KOÇ UNIVERSITY

College of Engineering

COMP 491 – Computer Engineering Design Project

Final Report

Multi-Class Motor Imagery Classification for Virtual Drone Control

Furkan Tuna 69730

Gürkan İnal 76765

Yiğit Meriç 69170

Alp Akkanlar 76022

Ata Tütek 76646

Yücel Yemez

Sacit Karamürsel

Spring 2024

Table of Contents

“

I. Abstract

The goal of this project is to develop a Brain-Computer Interface (BCI) that can accurately classify different MI tasks and interpret them into commands for controlling a virtual drone in real-time. Accurately classifying motor imagery (MI) remains a significant challenge and continues to be an active area of research. As the field advances, various approaches have been proposed. There are numerous state-of-the-art techniques for preprocessing, feature extraction, and classification of electroencephalography (EEG) motor imagery (MI) data. Ten different subjects have been used for data acquisition. Different classifiers have been used to compare the accuracy scores. A CNN was employed to classify left-hand, right-hand, and rest states, achieving 85% training accuracy and 70% test accuracy. The limited number of subjects, difficulty in distinguishing between motor imagery and rest states, and the presence of artifacts in EEG signals contributed to the relatively modest accuracy scores. Given that our primary objective is to operate the drone with a continuous data stream and achieve controllable flight, we trained the classifier on a different dataset encompassing various eye movements (looking left, looking right) and biting.The classifier achieved a training accuracy of 96% and a testing accuracy of 90% on this dataset. Due to the higher accuracy scores, we have decided to utilize motor movements for demonstrating our project within a virtual drone environment.

II. Introduction

2.1 Background

The human brain is an elaborate structure formed by billions of interconnected nerve cells called neurons. When stimulated, these neurons send signals that can be detected using electroencephalography (EEG). To detect these signals, electrodes on an EEG headset connected to a brain-computer interface (BCI) are used. BCIs are computer systems that enable communication between the brain and external devices, and can be worn externally like a helmet or implanted into the brain. Motor imagery (MI) is a mental process in which an individual imagines performing a movement without actually executing the movement. This mental simulation activates similar neural pathways in the brain that are involved in the actual physical movement. MI is a neuroparadigm that can be recorded using a BCI. Due to the noise and susceptibility to artifacts in MI and other EEG signals, developing an accurate classification model poses significant challenges. Moreover, the overlap of brain activity from other motor functions with MI signals affects the accuracy of the model, highlighting the need for improved sensors and signal processing techniques.As MI is a trait that can be learned, the capability to perform MI varies greatly among subjects. Furthermore, the specific brain locations of MI signals can differ, and the corresponding brain wave patterns may vary for each subject. Live classification of motor imagery (MI) signals can be conducted using either synchronous or asynchronous brain-computer interfaces (BCIs).

2.2 Motivation

The motivation for this project stems from the profound potential of motor imagery (MI) brain-computer interfaces (BCIs) to transform the lives of individuals with motor impairments. By enabling direct communication between the brain and external devices, BCIs offer a non-invasive means to restore lost motor functions, providing a critical tool for neurorehabilitation and neuroprosthetics. This project aims to harness the capabilities of MI-BCIs to create innovative solutions for those who are physically impaired, enhancing their independence and quality of life. Furthermore, the exploration of BCIs extends beyond healthcare applications. The ability to control devices using brain signals opens new avenues in various fields, including gaming, monitoring, and drone control. The focus of this project on drone control through MI-BCIs exemplifies the exciting potential for human-computer interaction in real-time, providing a dynamic and interactive user experience.

2.3 Limitations

This project faces several limitations, primarily related to data acquisition, the inherent challenges of working with EEG signals, and the variability in motor imagery (MI) performance among subjects. The data acquisition process was constrained by the limited number of subjects and recording sessions, affecting the generalizability of the results. EEG signals are inherently noisy and prone to artifacts from sources such as muscle activity and eye movements, complicating accurate classification despite advanced preprocessing techniques. Additionally, MI performance varies significantly among individuals, influenced by factors such as familiarity with the task, cognitive state, and individual neural architecture. This variability poses a challenge in developing a universal model, as different subjects may exhibit different brain activity patterns for the same MI task. These limitations underscore the need for continued research and development, including advancements in data collection methods, improved signal processing techniques, and adaptive modeling approaches to account for individual differences in MI capabilities.

2.4 Theoretical Background

Neurons communicate and transfer information by employing and encoding different types of electrical signals. The oscillations of these electrical signals in the brain can be divided into five distinct frequency bands: Delta (δ), Theta (θ), Alpha (α), also known as the mu rhythm, Beta (β), and Gamma. Their frequency range can be seen in Figure 2.1.

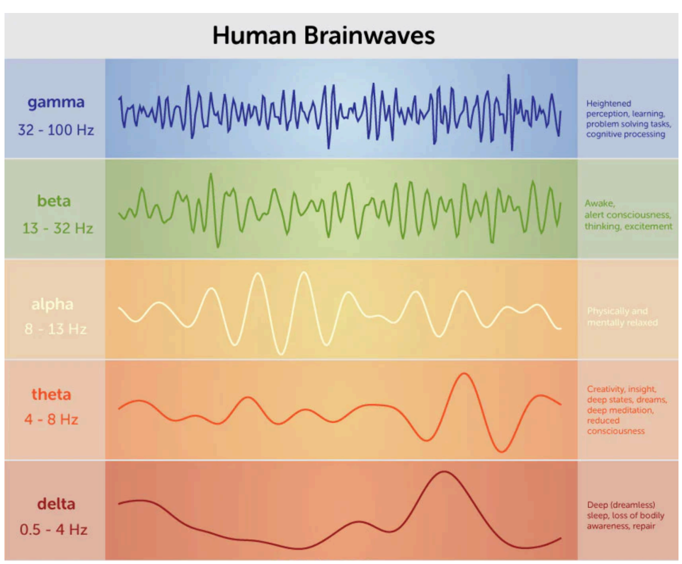


Figure 2.1: Frequency bands of brain signals. The illustration is from [1].

The motor cortex is a critical region of the brain responsible for the planning, control, and execution of voluntary movements. Brain regions can be seen in Figure 2.2. The human brain is segmented into four primary regions: the frontal lobe, parietal lobe, occipital lobe and temporal lobe [2]. Located in the posterior portion of the frontal lobe, it is divided into several areas, including the primary motor cortex (M1), the premotor cortex, and the supplementary motor area (SMA). The primary motor cortex, situated on the precentral gyrus, is directly involved in generating neural impulses that control the execution of movement. The premotor cortex and SMA are involved in the planning and coordination of movements. Neurons within the motor cortex are organized somatotopically, meaning different regions correspond to specific parts of the body, a concept often illustrated by the motor homunculus. This precise organization allows for the fine-tuned control of complex motor actions. The motor cortex interacts extensively with other brain regions, including the basal ganglia, cerebellum, and sensory cortices, to integrate sensory feedback and modulate motor commands, ensuring smooth and coordinated movements. Understanding the motor cortex's structure and function is essential for developing effective brain-computer interface (BCI) systems and for advancing treatments for motor impairments.

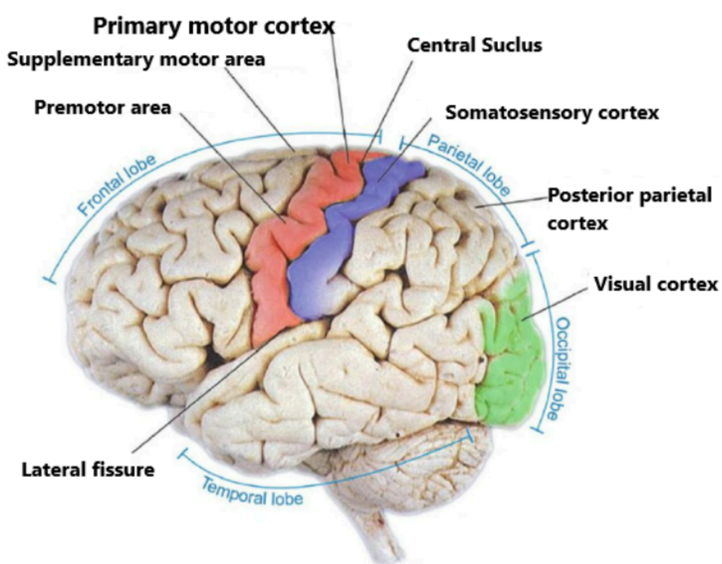


Figure 2.2: The human brain and its significant regions. Illustration adapted from [2].

This diagram depicts the areas of the motor cortex that correspond to the control of different parts of the body. The motor homunculus demonstrates that the primary motor cortex is organized in a manner where specific regions control distinct body parts, with the size of each body part in the illustration reflecting the degree of motor control or the number of neurons dedicated to that part in Figure 2.3. The exaggerated sizes of the hands, face, and tongue in the homunculus indicate the fine motor skills and higher sensory acuity required by these body parts, necessitating more cortical neurons to manage their movements [3].

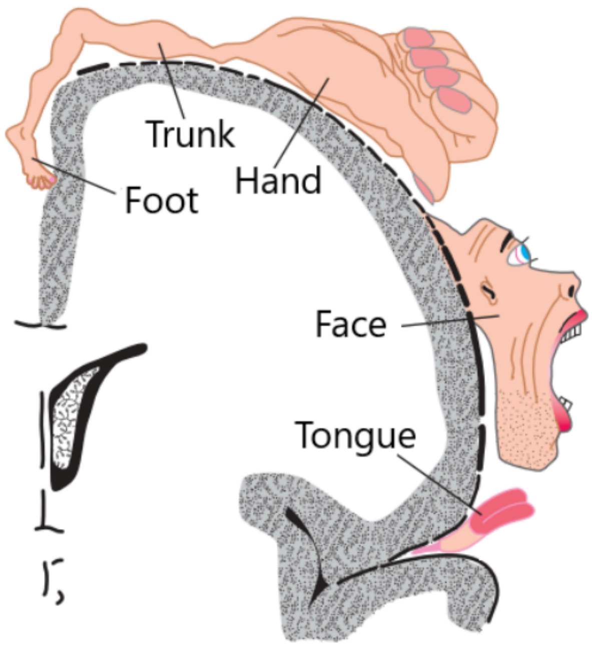


Figure 2.3: The primary motor cortex is structured with specific areas representing different body parts. Illustration is adapted from [3].

2.5 Electroencephalography

EEG allows for non-invasive exploration of the human brain, with signals collected through scalp-placed electrodes. The sensors capture the direct cortical activity, specifically the fluctuations in the amplitudes of electrical impulses. In a typical adult, these amplitudes typically range from 10μV to 100μV [3]. The parameters extracted via EEG are influenced by numerous factors, such as the subject's itself, age and mental health. EEG is widely employed in diagnosing neurological conditions such as Alzheimer's disease, schizophrenia, multiple sclerosis, stroke, and migraines .

Electrode placement is a critical aspect of electroencephalography (EEG) that significantly impacts the quality and accuracy of the recorded signals. Proper positioning of electrodes on the scalp is essential to ensure reliable data acquisition, as it determines the ability to capture specific brain activities corresponding to different regions. The standardized 10-20 system is commonly used for electrode placement, providing a systematic approach to positioning electrodes based on the anatomical landmarks of the skull. In the 10-20 system, electrode placement corresponds to specific areas of the cerebral cortex, with "10-20" indicating the spacing between neighboring electrodes. The 10-20 placing scheme is illustrated in figure 2.4.

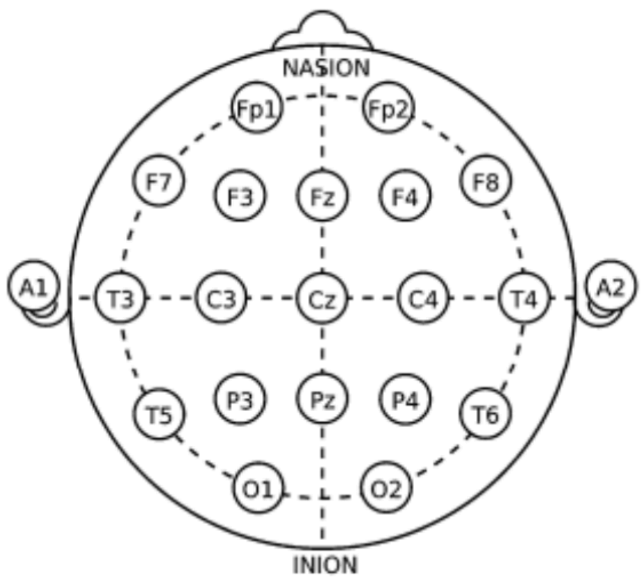


Figure 2.4: Illustration of the 10-20 placing scheme. Illustration is reused from [4].

EEG artifacts can be divided into physiological and non-physiological types. Physiological artifacts are generated by the body, with common examples being eye blinks, eye movements, head movements, heartbeats, and muscular noise. Non-physiological artifacts originate from the environment surrounding the subject. These artifacts are typically characterized by their anterior location, bilateral and synchronized appearance, and their linear summation with the EEG signal. To identify ocular artifacts, additional electrodes can be placed above and below the eyes. Non-physiological artifacts in EEG recordings are extrinsic disturbances that originate from the environment surrounding the subject rather than from the subject's own biological activity. These artifacts can significantly affect the quality and accuracy of EEG data. Common sources of non-physiological artifacts include electrical interference from nearby electronic devices, fluctuations in power supply, electromagnetic fields, and movements or vibrations in the environment.

III. System Design

3.1 Overview

The system design of our project, Mental Aviate, focuses on developing a Brain-Computer Interface (BCI) system for virtual drone control using specific physical actions detected through EEG signals. These actions include biting, moving eyes to the left, moving eyes to the right, and resting. The system is divided into several critical components: EEG Acquisition, Signal Processing, Feature Extraction and Classification, and the Real-Time Control Interface, which can be seen in Figure 3.1.

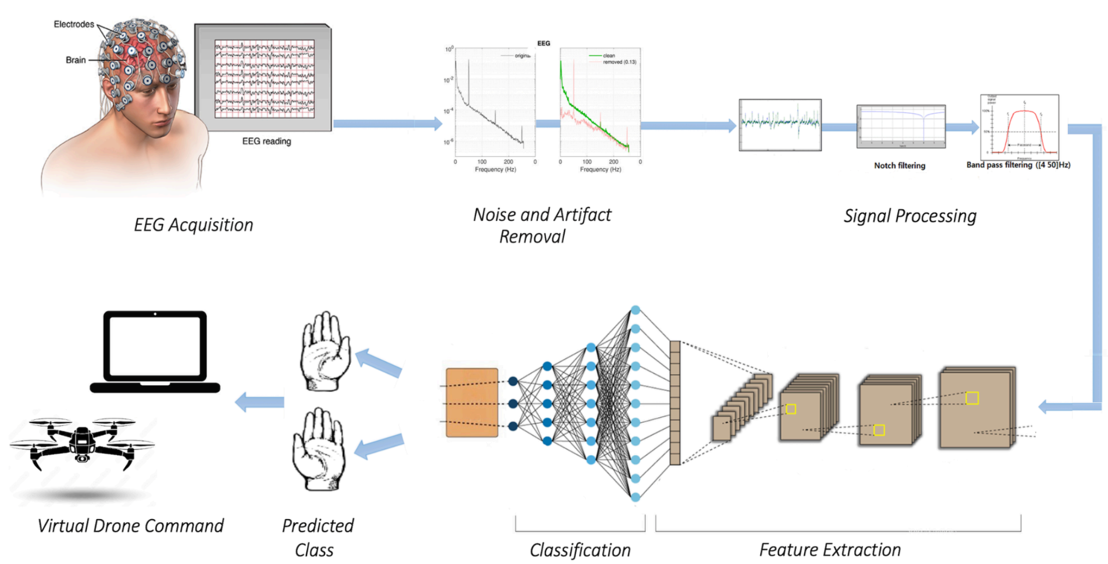


Figure 3.1: Pipeline including acquisition, preprocessing, virtual environment and classifier.

3.2 EEG Data Acquisition

We utilize an 8-electrode EEG headset for data acquisition. The electrodes are positioned according to the 10-20 system to capture the brain's electrical activity associated with specific actions. The EEG headset records high-resolution signals, which are crucial for accurate interpretation of these commands. The electrode placement is strategically designed to capture the relevant signals for biting and eye movements, with the 10-20 system standardizing this placement to enhance the reliability of the data acquisition process.

3.3 Signal Processing

The EEG signals undergo several preprocessing steps to ensure data quality and usability. Raw EEG signals often contain noise and artifacts from muscle movements, eye blinks, and external electrical sources. These unwanted components are removed using filtering techniques to ensure clean data for further processing. A bandpass filter ranging from 8-30 Hz is applied to retain frequencies relevant to the targeted actions while eliminating low-frequency drifts and high-frequency noise. Additionally, a notch filter centered at 50 Hz (48-52 Hz) is used to remove powerline interference, which is a common source of noise in EEG recordings.

3.4 Feature Extraction and Classification

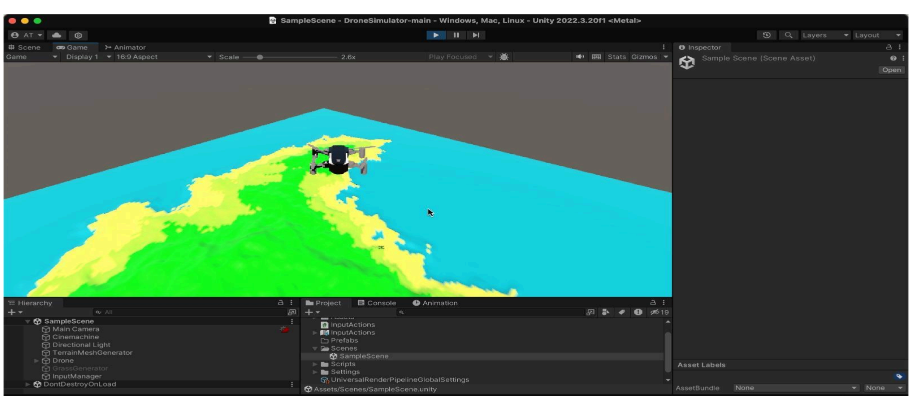
After preprocessing, the EEG signals are processed through a Convolutional Neural Network (CNN) to classify the detected actions. The CNN is designed to extract relevant features from the EEG signals and make accurate predictions regarding the user's intended action. The CNN architecture includes multiple convolutional layers that automatically learn to identify significant patterns in the EEG data. These layers extract hierarchical features that are crucial for distinguishing between different actions. The extracted features are classified into four categories: biting to fly and move the drone forward, moving eyes to the left to rotate the drone left, moving eyes to the right to rotate the drone right, and resting. The CNN is trained using a labeled dataset of EEG recordings corresponding to these actions, ensuring the model can generalize well to new data.

3.5 Real-Time Control Interface

The classified data is transmitted in real-time through a Lab Streaming Layer (LSL) interface to control the drone. The LSL ensures low-latency communication between the EEG processing system and the drone, allowing for responsive and accurate navigation based on the user's detected actions. The predicted class from the CNN is converted into specific commands for the drone. For instance, biting results in a command to make the drone fly and move forward, moving eyes to the left results in a command to rotate the drone left, moving eyes to the right results in a command to rotate the drone right, and resting maintains the drone's current position. The system's architecture ensures minimal delay between the user's actions and the corresponding drone movement, providing a seamless control experience.

3.6 System Workflow

The system workflow begins with the user wearing the EEG headset, and the electrodes capturing brain activity. The raw EEG signals undergo preprocessing to remove noise and artifacts. The cleaned signals are then filtered to isolate relevant frequency bands. These filtered signals are fed into the CNN for feature extraction. The CNN classifies the extracted features into specific actions, and the classified states are transmitted via LSL to control the drone's movements. This detailed system design ensures that our project effectively interprets specific actions from EEG signals and translates them into precise drone commands, enabling intuitive and responsive drone navigation.This detailed system design ensures that our project effectively interprets specific actions from EEG signals and translates them into precise drone commands, enabling intuitive and responsive drone navigation.



“

IV. Analysis and Results

4.1 Project Deliverables

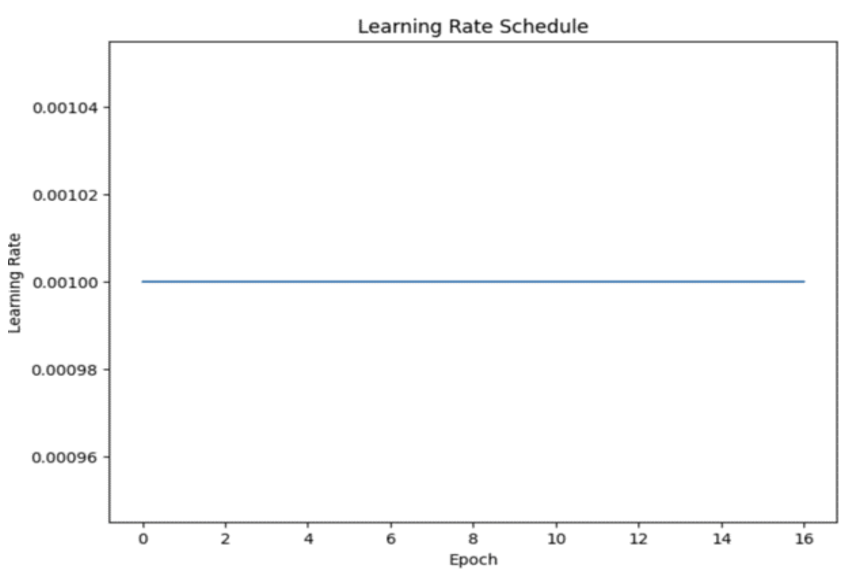
This project aimed to develop a Convolutional Neural Network (CNN) for classifying EEG signals into four categories: bite, left, rest, and right. The primary deliverables included:

1. **Data Preprocessing**: Merging and reshaping EEG data.
2. **Model Development**: Building a CNN model.
3. **Training and Validation**: Training the model with appropriate callbacks and evaluating its performance.
4. **Performance Analysis**: Visualizing various metrics and performance indicators.
5. **Unity Simulation**: Implementing a virtual drone control simulation using the classified EEG signals to validate the model's practical application.

Analysis of Results

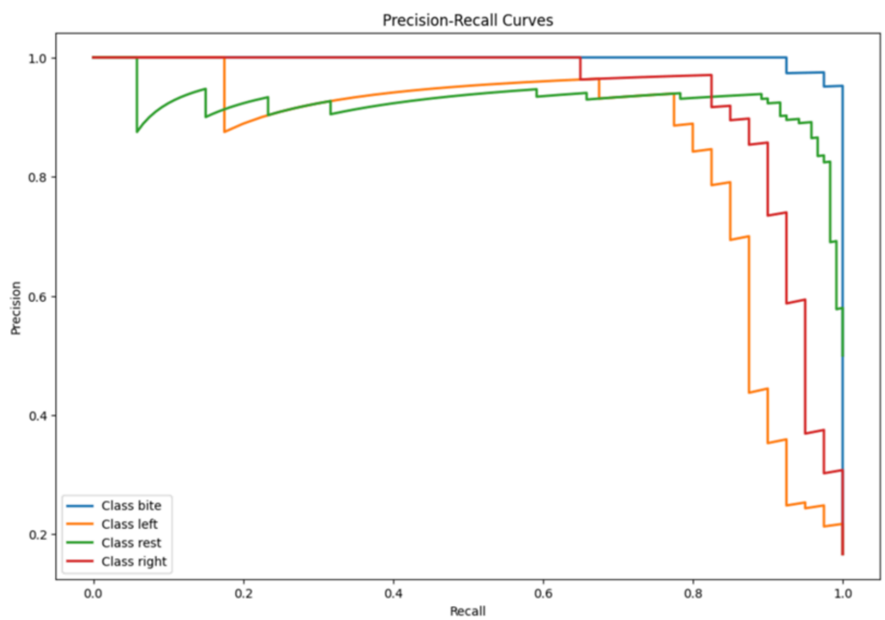
4.2 Learning Rate Schedule

The learning rate schedule (Figure 4.2) indicates that the learning rate remained constant at 0.001 throughout the training process. This steady learning rate suggests that the optimizer did not adjust the learning rate based on the training progress, which could be beneficial in some scenarios where a consistent learning rate is required.



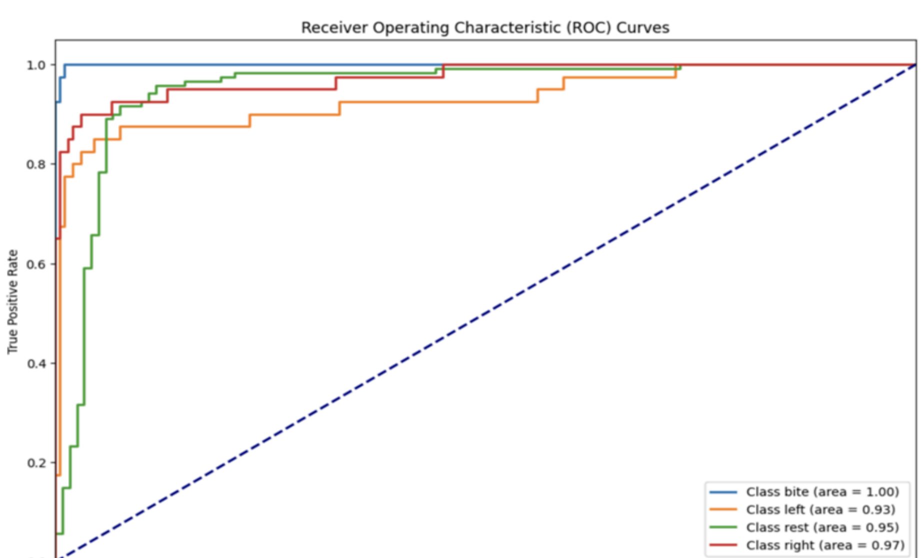
4.3 Precision-Recall Curves

The precision-recall curves (Figure 2) for each class show how well the model distinguishes between different categories. The curve for the 'bite' class is the highest, indicating that the model has the highest precision and recall for this class. The other classes, 'left', 'rest', and 'right', show lower precision and recall, with the 'rest' class being the most challenging for the model to classify correctly.



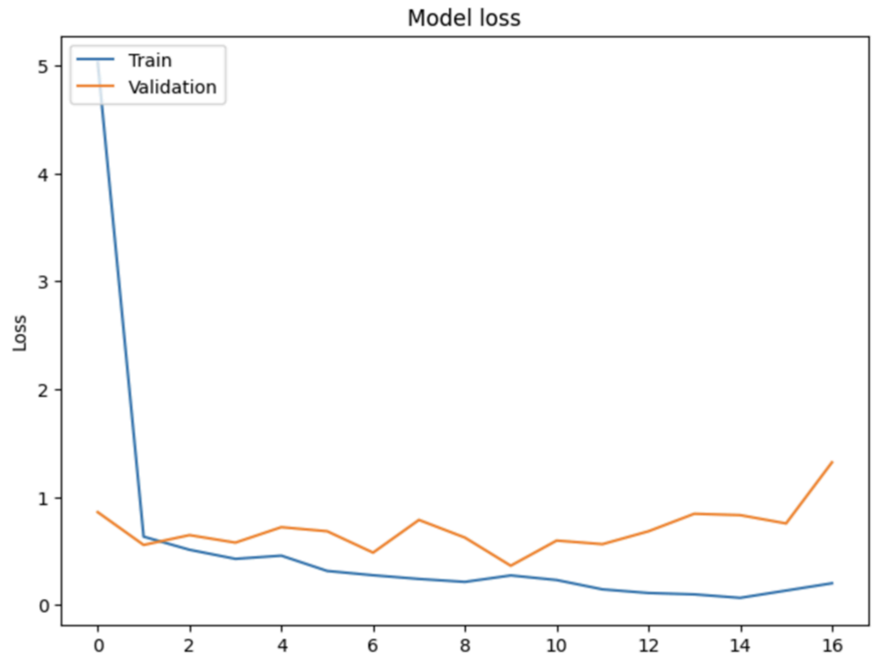
4.4 ROC Curves

The ROC curves (Figure 3) provide another perspective on the model's performance. The 'bite' class again shows excellent performance with an area under the curve (AUC) of 1.00, indicating perfect classification. The 'left', 'rest', and 'right' classes have AUCs of 0.93, 0.95, and 0.97, respectively, which are also strong indicators of good model performance.



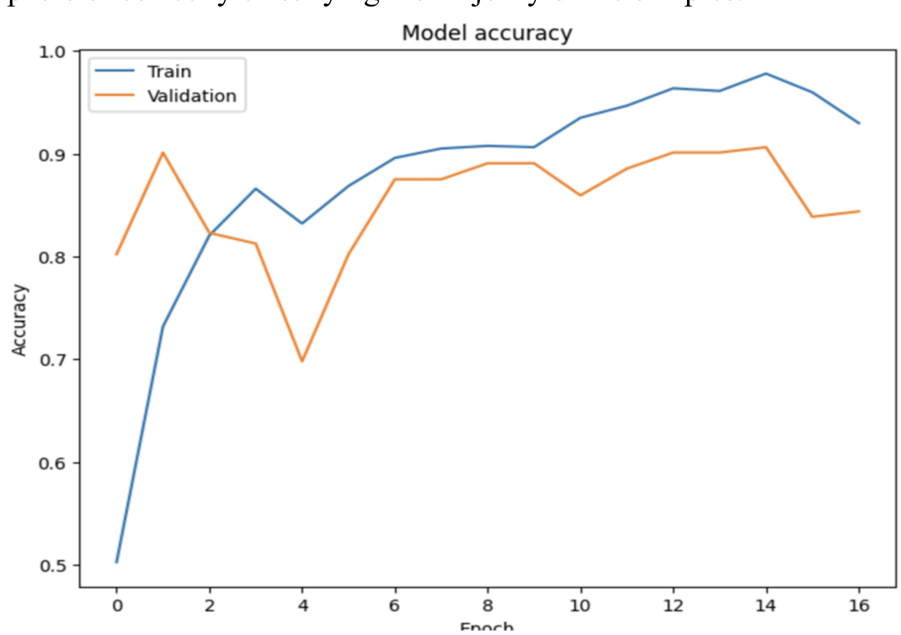
4.5 Model Loss

The model loss plot (Figure 5) shows the training and validation loss over the epochs. The initial epochs show a sharp decrease in loss, indicating that the model is learning effectively. However, the validation loss remains somewhat volatile, suggesting that the model might be overfitting to the training data after a certain point.



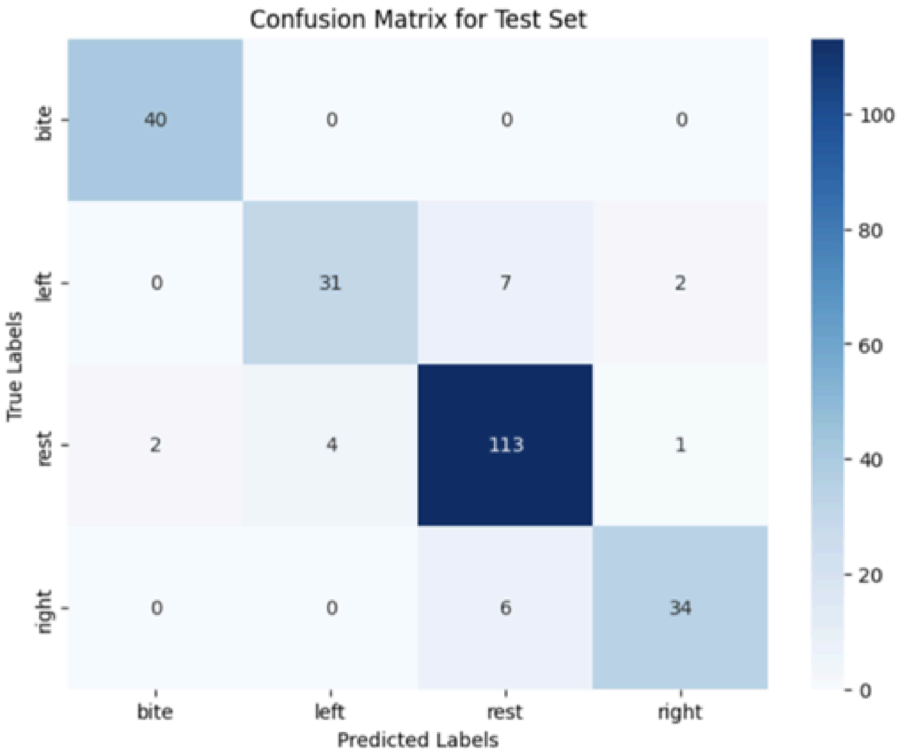
4.6 Model Accuracy

The model accuracy plot (Figure 6) shows the training and validation accuracy over the epochs. The training accuracy steadily increases, while the validation accuracy fluctuates, again indicating potential overfitting. However, the overall high accuracy values indicate that the model is capable of correctly classifying the majority of the samples.



4.7 Confusion Matrix

The confusion matrix (Figure 7) provides a detailed look at the model's performance on the test set. The matrix shows that the model performs well on 'bite' and 'right' classes, with very few misclassifications. However, there are some misclassifications between 'left' and 'rest' classes, indicating that these two classes are harder to distinguish for the model.



The project's outcomes demonstrate how well a Convolutional Neural Network (CNN) can classify EEG signals into the bite, left, rest, and right categories. The model has demonstrated a strong ability to generalize from the training data to new, unseen data, as evidenced by its overall test accuracy of 90.83%.

The precision-recall and ROC curves highlight the strengths and weaknesses of the model across different classes. The 'bite' class, with an AUC of 1.00 and high precision-recall values, shows that the model can identify this category with near-perfect accuracy. This might be due to distinctive signal patterns associated with the 'bite' action that the CNN is able to capture effectively. In contrast, the 'left' and 'rest' classes present more challenges, as evidenced by their relatively lower precision-recall values and AUCs of 0.93 and 0.95, respectively. This suggests that the EEG signals for these actions may share more similarities, making them harder to distinguish.

The loss and accuracy plots reveal important insights into the model's learning dynamics. The sharp decline in training loss and the corresponding increase in training accuracy during the initial epochs demonstrate that the model quickly learns to fit the training data. However, the fluctuations in validation loss and accuracy indicate that while the model is learning, it might be overfitting to the training set after a certain point. This overfitting can occur when the model starts to memorize the training data rather than learning generalizable features, which is a common issue in machine learning and deep learning models.

The confusion matrix provides a detailed view of the model's performance on the test set. The high number of correct predictions for the 'bite' and 'right' classes shows the model's robustness in these categories. However, the model's difficulty in distinguishing between 'left' and 'rest' indicates a need for further refinement. This difficulty could stem from overlapping signal characteristics or insufficient feature extraction from the CNN layers.

The steady learning rate observed in the learning rate schedule suggests that while a constant learning rate can stabilize training, experimenting with adaptive learning rates could potentially improve model performance. Techniques such as learning rate annealing or using optimizers with dynamic learning rates, like Adam or RMSprop, might help the model converge more effectively and reduce overfitting.

Additionally, the project included a Unity simulation for controlling a virtual drone using the classified EEG signals. This simulation was an integral part of the project, demonstrating the practical application of the model in a real-time environment. The simulation showed that the model's predictions could be effectively translated into control commands for the drone, validating the model's utility in practical scenarios.

In summary, the project successfully demonstrated the application of a CNN for classifying EEG signals with high accuracy. The results underscore the model's ability to learn and generalize from EEG data, particularly for distinct signal patterns like those in the 'bite' class. Future work should focus on addressing overfitting, enhancing feature extraction capabilities, and experimenting with adaptive learning rates to improve classification performance, especially for more challenging categories like 'left' and 'rest'. This could involve deeper network architectures, more sophisticated preprocessing techniques, or integrating additional types of data to provide the model with more context. Moreover, expanding the Unity simulation to include more complex control scenarios could further validate and refine the model's real-world applicability.

V. Conclusion

To conclude, through our classifier deployment and virtual drone implementation, we have implemented a mind-controlled drone. There, we achieved an adequate accuracy level in terms of classifying the input brain signals through which we can control the drone.

In such a procedure, we achieved our Project objectives through an EEG headset, our system architecture & our virtual drone implementation. There, we have managed to deliver a BCI system for drone control using Motor-Imagery classification. There, although we had several issues upon drone shipping; we have managed to overcome such problems by implementing our classification model upon the virtual drone formed in Unity.

Through our implementation of a CNN neural network & our train/test scores as discussed in the sections above, we can control our virtual drone only using the user's brain signals which completes our objectives of our project. Therefore, as a result we have achieved all our goals by implementing a mind-controlled drone that can be utilized in many sectors in real-life applications. With our implementation, we have completed all our objectives within our project without any missing section in an efficient, accurate and sustainable manner. Further, along our path of achieving our end-product, we have developed many systems with distinct accuracy levels in which we developed different data acquiring techniques for users' brain signals. However, there, we eliminated each system one-by-one to gather the best and most precise & accurate design in terms of both data acquiring and classifier modeling.

Although we have achieved all our objectives as described above, there are some possible design improvements. In such a sense, firstly, due to our drone shipping problems, we have implemented our virtual drone in Unity. There, this approach made us conquer incoming unexpected problems in the most efficient manner in very little time. On the other hand, directing our system implementation -classifier and real time deployment- into a real, physical drone would have made the implementation much more refulgent. In such a way, our system implantation would have been enhanced through the deployment of our system design onto a physically flying drone. Moreover, another system improvement can be achieved through more data acquisition. There, there is no limit for collecting more data to improve the CNN classifier in terms of test/train accuracy levels. Therefore, although we have collected lots of data to improve our classifier model; with even more data gathering, we can enhance our system design.

In terms of future real-world implementations of our project, we can utilize our classification model also on a physical drone which can be used in many industries including transportation, shipping, and many more. In every sector introducing a vehicle to let an object move from a place to another, our implementation of a mind-controlled drone can form a great opportunity. There, in terms of transportation, we can utilize our mind-controlled drone project for:

Personal Mobility: Accessibility (for individuals with mobility impairments), Emergency Response (for quickly accessing the inaccessible areas)

Shipping and Delivery: Efficient Logistics (for companies' stream live delivery processes)

Agriculture: Crop Monitoring & Targeted Treatment

Security & Public Safety: Law Enforcement, Event Security

Healthcare: Telemedicine (for transporting diagnostic tools or becoming communication relays between doctors and patients), Rapid Response (for remote or disaster-stricken areas)

Research: Wildlife conservation, Climate Studies

Entertainment: Filmmaking & Live Broadcasting

Custom Innovative Applications: Unique Applications (for distinct industries)

There, we should also note that in such real-world applications, there are several technical and ethical challenges. There, in terms of technical issues, signal interface in various environments, and battery life of drones encapsulate the main problems. Additionally, in terms of ethical issues, privacy concerns of users' brain signals, and additional regulations for aviation

& safety occupy the most significant challenges.

Consequently, to wrap everything up, through our implementation of a CNN neural network upon a virtual drone that is formed in Unity, we have achieved our goal of developing a mind-controlled drone. There, through adequate levels of accuracies in terms of testing, we have managed to classify the users' brain signals to control our virtual drone. By utilizing an EEG headset, our system architecture and our virtual drone implementation, we managed to deliver a BCI system design to control a drone with Motor-Imagery classification. There, due to our problem of drone shipping, we have decided to implement such classification upon our virtual drone that is formed in Unity. Through such a path, we did not only complete all of our objectives for the project by command handling through brain signals; but also, we learned how to solve problems in critical scenarios in real life problems.

VI. References

[1]choosemuse.com/blogs/news/a-deep-dive-into-brainwaves-brainwave-frequencies-explained-2

[2] Bernhard Graimann, Brendan Z Allison, and Gert Pfurtscheller. *Brain-computer interfaces: Revolutionizing human-computer interaction*. Springer Science & Business Media, 2010.

[3] Larry Squire,Darwin Berg,FloydEBloom,Sascha Du Lac,Anirvan Ghosh, and Nicholas C Spitzer. *Fundamental Neuroscience, Third Edition*. Academic press, 2008.

[4] Li Hu and Zhiguo Zhang. *EEG signal processing and feature extraction*. Springer, 2019.

[5] W. D. Hairston et al., "Usability of four commercially-oriented EEG systems," Journal of Neural Engineering, vol. 11, no. 4, Art. no. 046018, 2014. [Online]. Available: https://iopscience.iop.org/article/10.1088/1741-2560/11/4/046018/pdf.

[6] P. Wierzgała, D. Zapała, G. M. Wójcik, and J. Masiak, "Most Popular Signal Processing Methods in Motor-Imagery BCI: A Review and Meta-Analysis," Frontiers in Neuroinformatics, vol. 12, Art. no. 78, Nov. 2018. [Online]. Available: Frontiers | Most Popular Signal Processing Methods in Motor-Imagery BCI: A Review and Meta-Analysis (frontiersin.org).

[7] V. Mondini, A. L. Mangia, and A. Cappello, "EEG-Based BCI System Using Adaptive Features Extraction and Classification Procedures," Computational Intelligence and Neuroscience, vol. 2016, Art. no. 4562601, 14 pages, 2016. [Online]. Available: https://downloads.hindawi.com/archive/2016/4562601.pdf.

[8] P. Peining, G. Tan, and A. A. P. Wai, "Evaluation of Consumer-Grade EEG Headsets for BCI Drone Control," School of Computing, National University of Singapore and Neural & Biomedical Technology Department, Institute for Infocomm Research, Year. [Online]. Available: https://ircset.org/anand/2017papers/IRC-SET_2017_paper_S6-4.pdf.

[9] A.-L. Rusnac and O. Grigore, "CNN Architectures and Feature Extraction Methods for EEG Imaginary Speech Recognition," Sensors, vol. 22, no. 13, Art. no. 4679, Jul. 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9268757/.

[10] E. Antoniou et al., "EEG-Based Eye Movement Recognition Using Brain–Computer Interface and Random Forests," Sensors, vol. 21, no. 7, Art. no. 2339, Apr. 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8036672/.

VII. Appendix

7.1 Source Code

The source code for the project "COMP491-MentalAviate" is available on GitHub and can be accessed via the following link:

https://github.com/aakkanlar/COMP491-MentalAviate

And the codes for the unity drone simulation is available on GitHub and can be accessed via the following link:

https://github.com/gurkaninal/DroneSimulator