UNIVERSITÉ DU QUÉBEC À MONTRÉAL

# DETECTION OF "FLOW" STATE IN REAL-TIME BY ELECTROENCEPHALOGRAPHY AND MACHINE LEARNING DURING A GAMING EXPERIENCE

DISSERTATION

PRESENTED

AS PARTIAL REQUIREMENT

TO THE MASTERS IN COMPUTER SCIENCE

BY

SHAGHAYEGH YOUSEFPOUR LAZARJANI

FEBRUARY 2023

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

# DÉTECTION DE L'ÉTAT DE « FLOW » EN TEMPS RÉEL PAR ÉLECTROENCÉPHALOGRAPHIE ET APPRENTISSAGE MACHINE DURANT UNE EXPÉRIENCE LUDIQUE

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN INFORMATIQUE

PAR
SHAGHAYEGH YOUSEFPOUR LAZARJANI

FÉVRIER 2023

# ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Dr. Mounir Boukadoum, for his patience, kindness, wisdom, and help throughout my Master's. Being your student has been a rewarding experience.

Next, I would like to thank my supervisors at Ubisoft Montreal, Dr. Pierre Chalfoun, and Daniel Rivas. It was such an invaluable learning experience to work with you both, filled with tenacity and discipline.

In the next place, I would like to express my gratitude to the jury who have accepted to read and judge my Master's thesis. I am glad I can share what I did during the past years with you.

I also would send my deepest thankfulness to Mehri, Hossein, and Amin, for being such wonderful companions of mine and for helping me get through, literally, every moment of my life, and become who I am today. It would have been impossible without you three.

At last, I want to thank UQAM, Mitacs, and La Forge Ubisoft for providing me with the chance to study, advance and prosper.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# LIST OF SYMBOLS AND ACRONYMS

EEG ..... Electroencephalography

ECG ..... Electrocardiogram

EMG ..... Electromyogram

ERG ..... Electroretinogram

MEG ..... Magnetoencephalography

EDA ..... Electrodermal Activity

PPG ..... Photoplethysmography

fMRI ..... Functional Magnetic Resonance Imaging

CT ..... Computed Tomography

PET ..... Positron Emission Tomography

EMD ..... Empirical Mode Decomposition

VMD ..... Variational Mode Decomposition

DWT ..... Discrete Wavelet Transform

ASR ..... Artifact Subspace Reconstruction

PCA ..... Principal Component Analysis

wICA ..... Wavelet-enhanced Independent Component Analysis

IMF ..... Intrinsic Mode Function

IIR ..... Infinite Impulse Response

USE ..... User System Experience

FT ..... Fourier Transform

SFT ..... Short-Time Fourier Transform

DFT ..... Discrete Fourier Transform

ARMA ..... Auto-Regressive Moving Average

kNN ..... k-Nearest Neighbors

SVM ..... Suppport Vector Machine

ANN ..... Artificial Neural Network

CNN ..... Convolutional Neural Network

Acc ..... Accuracy

Pre ..... Precision

Rec ..... Recall

F1 ..... F1-score

RMS ..... Root-Mean-Square

# RÉSUMÉ

L'une des principales raisons de jouer aux jeux vidéo est la joie d'être dans un état de haute concentration qui se sépare du monde qui l'entoure. Une telle sensation d'implication s'appelle l'état de *flow*. Dans ce mémoire, il a été tenté de détecter un tel état à l'aide de données d'électroencéphalogramme multicanal acquises auprès des joueurs lors d'un jeu. Nous essayons de le faire par le biais d'un processus de nettoyage des données, de traitement des données, d'extraction et de sélection de caractéristiques, et enfin de classification, le tout en utilisant une combinaison d'ASR, de wICA, d'EMD/VMD, de DWT, de caractéristiques statistiques et enfin, de SVM et de réseaux de neurones. , respectivement. Les résultats de tous ces efforts n'étaient pas souhaitables pour l'industrie, ce qui nous a conduits à la deuxième étape de ce travail de recherche, à savoir pourquoi. Après avoir découvert que les résultats malheureux de la détection de flux étaient dus à des données corrompues et bruyantes, nous démontrons que la méthodologie qui a été suivie repose sur un raisonnement logique solide et le fait qu'elle peut être appliquée à la fois à une seule et à plusieurs -espace de données dimensionnel, le tout en appliquant une méthode à motif similaire sur des données EMG à canal unique, en essayant de classer les signaux sains et non sains. Cette dernière tentative aboutit à d'excellents résultats qui seront rapportés plus loin dans le mémoire.

Mots-clés : électroencéphalogramme, jeu vidéo, traitement du signal EEG, état de flux, engagement du joueur, classification EEG, apprentissage automatique

# ABSTRACT

One of the main reasons to play video games is the joy of being in a highly-concentrated state that separates oneself from the surrounding world. Such sensation of involvement is called the *flow* state. In this thesis, it has been tried to detect such a state using multi-channel electroencephalogram data that was acquired from the players during gameplay. We attempt to do so through a process of data cleaning, data processing, feature extraction and selection, and lastly, classification, all using a combination of ASR, wICA, EMD/VMD, DWT, statistical features, and lastly, SVM and neural networks, respectively. The results of all these efforts were not industry-desirable, which led us to the second step of this research work, investigating why. After finding out that the unfortunate results of flow detection were due to corrupted and noisy data, we show that the methodology that was followed, has a strong logical reasoning behind it, and the fact that it can be applied to both single- and multi-dimensional data space, all by applying a similar-patterned method on single-channel EMG data, trying to classify healthy and not-healthy signals. This last attempt leads to excellent results that will be reported later in the thesis.

Keywords: electroencephalogram, video game, EEG signal processing, flow state, player engagement, EEG classification, machine learning

# CHAPTER I

# INTRODUCTION

Today, billions of people around the world enjoy playing video games. Video game companies seem to be greatly interested in holding onto this billion-dollar market and competing with each other to attract more gamers. A solid story, originality, appealing graphics and visual effects, and having enough, yet doable, challenges are a number of characteristics of a successful video game that might be able to draw the gamers in. Immersion in games is referred to as "losing the notion of time", "being cut off from the actual world", or "feeling you are in the game", according to the people who were interviewed by (Brown & Cairns, 2004) and (Jennett et al., 2008).

As named by Csikszentmihalyi, a flow state is a mental state in which the person is so involved and immersed in the activity they are doing, accompanied by a feeling of enjoyment and full concentration (Csikszentmihalyi & Csikzentmihaly, 1990). There are hundreds of research works done on the detection of flow state in video games. (Cai et al., 2022) conceptualize flow and offer flow measurement techniques in the video game context by conducting various questionnaire-based interviews with participants. (Granato et al., 2017) and (Ye et al., 2020) use physiological data acquired from the players, such as galvanic skin response, heart rate, respiration rate, and facial electromyogram, along with the machine learning approaches to detect flow. In (Wu et al., 2021), electroencephalogram data was gathered from the participants doing game-based activities. The EEG data was analyzed and then classified into flow and not-flow. (Plotnikov et al., 2012) also use four-channel EEG to statistically distinguish a flow from a boredom condition by analyzing the different wavelengths of the signals and using machine learning approaches.

In this thesis, the original intent was to find the link between the flow state and the EEG data that were acquired from the players during sessions of gameplay.

**Research Question**: Is there any relation or link between EEG data and mental states like flow?

So, the objective of this research work is:

**Objective**: Development of algorithmic methods for the detection and characterization of mental states related to "flow" in high-level players during a gaming experience under industrial test conditions.

In order to reach the objective and answer the question, the data were gathered and provided by Ubisoft Montreal and the whole efforts were made for a closed-source project for Ubisoft Montreal.

The EEG data were preprocessed using artifact-removal algorithms (such as ASR and wICA) and processed using decomposition algorithms (such as EMD, VMD, DWT, etc.) to find the meaningful rhythms underlying the signals. After exploring several feature extraction methods, machine learning approaches were used for classification purposes, such as SVM and neural networks. All the procedures are explained in detail in chapter 3.

Despite the considerable amount of efforts that have been made during this research to detect flow state via EEG data, industry-desirable results were not obtained. To investigate the reason, we made further attempts that led to the understanding that the EEG data were not of proper quality. In addition, to show that the methodology we used did follow strong rationality, and to show that as we applied it to bi-dimensional EEG data space, it is also valid for single-dimension data space, we applied it to single-channel EMG data to classify the signals into healthy and not-healthy classes. EMG data was used because firstly, EMG, like EEG, is a type of physiological data, and is non-stationary signal with channels. All the results and attempts that were made are clearly described in chapters 4 and 5.

In the last chapter (6), we look into the limitations of this work and the possible future directions. We talk about how the data-acquisition tools and hardware is pivotal in the quality of the data that is being gathered, and how the cleanliness of the data plays an important role in the experiment being done on it. We also discuss that using other physiological data that was available could be an alternative way in this research to detect flow state, as being widely used in the literature.

# CHAPTER II

# LITERATURE REVIEW

This thesis is closely related to the previous studies that focus on using EEG to detect various mental states. Hence, this chapter will give an in-depth view of the literature regarding this topic.

## 2.1 Flow Detection in Video Games

As video game companies always want to make their products more attractive to their customers, finding how players engage more with video games has become an important subject of research. Detecting the so-called state of flow, and finding the characteristics which lead to it have been approached in various ways in the literature. For instance, (Burns & Tulip, 2017) show the feasibility of flow detection using facial expression. By using a webcam and a facial expression analytical software, they illustrate that the affection shown by the players during the flow state could be a compound cognitive state derived from a range of transitory basic emotions, such as joy, flat, or frustration. While some researchers find the flow state by monitoring the players' reactions, others gather data by letting the players speak for themselves. (Borderie & Michinov, 2016) suggest a method to detect flow episodes by observing players' behaviors during a session of gameplay, and then validating the labeled episodes by interviewing the players. They claim intense focus followed by the expression of joy or frustration is a useful pattern for flow detection. Taking a similar turn on this topic, (Schoenau-Fog et al., 2011) investigate player engagement by developing surveys focusing on the desire to continue playing to discover components, triggers, and categories involved in the player's flow. They manage to conclude the four main components of flow as objectives, accomplishments, activities and affects.

Some other researchers have developed models to measure players' flow. For instance, (Cowley et al., 2008) design a unified and effective approach to analyze a game and develop an insight into the flow of the game by using Cowley's user-system-experience model[1]. (Cowley et al., 2006) formulate the interaction between the game and the player based on learning that happens during the game. The authors believe that it is the learning cycle in the game that causes most of the flow, and absence or lack of learning would remove it. They state that the players who are already familiar with the game, or the ones who fail to master it, will have little or no flow.

---

[1]USE is a model for describing the user's experience in playing a game, proposed by (Cowley et al., 2006), since the quality of the user's experience is a vital factor in the video game industry.

So many researchers have developed models and procedures to measure flow based on players' feedback; using surveys or questionnaires, and it sure was a good start. However, self-evaluation might not always be an available option in research. Sometimes the players may not remember truly how they felt during the game or may evaluate their feelings carelessly. Hence, it was important and inevitable that other approaches emerged that would rely less on feedbacks or self-evaluations. For instance, (Klasen et al., 2012) characterize the flow state by particular neural activation patterns, by using the fMRI data gathered from the free play of a video game. They recognized and categorized multiple factors that lead to flow, such as; control over activity, having clear goals, concentration, the balance between ability and challenge, etc. Then, they managed to recognize the neural patterns resulting from the flow factors in the fMRI images.

As proposed by (Klasen et al., 2012), brain activities are widely used for detecting mental states, as our body is able to respond more accurately than what we might think is the answer. That is why the data collected by non-invasive techniques[2], such as CT scan, PET scan, MRI, EEG, MEG, etc. are frequently used in these types of studies for more precise results and findings. In the following section, a review of mental state detection using neurological data will be presented.

### 2.2 EEG Applications in Mental States Detection and Classification

The human brain cells communicate via electrical impulses. As a result, electrical activities are always present in the brain or on the scalp and can be measured via electrodes. What electrodes record, which visually appears as wavy lines, is called the electroencephalogram. EEG is a non-invasive method of recording the electrical activity of the brain. EEG signals have a high temporal resolution, are objective, and are reliable. Hence, they are being used in many studies to identify different mental states such as concentration, emotion, vigilance, attention, and so forth. For instance, (Albuquerque et al., 2019) utilize spectro-temporal features of EEG to assess the mental workload of people who work in hectic and fast-paced environments, such as nurses and firefighters. They extract power-spectral density (PSD) and amplitude modulation rate of change (AMRC) as features from their 8-channeled EEG data so that they would later be able to predict mental workload using the participant's activity levels. They classified once each feature individually, and once together, along with each activity level individually and once together. Using a random forest classifier with one-way ANOVA feature selection, 82% for accuracy is the average accuracy they get with PSD, AMRC, and both together, for all activity levels.

In the case of (Bashivan et al., 2016), they were interested in wearable EEG devices and to what extent they can be used for mental state detection. They gathered 4-channel EEG data from participants watching logical vs. emotional videos. They were able to show EEG's potential in distinguishing big but subtle differences using power spectral features, along with DWT level-7. For classification, they used logistic regression, SVM, random forest, and deep belief network. In the end, they succeeded in detecting the mental states motivated by the videos and obtained 31.5% for False-Positive Rate and 19.8% for False-Negative Rate.

[2]A non-invasive procedure is defined as a procedure where no skin break, or body cavity, or any contact with the outer membrane of the body is done.

In addition, (Myrden & Chau, 2017) show viability to monitor the changes in the mental states of the participants doing challenging mental tasks, using their EEG cortical activity. After applying ICA algorithm to remove the eye-blinking artifacts, a fast Fourier Transform was computed to obtain the spectral power of each electrode. Later, a fast correlation-based filter was employed for feature selection which were fed to an SVM classifier that resulted in the accuracy of $74.8 \pm 9.1\%$, $71.6 \pm 5.6\%$, and $84.8 \pm 7.4\%$ respectively for the fatigue, frustration, and attention states.

(Jung et al., 2022) try to classify emotion states using EEG in a virtual reality environment. The participants are exposed to low- and high-arousal VR stimuli, as well as a social anxiety situation. Then, the VR-evoked EEG is preprocessed by applying band-pass filters on them, and using ICA for artifact removal. The feature selection phase comprised power spectral density, differential and rational asymmetry, and fractal dimension, which was followed by the feature selection step by using lasso and ridge regression models that turned the 54-feature set into 14. These features were then fed to three classifiers: an XGBoost, a support vector classifier, and a logistic regression. The results of each two-class classification are mentioned in their paper.

In another research work, (Rahman et al., 2022) show their ability in determining three mental states -calmness, concentration, and impartiality- by analyzing the EEG signals. The authors use a Kaggle dataset comprising the EEG data of three mental states, scale it, and select the k-best features from it. Finally, by applying the classifiers on the feature set, they get the following results: 95.36% (SVM), 95.24% (Gradient Boosting), 93.10% (both XGBoost and KNN).

### 2.3 Signal Processing Methods

Biological signals, like EEG, are mostly contaminated by many unwelcome signals that can be caused by other parts of the body, or even environmental stimuli. For example, movements in the body parts create additional signals in the brain, or the hair on the scalp that may hinder the electrodes from properly recording EEG. As a result, an EEG signal might be noisy in many ways, and these non-cerebral-originated data could mislead any procedure that relies on clear data. This is when the denoising algorithms come in handy. (Plechawska-Wojcik et al., 2018) denoise EEG signals of the motor imagery data with principal component analysis (PCA), independent component analysis (ICA), and artifact subspace reconstruction (ASR). They report the ASR algorithm to be doing better than the others in artifact removal and signal loss. In (Maanen et al., 2022)'s work, they define a specific implementation of ASR for single-board computers in order to contribute to the online and real-time interpretation of EEG, as ASR is considered as a strong artefact removal method. Moreover, (Inuso et al., 2007) show the power of artifact removal of a combinatory algorithm of wavelet-based ICA, called wICA (Azzerboni et al., 2002), applied to an EMG dataset. Compared to ICA and another method with ICA plus noisy-channel removal, wICA is the most successful in removing the artifacts by 92% to 86% ratio.

In any research work that deals with signals, after preprocessing the signals to boost their quality, the natural next step would be processing the signal itself. Processing the signals enables us to capture the underlying information, which is the key to later revealing the secrets of a phenomenon or a mechanism. Here, some of the signal processing methods that have been put into service for classification purposes will be discussed, not limited to EEG signals or mental states (but widely

used). These methods have been reported to have a significant impact on the processing of signals, and have demonstrated great potential, especially being employed with state-of-the-art classification methods.

Currently, Empirical Mode Decomposition (EMD) (Huang et al., 1998) and Variational Mode Decomposition (VMD) (Dragomiretskiy & Zosso, 2013) are among the popular decomposition algorithms. These methods are able to decompose a signal into its simpler mode functions, the sum of which, would form the original signal. These methods have shown such potential for decomposing non-linear and non-stationary signals, such as EEG, that there exist many varieties of them. Ensemble EMD (EEMD) (Wu & Huang, 2009), Complementary Ensemble EMD (CEEMD) (Liu et al., 2018), Bidimensional EMD or VMD are all variations of these algorithms. As an example, (Bagheri et al., 2014) propose a way to analyze electroretinogram[3] signals to detect and classify retinal diseases using EMD and neural networks (ANN). ERG signals are first decomposed by EMD, then, some statistical features are computed out of the IMFs forming the feature set that is fed later to a neural network for classification. This leads to 90% to 100% success rate for 100 to 500 milliseconds of the IMFs.

In another interesting work, (Lahmiri & Boukadoum, 2013) apply EMD on MRI images, compute the statistical features from the IMFs, and then classify the brain images of Alzheimer's disease vs. healthy, utilizing an SVM which results into 97.280.0329% accuracy. (Lahmiri & Boukadoum, 2014) have taken this further, to combine EMD with Student's probability density function creating a stronger and more accurate algorithm to decompose, and later classify by a neural network, synthetic signals, ECG, and S&P500 financial time series. This work shows faster processing time and lowers loss function values with a little twist in EMD. The RMSE in the proposed approach is lower than EMD for all of the signals. In a later work, (Lahmiri & Boukadoum, 2015) take the Student-adjusted 2-D EMD and apply it to images of the retina with blood hemorrhage. After calculating features to represent the pathology grade, an SVM classifies them into low and high-density blood hemorrhage. This method obtained 93.5% sensitivity, and 92.5% specificity. Furthermore, in (Keerthi Krishnan & Soman, 2021), the authors make use of the IMFs derived from VMD, and form EEG spectrum images of the motor imaginary EEG, by applying a short-time Fourier transform to them. The EEG spectrum images are later fed into multiple CNNs, among which, EEGNet (Lawhern et al., 2018) obtains the highest performance of 90.41% averaged over the four datasets employed.

In addition to the mode decomposition methods, Discrete Wavelet Transform (DWT) (Sundararajan, 2016) is also being abundantly used as an alternative decomposition method. In DWT, a given signal is decomposed into a set of time-series of coefficients that demonstrate the time evolution of the relevant frequency band (Hosseinzadeh, 2020). (Adam et al., 2018) try to find cardiac abnormalities in electrocardiogram signals for cardiovascular disease diagnostic, employing DWT. The feature set is made of the fifth-level DWT outputs along with some non-linear features being fed to the kNN classifier results into 99.27% acc. Further, (Murugappan et al., 2010) compute the significant features from the EEG data which later will undergo another feature computation process. Several time-frequency features will be obtained and fed to a Fuzzy model (Bezdek et al., 1984) for emotion classification. Also, in (Mohammadpour et al., 2017)'s work, DWT is applied to each 5-second chunk of EEG signal with a 1-second overlapping window to obtain the feature

[3]electrical activities of the retina to the light stimuli.

vectors. After reducing the dimension of the feature vectors using PCA, several classifiers are tested on them, among which, neural networks outperform others with 55.58% accuracy. Also, (Bajada & Bonello, 2021) show that EEG-based emotion recognition is possible with even a small number of channels. They use DWT to extract time-series features from 4- or 5-channel EEG signals. By adding Entropy and Energy features, they are able to improve the classification accuracy of the SVM to 95%.

In this research work, many of the methods, tools and algorithms discussed above were used in order to make progress and proceed during the work. Based on the literature that was reviewed, it was decided to benefit from these methods that fit properly to our problem, in order to obtain the desired results, as they claimed they were getting. At this point, before heading to the next chapters, it is important to take a look at these techniques.

## 2.4 A Summary of Approaches and Definitions

In this section, some of the terms and algorithms that are used in this thesis, and might not be trivial to every reader, will be defined and explained, so that their definition in the middle of the other chapters would not interrupt the flow of the work. All the following notions are used in this research work:

1. **Smoothness**: Smoothness (Gonzalez, 2009) of a signal is defined as its variance. The less the variance is, the smoother the signal is, and vice versa.

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

   where $x_i$ is each data point of a signal, $\bar{x}$ is the mean of all the data points in the signal, and $n$ is the number of data points in the signal.

2. **Skewness**: Skewness, or the third moment, is the measure of the symmetry of a distribution. Any normally-distributed data have a skewness of zero. Left- and right-skewed distributions have negative and positive skewness, respectively.

$$\tilde{\mu_3} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{(n-1) * \sigma^3}$$

   where $x_i$ is each data point of a signal, $\bar{x}$ is the mean of all the data points in the signal, $n$ is the number of data points in the signal, and $\sigma$ is the standard deviation.

3. **Kurtosis**: Kurtosis, or the fourth moment, is the measure of the flatness of the distribution curve.

$$\tilde{\mu_4} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^4 / n}{\sigma^4}$$

   where $x_i$ is each data point of a signal, $\bar{x}$ is the mean of all the data points in the signal, $n$ is the number of data points in the signal, and $\sigma$ is the standard deviation.

4. **Line-length**: Line-length (LL) (Esteller et al., 2004) is a measure that computes the changes

in the amplitude-frequency of a decomposed sub-band signal, because LL is proportional to both amplitude and frequency variations. (The sub-band is derived from an EMD-DWT approach.)

$$LL_x = \frac{1}{N_x - 1} \sum_{n=1}^{N_x - 1} abs[a_x(n+1) - a_x(n)]$$

where $a_x(n)$ is the $x^{th}$ decomposed sub-band signal, $N_x$ is the total length of $a_x(n)$, and $n$ is the index.

5. **Log-energy-entropy**: Log-energy-entropy (LEnt) (Das & Bhuiyan, 2016) characterizes the non-linear dynamics of EEG signals.

$$LEnt_x = \sum_{n=1}^{N_x} log([a_x(n)]^2)$$

where $a_x(n)$ is the $x^{th}$ decomposed sub-band signal, and $N_x$ is the total length of it.

6. **Norm-entropy**: Norm-entropy (NEnt) (Göksu, 2018) of the $x^(th)$ sub-band signal is defined as

$$NEnt_x = \sum_{n=1}^{N_x} [a_x(n)]^p$$

where $p = 1$, $a_x(n)$ is the $x^{th}$ decomposed sub-band signal, and $N_x$ is the total length of it.

7. **$L^p$ norm**: When a Fourier Transform is applied to a signal of length $L$, and the resulted decomposed signals are represented as $y(n)$, $L^p$ norm (Mehla et al., 2021) of the resulted Fourier intrinsic band functions is defined as:

$$L^p = (\sum_{n=0}^{L-1} |y(n)|^p)^{1/p}$$

for various values of $p$.

8. **Artifact Subspace Reconstruction**: Artifact Subspace Reconstruction (ASR) (Chang et al., 2019) cleans the data from the artifacts by taking chunks of the signal and evaluating it to see whether there are any. Firstly, in order to preserve the EEG-specific characteristics, such as alpha frequency, the signal undergoes an IIR filter[4] to remove the alpha band so it never gets removed during the process. Then, the algorithm segments the signal into 1-second windows, calculates each window's root-mean-square[5], z-scores them, and pick the windows with their z-score being within a certain threshold as the clean data. Next, by computing the covariance matrix of the clean data, the algorithm *learns* the principle component space of the data and is able to judge whether each chunk of data is clean or noisy. If noisy, it removes

[4]Infinite Impulse Response filter is a digital filter that can be applied on linear time-invariant signals with their impulse response continues infinitely never being exactly zero.

[5]$rms = \sqrt{1/n \sum_i x_i^2}$ where $n$ is the number of measurements, and $x_i$ is each value.

the noise to the point that the signal is closer to the previously learned clean data. Figure 2.1 shows a chunk of our data (1 channel) being denoised by ASR algorithm.

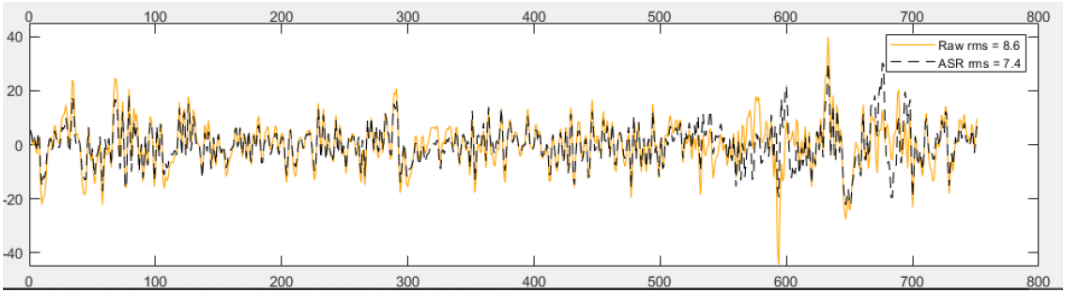


Figure 2.1: ASR artifact removal algorithm. One-channel raw EEG data (in yellow) and the denoised signal using ASR (in black). The root-mean-square of each signal is also stated in the top-right corner.

9. **Wavelet-Independent Component Analysis** wICA (Inuso et al., 2007) is an algorithm used for artifact removal and it is based on Wavelet Transform and Independent Component Analysis. In this algorithm, DWT first is applied to each channel of the signal, decomposing it into its four main bands: delta, theta, alpha, and beta. After all the bands of all the channels are gathered together, ICA is applied to them to cancel the artifactual traces. Then, to create the original signal, an inverse ICA is applied to the clean band signals, followed by an inverse DWT. The original signal is constructed, but clean! Figure 2.2 shows an example of the application of wICA on a channel of our EEG data.

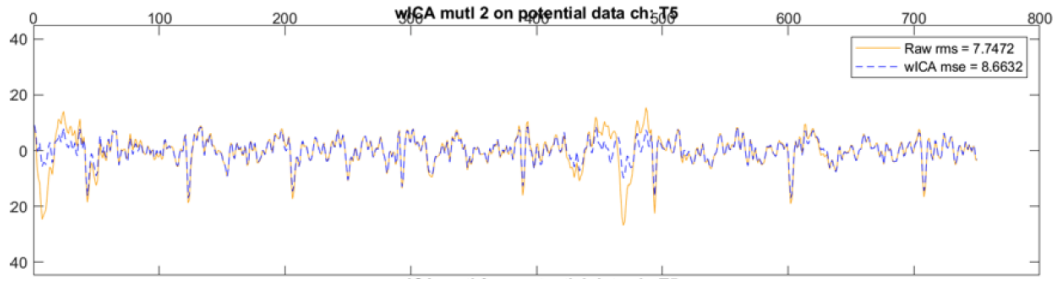


Figure 2.2: wICA artifact removal algorithm. One-channel raw EEG data (in yellow) and the denoised signal using wICA (in blue). The root-mean-square of each signal is also stated on the top-right corner.

10. **Empirical Mode Decomposition**: Empirical Mode Decomposition (EMD) (Huang et al., 1998) is a method that decomposes a signal into its Intrinsic Mode Function (IMF) components in a way that the summation of these IMFs and the residue would form the original signal. The IMF components have to have these two properties:

    - The mean of an IMF's values equals to zero,
    - Each IMF has only one extremum between every consecutive zero crossing.

    EMD decomposes the signal $s$ into $n$ IMFs and a residue $r$ according to the following steps (Lahmiri & Boukadoum, 2013):

- Firstly, the residue is considered equal to the signal $r_0(t) = s(t)$.

(a) Find all the local extrema of $s$,

(b) Compute the upper and lower envelopes of the signal by interpolation on the maxima and minima,

(c) Define the average of the upper and lower envelopes as the envelope's mean (EM),

(d) Compute the detail as $d(t) = s(t) - EM(t)$,

(e) check:

i. if $d(t)$ meets both the properties of an IMF, then $IMF_i = d(t)$ and then $r(t) = s(t) - IMF_i$

ii. otherwise, put $s(t) = d(t)$.

(f) Iterate over the previous steps until the residue $r(t)$ meets the stopping criteria[6].

Figure 2.3 shows EMD decomposition on one channel of our EEG data.

11. **Variational Mode Decomposition**: Variational Mode Decomposition (VMD) (Dragomiretskiy & Zosso, 2013) is a decomposition algorithm that is more optimal than EMD in finding the mother wavelet based on which, the signals are being decomposed. In VMD, when a real-valued signal is decomposed into a discrete number of modes, both the signal and the modes must be integrable and square-integrable up to two derivatives. In VMD,

- for each IMF, use Hilbert Transform (Hahn, 1996) to get the frequency spectrum,
- for each IMF, shift its frequency spectrum to baseband to its estimated center frequency,
- now, the bandwidth is estimated using $H^1$ Gaussian of the demodulated signal which is the squared $L^2$-norm of the gradient.

After computing the IMFs, they go through an optimization process so that the most optimal modes are resulted. For more detail on the optimization and re-construction procedures, please refer to (Dragomiretskiy & Zosso, 2013). An example of VMD algorithm on one channel of our EEG data is shown in figure 2.4.

12. **Discrete Wavelet Transform**: Discrete Wavelet Transform (DWT) (Sundararajan, 2016) is a decomposition algorithm that calculates how much of a wavelet there is in a signal for a specific scale[7] and location[8]. In this method, a wavelet of a scale is convolved over the entire signal, being multiplied to it at each location, resulting into coefficients of each time step. This process is repeated with multiple scales and locations with a finite set of wavelets. Figure 2.5 displays several wavelet families that can be used as the mother wavelet.

---

[6]Stopping criteria can be either the residual $r(t)$ not having the properties of an IMF anymore if continue decomposing, or any other criteria stated by the user such as a specific number of IMF.

[7]Scale of a signal represents its frequency properties. It is a measure of how stretched or slacked the wavelet is. The more stretched the wavelet is, the lower the frequency would be.

[8]The value of the wavelet at each time-point.

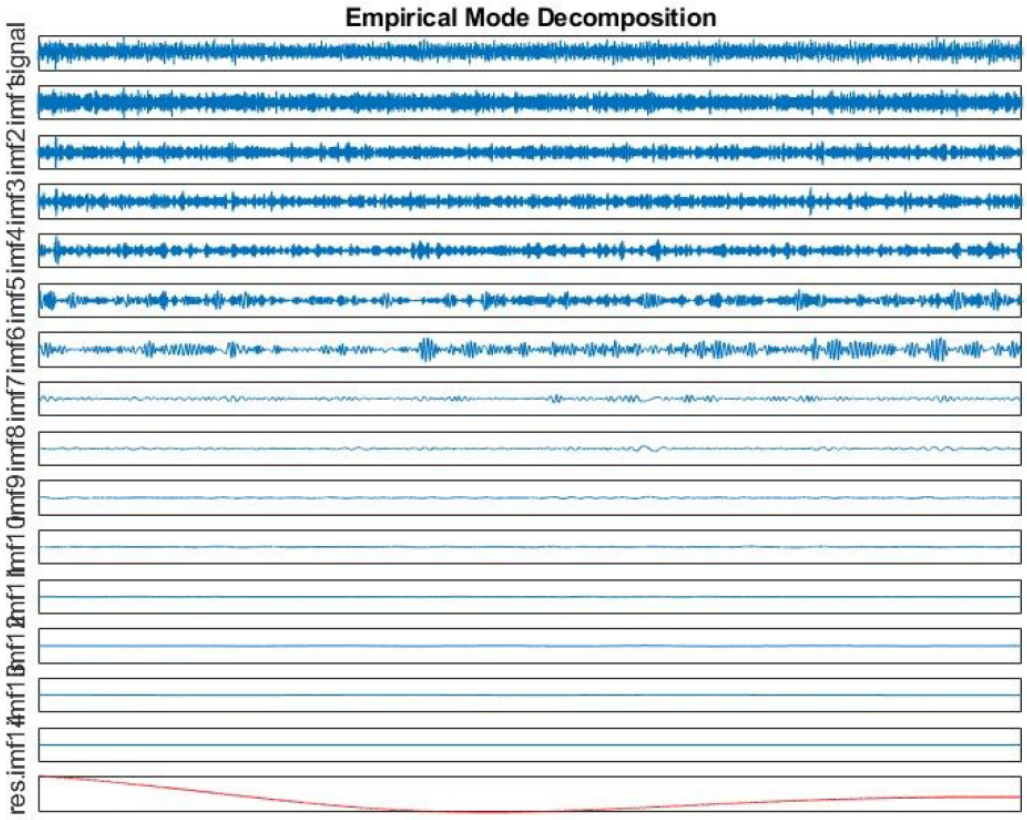


Figure 2.3: EMD. A single-channel EEG signal is decomposed into 14 IMFs and a residue. The x-axis represents time-resolution data points (time in seconds $*$ sampling rate (125 in our case)), and the y-axis shows the amplitude of the signals in $\mu V$.

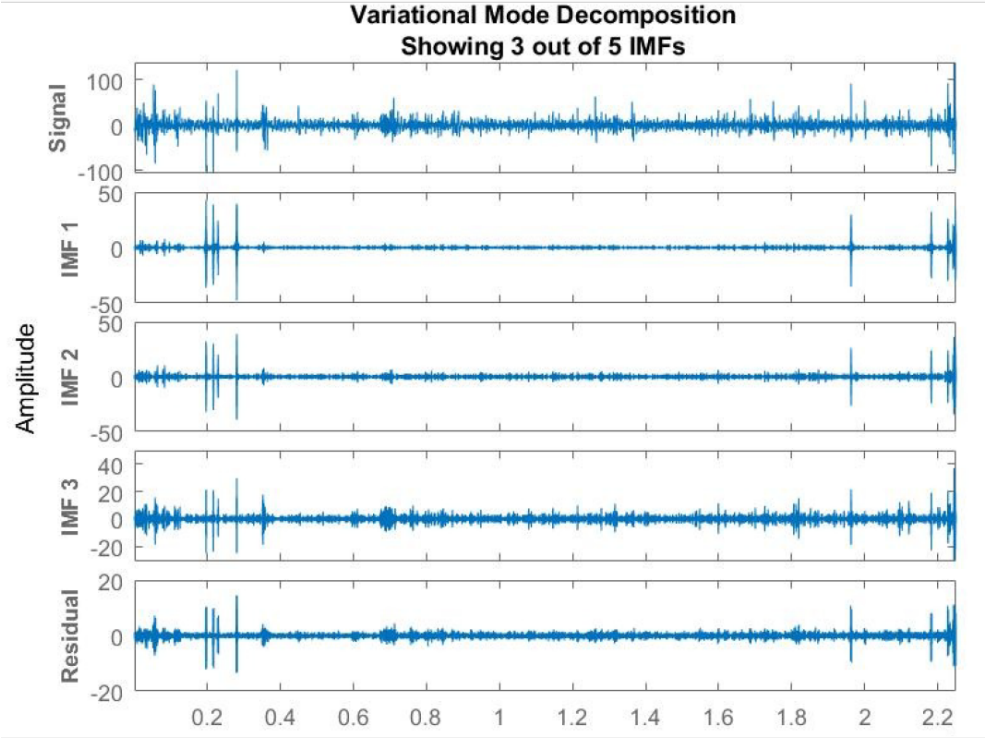


Figure 2.4: VMD. A single-channel EEG signal is decomposed into multiple IMFs, three of which are shown here, and a residue. The x-axis represents time-resolution data-points (time in seconds $*$ sampling rate (125 in our case)), and the y-axis shows the amplitude of the signals in $\mu V$.

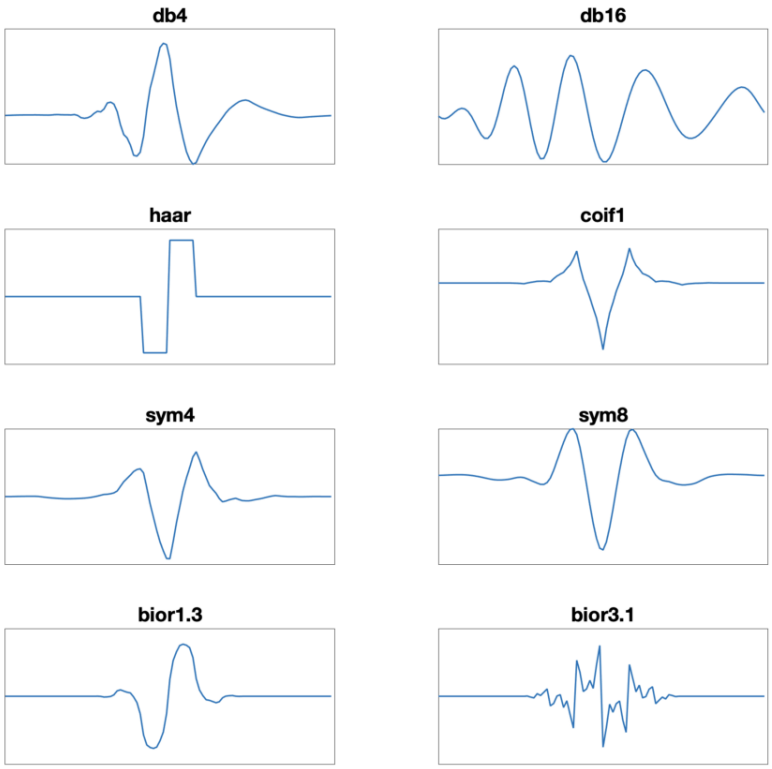


Figure 2.5: Wavelet families used in DWT. An example of several families of wavelets that are used in DWT to decompose signals using them: Daubechies, Haar, Coiflet, Symlet, Biorthogonal.[9]

In bi-dimensional DWT, using the mother wavelets, low-pass and high-pass filters are applied to both rows and columns, after downsampling, resulting in four coefficient matrices, as shown in figure 2.6. An example of 2D-DWT is demonstrated with a picture of a vase in (Gonzalez, 2009) in figure 2.7.

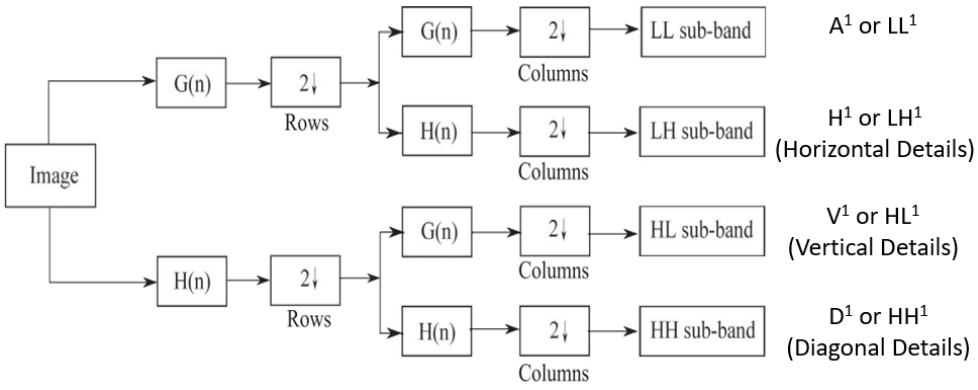


Figure 2.6: 2-D DWT. Firstly, low- and high-pass filters are applied to the rows, being downsampled, and then, both of these filters are applied to the columns of the resulted matrices, leading to four coefficient matrices: Low-Low (approximation), Low-High (horizontal), High-Low (vertical), and High-High (diagonal).[10]

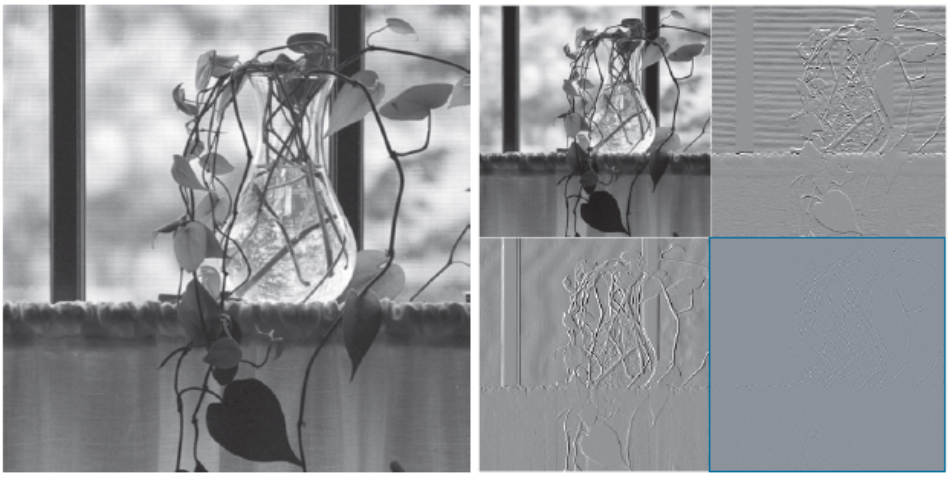

Figure 2.7: 2-D DWT on a picture of a vase (Gonzalez, 2009). The original picture of a vase is shown on the left, and the four outputs of the 1-level 2D-DWT algorithm are shown on the right: approximation (top-left), horizontal (top-right), vertical (bottom-left), and diagonal (bottom-right).

13. **Non-parametric statistical test**: The non-parametric statistical test proposed by (Maris

[9]https://miro.medium.com/max/1050/1*mkdL9Wjoj2MjbPtkrpoZjA.png

[10]https://miro.medium.com/max/974/1*JOE8S-9aIoMYdFeIdeoCyg.png

& Oostenveld, 2007) performs by taking the steps bellow:

- Pick samples of the two experimental conditions in one set,
- Random partition: draw as many trials of this set as there are in the first condition, and put them in subset 1, and the rest in subset 2.
- Calculate the test statistic on the set that resulted in the random partition step.
- Repeat the previous two steps many times so that a histogram of the test statistic can be made.
- From the histogram, compute the proportions ($p$-value) of the partitions that led to the greater test statistic.
- If the $p$-value is smaller than 0.05, we can infer the data in the two experimental conditions are significantly different.

14. **Auto-Regressive Moving-Average**: ARMA model (Box et al., 2015) describes a stochastic process in terms of two polynomials: auto-regression and moving average. An ARMA model is capable of modeling an unknown process with the minimum number of parameters, hence, it is capable of providing a practical and efficient linear model of stationary time-series (Zhang & Moore, 2014). Let $ARMA(m, n)$ be an ARMA model that characterizes a signal's time evolution:

$$ARMA(m, n) = c + \sum_{i=1}^{m} \Phi_i ARMA(t - i) + \sum_{j=1}^{n} \Psi_j \varepsilon(t - j)$$

where $m$ and $n$ are the orders of the AR and MA components, respectively, $c, \Phi$, and $\Psi$ the parameters to be estimated by the model, and $\varepsilon$ is Gaussian noise.

15. **EEGNet**: EEGNet is a convolutional neural network that has been proposed by (Lawhern et al., 2018) for EEG-based usage. As illustrated in figure 2.8, the model consists of three blocks. In the first block, there are two consecutive convolutional steps: firstly, $F_1$ feature maps are computed with the size being half of the input EEG's sampling rate, which contains the signal at various band-pass frequencies. Secondly, there is a deep-wise convolution layer to learn spatial filters of size $(1, C)$ ($C$ being the number of classes), per each temporal filter. The depth parameter $D$ takes care of the number of spatial filters. After Batch Normalization, there is a Dropout layer with the probability of 0.5 for within-subject classification, and 0.25 for cross-subject.

    In the second block, a separable convolution is used which is a depth-wise convolution followed by $F_2$ point-wise convolution. That helps decrease the number of parameters to fit, and a kernel learns each feature map and then optimally merges them. In this case, different feature maps may contain information from different time scales of the EEG signal. In the end, an Average Pooling layer is used for decreasing the dimension.

    In the last block, classification, the features are passed to a *softmax* classifier in order to be classified into $N$ classes.

16. **ReliefF**: To define the ReliefF algorithm, first the Relief algorithm has to be presented. Relief (Kira & Rendell, 1992) is a feature selection algorithm based on feature score and ranking. Imagine we have a dataset of $n$ instances, with $f$ number of features, in two classes. Firstly,

| Block | Layer | # filters | size | # params | Output | Activation | Options |
|---|---|---|---|---|---|---|---|
| 1 | Input | | | | (C, T) | | |
| | Reshape | | | | (1, C, T) | | |
| | Conv2D | $F_1$ | (1, 64) | $64 * F_1$ | ($F_1$, C, T) | Linear | mode = same |
| | BatchNorm | | | $2 * F_1$ | ($F_1$, C, T) | | |
| | DepthwiseConv2D | D * $F_1$ | (C, 1) | $C * D * F_1$ | (D * $F_1$, 1, T) | Linear | mode = valid, depth = D, max norm = 1 |
| | BatchNorm | | | $2 * D * F_1$ | (D * $F_1$, 1, T) | | |
| | Activation | | | | (D * $F_1$, 1, T) | ELU | |
| | AveragePool2D | | (1, 4) | | (D * $F_1$, 1, T // 4) | | |
| | Dropout* | | | | (D * $F_1$, 1, T // 4) | | $p = 0.25$ or $p = 0.5$ |
| 2 | SeparableConv2D | $F_2$ | (1, 16) | $16 * D * F_1 + F_2 * (D * F_1)$ | ($F_2$, 1, T // 4) | Linear | mode = same |
| | BatchNorm | | | $2 * F_2$ | ($F_2$, 1, T // 4) | | |
| | Activation | | | | ($F_2$, 1, T // 4) | ELU | |
| | AveragePool2D | | (1, 8) | | ($F_2$, 1, T // 32) | | |
| | Dropout* | | | | ($F_2$, 1, T // 32) | | $p = 0.25$ or $p = 0.5$ |
| | Flatten | | | | ($F_2$ * (T // 32)) | | |
| Classifier | Dense | N * ($F_2$ * T // 32) | | | N | Softmax | max norm = 0.25 |

Figure 2.8: EEGNet architecture (Lawhern et al., 2018). $F_1$: number of temporal filters, $F_2$: number of point-wise filters, D: number of spatial filters, N: number of classes, C: number of channels, T: number of time points, $p$: dropout layer probability $p = 0.25 or p = 0.5$ are the values of $p$ for cross-subject or within-subject classification, respectively.

all features should be between $[0 - 1]$. Secondly, for each feature vector $X$, we define near-hit and near-miss vectors as the feature vectors in the dataset that are closest to $X$ -by Euclidean distance[11]- and belong to the same and different classes as $X$, respectively. Then, an all-zero feature vector of size $f$ is defined as a weight vector $W$. Weights in $W$ are iteratively calculated by the formula below:

$$w_i = w_i - (x_i - nh_i)^2 + (x_i - nm_i)^2$$

where $w_i$ is an element of weight vector $W$, $x_i$ is a feature of the feature vector $X$, and $nh_i$ and $nm_i$ are elements of the near-hit and near-miss vectors, respectively. After $i$ iterations, each $w_i$ is divided by $i$ and if the result is more than a threshold $\theta$, then the feature gets selected.

ReliefF (Robnik-Šikonja & Kononenko, 1997) is an updated version of Relief that uses Manhattan L1-norm[12] instead of Euclidean distance, and absolute differences in the weight calculation formula instead of square of them. ReliefF can be used for multi-class classification in a way that it searches for $k$ near misses from different classes and average their contribution in updating the weight matrix $W$.

[11] $E_d = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$

[12] $M_d = \sum_{i=1}^{n} |x_i - y_i|$

# CHAPTER III

# EXPERIMENTAL SETUP AND METHODS

This chapter presents all the efforts that were done during a closed-source project at Ubisoft Montréal. It is worth mentioning that all the figures used in this section are based on the actual data that were used in this project, and since there is not access to this project anymore, the original figures are employed for illustration purposes.

## 3.1 Introduction

Video game companies are always looking for ways to liven up their products. A fascinating storyline, mesmerizing visual or sound effects, or fun and engaging features, are just a few examples these companies are interested in adding to their games, in order to attract more and more players. As a video game company, Ubisoft is no exception. So, the purpose of this project is to optimize player engagement in competitive video games, to a state where the player is entirely immersed in the game, where the player is at the highest level of engagement and immersion in the game, the *flow* state.

To address the objective and research questions of this project, an experiment was designed and several steps were followed that will be thoroughly explained below.

## 3.2 Data Collection

To be able to measure the state of flow of the players, a cohort of 27 players was recruited to play the first-person shooter battle royal game Apex Legends[1] at the rate of two consecutive half-days per participant. 16-channel Electroencephalographic (EEG), dermal conductance (EDA), photoplethysmography (PPG) [2], and eye tracking data were recorded throughout the experiment. The EEG data were recorded using OpenBCI 16-electrodes EEG caps, with passive Ag/AgCl electrodes, at a 125 Hz sampling rate (figure 3.1). For the eye tracking data, Tobii 4c sensors at 90 Hz sampling

[1] https://www.ea.com/games/apex-legends

[2] PPG is a non-invasive method that measures variations of blood circulation at skin level to monitor heart rate activities.(Castaneda et al., 2018)

rate, and for EDA/PPG, Shimmer 3 GSR+ sensors with both Skin Conductance and PPG sampled at 128 Hz were employed.

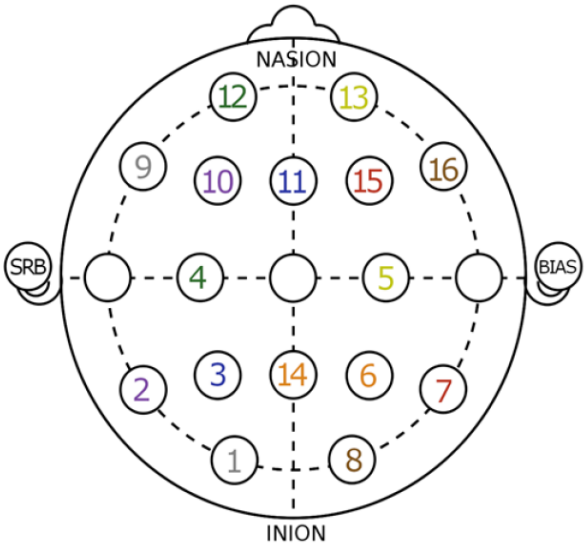


Figure 3.1: Electrodes placement. This top axial view of the brain shows the placement of the electrodes on the brain. Each electrode represents the channel number of the EEG signal. Nasion and SRB are the front and left side of the head, respectively.

In order to establish links between these physiological data and a subjective state of flow, the experience of the players was annotated in several ways:

1. **Retroactive questionnaire by segment**: Once a session of approximately one hour was over, a moderator guided the participant in an accelerated viewing of their gameplay. For each slice of three minutes (a segment) spent in an Apex Legends match, a short questionnaire was completed describing the experience lived on five axes related to the psychological construct of flow.

2. **Simultaneous task of fixing visual probes**: During gameplay, semi-transparent colored circles (probes) were displayed over the game screen at semi-random intervals in the edge regions of the screen. Participants were instructed to stare at these probes as quickly as they appeared while maintaining their in-game performance. hypothesizing that reduced fixation performance could indicate a high degree of immersion in the game.

3. **Per-match and global questionnaires**: At the end of each Apex Legends match, a questionnaire on the motivation to continue playing was administered. A final questionnaire was also completed at the end of the two half-days [on how likely the players are to play again].

One session of data gathering is shown in figure 3.2. The session begins with 1 hour of continuous gameplay, followed by 30 minutes of reviewing the game and answering the questionnaire. During the first hour, players play multiple matches. Depending on their skill level, their character might get killed soon which results in match termination in a short period of time, or they might actually


Figure 3.2: One session of data collection. Each session comprises one hour of continuous gameplay, followed by 30 minutes of reviewing and completing the questionnaire on motivation to continue.

live and win the match, consequently, play by the maximum time of each match, 20 minutes. So, the duration of the matches may vary up to 20 minutes, depending on the player's skill level. At the end of each match, the players are asked how likely they are to continue playing. Each match is then broken down into 3-minute segments. During the review time at the end of each session, the players are shown a record of their 1-hour gameplay and asked to fill up the questionnaire per each segment. Throughout the game, there are semi-transparent colored circles appearing on semi-random parts of the screen. These are called *probes*. These probes appear at random places within the quadrants of the screen and stay there for a few seconds (figure 3.3). If the player sees the probe, it is marked as *seen* or *hit*. Otherwise, it disappears after a few seconds and is marked as *not-seen* or *miss*. And as mentioned before, the biometric data was collected throughout the gameplay session. Tables 3.1 and **??** show examples of EEG and probe datasets, respectively.

Table 3.1: Scheme of an EEG dataset. Player number: 01 to 27, session number: 1 or 2, session start time: date and time when the recording of the session starts, session end time: date and time when the recording of the session end, and the EEG data: a matrix of EEG with 16 rows and 1 hour of data points.

| EEG dataset | | | | |
|---|---|---|---|---|
| Player number | Session number | Start of session | End of session | EEG data |
| 01 | 1 | 2019-10-08 13:08:54 | 2019-10-08 14:15:08 | [16-channel EEG data] |
| 02 | 2 | 2019-09-18 11:20:10 | 2019-09-18 12:17:31 | [16-channel EEG data] |
| 03 | 1 | 2019-10-09 15:33:10 | 2019-10-09 16:40:51 | [16-channel EEG data] |

Table 3.2: Scheme of a probe dataset. Player number: 01 to 27, session number: 1 or 2, probe number: 1 to the number of probes that appeared during the session, probe time: at what time the probe appeared on the screen, and the label of the probe: 0 indicating not-seen or miss, 1 indicating seen or hit.

| Probe dataset | | | | |
|---|---|---|---|---|
| Player number | Session number | Probe number | Probe time | Label |
| 01 | 1 | 01 | 13:09:03 | 1 |
| 01 | 1 | 02 | 13:09:013 | 1 |
| 01 | 1 | 03 | 13:09:30 | 0 |

## 3.3 Research Questions

As mentioned before, the question we were interested in answering was:

**Research Question**: Is there any relation or link between EEG data and mental states like flow?

To answer this question with EEG data on hand, we also made a hypothesis to be able to move forward in our research path in a more specific, detailed, and focused way. It was considered that if a player is deeply immersed in the game and is highly focused on the events that are going on within the game, they probably should not notice the probes appearing on the screen. The probes were being displayed within the corner quadrants of the screen, not right in the middle, so, if the players are concentrated on their character and the shooting actions, they might not notice the probes appearing on the outer part of the focus zone. Consequently, if there were fewer missed probes during a player's gameplay, they were probably not so much engaged and immersed in the game.

**Hypothesis**: The more probes the players looked at, the more immersed they were in the game.

## 3.4 Methodology

Based on what has been discussed about the problem we were trying to solve, it could be agreed upon that there exists a classification problem. With EEG data on hand, we are interested in classifying them into two classes hit vs. miss. In the general run of things, these types of problems are dealt with by processing the data, extracting features, selecting the best among them, and feeding them to a classifier. We, in particular, followed the same routine in this thesis. In bellow, the set of methods that were employed to address the problem in each of the steps are explained. It is also worth mentioning that all the coding was done in MATLAB, except for the classification part which was done in Python.

Figure 3.3: Probe. A red probe is displayed at a random place within the top-right quadrant of the screen. This scene is a snapshot of the game being played by player 08, session 1, segment 3/4.

### 3.4.1 Data Processing

The preprocessing phase starts with data cleaning and artifact removal. EEG signals have low amplitudes in order of $\mu V$, so even subtle artifacts can easily corrupt the signal, e.g. other neural activities, or the scalp and the hair getting in the way and obstructing the electrodes recording EEG, or even head or body movements. Detection and removal of these artifacts is not an easy job because they are both non-stationary data, or even the frequency bands of both signals might be overlapping. ASR algorithm has shown to be a powerful artifact removal approach that is able to omit large-amplitude or transient artifacts. Hence, we primarily used ASR to denoise the raw EEG data recorded from the players. Later on during the project, after taking an in-depth look at the EEG data once cleaned through ASR, while ASR did seem to lower the amplitude of the total signal, it was seemingly modifying the shape of the signal in relation to the noise. This was especially true when noise was dominant for a specific probe in a channel. We can see from figure 2.1 that the left part of the original EEG data (in yellow) was modified by ASR significantly more on the right side due to the higher presence of noise. That was when we employed wICA instead of ASR, an automated wavelet-based denoising approach on the data and compared it against the raw EEG signal. Figure 2.2 illustrates the expected results across all probes. It can be noticed that this approach, while less aggressive than ASR, does seem to do a better job at keeping the shape of the signal through the noise, while not altering the signals with better quality.

After running the wICA across all probes, we computed statistics of probe data quality across all players in order to evaluate an approach that would help us keep the best possible quality while removing the most noise. We identified some of the noisiest channels by calculating the peak-to-peak difference between the highest and the lowest amplitude of the channel, seeing that they were the ones causing the most noise across probes. We removed the brain regions in pairs so to keep a balanced set at the end. Further to noise removal, normalization of the data can also affect the

performance of the whole model as the dataset is more consistent, less redundant, and easier to calculate since all data points are within a specific range.

Then, to segment EEG, we were interested in the 5-second slices of the EEG signal that happened prior to the appearance of the probe. Because in this case, the player should be in a more concentrated state than during or after the probe display since probe appearance, if seen by the player, distracts the player. So, the one-hour EEG sessions were divided into 5-second-prior-to-the-probe chunks of EEG. From now on, it would be referred to as the probe EEG for ease of reference. Each probe EEG has 16 channels and 5 seconds$\times 125Hz$ sampling rate, 625 data points in time.

The second phase of data processing was to decompose the signals in order to find the rhythms and important information the signals may carry. We believed that EMD and VMD algorithms are the best choices to do so. As the EEG signal is non-stationary, the mode functions derived from the decomposition contain the dynamic characteristics of the signal. First, EMD was applied to every channel of the EEG signal, but later on, we decided that VMD gives a better result in decomposing the signals, as it gives the optimal IMFs, so it captures the essence of the signal in a better way. As the result of having too many IMFs per EEG signal, we then replaced VMD with 2D-VMD to be able to decompose all of the channels at once. Figure 3.4 shows an example of 2D-VMD decomposition. Each IMF also has 16 channels, for a closer look, figure 3.5 shows us each channel of an IMF in a bi-dimensional VMD decomposition.

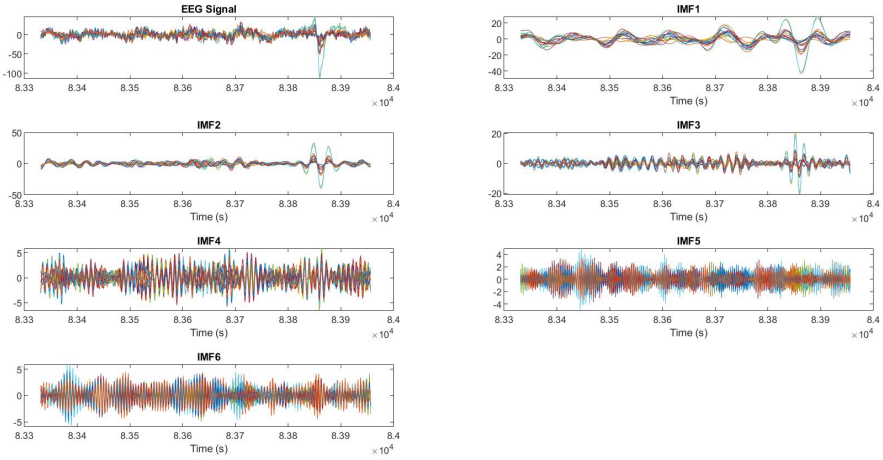


Figure 3.4: Bi-dimensional VMD on 16 channels of an EEG segment. Each channel is drawn with a unique color. The y-axis represents the amplitude of the signal in $\mu V$.

In addition to EMD and VMD, we used 2D-DWT algorithm for decomposition and size reduction. As DWT outputs are able to capture certain band-passes (which are informative) in the signal, and each output is half in size compared to the input signal, DWT is a great choice for both decomposition and size reduction of the EEG signals. 2D-DWT applies low- and high-pass filters on both rows and columns of the input matrix which in our case, would be capturing features such as fast- or slow-moving phenomena across time and channels. Another advantage of DWT is data augmentation. With each time 2D-DWT application on a signal, there will be four times as many data as input, e.g. level one of 2D-DWT on a signal gives four output signals.

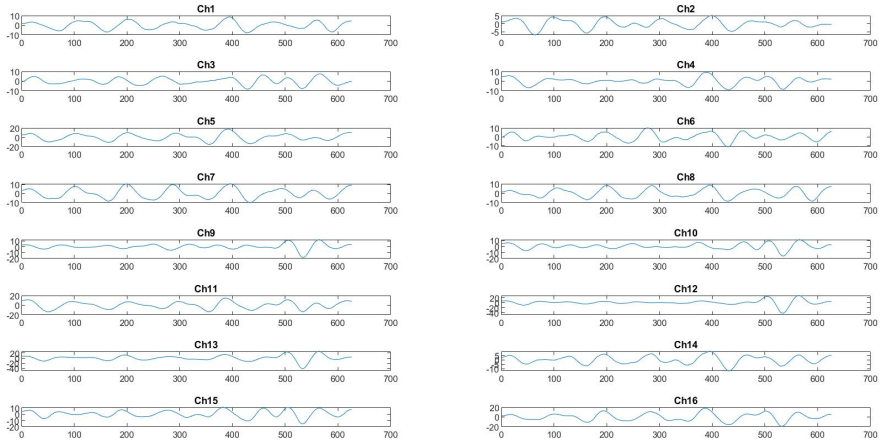


Figure 3.5: Channels of IMF-1 of a 2D-VMD decomposition. If all the channels of the first IMF shown in figure3.4 are plotted separately, it would look like the ones in this figure. Only 5 seconds of the channels are plotted. The axes show time and amplitude.

### 3.4.2 Feature Extraction

As some works in the literature were reporting promising results and based on our own knowledge that statistical features are able to capture the importance of the signals, we extracted two different sets of statistical features in two different approaches. Set $A$ =[mean, standard deviation, smoothness, skewness and kurtosis] (Lahmiri & Boukadoum, 2013), and set $B$ =[line-length, log-energy-entropy, norm-entropy, and $L^p$ norm] (Mehla et al., 2021) and (Anuragi et al., 2021).

Further down the road, an interesting point captured our attention. As shown in figure 3.1, each channel of EEG captures the activity of a particular region of the brain. And since all brain regions never have the exact same pattern activity simultaneously, having a signal with the same values through all channels at time step $t$, is not natural, and is probably caused by noise. Synchrony is a name that is given to this situation when all the channels of a signal have similar values durong a period of time. And because synchrony is mostly caused by noise or artifacts, we aim to get rid of them in order to have valuable parts. Hence, at some stage, as a way to embolden informative and essential parts of the EEG signal, we set the synchronous parts of EEG to zero and later remove the all-zero parts from the signal to reduce the sparsity.

At some point during the project, we decided to consider the probe signals as images. An image is basically a matrix, and so is a $16 \times 625$ probe signal. In this work, we chose to make one single image per each probe signal's IMFs, by stacking the IMFs resulting from the decomposition algorithm on top of each other, to form a larger image with all the IMFs included. So, after min-max normalizing the IMFs, all the IMFs were stacked on top of each other to form a matrix. In this case, we have all the information in all the IMFs all at once within one single image. Then, these probe images could be used with a CNN for classification purposes. Figure 3.6 shows some random examples of these probe images plotted in a heatmap. Heatmaps enable us to have an understanding of the amplitude

of the signal at each time step with help of the color map.

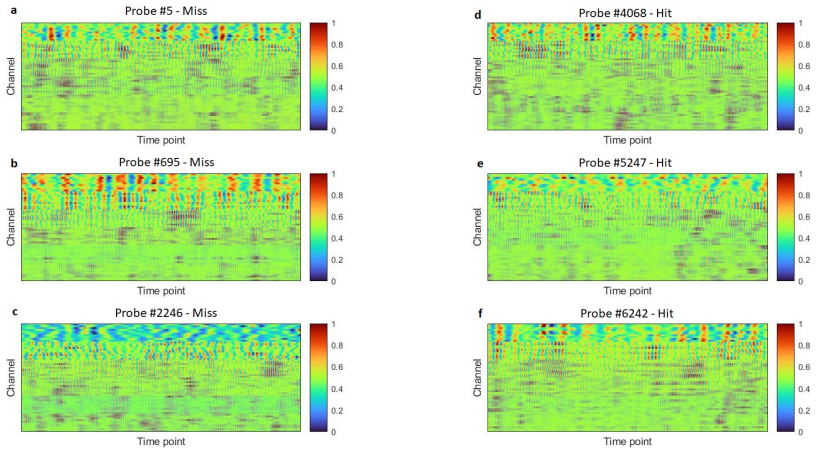


Figure 3.6: Heatmap of probe images. Probe images are plotted in heatmap form. $(a), (b), (c)$ show some random probes with miss label, and $(d), (e), (f)$ show some hit-labeled random probes.

### 3.4.3 Feature Selection

At the stage where the statistical features were calculated, to be able to select the most distinctive ones, they were ranked using the ReliefF algorithm (section 2.4), and the top 5 or 10 with the highest scores among others were selected based on our needs. The details of the procedure will be explained in the next chapter.

In some other parts of the research where we were exploring other techniques and using approaches other than statistical features, the features were selected based on their own characteristics and our needs at that point. For instance, after applying DWT to the EEG signals, the *approximation* coefficients were selected as the feature set, because we needed the closest feature to the actual signal.

All of the feature selection choices will be clarified in chapter 4.

### 3.4.4 Classification

In this thesis, the SVM was mainly used for feature classification and EEGNet for probe image classifications. Both of these tools have been proven to have great classification capabilities in the literature. Hence, SVM was employed due to the nature of the problem being a 2-class classification and SVM is a very popular means for that. Also, EEGNet is an EEG-based CNN especially designed for EEG classification that makes a great choice for classifying probe signals in form of an image.

In addition, in order to address over-fitting and feature selection in EEGNet, we used L1[3] and L2[4] regularization.

In the next chapter, the detail of how the approaches were used, along with the specific parameters will be explicated. In addition, the results obtained per each method are also reported.

[3]**Lasso Regression** shrinks the less important feature's coefficient to zero by adding the absolute value of the magnitude of the coefficients to the loss function as penalty.

[4]**Ridge Regression** shrinks the less important feature's coefficient to zero by adding the squared magnitude of the coefficients to the loss function as penalty.

# CHAPTER IV

# RESULTS

In the previous chapter, all the different methods and algorithms that were used in this project were mentioned. In this chapter, all the different approaches we carried out using the aforementioned algorithms will be explained, along with their parameters and outputs.

In the first place, the raw EEG data acquired from the participants were denoised using ASR algorithm. Then, the signals were segmented into 5-second chunks prior to the appearance of the probes. Then, each segment was decomposed into its mode functions using EMD. We noticed -by experience- that six IMFs could properly capture the essential information within the signals and leave a less-informative residue, so, each channel of the probe EEG chunks was decomposed into six IMFs and a residue that we decided to keep. Next, five statistical features were calculated for each of the resulting IMFs: mean, standard deviation, smoothness, skewness, and kurtosis. So, each of the 16 channels of probe EEG segments were decomposed into 6 IMFs plus a residue, from which the mentioned statistical features were computed. Figure 4.1 shows a scheme of the resulted dataset.

| ch1_IMF1_ mean | ch1_IMF1_s tddev | ... | ch2_IMF1_ mean | ... | ch16_IMF6 _kurtosis | PlayerNb | SessioNb | ProbeNb | Hit |
|---|---|---|---|---|---|---|---|---|---|
| 5.4404 | 1.2355 | ... | 6.3701 | ... | 2.087 | 1 | 1 | 1 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 0.0014 | 3.4358 | ... | 3.2514 | ... | 3.3485 | 49 | 2 | 147 | 1 |

Figure 4.1: A scheme of the dataset of the statistical features. Every feature is computed for every channel of every IMF. chX_IMFY_mean: mean value of the $Y^{th}$ IMF of the $X^{th}$ channel of the EEG signal corresponding to a specific probe in a particular session of a single player. The label of the probe: 1 for seen and 0 for not-seen, is under the column *Hit*.

**Results**: The features were split to train and test sets with the ratio of 0.2 test data and 0.8 train data. The raw features were fed into an SVM with the *rbf* kernel to catch the similarities, as well as any non-linear relation between the data points. The $C$ parameter of *rbf* was set to 1 for correctly classifying the train data. The accuracy of SVM was 0.58. Next, we applied *z-score* normalization[1] to the features, and added 5- and 10-fold cross-validation, which all resulted into 0.55 accuracy, 0.41 precision, and 0.41 F-score averaged over the folds in the best case.

[1] Z-score is a measure of closeness to the mean in terms of standard deviation. $score = \frac{x-m}{\sigma}$, where $x$ is a data point, $m$ and $\sigma$ are the mean and standard deviation of the sample, respectively.

In our second attempt, firstly, we decomposed all the probe EEG signals using VMD (applied VMD on each channel individually). We kept six IMFs derived from VMD, and ignored the residue as the signals seemed not informative. Then, another set of statistical features was calculated: Line-length, log-energy-entropy, norm-entropy, and $L^p$ norm, with $p = 0.1, 0.25, 0.5, 1, 2, 3, 5$. Four new features with ten different values were computed for each channel of each IMF, in total $10 \times 6 \times 16 = 960$ features. Since this many features are a lot for a classifier to take in, we computed the average of each of these features across the channels, so that each IMF would be having a feature vector of size 10, instead of 160. Later, they were ranked by their importance and effectiveness using ReliefF and the ten best features were fed to an SVM[2]. Depending on the data in the train and test split each time, because they were chosen randomly, the ten best features in the ranking were different each time.

**Results**: With the same train-test split of $0.8 - 0.2$, an SVM was applied to the feature data. All three kernels: *linear*, *polynomial*, and *rbf* were used separately with the $C$ parameter of 1 and 0.1, leading to the accuracy of 0.56, 0.56, and 0.55, respectively, in the best case (averaged over 5 folds of cross-validation, and unfortunately, the F-score is not available). We repeated this method with different data samples and ranked features each time, but the performance of the SVM did not differ significantly.

At this point in the research, we decided to run wICA on the raw EEG data instead of ASR, to remove the artifacts. After cleaning the data with wICA, EEG signals were all z-score normalized across all the players, in a way that each data point was normalized with regard to the corresponding data points of the other players. Then, the signals were again segmented into 5-second probe chunks. Afterward, 2D-VMD was applied to the 16-channel probe EEG segments resulting in 6 16-channel IMFs. The same set of features as the previous attempt was calculated on the IMFs and the ten best of them were chosen based on the ranking given by ReliefF.

**Results**: With the same train-test split of $0.8 - 0.2$, an SVM was applied to the feature data. With *rbf* kernel (because there was no significant difference in the results choosing different kernels), the $C$ parameter of 1 and 0.1, along with 5-fold cross-validation, and the best result was the accuracy of 0.57.

To further remove noise, the six noisiest channels: 16, 9, 12, 13, 1, and 8 were identified based on statistics of probe data quality across all players and were removed. Later, the 10-channel EEG signals were segmented into 5-second probe signals and then decomposed via 2D-VMD into 6 IMFs. The IMFs were min-max normalized and stacked on top of each other to form a matrix. Figure 3.6 shows some examples of these matrices in heatmap form. To see whether there is any difference between the probes of the two classes, the average of all the Hit and Miss probes was calculated and plotted (4.2). As there were some visual differences between the two groups, We decided to statistically analyze the images of these two classes to find the clusters of pixels where the images of these two groups are different.

**Results**: After running a non-parametric statistical test on about 1 second of these images (because the test would take too long to execute), numerous clusters were found. In figure 4.3, to be able

[2] An example of the ten best features in the ranking could be like: [ch1_IMF5_mean, ch5_IMF2_kurtosis, ch11_IMF3_smoothness, ..., ch1_IMF7_mean]

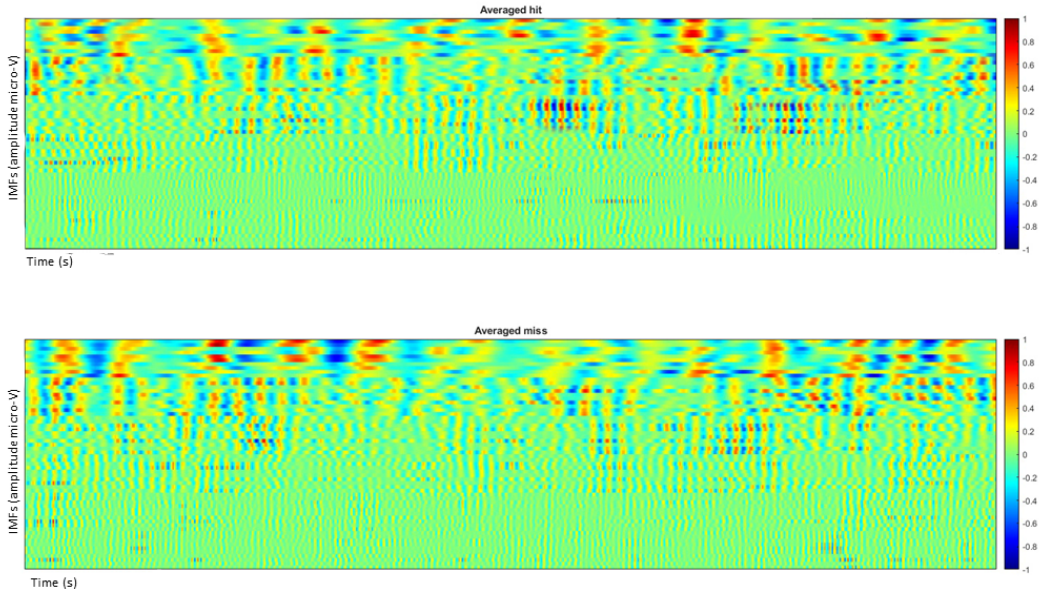


Figure 4.2: Average of all the Hit and Miss probe images. Average of all the IMFs was computed and then stacked on top of each other to form the average images.

to have a better visual understanding, the values where the different clusters are positioned are set to $-100$ (blue regions), and the rest of the image stays intact (red regions), so there would be a distinct color difference between the values which enables us to visually differentiate them better.

Each probe (whole 5 seconds) was averaged along the columns to get 1D vectors of size 625 (5 seconds $\times 125$ sampling rate). Then, the same statistical non-parametric test was applied to these 1D-vectors, but unfortunately no different clusters were recognized by this algorithm.

In the next step, as $60 \times 625$[3]- pixel images were too complicated for a CNN, we compressed the probe images using 2D-DWT. As shown in figure 4.4, each probe image was decomposed and compressed into four matrices using "*db4*" (Daubechies family) wavelet. We also went further in compression and applied 2D-DWT once more on the approximation matrix resulted from the first-level 2D-DWT, to get four other matrices that are even smaller in size. After all, the dataset consisted of 1000 approximation matrices as the training set, and 500 of the same matrices as test, per each class (3000 matrices in total).

**Results**: By splitting the data into 2000 and 1000 matrices that were chosen randomly as the train and test datasets, respectively, we fed them to EEGNet with all the parameters being the default values. Table 4.1 shows the results. As the next effort, some changes were made to the EEGNet input parameters: we set the dropout rate to 0.1, the batch size to 128, the number of epochs to 5000, removed the average pooling layers (because the pooling layers were omitting the minute differences there were in the probe images), and added L1 and L2 regularization with both L1 and L2 values as 0.0001 as the kernel regularization (from this point to the end of this research, EEGNet's

[3]rows: 6 IMFs $\times 10$ channels, and columns: 5 seconds $\times$ 125 sampling rate

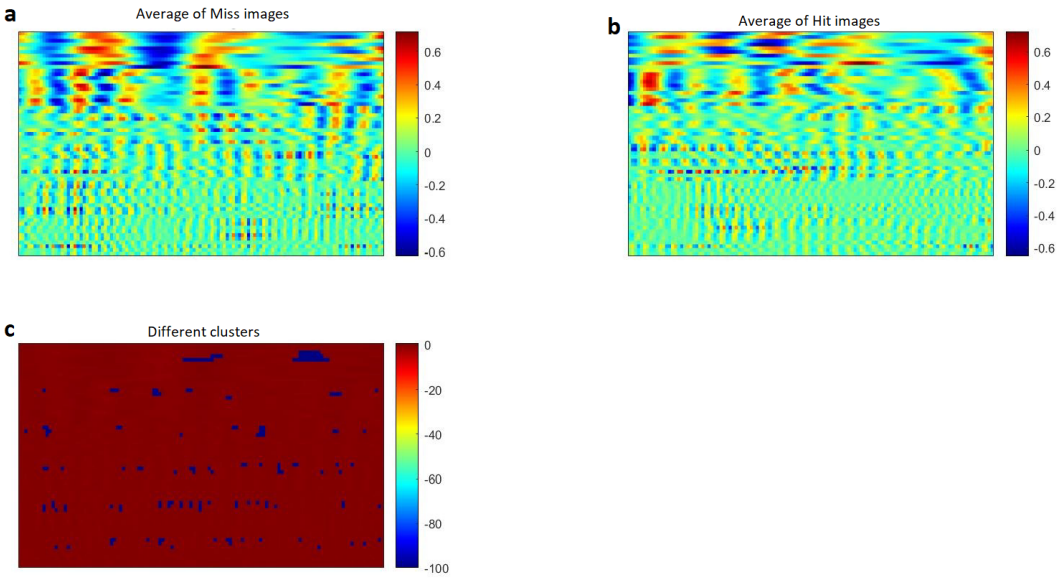


Figure 4.3: Different clusters between the Hit and Miss images identified by the non-parametric statistical test. *a* and *b* show 1 second of average images of all the Miss and Hit probes, respectively. The rows are the channels of the IMFs and the columns represent the data points during time. *c*: The values of the identified clusters were set to $-100$ in order to create a big color difference between the clusters (blue) and the rest of the image (red).

parameters stayed the same as explained here). Then, we put in all the four outputs of 2D-DWT to EEGNet, one-by-one, for each level. From now on, all the values reported are the average of the values obtained by 5-fold cross-validation. Table 4.2 shows the performance of EEGNet in this effort.

Table 4.1: EEGNet performance metrics on level-1 and level-2 DWT approximation matrices.

| DWT component | Train ACC | Test ACC | Train loss | Test loss | F-score |
|---|---|---|---|---|---|
| Apprx - L1 | 0.82 | 0.53 | 0.39 | 0.88 | 0.54 |
| Apprx - L2 | 0.78 | 0.52 | 0.47 | 0.80 | 0.51 |

Table 4.2: EEGNet performance metrics on all level-1 and level-2 DWT output matrices.

| DWT component | Train ACC | Test ACC | Train loss | Test loss | F-score |
|---|---|---|---|---|---|
| Apprx - L1 | 0.82 | 0.53 | 0.39 | 0.88 | 0.54 |
| Vertical - L1 | 0.80 | 0.51 | 0.42 | 0.89 | 0.51 |
| Horizontal - L1 | 0.82 | 0.51 | 0.39 | 0.90 | 0.52 |
| Diagonal - L1 | 0.81 | 0.50 | 0.39 | 0.89 | 0.52 |
| Apprx - L2 | 0.78 | 0.52 | 0.47 | 0.79 | 0.51 |
| Vertical - L2 | 0.79 | 0.51 | 0.46 | 0.81 | 0.51 |
| Horizontal - L2 | 0.78 | 0.51 | 0.47 | 0.80 | 0.52 |
| Diagonal - L2 | 0.79 | 0.50 | 0.45 | 0.83 | 0.50 |

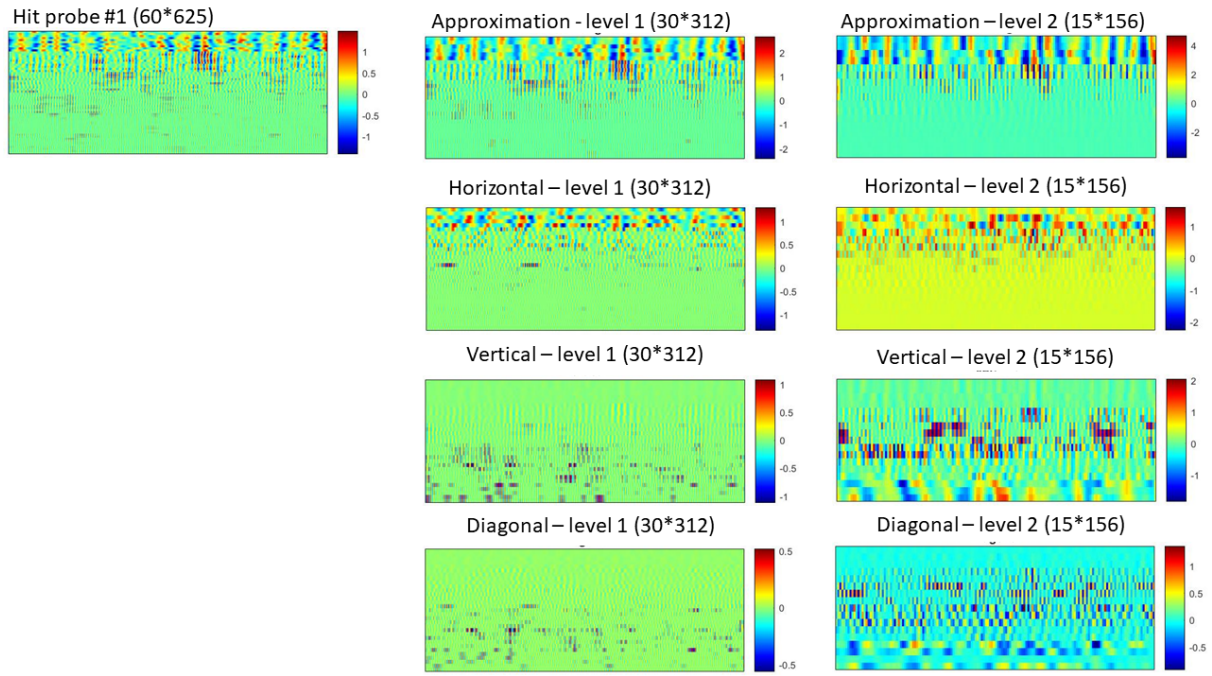


Figure 4.4: Two levels of 2D-DWT on a probe. The first-level application of 2D-DWT on a hit probe results into the four images (in the middle) that are also half in size. The second-level application of 2D-DWT on the first-level approximation matrix gives the four matrices (on the right) that are smaller in size by four times than the original probe.

In the next step, we used the probes that appeared only during combat scenes (aiming, shooting, etc.) in the game, because, hypothetically, during a combat, the player is probably at the highest level of concentration and flow. 20% of all the probes (1300 probes) were randomly selected and manually annotated which resulted into 230 combat probes in total (115 probes per Hit/Miss). These combat probes were then decomposed into 6 IMFs using 2D-VMD, and the IMFs were stacked on top of each other to form the combat images. After applying DWT on the combat images, each matrix was fed into EEGNet, separately, resulting in table 4.3's performance.

Table 4.3: EEGNet performance metrics on level-1 DWT matrices on combat probes.

| DWT component | Train ACC | Test ACC | Train loss | Test loss | F-score |
|---|---|---|---|---|---|
| Apprx - L1 | 1.0 | 0.45 | 0.08 | 2.16 | 0.47 |
| Vertical - L1 | 1.0 | 0.40 | 0.09 | 1.05 | 0.40 |
| Horizontal - L1 | 1.0 | 0.47 | 0.06 | 0.92 | 0.48 |
| Diagonal - L1 | 1.0 | 0.47 | 0.07 | 0.97 | 0.49 |

In another attempt, the average of all the Hit and Miss combat images was calculated and then z-score normalized, as shown in figure 4.5. After examining the average images, we chose to remove the parts of the signals where all the channels were synchronous. For synchrony removal, we decided to set a threshold and set all the values within the threshold to zero, so that we would be left with the salient spikes. To calculate the threshold, Hit and Miss datasets were split into train and test parts, randomly chosen. Then, the average of the training Hit images and Miss images were separately calculated so that any other data of the test images would not be leaking to the training set. Then,

- the difference between the average-Hit and average-Miss was obtained,
- got the absolute value of the first IMF (since the first IMF is the most similar IMF to the actual signal),
- split it to 5 consecutive windows,
- get the mean and standard deviation of each window,
- take the mean of all these values to get one mean and one standard deviation for the whole signal,
- and lastly, we set all the values between -(mean + standard deviation) and +(mean + standard deviation) to zero.

This is how we obtained the synchrony-removal mask for Hit and Miss combat images, as in figure 4.6 *(b)*. After applying each mask to its relative dataset, resulting images were fed to EEGNet.

**Results**: The dataset consisted of 87 training images and 22 images for test, per each class. Table 4.4 shows how the results turned out.

In order to reduce the problem space even more, the parts of the signal where the values were all zero, due to synchrony removal, were omitted. So the matrices would turn from figure 4.6 *(b)* to *c*. This would reduce the sparsity of the matrices. We tested EEGNet on the resulting matrices: table

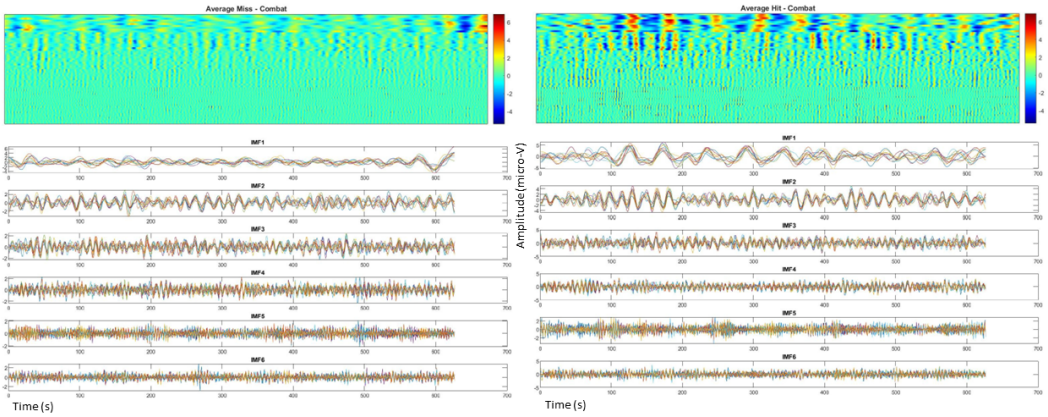


Figure 4.5: Average of all the 2D-VMD decomposed Hit and Miss combat probes in heatmap and signal views.

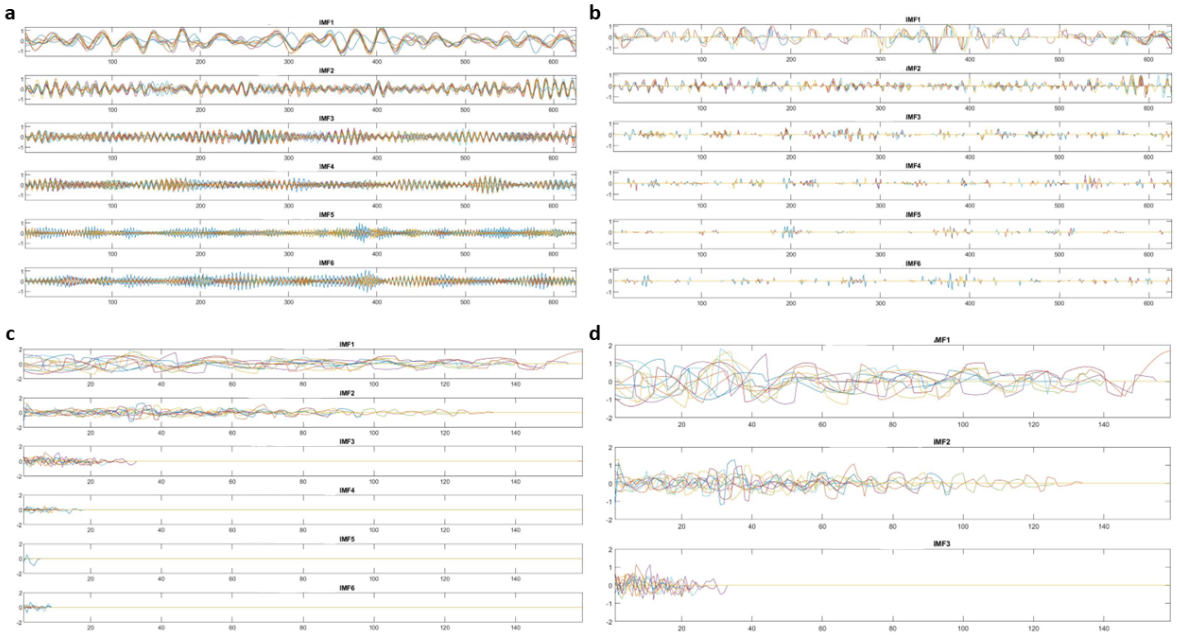


Figure 4.6: The process of synchrony removal, zero removal, and IMF removal of the combat probes. *(a)*: Miss combat probe number 50 of size $60 \times 625$. *(b)*: Matrix in *(a)* after synchrony removal, of size $60 \times 625$. *(c)*: Average of the Hit combat probes after removing all the zeros left from synchrony removal, of size $60 \times 159$. *(d)*: *(c)* after removing the last three IMFs due to sparsity, size $30 \times 159$. The x- and y- axes represent time(s) and amplitude($\mu V$).

Table 4.4: EEGNet performance metrics on asynchrious Hit and Miss combat probe images.

| Train ACC | Test ACC | Train loss | Test loss | F-score |
|---|---|---|---|---|
| 1.0 | 0.36 | 0.01 | 1.27 | 0.54 |

4.5. Because at this point, at least the last three IMFs were super sparse, thus not informative, they were removed, as in figure 4.6 *(d)*. EEGNet was again applied to the resulting matrices: table 4.6.

Table 4.5: EEGNet performance metrics on asynchrious Hit and Miss combat probe images after removing the zero parts.

| Train ACC | Test ACC | Train loss | Test loss | F-score |
|---|---|---|---|---|
| 1.0 | 0.40 | 0.02 | 1.42 | 0.54 |

Table 4.6: EEGNet performance metrics on asynchrious Hit and Miss combat probe images after removing the zero parts and the last three sparse IMFs.

| Train ACC | Test ACC | Train loss | Test loss | F-score |
|---|---|---|---|---|
| 0.99 | 0.45 | 0.08 | 1.02 | 0.54 |

As exhaustively described in this section, through all the steps that were taken, the results were not conclusive. In the next chapter, the potential reasons are discussed, and a new experiment will be undertaken for further proof.

# CHAPTER V

# DISCUSSION AND EXPERIMENTS WITH EMG

In the previous chapter, the methods we used were described, as well as the paths we went through to be able to find patterns that enable us to detect mental flow states in EEG recordings in the players' game-play. Yet, the desired results were not obtained. In this chapter, the reasons will be discussed.

In the very first attempt, when the SVM failed to properly classify the statistical features, the mutual information of the features was calculated that were all in the range of $[0 - 0.025]$ which is too low for any of the features to have any connection with the target. So, none of the features could help the classifier to make a distinction between the two classes. This was almost the case for the failure of the second set of statistical features, too. In this case, the mutual information between the ten selected features was obtained, and it was in the range of $[0 - 0.02]$ which is still lower than what is expected for making a distinct classification.

To explain the poor outcome of the non-parametric statistical test, even when we reduced the size of the images to 1D vectors, we believe that averaging the columns might have crushed the superpixels that were forming the different clusters. Thus, crushing the information they bore, resulted in finding no different clusters.

As for all the trials we did, using EEGNet to classify the Hit vs. Miss probes using probe images or in-combat probe images, there are a couple of explanations. First of all, there are lots of research work done using EEGNet for classification purposes, so, there are no doubts about EEGNet's capabilities as a neural network specifically developed for EEG signals. And as neural networks usually require lots of input data in order to find the patterns and learn the data, we were unfortunately short on the input data, as there were a limited number of in-combat probes manually annotated, or even the probes that happened during the game. In addition, large input space leads to confusing the neural network most of the time. Because there are lots of pixels to process, thus, there is a big amount of calculations to carry out. We tried to address this issue by reducing the problem space by using DWT in two orders, or synchrony and zero removal techniques, which unfortunately did not lead to any improvement in the outcome.

In this research, after months of trying various approaches and not being able to reach the anticipated results, we investigated the reason why. There are three possible rationales: (1) there is no relation between EEG and the flow mental state, (2) the data we acquired was not of proper quality, and (3) the approaches we applied were not suitable for this problem. As for the first assumption, a

reference to the literature would help. There have been tons of research works on EEG data and its use in detecting various mental states such as concentration, emotions, vigilance, etc. as we discussed in chapter 2. So, there surely exists a strong possibility that the flow state could be detected by EEG data. In our case, a potential alternative could be the use of other biometric data to empower the model to make classification decisions by injecting more information into the dataset. Because clearly, the EEG data on its own was not enough.

We believe that the reason for the failure of our attempts was reason number two, the data. In this experiment, the EEG data were acquired in a real-world video game setting, using some hardware and tools that were at hand at the time. The EEG cap could have been of not good quality, or the experiment setting could have been in a way that caused lots of noise and artifacts in the EEG data. We believe the EEG data was too noisy and too corrupted to find any patterns in it. As stated in the previous chapters, we tried different noise-removal algorithms such as ASR and wICA which have a good reputation in the literature, we used decomposition methods such as EMD and VMD which can also be used for de-noising purposes, yet, nary any desirable results. We did lots of data scaling and normalization, removing outlier probes, and employed powerful tools such as SVM and EEGNet that are known to be successful means in this field. We believe the fact that we could not get the optimal results in any of our experiments and trials is because the data acquired was too corrupt. To further investigate this assumption, another colleague of mine, Billal Belainine, after me, proved that Ubisoft's EEG data did not have the minimum quality for doing research on. His efforts, and the fact that none of my efforts, or the ones before me, resulted in any interesting or valid results, all show that the EEG data acquired by Ubisoft for this project was too noisy.

To prove the inferior quality of the EEG data gathered by Ubisoft, Bilal performed some statistical tests and had them compared with the same results from a known EEG dataset, SPIS-Resting-State dataset(Torkamani-Azar et al., 2020). This dataset was chosen as an example of a higher-quality EEG dataset. To find the attributes of the signals, Bilal computed their amplitude, entropy, mean, and standard deviation, based on the state of the art. He also manually divided the data into two classes; Hit (where the probes were seen) and in-combat. Then, to be able to read up on the relations between the attributes, he used Chi-Square and Information Gain tests. He ran his methodology on the signals of the SPIS-resting-state and Ubisoft's datasets, on the 10 seconds of signals before and after the probes, and on the IMFs resulting from the EMD algorithm. Moreover, the attributes were used in classifying the signals between the two classes using an SVM classifier with a linear kernel. The comparison between the correlation matrices (of the signals themselves, or their IMFs) derived from the statistical tests on both datasets, showed that in our dataset, the heatmap shows significant losses between IMF1 and IMF4; hence IMF4's heatmap is almost black (figure 5.1). However, the SPIS-Resting-State dataset has a gradual loss between IMF1 and IMF4 (5.2). This explains why certain data are problematic. That is to say, we could suspect that some data has been recorded incorrectly or that the electrodes have malfunctioned. For the classification, there was a big difference between the results obtained by our data with F-Masure of 71.1% and the results obtained by SPIS-Repos-State with 90%. With a difference of almost 19%, we can confirm that the classification fell far short of our expectations. The obtained results show either the data was badly recorded and noisy, or the EEG data and its attributes were not enough to draw a conclusion from, and other types of data could have been used.

As for the approaches we used in this project, it is important to have in mind that they were

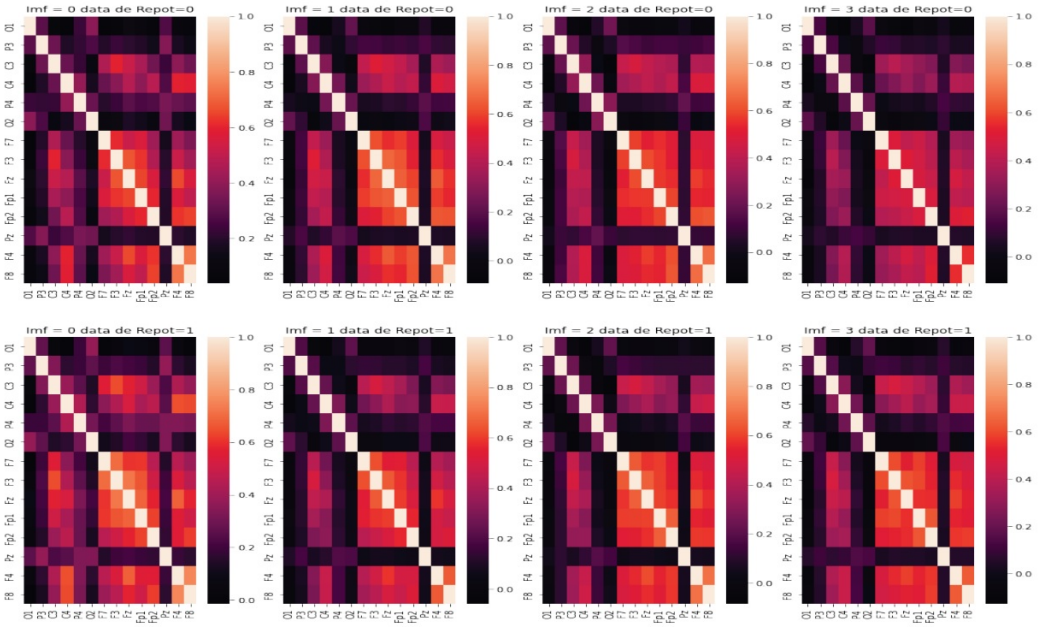


Figure 5.1: Heatmap of the covariance matrices represent the first 4 IMFs by two classes; eyes open and eyes closed.

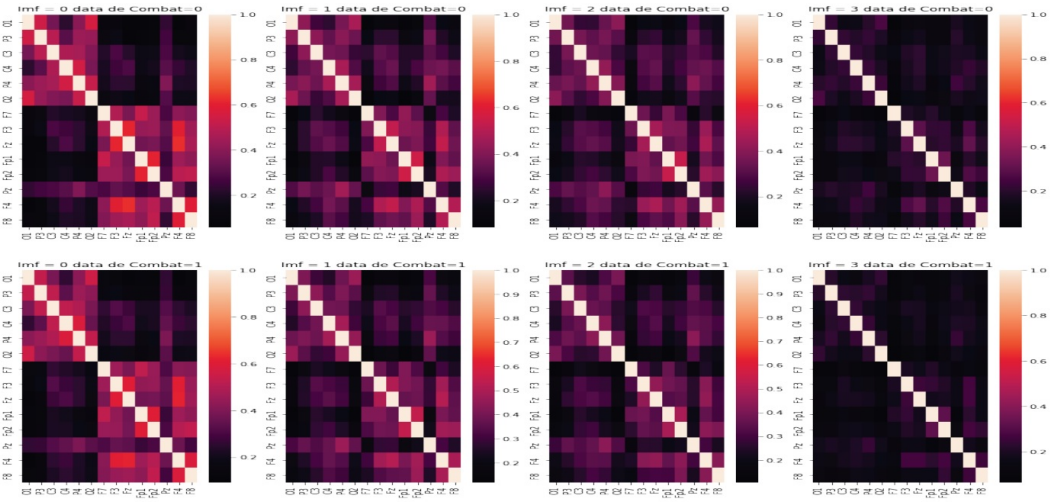


Figure 5.2: Heatmap covariance matrices represent the first 4 IMFs by two classes, the Combat state and the no-combat state.

suggested by the literature and they had shown promising and robust performance. In some cases in the literature, the dataset being used by the other works was single-channel EEG, or computer-generated EEG, which could make a great impact on the quality of the data. Also, the conditions are different in various works, and so was in ours. We followed a strong methodology to find a link between EEG and the mental state, all these methods are generic and not specific to EEG data. So, we decided to test the methodology. In our testing, we are considering the fact that all these methods are generic and not specific to EEG data, or to single- or multi-channel signals, so by applying some tweaks to the variables and some methods, we tried them with single-channel EMG data to show that the methodology was fine.

## 5.1 EMG classification

In order to make sure the results we got were due to bad data, and not wrong methodology, another experiment with EMG was carried out to prove the point. In this experiment, we tried to classify EMG signals into three classes: myopathy, neuropathy, and healthy, following the same logic of the methodology we followed in Ubisoft's project.

### 5.1.1 Introduction

In patients with muscle dysfunction or muscle weakness, the cause is normally due to a damaged muscle that cannot function properly (myopathy), or the neural system cannot function properly and send the neural massages correctly (neuropathy). An electromyogram is a test that assesses the function of muscles and the nerves that control them. In this project, we tried to classify the EMG signals of patients with either neuropathy or myopathy, or healthy people. To do so, and to show the strong logic of our methodology, the same path of statistical features, decomposition, dimension reduction, and classification was followed.

### 5.1.2 Quick Survey on EMG

In the literature, there are numerous research works regarding physiological signal classification dealing with EMG signals. (Shaw & Bagha, 2012) use EMG signals of healthy people vs. patients with neuromuscular disease. The power spectrum of the signals is computed as the feature set. Later, they use PCA to embolden the similarities and differences between the groups, while the dimensionality is decreased as well. A probabilistic neural network classifies the signals into three groups, with an average classification rate of 91.72% over the different lengths of patterns. In (Subasi et al., 2006), the authors record the EMG signals of healthy subjects, subjects suffering from myopathy, and neurogenic disease. They compute the coefficients derived from an autoregressive model which are later fed to a feedforward error backpropagation artificial neural network and a wavelet neural network. The former manages to get 88% as the classification success rate, while the latter gets 90.7%. In another work, (Koçer, 2010) use AR coefficients as their feature set, similar to the previous work. Their purpose is to classify the EMG signals obtained from healthy, neuropathic, and myopathic subjects. To do so, the data undergoes the AR model order-10, and the resulting

coefficients are used with a neuro-fuzzy system. In different trials, classification success of 80%, and 83.3% when the fuzzy system has 28 and 26 rules, respectively. After pruning the features and reducing the number of rules to 10, they get a classification rate of 90%.

Taking the same route as the predecessors, (Lahmiri & Boukadoum, 2017) propose a fast-for-clinical-application approach for healthy vs. neuropathic EMG classification. They employ an auto-regressive moving average (ARMA) model and take the coefficients that represent EMG's time-evolution essence. Later, the coefficients are passed to a Linear Discriminant Analysis classifier ((Balakrishnama & Ganapathiraju, 1998)) that can successfully classify them with 99.74% accuracy, 100% sensitivity and 99.66% specificity. In another similar work, (Lahmiri & Boukadoum, 2018) pay additional attention to the linear dynamics of EMG. As a result, they take the coefficients derived from the ARMA model along with the variance of its disturbances (variance of the residual) and give the features to an LDA because of its simplicity and interpretability. This approach shows more promising performance in accuracy, sensibility, and specificity compared to approaches with only one AR or MA coefficient and kNN or MLP (Multi-Layer Perceptron). They report $90.25\%, 83.67\%, 92.65\%$ for accuracy, sensitivity, and specificity, respectively. Furthermore, the use of Fourier transform, or EMD is also bold in EMG processing, as the deep learning techniques are for classification purposes since these methods are better able to capture the properties and relationships in the big data. This claim is properly demonstrated in (Bakircioğlu & Özkurt, 2020). The authors of this paper use the EMG of the forearm acquired while performing simple daily hand movements, and apply EMD, Fourier transform, and root-mean-square to it. The output of each of these methods is then given to different CNNs, which later prove to be a promising tool. The combination of EMD and CNN outperforms the rest of the methods with 95.9% accuracy.

As EMG contains information about the muscle, it is vastly used for muscle damage detection, as explained in the previous works, or motion detection for prosthetic arms. (Tam et al., 2019)'s paper presents a real-time hand gesture recognition method that is able to classify forearm muscle contractions and can be used in a prosthetic hand. The EMG that is derived from a 32-channel sensor, performing eight hand gestures, is mapped to a $4 * 8$ array every second. The resulting hand gesture maps are then cleaned by removing unwanted frequency bands, and amplitudes. Then a CNN classifies the gestures in real-time with an accuracy of 98.2%. In another work, (Gandolla et al., 2017) have designed an EMG-controller for a robotic arm that is able to classify the movement intention (before the movement happens) with decent performance for healthy subjects, and higher than expected for post-stroke patients. The EMG underwent low- and high-pass filters and then is classified by an ANN getting the accuracy of 76% for healthy subjects, and 90% for post-stroke patients.

### 5.1.3 Data

The EMG data were recorded from three patients: a 44-year-old healthy man, a 62-year-old with neuropathy, and a 57-year-old man with myopathy. More detail is available on PhysioNet's (Goldberger & Stanley, 2000) website[1]. Data were recorded at 50 kHz, and after applying a 20 Hz high-pass filter and a 5 kHz low-pass filter, it was downsampled to 4 kHz.

[1] https://physionet.org/content/emgdb/1.0.0/

### 5.1.4 Methodology

Firstly, the signals are chopped into smaller chunks in order to extract their statistical features. in regards to the fact that the dataset consists of single-channel EMG which is less complex, a basic regression model seemed to be a great choice. As inspired by many of the works in the literature, and the classification success rate they report ((Subasi et al., 2006), (Koçer, 2010), (Lahmiri & Boukadoum, 2017), (Lahmiri & Boukadoum, 2018)), an ARMA model was applied to the signal segments so that each chunk of signal could be shrunk into its coefficients. ARMA model is a great choice when there are unexpected shocks appearing in the signal, like there are spikes or abnormalities in the EMG signal for neuropathic or myopathic patients. Another of its advantages is capturing the bold and distinct characteristics of the signal and turning them into a number of coefficients, which is a great dimension reduction. It is also accurate and converges quickly. In another attempt, as inspired by (Bakircioğlu & Özkurt, 2020), DWT was also used to decompose and compress the EMG segments, as well as to extract the underlying information in the EMG signals. Later, the approximation coefficient vector was chosen to be fed to the classifier as this vector is the closest to the input vector. In addition, as used previously, EMD decomposed the EMG signals as well, so that we have a valid reference to compare the EEG work with EMG. For classification, we used SVM, kNN, a multi-layer neural network, and a CNN, so that we test both classic and modern means of classification on this problem.

### 5.1.5 Results

First, the long EMG recordings were segmented into 2-millisecond chunks, so that in this case, we would have 63 healthy, 137 myopathy, and 184 neuropathy signals. Then, the ARMA model was applied to each of these data chunks, in two orders, ARMA order one and order two, resulting in two and four AR and MA coefficients, respectively. In addition, for each ARMA application, we kept the intercept and the variance as well in the feature set. So, based on ARMA's equation in chapter 2.4, the feature set would look like $[c, \Phi, \Psi, \varepsilon]$ for ARMA order one, and $[c, \Phi_1, \Psi_1, \Phi_2, \Psi_2, \varepsilon]$ for ARMA order two. After extracting these feature vectors for each segment of EMG, they were min-max normalized, shuffled, and split randomly into balanced train and test sets using 5-fold stratified cross-validation, with the random state set to 7, and shuffle set to *True*. 5-fold cross-validation is done for all of the classifiers.

The first classifier we used was a kNN. In the range of $k = [4, 10]$, $k = 4$ showed to have slightly better results compared to other numbers of neighbors. The performance of the model averaged over the 5 folds are shown in table 5.1.

Table 5.1: kNN performance metrics on ARMA order 1 and 2 coefficients of the EMG data.

| ARMA order | Accuracy | Precision | Recall | F-score |
|---|---|---|---|---|
| Order-1 | 0.98 | 0.98 | 0.98 | 0.98 |
| Order-2 | 0.95 | 0.972 | 0.954 | 0.96 |

The second classifier that was employed was SVM with *rbf* kernel and $C$ parameter of 1. Table 5.2 shows the results averaged over all of the folds of cross-validation.

Table 5.2: SVM performance metrics on ARMA order 1 and 2 coefficients of the EMG data.

| ARMA order | Accuracy | Precision | Recall | F-score |
|---|---|---|---|---|
| Order-1 | 0.978 | 0.976 | 0.978 | 0.974 |
| Order-2 | 0.92 | 0.89 | 0.87 | 0.87 |

The third classifier was a neural network. The ANN's architecture is shown in figure 5.3. Also, the model's loss function is *categorical cross-entropy*, along with *adam* optimizer, with a batch size of 32 and 200 epochs. The results of the ANN, averaged over 5 folds, are illustrated in table 5.3. The accuracy and loss diagram is also shown in figure 5.4.

Table 5.3: ANN performance metrics on ARMA order 1 and 2 coefficients of the EMG data.

| ARMA order | Accuracy | Precision | Recall | F-score |
|---|---|---|---|---|
| Order-1 | 0.976 | 0.97 | 0.978 | 0.974 |
| Order-2 | 0.988 | 0.98 | 0.98 | 0.98 |

In another attempt, DWT was applied to the EMG segments resulting in two matrices per segment signal, using *db4* mother wavelet. We then selected the approximation coefficient matrix as the new feature set to use. It was then fed into a CNN. The CNN's architecture is shown in figure 5.5. Also, the model's loss function is *categorical cross-entropy*, along with *adam* optimizer with a learning rate of 0.01, with 200 epochs.

Table 5.4: CNN performance metrics on level-1 DWT approximation matrix of the EMG data. The accuracy and loss diagrams are shown in figure 5.6.

| DWT component | Accuracy | Precision | Recall | F-score |
|---|---|---|---|---|
| Apprx - L1 | 0.976 | 0.977 | 0.976 | 0.978 |

Lastly, the EMD decomposition algorithm was applied to EMG chunks of signal, decomposing them into six IMFs. As the first and second IMFs are closer to the actual signals, information-wise, these two were picked as the input to the classifier. In this case, since each EMG segment was turned into two IMFs, the number of data in the dataset was doubled compared to the previous methods. So, the total number of data for healthy, myopathy, and neuropathy datasets was 126, 274, and 368, respectively. Later, these data were fed into a CNN with the same structure as 5.5. The CNN classification has resulted in an accuracy of 0.853, with 0.85, and 0.849 per precision and recall, averaged over the 5 folds of cross-validation. Figure 5.7 shows the accuracy and loss diagrams of CNN.

### 5.1.6 Discussion

In this section, a comparison between this attempt and other similar studies will be provided, as illustrated in table 5.5. In this work, we, inspired by some of the works in the literature, tried to classify EMG using the ARMA model, DWT, EMD decomposition, and several classifiers. Compared to the results obtained by some of the other works in the literature with alike approaches.

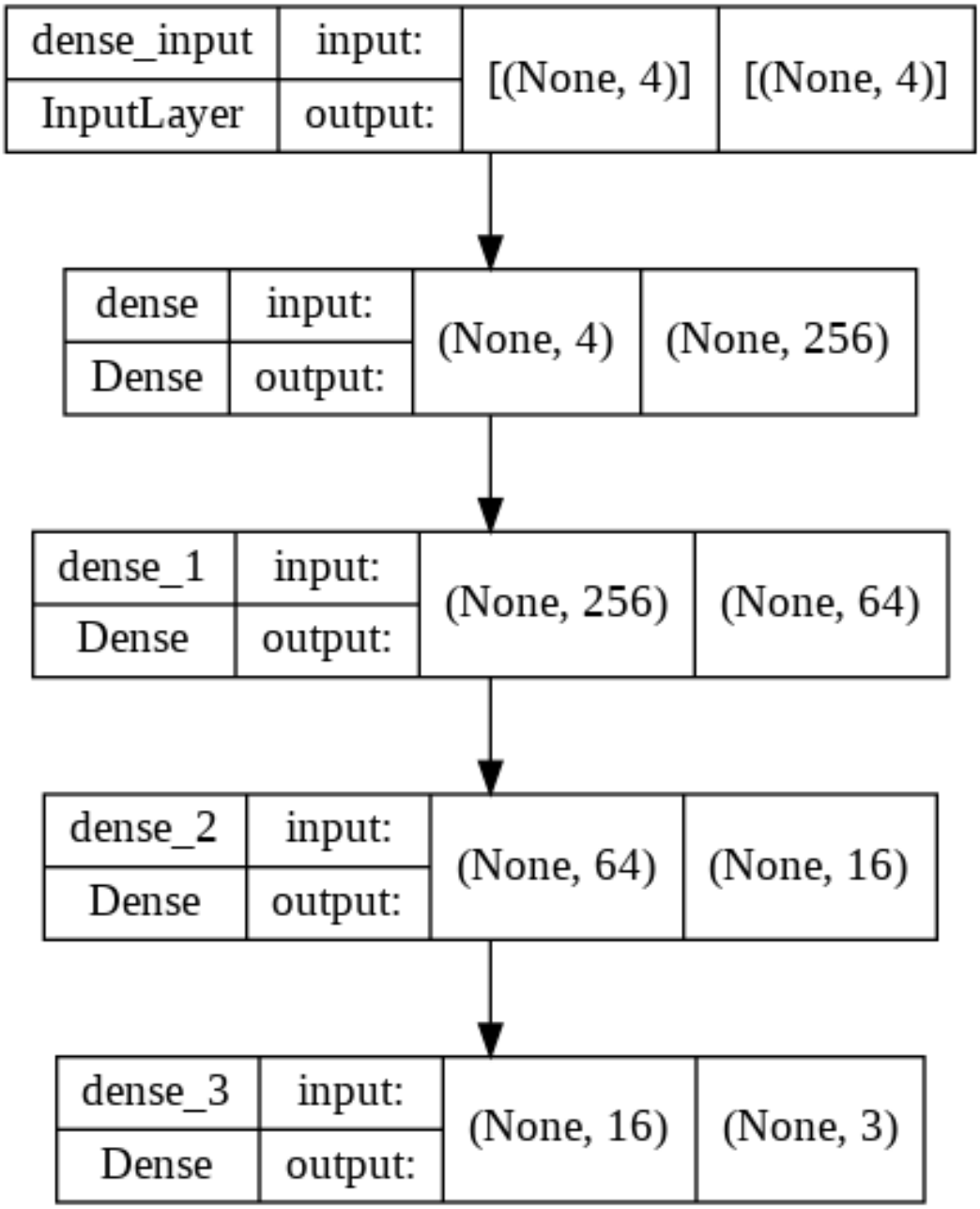


Figure 5.3: The architecture of the artificial neural network. The activation function of all the layers is ReLU function, and of the output layer is softmax.

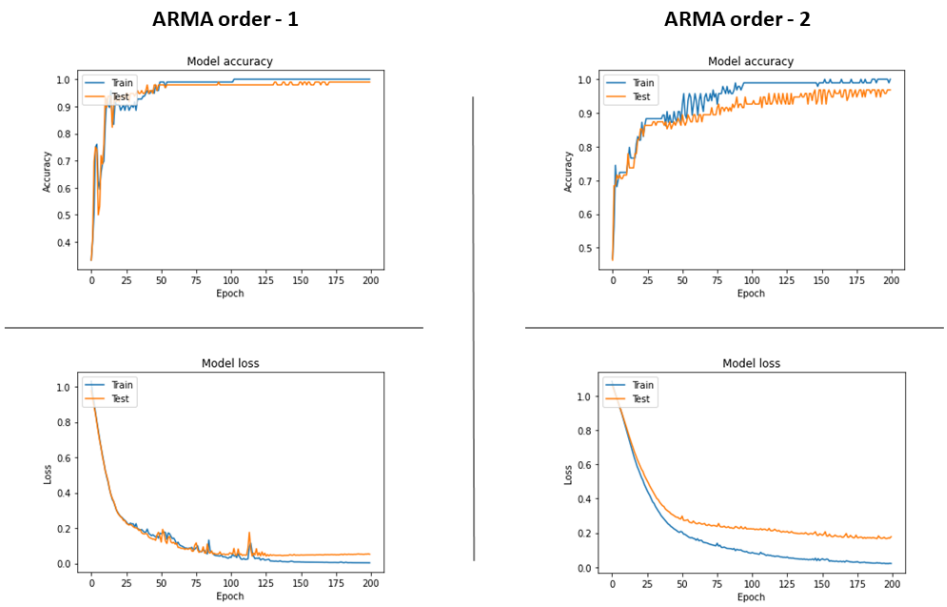


Figure 5.4: The accuracy and loss diagram of the artificial neural network on the ARMA order 1 and 2 coefficients of the EMG data, over 200 epochs. These diagrams show one random fold among the 5 folds of the cross-validation.

As illustrated in the table 5.5, our proposed model with DWT and CNN is performing better than (Bakircioğlu & Özkurt, 2020)'s proposed way of EMD decomposition and CNN, or (Shaw & Bagha, 2012)'s tactic of PCA and probabilistic neural network. Also, all of our attempts with ARMA and machine learning components outperform the other works which use ARMA; (Lahmiri & Boukadoum, 2017) and (Lahmiri & Boukadoum, 2018). In the case of other statistical features obtained by (Tam et al., 2019) along with CNN, our work is performing just as fine. For comparing (Bakircioğlu & Özkurt, 2020)'s EMD + CNN approach with our EMD + CNN approach, our result shows a lower accuracy. It is important to bear in mind that the two datasets in these works are not the same, and neither is the number of IMFs extracted via EMD. The performance gap could be due to these differences. Especially, in the case of our work, during the 5-fold cross-validation, there were some folds using which, the CNN had a 97% accuracy, hence, depending on which data is chosen to train and test the model, the accuracy may vary between 81% and 97%.

In this work, we tried to combine the classic statistical methods with state-of-the-art machine learning approaches, as a contribution to the field. We were able to classify the EMG data into three different classes with high performance which shows that the path we went down was a successful one. It illustrates that our expectations about selecting the right methods that are capable of efficiently extracting the bold and informative characteristics of the data, along with the right ML methods that are capable of learning the true nature of it, were met.

At this point, it is necessary to discuss the differences and resemblances between the research we conducted in this thesis on EEG and EMG, and compare the results each produced. Below, some of the aspects they have in common will be mentioned:

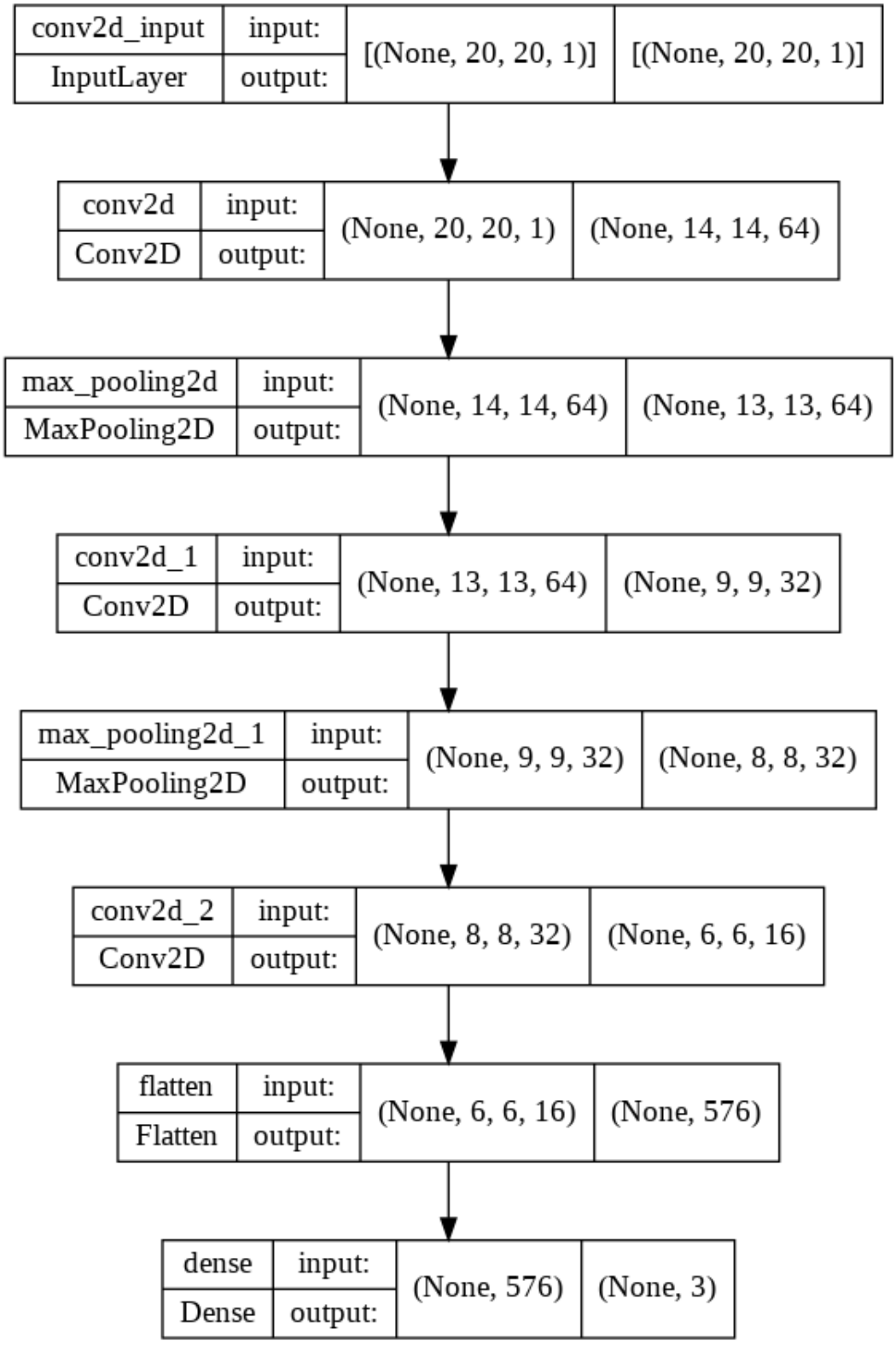


Figure 5.5: The architecture of the convolutional neural network. This architecture was proposed by (Bakircioğlu & Özkurt, 2020). The activation function of all the convolution layers is ReLU function, and of the output layer is softmax.

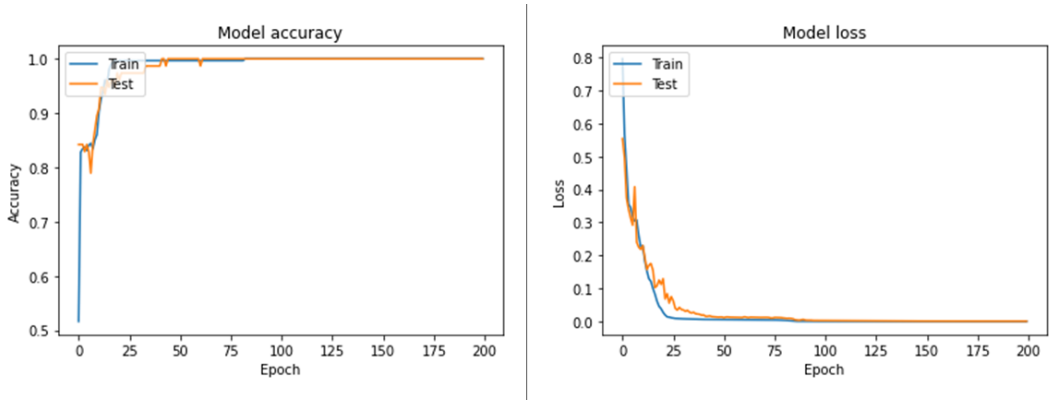


Figure 5.6: The accuracy and loss diagram of the convolutional neural network on the DWT level-1 coefficients matrix of the EMG data, over 200 epochs. These diagrams show one random fold among the 5 folds of the cross-validation.

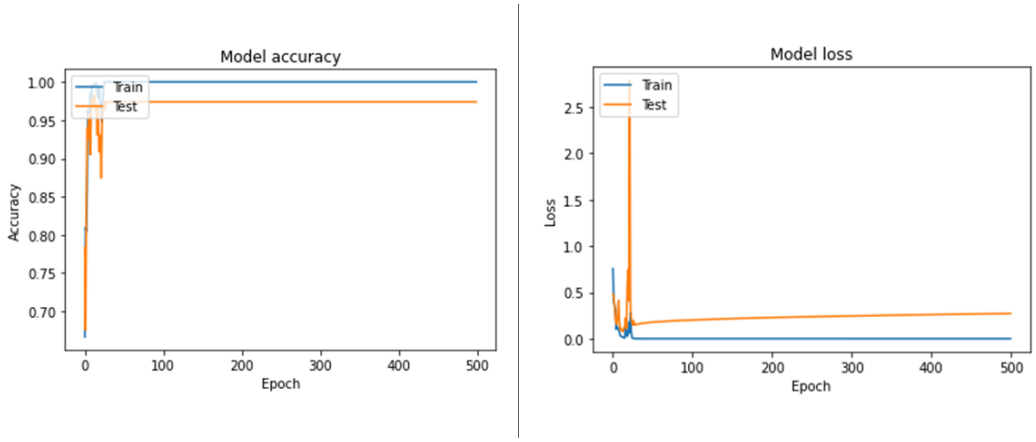


Figure 5.7: The accuracy and loss diagram of the CNN on the first two IMFs of EMD, over 500 epochs. These diagrams show one random fold among the 5 folds of the cross-validation.

Table 5.5: A comparison between the results obtained in this work vs. other research works in the literature with a close approach to us.

| Study | Method | Classifier | Accuracy |
|---|---|---|---|
| (Bakircioğlu & Özkurt, 2020) | EMD | CNN | 95.9% |
| (Lahmiri & Boukadoum, 2018) | ARMA(2,2) | LDA | 90.25% |
| (Tam et al., 2019) | frequency-time-space cross-domain | CNN | 98.2% |
| (Lahmiri & Boukadoum, 2017) | AR | LDA | 52.60% |
| (Shaw & Bagha, 2012) | PCA | PNN | 91.72% |
| This work | ARMA(1,1) | kNN | 98% |
| This work | ARMA(2,2) | kNN | 95% |
| This work | ARMA(1,1) | SVM | 97.8% |
| This work | ARMA(2,2) | SVM | 92% |
| This work | ARMA(1,1) | ANN | 97.6% |
| This work | ARMA(2,2) | ANN | 98.8% |
| This work | DWT L-1 | CNN | 97.6% |
| This work | EMD | CNN | 85% |

- Both of these data are non-stationary physiological time series that share every property that signals in general have in common, such as sampling rate, frequency-based features, time-based features, etc. Hence, every method that is appropriate for non-stationary signal processing can be applied to both.
- Regarding the methodology that was followed in both of these works, both of these data were normalized. The EMG signals were Min-Max normalized and the EEG data were both Min-Max and Z-score normalized, being in the signal form, probe image form, or IMFs.
- Statistical features were derived from both EEG ad EMG. In order to capture the information within these data in statistical terms, and in order to reduce the problem space from hundreds of data points in a signal to a couple of statistics in the form of a feature vector, we computed some relevant statistical features. To do so even further, the decomposition algorithms were applied to both of these data, as well.
- And lastly, as per the classification phase, machine learning methods were used in both cases.

All these common aspects between the two data types led us to use the same methodology and line of thinking for both. Although the approaches utilized were not the same in detail, the way of thinking and reasoning were the same for both of these projects: non-stationary physiological data at hand, decomposing, and processing them to get the informative features to use with classification methods. As mentioned, the path of research was not the same step-by-step: for example, we tried to denoise the EEG data, as it was super noisy, bringing a few algorithms into play, yet, the EMG data did not need denoising since the dataset was already preprocessed before becoming available to public access. However, the logic was the same.

In the end, by carrying out the experiment with classifying the EMG data, I believe we were successful in showing that the reasoning and the methodology that was used throughout this whole project were not only valid and strong but also applicable to other similar-in-nature problems with even a different data space. Thus, it truly is a shame that the data from the Ubisoft project did not lead to conclusive results, otherwise, with the use of such promising approaches, appealing results could have been obtained.

Lastly, after discussing the outcomes of both of the projects, and presenting a detailed comparison between the two, a conclusion of this thesis would be given in the next chapter.

# CHAPTER VI

# CONCLUSION

## 6.1 Summary

In this thesis, all the works and efforts that were done on a classification project with the objective of finding the flow mental state during gameplay were thoroughly explained. In this project, the mental activities of the gamers were recorded during sessions of gameplay, and the objective was to detect when during the game, the players were concentrated and immersed into the game. This state of mind is called the flow state.

To detect the flow state, we used the EEG signals that were acquired from the players while gaming. To be able to extract meaningful characteristics out of these signals, we followed a couple of methods: we processed the EEG signals to remove as much noise as possible (using ASR, wICA) and decomposed them into their constituent components so that capturing their informative rhythms would be easier (using EMD, VMD, DWT). After extracting their meaningful features, such as statistical features, or features of frequency and time dimensions, we tried classifying them using classic classification models; such as SVM, as well as some state-of-the-art models proposed by the literature; such as EEGNet.

After trying so many paths that are considered to be strong methodology, going after several approaches for feature extraction, finding differences, and utilizing the best means of classification, the expected reults were not obtained. To further investigate the reason, we came up with the answer which was the corruption and poor quality of the EEG data. We managed to prove this by explaining that

- several people have spent quite some time on this project and not getting desirable results,
- showing that all the methodologies proposed by the literature with promising performance that were employed in this project led to failure,
- one of my colleagues was able to show that the data is extremely noisy, by following a step-by-step path proposed by other research works, and the fact that everything was leading to success for the proposed dataset, actually led to a failure for the Ubisoft EEG dataset.
- In addition, we took another step to show that meaningful and logical methodology does work when the data is clean and of proper quality, by applying a similar approach to an EMG

dataset. We did the same routine on the EMG dataset: normalization of the data, feature extraction, decomposition, and classification, and all the efforts led to a great performance.

## 6.2 Limitations

In this project, there were several limitations that came to our attention:

1. Firstly, as mentioned before, the biggest limitation, the data was not of proper quality. If the data was alright, and we could have succeeded during the first steps, we could have probably taken this research to other levels, such as creating a model that was able to detect flow mental state for a given input, or we could transfer the model to a version that was compatible with other data from other games. For example, to have a model that was able to detect the flow state when the EEG data of any game was being input into the model.
2. In the next place, the hypothesis we had at the beginning of the project could change. We were assuming that if the players missed more probes, then hypothetically, they are more focused on the game. We based all of our efforts on this hypothesis, and even in the final steps of the project, the hypothesis of being in a combat probably needs more concentration, yet both of these hypotheses could change. We probably should have explored other ways of looking at the problem, to have a thorough view and to make the most of the resources we had.
3. While implementing the ways proposed by the research works in the literature and not obtaining the promising results they claimed they were getting, we noticed that most of these works are using existing EEG datasets, and most of those promising outcomes were being derived from single-channel EEG data. This rises two issues: first, using existing datasets increases the chance of good performance, because these datasets are usually of good quality and are significantly less noisy than the data acquired by Ubisoft, being acquired in better data-collection conditions with better tools and more sophisticated hardware. Second, single-channel EEG lowers the complexity of data processing and feature extraction and simply, every step of the process. Having 16 channels was challenging: it lied questions like do we decompose each channel or all channels together, do we compute the features of each channel or all together, if we consider channels individually, are we capturing the important information of the signal or not? In addition, run-time and memory were greatly influenced by the choice of a multi- or single-channel signal.
4. Another limitation was time. If more time could be spent on this project, data-tool combinations could be explored. Other types of biometric data that were available could be used, or tools we were using could be changed to focus on some other approaches that are usually used with other types of data.

## 6.3 Future Direction

There are several directions that could be followed in order to ameliorate the outcome of this project and they could be viewed as some candidate methods to pursue.

- As mentioned before, there are other physiological data that were acquired during the data collection procedure: EDA, PPG, and eye tracking data. In addition, some other physiological measures can be calculated such as cognitive engagement, cognitive load, heart rate variability, EDA energy, etc. These data could be used along with the EEG data or on their own, to further investigate the ways to solve this problem.
- Another path that could be followed is to continue the last effort we did with the in-combat probes. Synchrony removal and removing the zeros to reduce the problem space can also be explored with all the probes, not only the in-combat ones. Maybe a useful outcome could be derived from this approach.
- As another alternative to the classification means, probably some simpler classifiers could be utilized. Since the data is not enough for EEGNet, and as neural networks require quite a large amount of data, then some simpler classifiers might have better performance. Instead of CNNs, or any other neural network, a linear regressor,

# BIBLIOGRAPHY


Adam, M., Oh, S. L., Sudarshan, V. K., Koh, J. E., Hagiwara, Y., Tan, J. H., San Tan, R. & Acharya, U. R. (2018). Automated characterization of cardiovascular diseases using relative wavelet nonlinear features extracted from ecg signals. *Computer methods and programs in biomedicine*, *161*, 133–143.

Albuquerque, I., Rosanne, O., Gagnon, J.-F., Tremblay, S. & Falk, T. H. (2019). Fusion of spectral and spectro-temporal eeg features for mental workload assessment under different levels of physical activity. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 311–314. IEEE.

Anuragi, A., Sisodia, D. S. & Pachori, R. B. (2021). Automated fbse-ewt based learning framework for detection of epileptic seizures using time-segmented eeg signals. *Computers in Biology and Medicine*, *136*, 104708.

Azzerboni, B., Finocchio, G., Ipsale, M., Foresta, F. L. & Morabito, F. C. (2002). A new approach to detection of muscle activation by independent component analysis and wavelet transform. In *Italian Workshop on Neural Nets*, pp. 109–116. Springer.

Bagheri, A., Persano Adorno, D., Rizzo, P., Barraco, R. & Bellomonte, L. (2014). Empirical mode decomposition and neural network for the classification of electroretinographic data. *Medical & biological engineering & computing*, *52*(7), 619–628.

Bajada, J. & Bonello, F. B. (2021). Real-time eeg-based emotion recognition using discrete wavelet transforms on full and reduced channel signals. *arXiv preprint arXiv:2110.05635*.

Bakircioğlu, K. & Özkurt, N. (2020). Classification of emg signals using convolution neural network. *International Journal of Applied Mathematics Electronics and Computers*, *8*(4), 115–119.

Balakrishnama, S. & Ganapathiraju, A. (1998). Linear discriminant analysis-a brief tutorial. *Institute for Signal and information Processing*, *18*(1998), 1–8.

Bashivan, P., Rish, I. & Heisig, S. (2016). Mental state recognition via wearable eeg. *arXiv preprint arXiv:1602.00985*.

Bezdek, J. C., Ehrlich, R. & Full, W. (1984). Fcm: The fuzzy c-means clustering algorithm. *Computers & geosciences*, *10*(2-3), 191–203.

Borderie, J. & Michinov, N. (2016). Identifying flow in video games: towards a new observation-based method. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, *8*(3), 19–38.

Box, G. E., Jenkins, G. M., Reinsel, G. C. & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Brown, E. & Cairns, P. (2004). A grounded investigation of game immersion. In *CHI'04 extended abstracts on Human factors in computing systems*, pp. 1297–1300.

Burns, A. & Tulip, J. (2017). Detecting flow in games using facial expressions. In *2017 IEEE Conference on Computational Intelligence and Games (CIG)*, pp. 45–52. IEEE.

Cai, X., Cebollada, J. & Cortiñas, M. (2022). Self-report measure of dispositional flow experience in the video game context: Conceptualisation and scale development. *International Journal of Human-Computer Studies*, *159*, 102746.

Castaneda, D., Esparza, A., Ghamari, M., Soltanpur, C. & Nazeran, H. (2018). A review on wearable photoplethysmography sensors and their potential future applications in health care. *International journal of biosensors & bioelectronics*, *4*(4), 195.

Chang, C.-Y., Hsu, S.-H., Pion-Tonachini, L. & Jung, T.-P. (2019). Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel eeg recordings. *IEEE Transactions on Biomedical Engineering*, *67*(4), 1114–1121.

Cowley, B., Charles, D., Black, M. & Hickey, R. (2006). User-system-experience model for user centered design in computer games. In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pp. 419–424. Springer.

Cowley, B., Charles, D., Black, M. & Hickey, R. (2008). Toward an understanding of flow in video games. *Computers in Entertainment (CIE)*, *6*(2), 1–27.

Csikszentmihalyi, M. & Csikzentmihaly, M. (1990). *Flow: The psychology of optimal experience*, volume 1990. Harper & Row New York.

Das, A. B. & Bhuiyan, M. I. H. (2016). Discrimination and classification of focal and non-focal eeg signals using entropy-based features in the emd-dwt domain. *biomedical signal processing and control*, *29*, 11–21.

Dragomiretskiy, K. & Zosso, D. (2013). Variational mode decomposition. *IEEE transactions on signal processing*, *62*(3), 531–544.

Esteller, R., Echauz, J. & Tcheng, T. (2004). Comparison of line length feature before and after brain electrical stimulation in epileptic patients. In *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 2, pp. 4710–4713. IEEE.

Gandolla, M., Ferrante, S., Ferrigno, G., Baldassini, D., Molteni, F., Guanziroli, E., Cotti Cottini, M., Seneci, C. & Pedrocchi, A. (2017). Artificial neural network emg classifier for functional hand grasp movements prediction. *Journal of International Medical Research*, *45*(6), 1831–1847.

Göksu, H. (2018). Eeg based epileptiform pattern recognition inside and outside the seizure states. *Biomedical Signal Processing and Control*, *43*, 204–215.

Goldberger, A., L. A. L. G. J. H. P. C. I. R. M. J. E. M. G. B. M. C. K. P. & Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. 101 (23), pp. e215–e220. Retrieved on 2022-08-30 from `https://www.physionet.org/`

Gonzalez, R. C. (2009). *Digital image processing*. Pearson education india.

Granato, M., Gadia, D., Maggiorini, D. & Ripamonti, L. A. (2017). Emotions detection through the analysis of physiological information during video games fruition. In *International Conference on Games and Learning Alliance*, pp. 197–207. Springer.

Hahn, S. L. (1996). *Hilbert transforms in signal processing*. Artech House Signal Processing.

Hosseinzadeh, M. (2020). Robust control applications in biomedical engineering: Control of depth of hypnosis. In *Control Applications for Biomedical Engineering Systems* pp. 89–125. Elsevier.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C. & Liu, H. H. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, *454*(1971), 903–995.

Inuso, G., La Foresta, F., Mammone, N. & Morabito, F. C. (2007). Wavelet-ica methodology for efficient artifact removal from electroencephalographic recordings. In *2007 international joint conference on neural networks*, pp. 1524–1529. IEEE.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T. & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, *66*(9), 641–661.

Jung, D., Choi, J., Kim, J., Cho, S. & Han, S. (2022). Eeg-based identification of emotional neural state evoked by virtual environment interaction. *International Journal of Environmental Research and Public Health*, *19*(4), 2158.

Keerthi Krishnan, K. & Soman, K. (2021). Cnn based classification of motor imaginary using variational mode decomposed eeg-spectrum image. *Biomedical Engineering Letters*, *11*(3), 235–247.

Kira, K. & Rendell, L. A. (1992). A practical approach to feature selection. In *Machine learning proceedings 1992* pp. 249–256. Elsevier.

Klasen, M., Weber, R., Kircher, T. T., Mathiak, K. A. & Mathiak, K. (2012). Neural contributions to flow experience during video game playing. *Social cognitive and affective neuroscience*, *7*(4), 485–495.

Koçer, S. (2010). Classification of emg signals using neuro-fuzzy system and diagnosis of neuromuscular diseases. *Journal of Medical Systems*, *34*, 321–329.

Lahmiri, S. & Boukadoum, M. (2013). An application of the empirical mode decomposition to brain magnetic resonance images classification. In *2013 IEEE 4th Latin American Symposium on Circuits and Systems (LASCAS)*, pp. 1–4. IEEE.

Lahmiri, S. & Boukadoum, M. (2014). Adjusted empirical mode decomposition with improved performance for signal modeling and prediction. In *2014 IEEE 5th Latin American Symposium on Circuits and Systems*, pp. 1–4. IEEE.

Lahmiri, S. & Boukadoum, M. (2015). Pathology grading in retina digital images using student-adjusted empirical mode decomposition and power law statistics. In *2015 IEEE 6th Latin American Symposium on Circuits & Systems (LASCAS)*, pp. 1–4. IEEE.

Lahmiri, S. & Boukadoum, M. (2017). An accurate automatic system for distinguishing neuropathy

and healthy electromyography signals. In *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–4. IEEE.

Lahmiri, S. & Boukadoum, M. (2018). Improved electromyography signal modeling for myopathy detection. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–4. IEEE.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P. & Lance, B. J. (2018). Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, *15*(5), 056013.

Liu, T., Luo, Z., Huang, J. & Yan, S. (2018). A comparative study of four kinds of adaptive decomposition algorithms and their applications. *Sensors*, *18*(7), 2120.

Maanen, P., Blum, S. & Debener, S. (2022). Mobile eeg artifact correction on limited hardware using artifact subspace recon-struction. *arXiv preprint arXiv:2204.13444*.

Maris, E. & Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, *164*(1), 177–190.

Mehla, V. K., Singhal, A., Singh, P. & Pachori, R. B. (2021). An efficient method for identification of epileptic seizures from eeg signals using fourier analysis. *Physical and Engineering Sciences in Medicine*, *44*(2), 443–456.

Mohammadpour, M., Hashemi, S. M. R. & Houshmand, N. (2017). Classification of eeg-based emotion for bci applications. In *2017 Artificial Intelligence and Robotics (IRANOPEN)*, pp. 127–131. IEEE.

Murugappan, M., Rizon, M., Nagarajan, R. & Yaacob, S. (2010). Inferring of human emotional states using multichannel eeg. *European Journal of Scientific Research*, *48*(2), 281–299.

Myrden, A. & Chau, T. (2017). A passive eeg-bci for single-trial detection of changes in mental state. *IEEE Transactions on neural systems and rehabilitation Engineering*, *25*(4), 345–356.

Plechawska-Wojcik, M., Kaczorowska, M. & Zapala, D. (2018). The artifact subspace reconstruction (asr) for eeg signal correction. a comparative study. In *International Conference on Information Systems Architecture and Technology*, pp. 125–135. Springer.

Plotnikov, A., Stakheika, N., De Gloria, A., Schatten, C., Bellotti, F., Berta, R., Fiorini, C. & Ansovini, F. (2012). Exploiting real-time eeg analysis for assessing flow in games. In *2012 IEEE 12th International Conference on Advanced Learning Technologies*, pp. 688–689. IEEE.

Rahman, A. A., Siraji, M. I., Khalid, L. I., Faisal, F., Nishat, M. M., Islam, M. R. et al. (2022). Detection of mental state from eeg signal data: An investigation with machine learning classifiers. In *2022 14th International Conference on Knowledge and Smart Technology (KST)*, pp. 152–156. IEEE.

Robnik-Šikonja, M. & Kononenko, I. (1997). An adaptation of relief for attribute estimation in regression. In *Machine learning: Proceedings of the fourteenth international conference (ICML'97)*, volume 5, pp. 296–304.

Schoenau-Fog, H. et al. (2011). The player engagement process-an exploration of continuation desire in digital games. In *DiGRA Conference*. Citeseer.

Shaw, L. & Bagha, S. (2012). Online emg signal analysis for diagnosis of neuromuscular diseases by using pca and pnn. *International Journal of Engineering Science and Technology (IJEST)*, *4*(10), 4453–4459.

Subasi, A., Yilmaz, M. & Ozcalik, H. R. (2006). Classification of emg signals using wavelet neural network. *Journal of neuroscience methods*, *156*(1-2), 360–367.

Sundararajan, D. (2016). *Discrete wavelet transform: a signal processing approach*. John Wiley & Sons.

Tam, S., Boukadoum, M., Campeau-Lecours, A. & Gosselin, B. (2019). A fully embedded adaptive real-time hand gesture classifier leveraging hd-semg and deep learning. *IEEE transactions on biomedical circuits and systems*, *14*(2), 232–243.

Torkamani-Azar, M., Kanik, S. D., Aydin, S. & Cetin, M. (2020). Prediction of reaction time and vigilance variability from spatio-spectral features of resting-state eeg in a long sustained attention task. *IEEE Journal of Biomedical and Health Informatics*, *24*(9), 2550–2558. http://dx.doi.org/10.1109/JBHI.2020.2980056

Wu, S.-F., Lu, Y.-L. & Lien, C.-J. (2021). Measuring effects of technological interactivity levels on flow with electroencephalogram. *IEEE Access*, *9*, 85813–85822.

Wu, Z. & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, *1*(01), 1–41.

Ye, X., Ning, H., Backlund, P. & Ding, J. (2020). Flow experience detection and analysis for game users by wearable-devices-based physiological responses capture. *IEEE Internet of Things Journal*, *8*(3), 1373–1387.

Zhang, Z. & Moore, J. C. (2014). *Mathematical and physical fundamentals of climate change*. Elsevier.

# Le dépôt final du mémoire de maîtrise

Imprimé par : Arnault Major-Cardinal (major-cardinal.arnault@uqam.ca), 24.02.2023 10:43:12

| | | | |
|---|---|---|---|
| **État** | ouvert | **Âge** | 2 j 1 h |
| **Priorité** | 3 normal | **Créé** | 22.02.2023 09:14:53 |
| **File** | FS::Maitrise en informatique | **Temp passé** | 0 |
| **Verrouiller** | verrouiller | | |
| **Code client** | YOUS26559606 | | |
| **Propriétaire** | majorcardina (Arnault Major-Cardinal) | | |
| **Type** | Non défini | | |

## Information client

| | |
|---|---|
| **Prénom:** | Shaghayegh |
| **Prénom légal:** | Shaghayegh |
| **Nom:** | Yousefpourlazarjani |
| **Connexion:** | af691922 |
| **E-mail:** | yousefpourlazarjani.shaghayegh@courrier.uqam.ca |
| **Code client:** | YOUS26559606 |
| **Titre:** | Personne étudiante au deuxième cycle |
| **Programme:** | 3281 |
| **Statut:** | Étudiant |

## Article # 3

| | |
|---|---|
| **De:** | "Yousefpourlazarjani, Shaghayegh" <yousefpourlazarjani.shaghayegh@courrier.uqam.ca> |
| **À:** | "Boukadoum, A. Mounir" <boukadoum.mounir@uqam.ca>, Faculté des sciences <mi@uqam.ca> |
| **Sujet:** | Re: Autorisation de dépôt final de mémoire |
| **Créé:** | 20.02.2023 13:03:02 par client |
| **Type:** | e-mail externe |
| **Pièce jointe:** | Memoire_maitrise_YOUS26559606.pdf (8.4 MBytes)<br>YOUS26559606_3281_SDU522_20230120_172111.PDF (57.9 KBytes)<br>pages_titres.pdf (82.9 KBytes) |

```
     Bonjour,

Voici joint à ce mail, la version finale de mon mémoire de maîtrise, le
formulaire SDU522 signé et daté, et le document pdf de la page de titre +
résumé.

Merci,
Shaghayegh

From: Boukadoum, A. Mounir <boukadoum.mounir@uqam.ca>
Sent: Monday, February 20, 2023 12:35 PM
```

To: Faculté des sciences <mi@uqam.ca>
Cc: Yousefpourlazarjani, Shaghayegh
<yousefpourlazarjani.shaghayegh@courrier.uqam.ca>
Subject: Autorisation de dépôt final de mémoire

Bonjour,

Ci-joint le formulaire signé d’approbation de corrections mineures apportées au mémoire de maîtrise de Mme Shaghayegh Yousefpourlazarjani.

Cordialement,

\|/
(0 0)
_____oOO__(_)__OOo_____
A. Mounir Boukadoum, Ing./PE, PhD
Professeur/Professor

Codirecteur/codirector  ReSMiQ ([1]resmiq.org)

Dep. Informatique

Co-Gen. Chair, IEEE NEWCAS 2023, Edinburgh (UK)
_____(_)___(_)_____

[1] http://resmiq.org/