BCI system helps to ease people life and provide more functions.

2.2 Previous Work

This thesis is the extended version of the conference paper [10] titled 'A Novel Binary BCI Systems Based on Non-oddball Auditory and Visual Paradigms'. The paper proposed novel binary BCI system that minimizes the external stimulus in both auditory and visual paradigm. Users barely hear the sound or visual stimulus. Three mental tasks defined: 1) non-intension (NI), 2) passive concentration (PC), and 3) active concentration (AC). There were used LDA and CNN for decoding algorithms. The decoding results for LDA were 78% visual cue-AC, 74% visual cue-PC, 78% auditory-AC, 56% auditory-PC, whereas CNN showed 78% for both active visual and auditory paradigm's accuracy. Auditory-PC and Visual-PC accuracy were 85.7 % and 87.5 % respectively.

2.2.1 User Experience in BCI

BCI requires the computer training as same as user training [11]. The user must learn produce stable brain signals for easily distinguish the target and non-target signals. The distinguishable signals accelerate the computer decoding the commands. Consequently, clear raw signals are high decoding accuracy, which lead to low cost of source McFarland and Wolpaw [12]. Neuper and Pfurtscheller [13] showed naive users cannot control the brain signals after some training.

Lea Pillette [14] indicates the three criteria for analyzing the performance, which are MI-BCI performance, User Experience, Experimenters' and participants' profile.

These criteria made from literature which focused on various mental state of participants. Hammer et al. [15] stated the user's states as motivation and attention can influence on decoding accuracy. They tried users to keep motivation and attention by giving non-alcohol drinking, sweet and period of rest. Wheeler and Petty [16] focused on psychosocial factors, which as example showed elderly people react more slowly or memory disfunction. Subsequently, Lea Pillette [14] took account characteristics as gender, age, ethnicity, professional status and/or behaviour as gaze, touch, verbal interaction Rosenthal [17]. They concluded social presence and emotional feedback are meant to increase the effort, motivation and engagement of the participants throughout the learning.

Alvaro Fernández-Rodríguez [18] analyzed the impact of background sound in auditory-based ERP-BCI. They stated users show the high cognitive workload from shared attention in auditory stimuli. Their study indicated the background noises or speech are important interference and should be avoided in experiment. However, the P300 amplitude showed high signals while user tried to focus on the target and ignore the background speech.

2.2.2 Decoding Models of BCI signals

During the last decade, many pieces of research on BCI have raised. Although the quality is not on the level which can be sold to society, that is why the work continued. The current works limited in the way of control because the works offer one or two mental actions to users. Moreover, the commands take more time to check the action to perform. It is not as natural as needed. In that way, the users prefer the natural way without BCI. The current goal is to update the algorithm to be natural and reliable. Min-Ho Lee et al. [6] shows the various methods to take the brain signals. Motor Imaginary (MI), Event-related potential (ERP) and Steady-State Visually Evoked Potential (SSVEP) tested on 54 participants. The authors checked the state of participants during using BCI methods. The distinctive feature is in SSVEP make eye fatigue more than MI and ERP, which should be not used in a home environment.

Lately, ERP studies [19, 20, 21] focused on various models of CNN, which showed conventional results. The authors showed succeeded result in various EEG used BCI tasks, such as emotion recognition [22], P300 detection [23], and memory performance prediction [24]. Eduardo Santamaría-Vázquez et al. [19] did comparative research on CNN models which listed as rLDA [25], xDAWN + RG [26], CNN-BLSTM [27], DeepConvNet [28], EEGNET [29]. These models classified ERP-based speller and proffer feature extraction at different temporal scales.

Saim Rasheed [30] did a review on machine learning models comparison in BCI field. The table 2.1 was created towards similarity of BCI task and classification method. The common classification methods, which had >80% accuracy level, were Linear Regression (LR), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM). For feature extraction were taken Principal Component Analysis (PCA), Riemannian geometry, Common Spatial Pattern (CSP), Particle Swarm Optimization (PSO), Wavelet Transform (WT), Fast Fourier Transform (FFT).

This thesis assayed the most convenient classification model on non-oddball paradigm. For increasing the classification accuracy, data are augmented due to small difference between target and non-target brain signals.

2.3 Current Limitations and Possible Solutions

The BCI have many limitations of usage. The current state of BCI is using enough environment where contribute subjects or devices needed for work for successful completion. For that reason, accuracy and performance can be much low than expected in various environment. Also, BCI users have to sit while using it, because the other motion can interfere with the main action. Consequently, the result can be wrong and the action might be dangerous for users in some cases. User Experience is important in the case of new technologies. In some times the wrong designed devices can lead to extremely dangerous consequences. Andéol Évain et al. [7] takes into

Reference	Year	Feature Extraction	Classification Method	Accuracy Level	BCI Task
[31]	2020	-	LR	95%	EEG signal categorization
[32]	2012	-	SVM	90.55%	RP signal categorization
[33]	2018	PCA	SVM	92.50%	BCI therapy stage classification
[34]	2019	Riemannian geometry, CSP and PSO	CNN	80.44%	EEG signal categorization
[35]	2018	WT	SVM	>90%	EEG signal categorization
[36]	2018	FFT	LDA, KNN, SVM	95%, 100%, 100%	EEG signal categorization
[23]	2010	-	CNN	95.50%	ERP signal classification
[37]	2016	-	SVM	92.50%	P300-based BCI operation
[38]	2020	-	LSTM	97.13%	EEG signal categorization
[39]	2006	PCA	SVM	>95%	EEG signal categorization

Table 2.1: Comparison of Feature Extraction and Classification Techniques.

account the importance of usability and experimented mental state of users such as frustration, motivation and fatigue while using fake SSVEP-based BCI application which shows fake feedback. The authors highlighted that a high error rate increases the user's frustration over time. The suggestion is favouring user comfort over the small improvement of accuracy.

The BCI accuracy depends on quality and quantity of acquired data, thus the incorrect data lead to low performance of the model. Nagasawa Tomoyuki, et al. [40] improved the fNIRS-BCI accuracy by augmenting the fNIRS data using generative models. Generative adversarial network (GAN) generates new data based on real dataset to overcome the lack of labeled classes, widely used for image-generation field [41]. GAN acts as a game: a discriminative model that learns to determine whether a sample true or fake. The generative model improves each next sample to win discriminative model that try to detect the generated data. This approach lead to precise new data. Competition between two models improves their methods until the fake data are indistinguishable from the real data [41].

Alvaro Fernández-Rodríguez suggested avoiding the negative effect of workload as background noises in auditory-based BCI. Alternative solutions are take account the background noises in bypassing of the feature extraction or integrate it into BCI as stimulus with external audio.

Chapter 3

Methodology

3.1 Participants and Data Acquisition

In total, fourteen healthy people took part in the study. Five participants were completely new to BCI tasks, while the other nine had a basic understanding of the technology. All of the participants said they were free of psychiatric or neurological illnesses and that their vision was normal or corrected. The environment of the experiment was equipped with a chair and monitor. The monitor was 19 inch LCD monitor (60 Hz refresh rate, 1280 x 1024 resolution). The aim of experiment and the process of gathering data were explained to the subjects.

The goal, technique, and entire experiment process were all disclosed to participants before to the experiment. Participants signed a contract for the use of their personal data.

Before experiment, the questionnaire about health condition was filled by each participant. During experiment, the participants were seated on chair with a vacuum earphone to hear sound with minimum background noises. One or two days were taken for one participant on training and testing states. After completing one trial of experiment, the participant was required to fill in the questionnaire about the experiences and health condition. The questionnaire was aimed to identify weak moments in BCI application.

EEG activity was recorded using a 32-channels ActiCap EEG amplifier (Fp1-2, F3-4, Fz, F7-8, FC5-6, FC1-2, T7-8, C3-4, Cz, CP1-2, CP5-6, TP9-10, P3-4, P7-8, Pz, PO9-10, O1-2, and Oz) (Brain Products, Munich, Germany). The electrodes utilised were Ag/AgCl electrodes that followed the international 10-20 system. With a forehead ground and an impedance of 10k Ohm or less, the reference is positioned on the nose. Before the arrival of the participants, the apparatus and cap with channels were arranged in advance. The sampling rate was 1000 Hz, and the DC artefact was removed with a 60 Hz notch filter. A 5th order Butterworth filter was used to band-pass filter the EEG signal between 0.5 and 30 Hz.

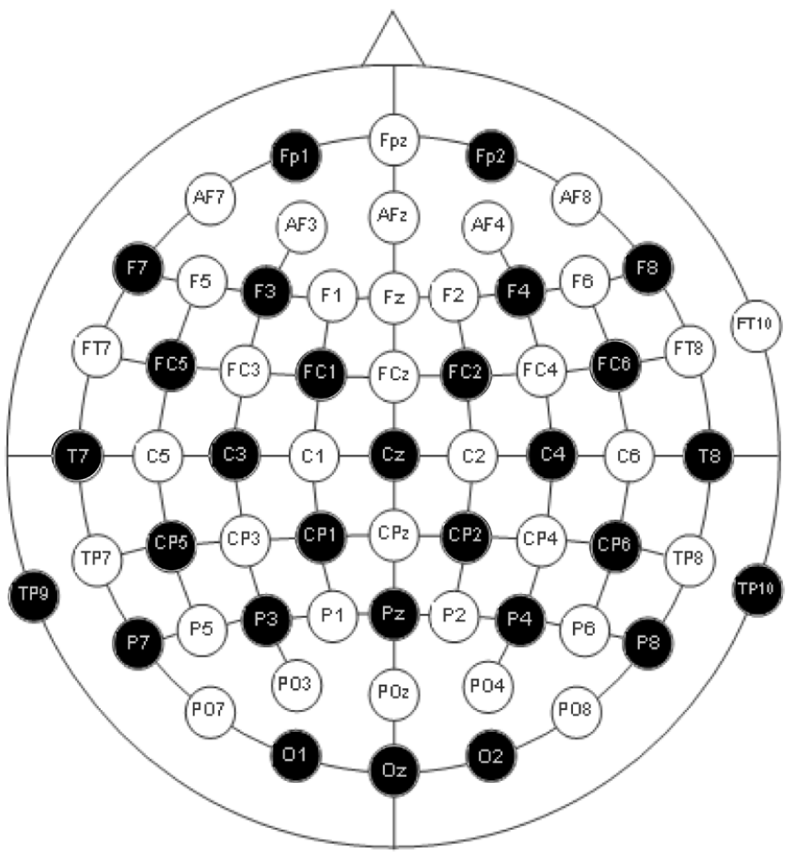

Figure 3-1: **EEG channels.** Used 32 channel ActiCap EEG amplifier (Fp1-2, F3-4, Fz, F7-8, FC5-6, FC1-2, T7-8, C3-4, Cz, CP1-2, CP5-6, TP9-10, P3-4, P7-8, Pz, PO9-10, O1-2, and Oz)

3.1.1 Ethical Issues in Experiment

There are several ethical issues were held in the experiment. The all participants were voluntary to this work. Participants are able to leave experiment anytime. They are informed about consent, therefore they understand the purpose, benefits, the experiment's details, consumed time. Confidentiality of participant's information is strongly hidden and cannot be shared to anyone without participant's agreement. There are no potential harm for participant during experiment, i.e., physical, social, psychological and all other types of harm are minimized.

3.2 Experimental Paradigms and Task Definition

The experiments concentrated on non-oddball visual-cue and auditory-cue conditions. Thus, the goal is removing the oddball characteristics so that the endogenous potentials could be fully derived. As a result, the visual and audio stimuli were provided to the individual with a same interval and were entirely predictable.

A training phase and a testing phase were both included in the experiment. The classifier parameters were estimated using all of the trials from the training phase. The decoding accuracy was validated using the testing data.

Three mental tasks were defined: 1) non-intention (NI), 2) passive concentration (PC), and 3) active concentration (AC). The number of trials in each of the three tasks was evenly distributed. The participant was provided with an identical stimulus in three tasks, but different mental tasks. Participants in the NI condition concentrated on the stimuli without any purpose, i.e. they were in a resting state. On the contrary, the subjects were instructed to attend to the target selection by passively or actively focusing on the stimulus in the PC and the AC conditions. The PC condition required the subject to simply concentrate on a given stimulus which has been the normal instruction in previous ERP-based BCI studies. In the AC condition, subjects were instructed to perform a sound imagery task. Prior to the experiment, the participants were exposed to beep noises at a frequency of 8000 Hz for roughly 1 minute in order to memorise the tone of the sound repetitively in the AC condition.

All experimental paradigms were created using Psychophysics Toolbox and OpenBMI (MathWorks; MA, USA). The Korea University Institutional Review Board examined and approved this study [1040548-KUIRB-16-159-A-2], and all individuals gave written informed permission before participating in the tests.

3.3 Non-oddball Visual Cue Experiment

We designed a bar-stimulus that continuously moves along its x-axis in one direction (right to left). A fixed cross-symbol was located in the center. In the training phase, individual bars had different colors according to their experimental condition: gray, blue, and red colors referred to NI, PC, and AC, respectively. Participants were instructed to fix their eyes on the cross-symbol and perform the designated tasks when the bar-stimulus exactly overlapped the fixed cross-symbol (see Figure 1a). The individual bars were equally spaced with an ISI of 1 s, i.e., the subject performed a certain task every second. Sufficient resting periods were given to the subjects after every 50 trials. At the end of the training phase, a total of 1080 trials comprised of 360 trials for each NI, PC and AC were collected. Two binary classifiers were then constructed: NI vs. PC and IG vs. AC.

In the test phase, ten bars with a gray-color were consecutively presented in a single attempt. Subjects were instructed to perform a specific task ten times designated by the given voice cue 5s before the visual stimuli presentation (see Figure 1b). For instance, subjects performed sound-imagery task ten times every second by following the visual stimuli when the sound-imagery cue was given. Contrarily, subjects unintentionally gazed at the visual stimulus when given the ignoring-cue. EEG data were acquired in real-time and fed to the classifiers. The classification result was given to the subject as feedback, therefore subjects see what specific action was selected by subject, since in the training phase system evaluates brain signals without the feedback. Subjects performed 10 attempts in each class, so 300 trials (10 attempts $\times$ 10 sequences $\times$ 3 classes) were collected in the test phase.

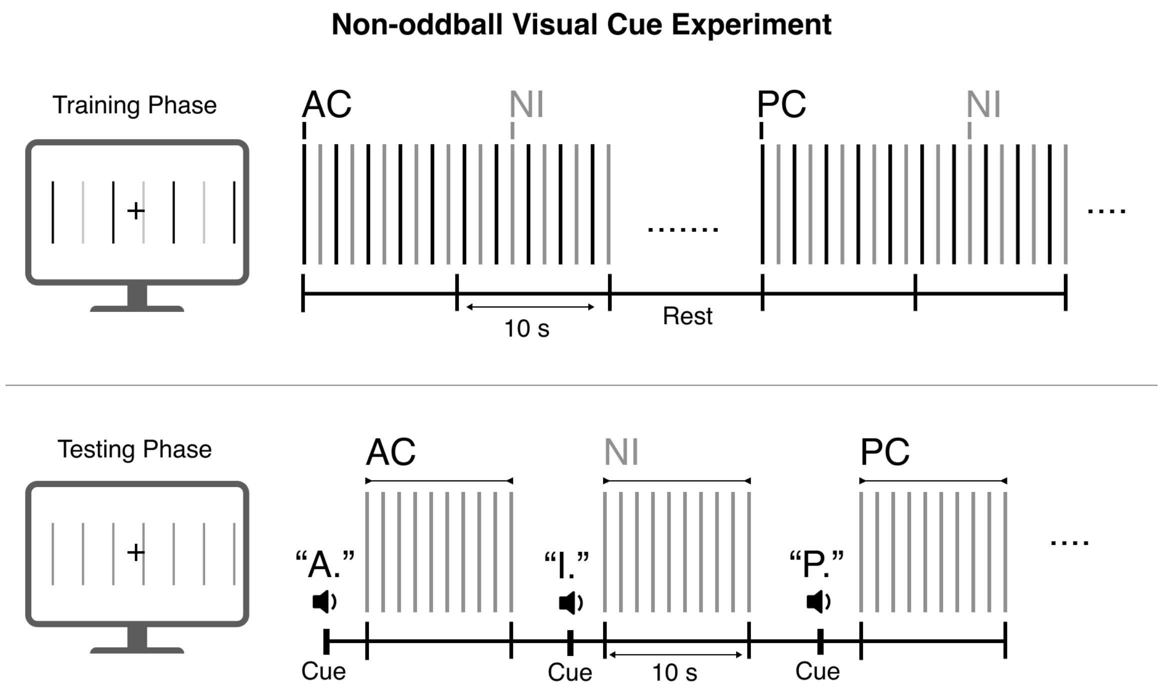

Figure 3-2: **Illustration of visual-cue experiment implemented in study:** training and testing phases.

3.4 Non-oddball Auditory Cue Experiment

In a visually blindfolded context, a non-oddball auditory experiment was conducted to evaluate ERP responses. The participant closed their eyes and sat in an armchair with armrests. The individuals had previously been exposed to an 8000 Hz high-pitched beep-type auditory stimuli. The volume of auditory stimulation was so low during the experiment that the individual could hardly hear it. The purpose of this condition is to recognise the time of the supplied auditory stimuli (i.e timing).

During the testing phase, the individual closed their eyes and completed 10 tasks as directed by a voice signal presented 5 seconds before the audio stimuli. In the testing phase, the participants made 10 tries in each class, for a total of 200 trials (10 attempts *times* 10 sequence *times* 2 class).

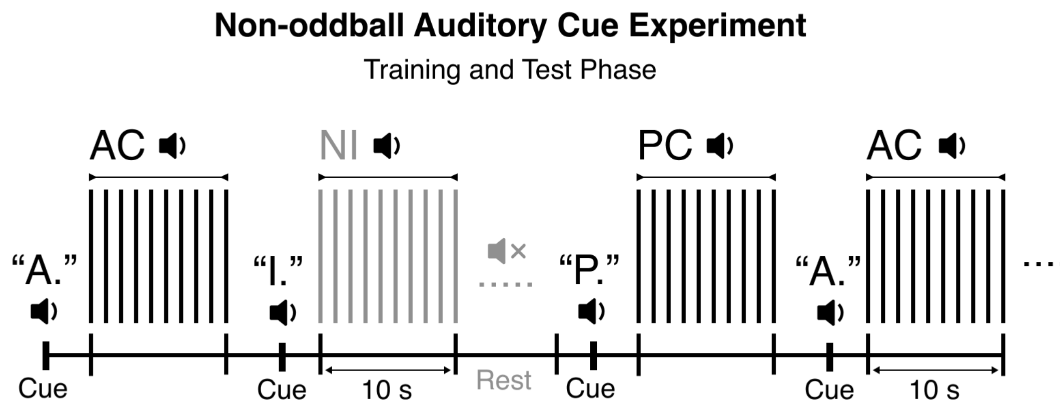

Figure 3-3: **Illustration of auditory-cue experiment implemented in study:** training and testing phases. Illustrated "A." stand for "Active", "I." - "Ignore", and "P." - "Passive" accordingly which are instructions from program to user.

3.5 Data Analysis and Performance Evaluations

Fourteen subjects' data were collected. Each EEG signal was typically down-sampled to 100 Hz. The EEG data was then segmented into individual trials ranging from -200 to 1000 ms before the stimulus start. The mean amplitudes in the -100 to 0 ms pre-stimulus period were subtracted to adjust for baseline.

All trials in the training and test phases were concatenated along with each condition to visually investigate the ERP responses for NI, PC, and AC tasks. Grand averaged ERP patterns and signed r-squared value [25] was applied to evaluate the temporal and statistical differences in the ERP patterns across all the subjects.

The standard linear model [7, 8, 42, 43] and the suggested CNN approach were built for performance evaluation based on the training data. During the test phase, certain target stimuli were presented 10 times (i.e., ten sequences), which is the conventional way for getting a valid result in ERP studies by averaging the current trials with the prior trials [43, 44, 45]. As an outcome, decoding accuracy in individual sequences ranging from one to ten was determined. The average of epochs accumulated throughout all 10 sequences, for example, was used to measure decoding accuracy at the tenth sequence.

A set of single-trial EEG is defined with $\mathrm{X} = \{\mathrm{x}_n\}_{n=1}^{N}$, $x \in R^{T \times D}$ (N - the total trial number, T - time-series data samples, D - channel numbers). $\Omega \in \{1, 2\}$ and ω_n define a set of class labels corresponding to target (AC) and non-target (NI) classes.

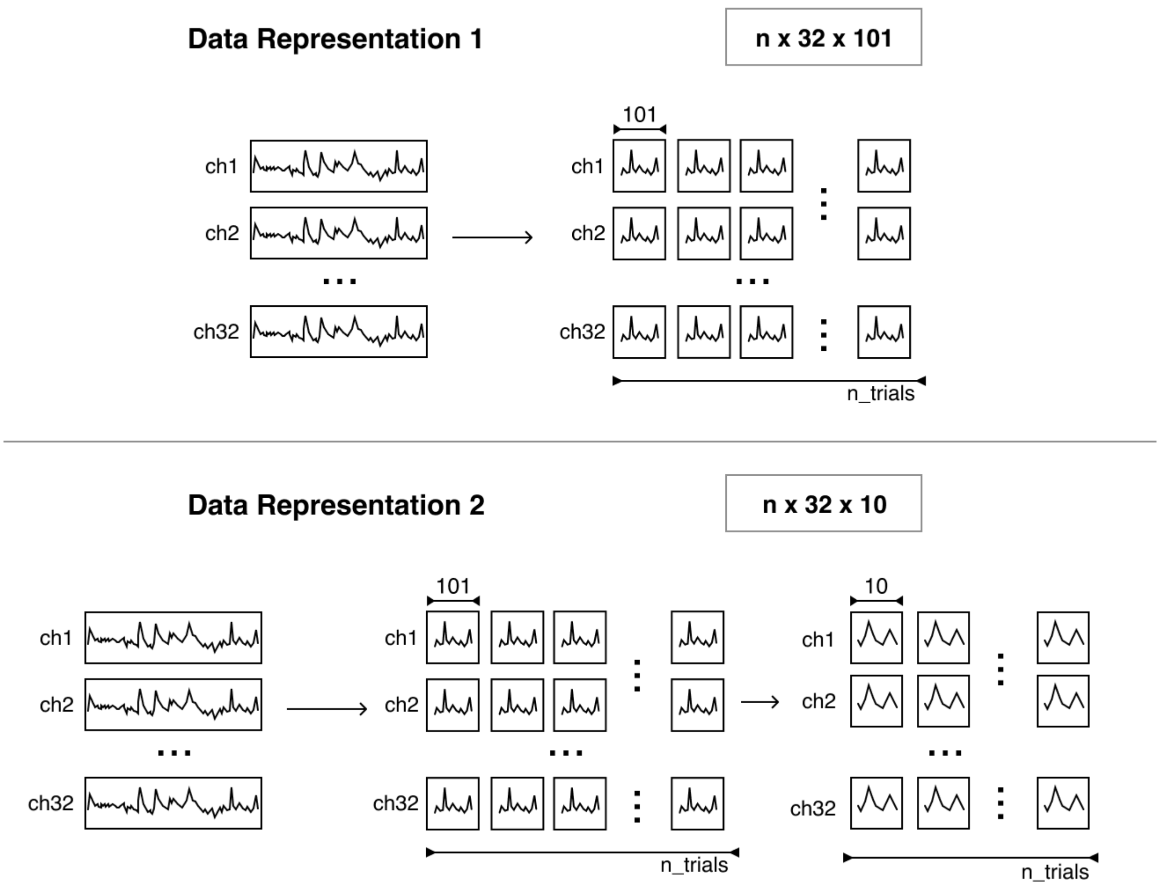

Figure 3-4: **Data Representation.** The model was trained in two cases. First model is trained to detect in each cue sends in first data representation, whereas second model is trained to detect each trials. (N - number of trials, 32 - channels, 101 - attempts, 10 - mean of attempt)

Data was presented in two cases. First case represents the common method of training data. Data is in initial condition which segmented by cue. Figure 3-4 indicates the process of data modification. The (n x 32 x 101) stands for N number of trials , D channels, T time-series data samples (1 second each). This method is applied for decoding each second attempt for evaluating immediately mental task.

The second case is adding averaging the time-series data samples into one sample. $\mathrm{X} = \{\mathrm{x}_n\}_{n=1}^{N}$, $\mathrm{x} \in R^{T \times D}$ turns to $\mathrm{X} = \{\mathrm{x}_n\}_{n=1}^{N}$, $\mathrm{x} \in R^{T \times \overline{D}}$, where $\overline{D}$ mean value of time-series data sample. This method is applied to detect signal after acquiring one trial. The second method of data representation is more robust comparing with the first method, but consumes more time.

3.6 Software Description

BrainVision Product was used to acquire the brain signals.

The project was created in Google Colab notebook using the Python programming language. For preprocessing and models, keras, numpy, and sklearn were utilized, as well as seaborn and matplot for visualization.

3.7 Model Development

3.7.1 Data Preprocessing

The initial data in the project dataset is raw, therefore, it was necessary to perform preliminary data processing before classifying it. The function that implements the preprocessing of the ERP data, firstly, changes the sampling rate of the given EEG signal to 100. Further, a band-pass filter was applied to select the target frequency in the range 0.5-40Hz, and the result was segmented in the time interval [-200 1000] ms, with subsequent baseline correction by subtracting the mean amplitudes in the [-200 0] ms pre-stimulus interval (the specific time interval [0 1000] ms was selected). It was decided to take ERP data from the following 32 channels: Fp1-2, F3-4, Fz, F7-8, FC5-6, FC1-2, T7-8, C3-4, Cz, CP1-2, CP5-6, TP9-10, P3-4, P7-8, Pz, PO9-10, O1-2, and Oz. The mean amplitudes in 10 discriminant time intervals were used to extract subject-dependent spatio-temporal characteristics.

3.7.2 LDA

The data was specified as $[0-100], [50-150], ..., [900-1000]$, with a 100 ms length and a 50 ms step size from the stimulus onset. Mean amplitude features in specified time intervals were calculated across all channels. After that, all features concatenated from the EEG trials X. $\mathrm{V} = \{\mathrm{v}\}_{n=1}^{N}$ defined the feature vector set and formed as $R^{(D\times k)\times N}$. The obtained feature set V was used to create a regularized linear discriminant analysis (RLDA) [46] classifier. $f(v) = w^T \cdot v + b$ defined the decision function (w - the hyper-plane for separation of binary classes, b is a bias term).

3.7.3 CNN Approach

The spectral filtered EEG trials were concatenated into third dimension which third dimension into set of $X = \{\mathrm{X}\}_{f=1}^{f_n}$. Consequently, $X \in R^{T\times D\times f_n}$ was defined as the input shape, where (T) stands for temporal , (D) spatial, and (f_n) frequency domains.

CNN

Layer (type)	Output Shape	Param #
conv2d_33 (Conv2D)	(None, 32, 101, 4)	68
conv2d_34 (Conv2D)	(None, 32, 101, 16)	208
average_pooling2d_21 (AveragePooling2D)	(None, 32, 25, 16)	0
dropout_33 (Dropout)	(None, 32, 25, 16)	0
conv2d_35 (Conv2D)	(None, 32, 25, 16)	8208
average_pooling2d_22 (AveragePooling2D)	(None, 32, 6, 16)	0
batch_normalization_10 (BatchNormalization)	(None, 32, 6, 16)	64
flatten_13 (Flatten)	(None, 3072)	0
dense_39 (Dense)	(None, 200)	614600
dropout_34 (Dropout)	(None, 200)	0
dense_40 (Dense)	(None, 100)	20100
dense_41 (Dense)	(None, 2)	202

Total params: 643,450
Trainable params: 643,418
Non-trainable params: 32

EEGNet

Layer (type)	Output Shape	Param #
input_3 (InputLayer)	[(None, 32, 101, 1)]	0
conv2d_5 (Conv2D)	(None, 32, 101, 8)	256
batch_normalization_5 (BatchNormalization)	(None, 32, 101, 8)	32
depthwise_conv2d_2 (DepthwiseConv2D)	(None, 1, 101, 16)	512
batch_normalization_6 (BatchNormalization)	(None, 1, 101, 16)	64
activation_2 (Activation)	(None, 1, 101, 16)	0
average_pooling2d_3 (AveragePooling2D)	(None, 1, 25, 16)	0
dropout_2 (Dropout)	(None, 1, 25, 16)	0
separable_conv2d (SeparableConv2D)	(None, 1, 25, 16)	384
batch_normalization_7 (BatchNormalization)	(None, 1, 25, 16)	64
activation_3 (Activation)	(None, 1, 25, 16)	0
average_pooling2d_4 (AveragePooling2D)	(None, 1, 6, 16)	0
dropout_3 (Dropout)	(None, 1, 6, 16)	0
flatten (Flatten)	(None, 96)	0
dense (Dense)	(None, 2)	194
softmax (Activation)	(None, 2)	0

Total params: 1,506
Trainable params: 1,426
Non-trainable params: 80

Figure 3-5: CNN models description

Figure 3-5 illustrated the architecture of CNN model, which model have 3 convolution layers with *[1, 16], [1, 3], [32, 1]* dimensions with 4, 16, 16 depth respectively. Each layers extract the spatial - temporal features from the EEG signal.

$$(n + 2p - f)/s + 1, \tag{3.1}$$

the formula calculates the output size, where n represents the number of filters, p - padding size, f - filter size, and s - the amount of stride. The CNN model has 4 ReLu layers to accelerate the training process by applying threshold operation. Moreover, there are 3 Dropout layers to prevent overfitting in the network.

The last layers flatten neurons and shorten to 128 shape and produce two classes which are AC and NI. The sigmoid activation works well with binary classifications, thus it was used in network. Hyperparameters of 0.001 learning rate batch size of 100 and 32 epochs were acquired through multiple analytical trial runs. A small learning rate ensures that our updates are not too big as the network learns the most appropriate weights. Meanwhile, a batch size affect the loss function and ensures that we review multiple entries before deciding how to backpropagate. The remaining hyperparameter of epochs simply allows us to iterate over the same dataset multiple times. Otherwise, there would be not enough training data to approximate on.

Figures 3-5 demonstrate a summary of EEGNet[25] with input [32,101,1]. To minimise the categorical cross-entropy loss function, the Adam optimizer is utilised to fit the model using conventional parameters. The training epochs were found to be the most effective at 100 after a series of testing. Additionally, the validation is completed by saving the optimal model weights that result in the least amount of validation loss.

3.7.4 ERP Data Augmentation

The existed problem lays on the high cost of human behavior tasks, thus the size of recorded EEG data is typically small. The CNN model can be efficiently increased with increasing the amount of training data by the data augmentation. In machine learning experiments, a variety of augmentation strategies have been proposed, including geometric modifications, adding noise to existing data, and creating new artificial data. ERP potential is continuous signals with context meaning across time (e.g., N200, P300). The solution [10] was to add Gaussian random noise to the task-relevant ERP signals based on these assumptions.

The x_n ERP signal can be calculated as $\mathrm{x}_n = p\,(t) + r_n\,(t)$, where p is the task-relevant component and r_n is the noise signal with $\mathcal{N}(0,\,\sigma^2)$. Since EEG signals are very unpredictable and non-stationary, it is nearly impossible to separate task-relevant components from noise, especially on a continuous signals. As a result, in the previous study, a heuristic or empirical parameter was used to estimate the noise signal [47]. In principle, adding Gaussian random noise to the task-relevant component p can produce an infinite number of synthetic ERP trials if the signal parameter of r_n is approximated. Depending on the ERP paradigm's characteristics, the r_n signal parameter approximately was estimated in this thesis.

The averaging process reduces (cancels) the magnitude of noise r_n across trials N, i.e., $r(t) := 1/N \sum_{n=1}^{N} r_n\,(t) \sim \mathcal{N}(0,\,\sigma^2/N)$. On the basis of the provided assumption, σ^2 of the noise signal was acquired as follows:

The averaging technique across the trials N reduces (cancels) the amplitude of noise r_n, i.e. $r(t) := 1/N \sum_{n=1}^{N} r_n\,(t) \sim \mathcal{N}(0,\,\sigma^2/N)$. The noise signal's σ^2 was calculated using the following assumptions:

- Firstly, all training trials are averaged by the task-relevant signal $\tilde{p}(t)$, i.e., $\tilde{p}(t) = \bar{X} = 1/N \sum_{n=1}^{N} \mathrm{x}_n$, where $r(t)$ is considered to be close to zero after an appropriate averaging technique.

- A set of trials was chosen at random from the X and then averaged.(denoted by $\bar{\mathrm{Z}} = 1/K \sum_{k=1}^{K} \mathrm{z}_k$).

- $\bar{\mathrm{Z}} - \bar{\mathrm{X}} = \tilde{p}(t) - p\,(t) + r_k\,(t)$ is calculated. According to the prior assumption, the output is $r_k\,(t) \sim \mathcal{N}(0,\,\sigma^2/K)$.

- The noise variable (i.e., σ^2) was then calculated by $K \cdot var(r(t))$.

The probability density function P was used to define the random variable v as regards:

$$P\left(v\right) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{(v-\mu)^2/2\sigma^2}, \tag{3.2}$$

where the μ is zero and the $\sigma^2 = K \cdot var(r(t))$.

Lastly, $\bar{\mathrm{X}}_t$ and $\bar{\mathrm{Z}}_t$ were added with Gaussian noise to create new artificial data, i.e., $\tilde{p}(t) + v(t)$, $p(t) + r_k\left(t\right) - r_k\left(t\right) + v(t)$. This recurrent augmentation approach performed until the CNN model was satisfied that it had enough training data.

Chapter 4

Result and Discussion

4.1 ERP Responses in Non-oddball Paradigm

The grand averaged ERP responses of the *NI* and *AC* tasks are shown in figure 4-1.

Figure 4-1 indicates the grand averaged ERP responses of the individual tasks (i.e., NI, PC, and AC). In both non-oddball visual and auditory conditions, highly discriminative ERP responses were investigated in the three tasks, and these patterns were in discord with previous knowledge of the typical ERP components. First of all, the ERPs gradually increased from the stimulus onset and peaked in the 300-400 ms interval. These ERPs then gradually decreased until 800 ms after stimulus onset. The sound imagery task induced the largest amplitude, and passively concentrating as well as the ignoring task barely induced ERP responses by the non-oddball paradigms. These ERP components were induced at the central cortex (Cz), however, the ERP responses were not evoked in the occipital or temporal lobes.

The means of peak amplitude in the interval of 300-400 ms for NV were 0.558 ($\pm$2.083uV), 2.277 ($\pm$3.461)uV, and 3.547 ($\pm$3.637)uV, and NA were 0.204 ($\pm$1.079)uV, 1.210 ($\pm$1.683)uV, and 2.257 ($\pm$3.090)uV for NI, PC, and AC conditions, respectively. We performed one-way, Bonferroni corrected ANOVA for the three conditions with the null hypothesis of equal means to the maximum amplitude within the two specific intervals. Before the statistical test, we validated the data distributions in all conditions with the Jarque-Bera test for normality. In both non-oddball visual and auditory paradigms, the active task showed a significant difference with the NI condition in the 300-400 interval ($p < 0.05$).

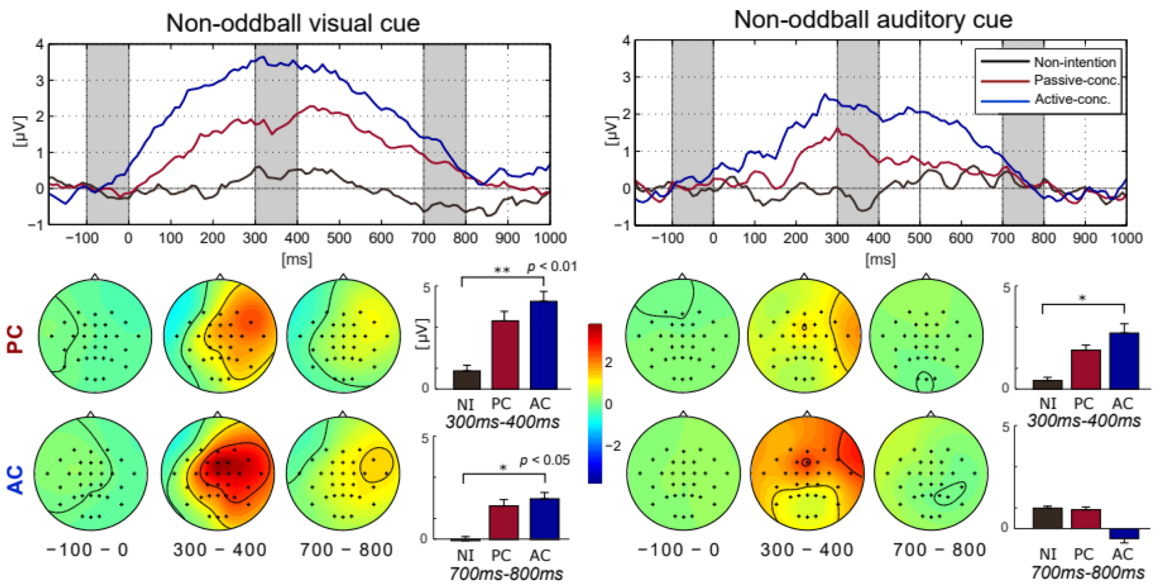

Figure 4-1: **Averaged ERP responses at Cz electrode for three sessions, i.e., non-oddball visual/auditory cue.** The scalp plots indicate the distribution of signal response for the three different conditions, i.e., NI, PC and AC in the non-oddball paradigms. The corresponding responses for the auditory cue experiment are less pronounced compared to the visual cue experiment.

4.2 Decoding Accuracy of Non-oddball Paradigm

The suggested non-oddball auditory paradigms' decoding accuracy using LDA is illustrated in Figure 4-2. To attain a 70% accuracy rate in the active tasks, four sequences were required, which is defined as an efficient communication rate in BCI research. [9]. In the confusion matrix has been seen the 23% predicted values of NI classified as AC.

The two CNNs were tested on non-oddball paradigm (table 4.1) which CNN1 showed more than 73.0 % accuracy and EEGNet showed more than 94.5% for the visual and auditory paradigm. In the CNN training procedure, the mean value of the previous sequences, which was the accurate value of paradigms, was employed.

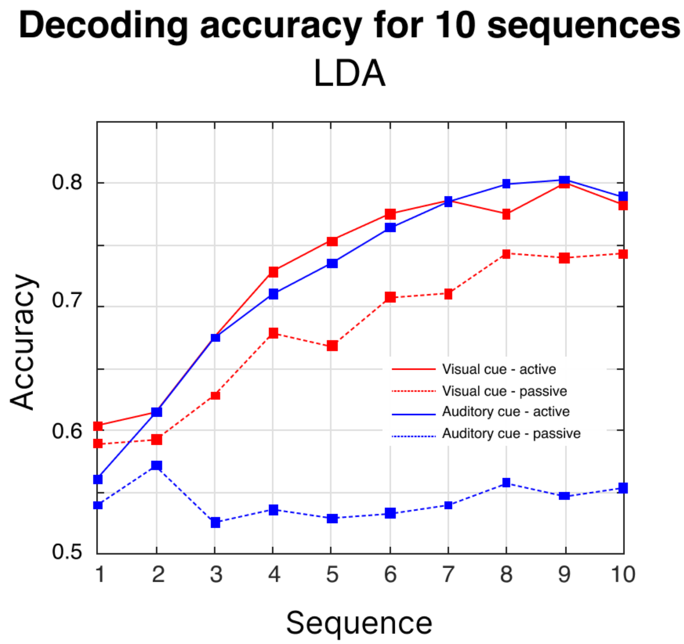

Figure 4-2: **LDA Confusion Matrix for AC and NI states**

CNN1 model (table 4.1) showed minimum 73.0%, 88.5% for AC vs NI, PC vs NI auditory paradigm respectively, and 97.0%, 81.5% for AC vs NI, PC vs NI visual paradigm. The figure 4-3 indicated the confusion matrix for all paradigms and training accuracy and loss, validation accuracy and loss. The test loss showed 8.74%, 13.5% for AC vs NI, PC vs NI auditory paradigm respectively, and 24.2%, 18.3% for AC vs NI, PC vs NI visual paradigm.

EEGNet model (table 4.1) showed minimum 98.5%, 98.0% for AC vs NI, PC vs NI auditory paradigm respectively, and 94.9%, 94.5% for AC vs NI, PC vs NI visual paradigm. The figure 4-3 indicated the confusion matrix for all paradigms and training accuracy and loss, validation accuracy and loss. The test loss showed 8.74%, 13.5% for AC vs NI, PC vs NI auditory paradigm respectively, and 24.2%, 18.3% for AC vs NI, PC vs NI visual paradigm.

The subject independent dataset were concatenated from 14 subjects dataset. The total number of dataset were 7560 trials for training dataset and 2800 trials for testing dataset.

CNN 1	AU		VI	
Tasks	**AC vs NI**	**PC vs NI**	**AC vs NI**	**PC vs NI**
Subject 1	0.995	0.960	0.990	0.985
2	0.990	0.980	0.980	0.985
3	0.995	0.995	0.985	0.980
4	0.985	0.980	0.985	0.930
5	0.995	0.995	0.995	0.995
6	**0.730**	0.990	1.000	0.985
7	0.990	**0.885**	0.990	0.985
8	0.990	0.995	1.000	0.980
9	0.965	0.980	**0.970**	0.950
10	0.980	0.995	0.990	0.980
11	0.985	0.995	0.980	**0.815**
12	0.990	0.985	0.985	0.985
13	0.990	0.990	0.995	0.975
14	0.990	0.985	0.995	0.960
Min	**0.730**	**0.885**	**0.970**	**0.815**
Max	**1.000**	**0.995**	**1.000**	**0.995**

EEGNet	AU		VI	
Tasks	**AC vs NI**	**PC vs NI**	**AC vs NI**	**PC vs NI**
Subject 1	**0.985**	0.990	0.995	1.000
2	0.995	0.990	0.970	1.000
3	1.000	1.000	0.990	0.975
4	1.000	**0.980**	**0.950**	1.000
5	0.995	0.995	0.990	0.995
6	1.000	0.985	1.000	1.000
7	0.995	0.990	0.995	0.990
8	1.000	1.000	1.000	0.995
9	**0.985**	0.980	1.000	0.965
10	1.000	1.000	1.000	**0.945**
11	0.990	0.985	0.995	0.995
12	1.000	**0.980**	0.985	1.000
13	0.995	0.995	0.990	0.975
14	0.995	0.995	0.995	0.970
Min	**0.985**	**0.980**	**0.950**	**0.945**
Max	**1.000**	**1.000**	**1.000**	**1.000**

Table 4.1: **CNN and EEGNet Results**: AC vs NI, PC vs NI decoding accuracy for both auditory and visual non-oddball paradigm. The bold text is minimum values

CNN model showed 99.2%, 99.1% for AC vs NI, PC vs NI auditory paradigm respectively, and 99.4%, 99.4% for AC vs NI, PC vs NI visual paradigm. The figure 4-5 indicated the confusion matrix for subject paradigm and training accuracy and loss, validation accuracy and loss. The test loss showed 3.31%, 4.35% for AC vs NI, PC vs NI auditory paradigm respectively, and 1.98%, 2.20% for AC vs NI, PC vs NI visual paradigm.

EEGNet model showed 94.4%, 96.4% for AC vs NI, PC vs NI auditory paradigm respectively, and 94.8%, 92.5% for AC vs NI, PC vs NI visual paradigm. The figure 4-5 indicated the confusion matrix for subject paradigm and training accuracy and loss, validation accuracy and loss. The test loss showed 14.8%, 12.5% for AC vs NI, PC vs NI auditory paradigm respectively, and 18.4%, 21.7% for AC vs NI, PC vs NI visual paradigm.

4.3 Discussion

This investigation advances real-world BCI with binary visual and auditory system. BCI system works with CNN model which classify signals over temporal features

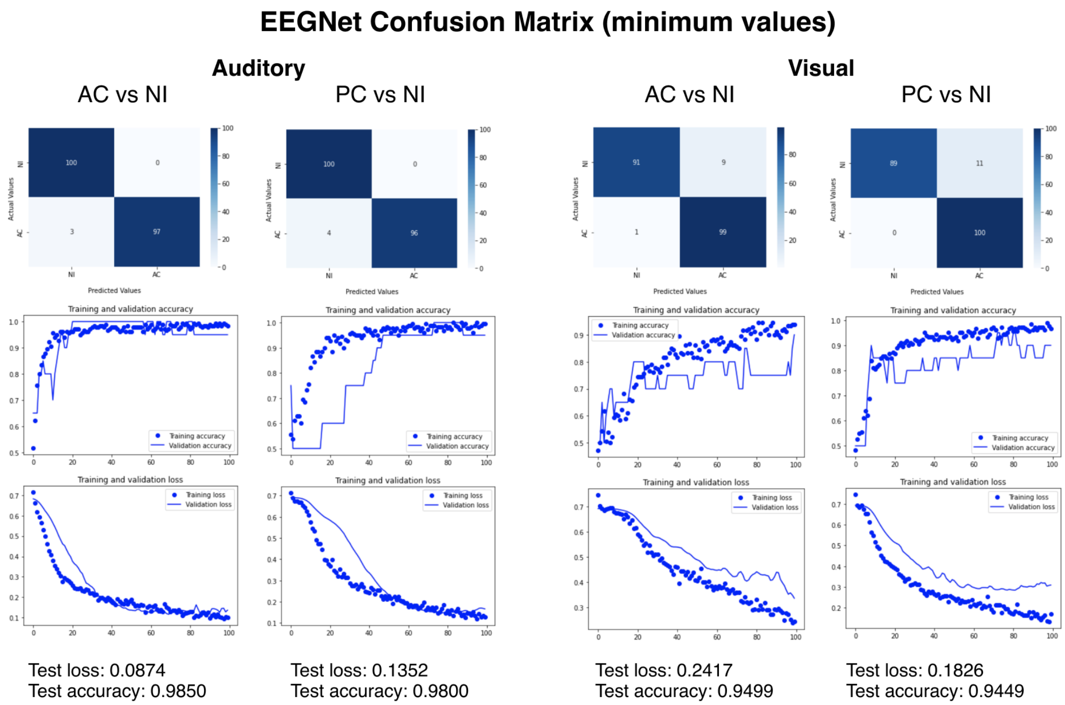

Figure 4-3: **EEGNet Confusion Matrix**: Minimum accuracy over subjects and loss

with small amount of EEG electrodes. Proposed CNN trained and tested with augmented data to maintain high decoding accuracy. The binary BCI system can be presented with minuscule stimulus effect since the volume of auditory and size of visual stimulus were barely heard/seen by subjects. The significant ERP responses were extracted from subject intention. The perceived ERP responses validate robust endogenous potentials. Therefore, we introduce a stimulus-free binary ERP paradigm. This proposed paradigm minimizes dependence on the external factors which are current limitation in oddball paradigm.

In our experiment, we present unobtrusive continuous visual and auditory stimuli where the magnitude of the auditory and visual stimuli is reduced to the point that they are barely noticeable. In prior investigations, high stimulation was used to induce a robust ERP response. Continuous low-amplitude stimulation does not create a reliable ERP response, but it does serve as a timing mechanism for the user. The process enables subjects to create endogenous potentials for systems in the form

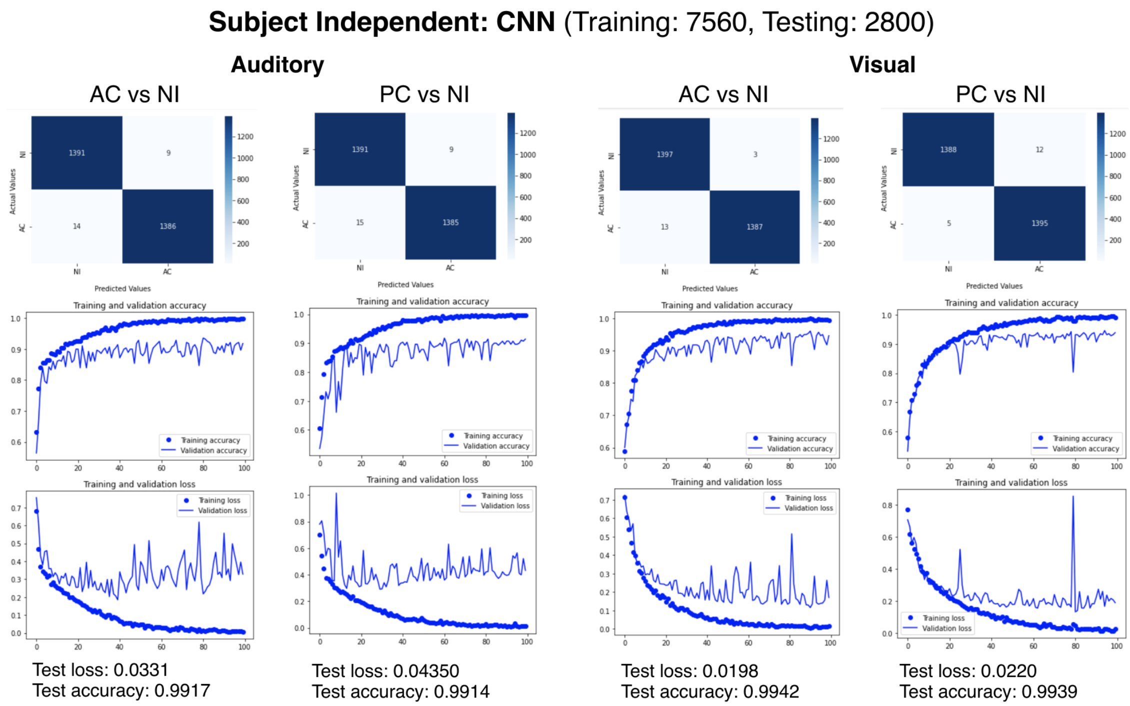

Figure 4-4: **CNN Results**: Subject Independent

of intention. This method is user-friendly and does not cause fatigue due to high loudness or flashing visuals. In addition, the proposed solution is advised for patients who are seriously disabled or who are sensitive to extreme external simulation. The proposed non-oddball paradigm can be integrated in daily life because of mental task. Program will perform command when user creates mental responses at appropriate level of stimulus, otherwise program ignores the rest of the brain signals.

The grand averaged ERP responses of AC and PC for the visual cue were greater than those resulting from the auditory cue. This is due to the different experimental protocols of each paradigm. Specifically, subjects were instructed to perform the mental task 10 times, once every second in the auditory paradigm, while they alternatively performed the mental tasks and ignoring state in the visual paradigm (see Figure 1). A sufficient inter-stimulus interval between the active mental tasks not only benefits mental task preparation but also leads to less jitter of the brain response and thus higher amplitudes in the average. The most remarkable result is

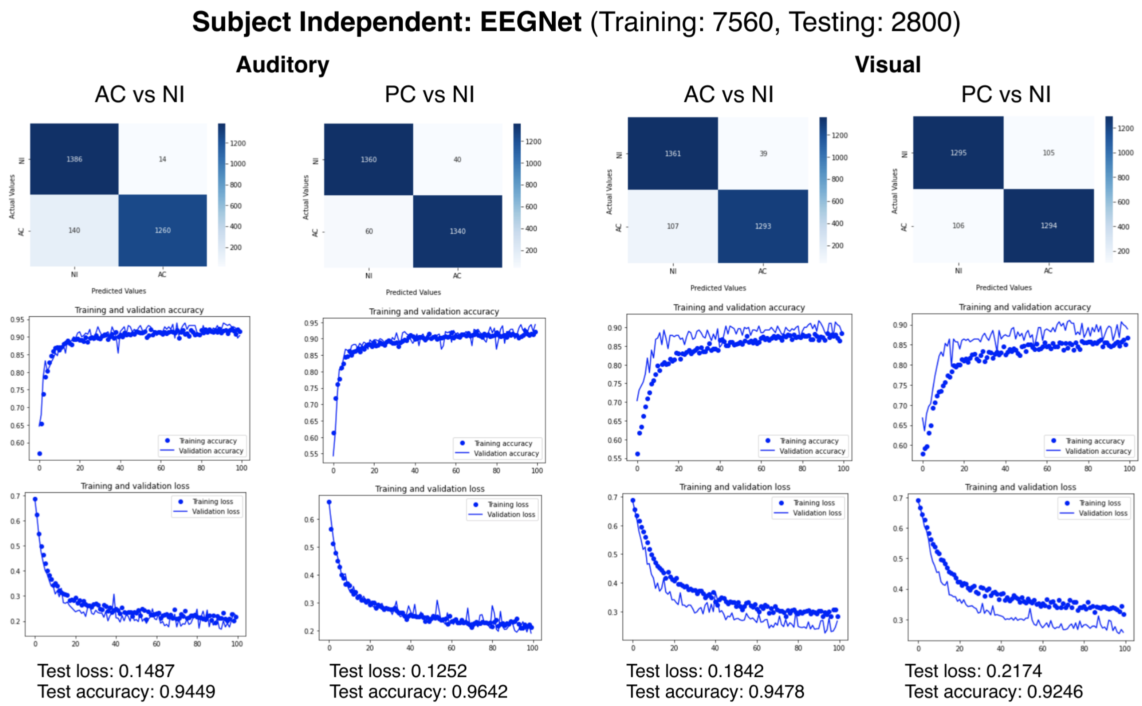

Figure 4-5: **EEGNet Results**: Subject Independent

that CNN1 model showed more than 73% of accuracy and EEGNet showed more than 94.5% for all cases. Both CNN models more than 92.5% for all tasks of the subject independent dataset. Although our CNN results differ slightly from Linear classifier model results. With a higher than average decoding accuracy, our system still retains a very reasonable performance rate for a BCI communication system [21].

Chapter 5

Conclusion and Future Work

The presented approach does not rely on robust stimulation by increasing external stimulus as conventional oddball-based ERP approaches. The presented design investigated to increase individual's intention that produces endogenous potentials with less auditory stimuli intense. The LPP response, which is the outcome of brain activity from an active cognitive process, was the major emphasis. This is a further step from the previous ERP studies [42, 44, 48, 49, 50]. Our study is important for next studies of BCI application, because proposed system: reduces unintentional command signalling and interference from peripheral stimulation by reducing the magnitude of the stimulus.

Tactile paradigm [51], calibration-free classifier model [52] and others studies [53] were investigated to improve the performance of BCI systems based on traditional ERP paradigm. This thesis will be useful in the future as an easy-to-implement technique to improve the performance of many existing ERP systems. We expect that other researchers will adopt our technique to develop unique and trustworthy BCI systems that may be used in a practical situation.

Bibliography

[1] Gert Pfurtscheller and Christa Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001.

[2] Nancy K Squires, Kenneth C Squires, and Steven A Hillyard. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neurophysiology*, 38(4):387–401, 1975.

[3] Sourav Kundu and Samit Ari. P300 based character recognition using convolutional neural network and support vector machine. *Biomedical Signal Processing and Control*, 55:101645, 2020.

[4] William Speier, Aniket Deshpande, Lucy Cui, Nand Chandravadia, Dustin Roberts, and Nader Pouratian. A comparison of stimulus types in online classification of the P300 speller using language models. *PloS one*, 12(4):e0175382, 2017.

[5] Z Kubová, J Kremlacek, J Szanyi, J Chlubnová, and M Kuba. Visual event-related potentials to moving stimuli: normative data. *Physiological Research*, 51(2):199–204, 2002.

[6] Qi Li, Zhaohua Lu, Ning Gao, and Jingjingc Yang. Optimizing the performance of the visual P300-speller through active mental tasks based on color distinction and modulation of task difficulty. *Frontiers in human neuroscience*, 13:130, 2019.

[7] Xinru Zhang, Jing Jin, Shurui Li, Xingyu Wang, and Andrzej Cichocki. Evaluation of color modulation in visual p300-speller using new stimulus patterns. *Cognitive Neurodynamics*, pages 1–14, 2021.

[8] Seul-Ki Yeom, Siamac Fazli, Klaus-Robert Müller, and Seong-Whan Lee. An efficient ERP-based brain-computer interface using random set presentation and face familiarity. *PloS one*, 9(11), 2014.

[9] Qi Li, Shuai Liu, Jian Li, and Ou Bai. Use of a green familiar faces paradigm improves P300-speller brain-computer interface performance. *PloS one*, 10(6):e0130325, 2015.

[10] Madina Saparbayeva, Adai Shomanov, and Min-Ho Lee. A novel binary bci systems based on non-oddball auditory and visual paradigms. In *International Conference on Neural Information Processing*, pages 3–14. Springer, 2021.

[11] Maureen Clerc, Laurent Bougrain, and Fabien Lotte. *Brain-computer interfaces 2: technology and applications*. John Wiley & Sons, 2016.

[12] Dennis J McFarland and Jonathan R Wolpaw. Brain–computer interface use is a skill that user and system acquire together. *PLoS biology*, 16(7):e2006719, 2018.

[13] Christa Neuper, Reinhold Scherer, Miriam Reiner, and Gert Pfurtscheller. Imagery of motor actions: Differential effects of kinesthetic and visual–motor mode of imagery in single-trial eeg. *Cognitive brain research*, 25(3):668–677, 2005.

[14] Léa Pillette, Aline Roc, Bernard N'Kaoua, and Fabien Lotte. Experimenters' influence on mental-imagery based brain-computer interface user training. *International Journal of Human-Computer Studies*, 149:102603, 2021.

[15] Eva Maria Hammer, Sebastian Halder, Benjamin Blankertz, Claudia Sannelli, Thorsten Dickhaus, Sonja Kleih, Klaus-Robert Müller, and Andrea Kübler. Psychological predictors of smr-bci performance. *Biological psychology*, 89(1):80–86, 2012.

[16] S Christian Wheeler and Richard E Petty. The effects of stereotype activation on behavior: a review of possible mechanisms. *Psychological bulletin*, 127(6):797, 2001.

[17] Robert Rosenthal. On the social psychology of the psychological experiment: 1, 2 the experimenter's hypothesis as unintended determinant of experimental results. *American Scientist*, 51(2):268–283, 1963.

[18] Álvaro Fernández-Rodríguez, Ricardo Ron-Angevin, Ernesto J. Sanz-Arigita, Antoine Parize, Juliette Esquirol, Alban Perrier, Simon Laur, Jean-Marc André, Véronique Lespinet-Najib, and Liliana Garcia. Effect of distracting background speech in an auditory brain–computer interface. *Brain Sciences*, 11(1), 2021.

[19] Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Fernando Vaquerizo-Villar, and Roberto Hornero. Eeg-inception: A novel deep convolutional neural network for assistive erp-based brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(12):2773–2782, 2020.

[20] Takumi Kodama and Shoji Makino. Convolutional neural network architecture and input volume matrix design for erp classifications in a tactile p300-based brain-computer interface. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3814–3817. IEEE, 2017.

[21] Jongmin Lee, Kyungho Won, Moonyoung Kwon, Sung Chan Jun, and Minkyu Ahn. Cnn with large data achieves true zero-training in online p300 brain-computer interface. *IEEE Access*, 8:74385–74400, 2020.

[22] Yu-Xuan Yang, Zhong-Ke Gao, Xin-Min Wang, Yan-Li Li, Jing-Wei Han, Norbert Marwan, and Jürgen Kurths. A recurrence quantification analysis-based channel-frequency convolutional neural network for emotion recognition from eeg. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(8):085724, 2018.

[23] Hubert Cecotti and Axel Graser. Convolutional neural networks for p300 detection with application to brain-computer interfaces. *IEEE transactions on pattern analysis and machine intelligence*, 33(3):433–445, 2010.

[24] Xuyun Sun, Cunle Qian, Zhongqin Chen, Zhaohui Wu, Benyan Luo, and Gang Pan. Remembered or forgotten?—an eeg-based computational prediction approach. *PloS one*, 11(12):e0167497, 2016.

[25] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe, and Klaus-Robert Müller. Single-trial analysis and classification of erp components—a tutorial. *NeuroImage*, 56(2):814–825, 2011.

[26] Marco Congedo, Alexandre Barachant, and Rajendra Bhatia. Riemannian geometry for eeg-based brain-computer interfaces; a primer and a review. *Brain-Computer Interfaces*, 4(3):155–174, 2017.

[27] Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Javier Gomez-Pilar, and Roberto Hornero. Deep learning architecture based on the combination of convolutional and recurrent layers for erp-based brain-computer interfaces. In *Mediterranean Conference on Medical and Biological Engineering and Computing*, pages 1844–1852. Springer, 2019.

[28] Robin Tibor Schirrmeister, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. Deep learning with convolutional neural networks for eeg decoding and visualization. *Human brain mapping*, 38(11):5391–5420, 2017.

[29] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, 15(5):056013, 2018.

[30] Saim Rasheed. A review of the role of machine learning techniques towards brain–computer interface applications. *Machine Learning and Knowledge Extraction*, 3(4):835–862, 2021.

[31] Haibo Yi. Efficient machine learning algorithm for electroencephalogram modeling in brain–computer interfaces. *Neural Computing and Applications*, pages 1–11, 2020.

[32] Adrien Combaz, Nikolay Chumerin, Nikolay V Manyakov, Arne Robben, Johan AK Suykens, and Marc M Van Hulle. Towards the detection of error-related potentials and its integration in the context of a p300 speller brain–computer interface. *Neurocomputing*, 80:73–82, 2012.

[33] Rosaleena Mohanty, Anita M Sinha, Alexander B Remsik, Keith C Dodd, Brittany M Young, Tyler Jacobson, Matthew McMillan, Jaclyn Thoma, Hemali Advani, Veena A Nair, et al. Machine learning classification to identify the stage of brain-computer interface therapy for stroke rehabilitation using functional connectivity. *Frontiers in neuroscience*, 12:353, 2018.

[34] Ikhtiyor Majidov and Taegkeun Whangbo. Efficient classification of motor imagery electroencephalography signals using deep learning methods. *Sensors*, 19(7):1736, 2019.

[35] Catur Atmaji, Agfianto Eko Putra, and Irvan Albab Tontowi. Three-class classification of eeg signals using support vector machine methods. In *2018 4th International Conference on Science and Technology (ICST)*, pages 1–4. IEEE, 2018.

[36] Mamunur Rashid, Norizam Sulaiman, Mahfuzah Mustafa, Sabira Khatun, and Bifta Sama Bari. The classification of eeg signal using different machine learning techniques for bci application. In *International Conference on Robot Intelligence Technology and Applications*, pages 207–221. Springer, 2018.

[37] Vibhore Bhatnagar, Nikhil Yede, Rahul Singh Keram, and RK Chaurasiya. A modified approach to ensemble of svm for p300 based brain computer interface. In *2016 International Conference on Advances in Human Machine Interaction (HMI)*, pages 1–5. IEEE, 2016.

[38] Xiao Zheng, Wanzhong Chen, Yang You, Yun Jiang, Mingyang Li, and Tao Zhang. Ensemble deep learning for automated visual classification using eeg signals. *Pattern Recognition*, 102:107147, 2020.

[39] Manoj Thulasidas, Cuntai Guan, and Jiankang Wu. Robust classification of eeg signal for brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(1):24–29, 2006.

[40] Tomoyuki Nagasawa, Takanori Sato, Isao Nambu, and Yasuhiro Wada. Improving fnirs-bci accuracy using gan-based data augmentation. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1208–1211. IEEE, 2019.

[41] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[42] R Ron-Angevin, L Garcia, Á Fernández-Rodríguez, J Saracco, JM André, and V Lespinet-Najib. Impact of speller size on a visual P300 brain-computer interface (BCI) system under two conditions of constraint for eye movement. *Computational Intelligence and Neuroscience*, 2019, 2019.

[43] Min-Ho Lee, O-Yeon Kwon, Yong-Jeong Kim, Hong-Kyung Kim, Young-Eun Lee, John Williamson, Siamac Fazli, and Seong-Whan Lee. Eeg dataset and openbmi toolbox for three bci paradigms: an investigation into bci illiteracy. *GigaScience*, 8(5):giz002, 2019.

[44] Markus A Wenzel, Inês Almeida, and Benjamin Blankertz. Is neural activity detected by ERP-based brain-computer interfaces task specific? *PloS one*, 11(10):e0165556, 2016.

[45] Min-Ho Lee, John Williamson, Dong-Ok Won, Siamac Fazli, and Seong-Whan Lee. A high performance spelling system based on EEG-EOG signals with visual feedback. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(7):1443–1459, 2018.

[46] Jerome H Friedman. Regularized discriminant analysis. *Journal of the American Statistical Association*, 84(405):165–175, 1989.

[47] Fang Wang, Sheng-hua Zhong, Jianfeng Peng, Jianmin Jiang, and Yan Liu. Data augmentation for eeg-based emotion recognition with deep convolutional neural networks. In *International Conference on Multimedia Modeling*, pages 82–93. Springer, 2018.

[48] John Polich and Heather K McIsaac. Comparison of auditory P300 habituation from active and passive conditions. *International Journal of Psychophysiology*, 17(1):25–34, 1994.

[49] Jennifer Y Bennington and John Polich. Comparison of P300 from passive and active tasks for auditory and visual stimuli. *International Journal of Psychophysiology*, 34(2):171–177, 1999.

[50] Mathew Salvaris and Francisco Sepulveda. Visual modifications on the P300 speller BCI paradigm. *Journal of neural engineering*, 6(4):046011, 2009.

[51] Jing Jin, Zongmei Chen, Ren Xu, Yangyang Miao, Xing yu Wang, and Tzyy-Ping Jung. Developing a novel tactile P300 brain-computer interface with a cheeks-stim paradigm. *IEEE Transactions on Biomedical Engineering*, 2020.

[52] Jing Jin, Shurui Li, Ian Daly, Yangyang Miao, Chang Liu, Xingyu Wang, and Andrzej Cichocki. The study of generic model set for reducing calibration time in p300-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019.

[53] Albina Li, Kanat Alimanov, Siamac Fazli, and Min-Ho Lee. Towards paradigm-independent brain computer interfaces. In *2020 8th International Winter Conference on Brain-Computer Interface (BCI)*, pages 1–6. IEEE, 2020.