Faculty of Computer Science and Information Technology

CLASSIFYING ATTENTION STATES FROM BRAIN SIGNALS USING CONVOLUTIONAL NEURAL NETWORK (CNN)

Angelina Anak Heroldson

Bachelor of Software Engineering with Honors

2025

CLASSIFYING ATTENTION STATES FROM BRAIN SIGNALS USING CONVOLUTIONAL NEURAL NETWORK (CNN)

ANGELINA ANAK HEROLDSON

This project is submitted in partial fulfillment of the
requirements for the degree of
Bachelor of Software Engineering with Honors

Faculty of Computer Science and Information Technology

UNIVERSITI MALAYSIA SARAWAK

2025

CLASSIFYING ATTENTION STATES FROM BRAIN SIGNALS USING CONVOLUTIONAL NEURAL NETWORK (CNN)

ANGELINA ANAK HEROLDSON

Projek ini merupakan salah satu keperluan untuk
Ijazah Sarjana Muda Kejuruteraan Perisian dengan Kepujian

Fakulti Sains Komputer Dan Teknologi Maklumat

UNIVERSITI MALAYSIA SARAWAK

2025

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE CLASSIFYING ATTENTION STATES FROM BRAIN SIGNALS USING CONVOLUTIONAL NEURAL NETWORK (CNN)

ACADEMIC SESSION: 2024/2025

ANGELINA ANAK HEROLDSON

hereby agree that this Thesis* shall be kept at the Centre for Academic Information Services, Universiti Malaysia Sarawak, subject to the following terms and conditions:

1. The Thesis is solely owned by Universiti Malaysia Sarawak
2. The Centre for Academic Information Services is given full rights to produce copies for educational purposes only
3. The Centre for Academic Information Services is given full rights to do digitization in order to develop local content database
4. The Centre for Academic Information Services is given full rights to produce copies of this Thesis as part of its exchange item program between Higher Learning Institutions [or for the purpose of interlibrary loan between HLI]
5. ** Please tick (√)

☐	CONFIDENTIAL	(Contains classified information bounded by the OFFICIAL SECRETS ACT 1972)
☐	RESTRICTED	(Contains restricted information as dictated by the body or organization where the research was conducted)
☑	UNRESTRICTED	

Validated by

(AUTHOR'S SIGNATURE)

(SUPERVISOR'S SIGNATURE)

Dr Liew Siaw Hong
Lecturer
Faculty of Computer Science and Information Technology
Universiti Malaysia Sarawak

Permanent Address

LORONG 12K2C TAMAN UNI GARDEN 94300 KOTA SAMARAHAN, SARAWAK.

Date: 20/06/2025

Date: 20/06/2025

Note * Thesis refers to PhD, Master, and Bachelor Degree
** For Confidential or Restricted materials, please attach relevant documents from relevant organizations / authorities

DECLARATION

I hereby declare that this project is my original work. I have not copied from any other student's work or from any other sources except where due to reference or acknowledgement is not made explicitly in the text, nor has any part had been written for me by another person.

……………………………………

(Angelina Anak Heroldson) 20th June 2025

Matric No: 77829

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Dr. Liew Siaw Hong, for her guidance and supervision throughout my final year project. Second, I would like to thank my examiner, Dr. Noralifah Binti Annuar, for providing useful feedback and advice on my final-year project. I would also like to thank my Final Year Project Coordinator, Professor Dr. Wang Yin Chai, who gave helpful guidance and tips during lecture.

Furthermore, I would like to thank my university, Universiti Malaysia Sarawak (UNIMAS), and my faculty, the Faculty of Computer Science and Information Technology, for providing me with this final-year project opportunity. I am extremely appreciative because this project allows me to explore and learn the domain of Artificial Intelligence (AI). Special thanks to my family and friends for their support in completing this project on time.

TABLE OF CONTENTS

DECLARATION I
ACKNOWLEDGEMENT II
TABLE OF CONTENTS III
LIST OF FIGURES V
LIST OF TABLES VI
ABSTRACT VII
ABSTRAK VIII
CHAPTER 1: INTRODUCTION 1
1.1 Introduction 1
1.2 Problem Statement 2
1.3 Scope 3
1.4 Objectives 3
1.5 Methodology 3
1.5.1 Data Description 4
1.5.2 Data Preprocessing and Preparation 4
1.5.3 Feature Extraction and Modelling 5
1.5.4 Performance Measurement 6
1.6 Significance of Project 7
1.7 Project Schedule 7
1.8 Expected Outcome 8
1.9 Project Outline 8
1.10 Summary 9
CHAPTER 2: LITERATURE REVIEW 10
2.1 Introduction 10
2.2 Background on EEG and Attention States 10
2.2.1 Protocols 12
2.2.2 Electrode Placements 13
2.3 Overview of Attention Classification Techniques 15
2.3.1 K-Nearest Neighbour (KNN) Model 15
2.3.2 Support Vector Machine (SVM) Model 17
2.3.3 Models Comparison 18
2.4 CNN Architecture for EEG Signal Analysis 25

2.5 Summary 26
CHAPTER 3: METHODOLOGY 27
3.1 Introduction 27
3.2 Experimental Design and Framework 27
3.3 Dataset Description and Preparation 28
3.4 CNN Model Architecture 30
3.5 Performance Evaluation 32
3.6 Benchmarking 33
3.7 Summary 34
CHAPTER 4: IMPLEMENTATION, RESULT, AND DISCUSSION 35
4.1 Introduction 35
4.2 CNN Implementation and Training 35
4.2.1 Software and Tools 35
4.2.2 Detailed CNN Architecture 36
4.2.3 Training Procedure and Callbacks 39
4.3 Results and Comparative Analysis 40
4.3.1 CNN Classification Performance 40
4.3.2 K-NN and SVM Classification Performance 44
4.3.3 Cross-Model Comparison and Discussion 45
4.4 Summary 47
CHAPTER 5: CONCLUSION 49
5.1 Introduction 49
5.2 Contribution 49
5.3 Limitation 50
5.4 Future Works 50
REFERENCES 52
APPENDICES 55
Appendix A: Project Schedules for FYP1 and FYP2 56

LIST OF FIGURES

Figure 1.1 Typical CNN Architecture Employed for EEG Signal Analysis 5
Figure 2.1 Six Primary Brain Region.. 12
Figure 2.2 International 10–20 Electrode Placements 14
Figure 2.3 Example of KNN Graph 16
Figure 2.4 Example of SVM Graph 17
Figure 3.1 Framework of the Proposed Methodology 28
Figure 3.2 Visual Stimuli Presentation 29
Figure 3.3 CNN Model Architecture Visual Representation 31
Figure 3.4 Confusion Matrix 32
Figure 4.1 CNN Model Accuracy Learning Curve Graph 41
Figure 4.2 CNN Model Loss Learning Curve Graph... 42
Figure 4.3 CNN Model Confusion Matrix 43
Figure 4.4 KNN Model Confusion Matrix 44
Figure 4.5 SVM Confusion Matrix 45
Figure 4.6 Bar Chart of Model Metrics Comparison 46

LIST OF TABLES

Table 2.1 EEG Signal Rhythms 11

Table 2.2 Comparison from Previous Researches 24

Table 4.1 CNN Model Parameter Settings 37

ABSTRACT

The ability to monitor and classify attention states has profound implications in areas like education, healthcare, workplace productivity, and transportation safety. This study explores the classification of human attention states from brain signals using convolutional neural networks (CNNs) with a focus on distinguishing between focused and distracted states under two environmental conditions. EEG (Electroencephalography) signals were employed as the primary data source as it offers insights into neural patterns associated with cognitive states. EEG signals in the alpha band (8–13 Hz) were recorded from 45 healthy adults under both quiet and simulated noisy (70 dB) conditions. After preprocessing including band-pass filtering, artifact correction, and segmenting signals into 21-channel, one-second epochs, the data were fed into a three-layer convolutional neural network (CNN) featuring successive convolution, max-pooling, batch normalization, and dropout operations. The network converged into 37 epochs and achieved 91% accuracy, 93% precision, 91% recall, and a 0.92 F1-score. Comparative benchmarks against k-Nearest Neighbors (76% accuracy) and Support Vector Machine with PCA (87% accuracy) demonstrated CNN's superior consistency, reduced error rates, and robust feature learning. Detailed evaluation via confusion matrices and learning curves provided insights into class-specific performance and model convergence behavior. These results confirm the feasibility of real-time EEG-based attention monitoring and establish a reproducible implementation and evaluation pipeline for future cognitive state classification research.

ABSTRAK

Keupayaan untuk memantau dan mengklasifikasikan tahap perhatian mempunyai implikasi yang mendalam dalam bidang seperti pendidikan, penjagaan kesihatan, produktiviti tempat kerja, dan keselamatan pengangkutan. Kajian ini meneroka klasifikasi keadaan perhatian manusia menggunakan isyarat otak (EEG) dengan rangkaian neural konvolusional (CNN), memfokuskan pada pembezaan antara keadaan fokus dan teralih di bawah dua keadaan persekitaran. Isyarat EEG dalam jalur alfa (8–13 Hz) direkodkan daripada 45 orang dewasa sihat di kedua-dua keadaan sunyi dan simulasi bising (70 dB). Selepas pra-pemprosesan, termasuk penurasan jalur lebar, pembetulan artifak, dan pembahagian isyarat kepada segmen satu saat dengan 21 saluran, data tersebut dimasukkan ke dalam rangkaian neural konvolusional tiga lapisan yang mempunyai operasi konvolusi berurutan, pengepalaan maksimum (max-pooling), penskalaan kelompok (batch normalization), dan dropout. Rangkaian tersebut mencapai penumpuan setelah 37 epok dan mencatat ketepatan 91%, ketulenan (precision) 93%, panggilan (recall) 91%, dan skor F1 0.92. Penanda aras perbandingan terhadap k-Nearest Neighbors (76% ketepatan) dan Sokongan Vektor Mesin dengan PCA (87% ketepatan) menunjukkan keunggulan CNN dari segi konsistensi, kadar kesilapan yang lebih rendah, dan pembelajaran ciri yang kukuh. Penilaian terperinci melalui matriks kekeliruan dan lengkung pembelajaran memberikan pandangan tentang prestasi khusus kelas dan tingkah laku penumpuan model. Keputusan ini mengesahkan kebolehlaksanaan pemantauan perhatian masa nyata berasaskan EEG dan mewujudkan saluran pelaksanaan dan penilaian yang boleh diulang untuk penyelidikan klasifikasi keadaan kognitif masa depan.

CHAPTER 1: INTRODUCTION

1.1 Introduction

The ability to pay attention is important as it helps us focus our awareness on our environment, making important decisions, learn effectively, perform tasks and communicate well. Attention state can vary significantly in response to environmental factors like noise and distractions. In recent years, the ability to monitor and understand human attention has gained significance across various fields, from education, healthcare and workplaces. With advancements in technology, researchers are increasingly exploring brain-computer interface (BCI) technologies to assess cognitive states, including attention, through electroencephalography (EEG). EEG records electrical activity in the brain by placing electrodes on the scalp. These signals result from ionic currents generated in neurons and vary in patterns, reflecting different brain states (Liew et al., 2015). However, it remains challenging to achieve accurate attention classification using EEG data due to the sensitivity of EEG signals to noise and cognitive distractions (Landau et al., 2020).

Convolutional Neural Networks (CNNs) are deep-learning architectures designed to learn hierarchical features directly from raw and high-dimensional inputs (Tao et al., 2021). A typical CNN consists of multiple convolutional layers with each applying trainable filters that slide over the input and followed by pooling layers that down sample feature maps and fully connected layers that output class probabilities via softmax (Rajwal & Aggarwal, 2023). In essence, convolutions act as learnable filters that detect local patterns (e.g., edges or textures) via dot products, while pooling reduces spatial resolution to improve computational efficiency and robustness to small translations. CNNs excel at extracting increasingly abstract representations. The early layers capture low which is the level features (edges, gradients), mid-level layers combine these into textures or shapes, and deep layers form high-level concepts (objects or faces). Parameter sharing (applying the same filter across all positions) and sparse connectivity (filters cover only local regions) reduce the total number of trainable parameters compared to fully connected networks. Hierarchical design of CNN enables automatic feature learning and eliminates the need for manual feature engineering (Rajwal & Aggarwal, 2023).

In a study by Atilla and Alimardani (2021), the authors explore the use of conventional methods, which often rely on traditional machine learning techniques, such as Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN). These models require extensive preprocessing and are generally designed for controlled settings. The authors found that using CNNs on raw EEG data significantly improved classification accuracy, achieving up to 89% accuracy under realistic driving conditions. This highlights CNNs' ability to automatically extract critical features from EEG. Previous study, such as (Craik et al., 2019), have reviewed the potential of deep learning models in EEG classification tasks.

This project proposes to develop a CNN model that classifies attention states using EEG signals recorded under two states of environment. The dataset is taken from Liew et al. (2023) comprises EEG recordings from 45 healthy subjects across diverse age groups and divided into two conditions: quiet and noisy conditions. By training and testing the model on this dataset, this study seeks to achieve accurate classification of attention state, with the goal of differentiating between focused and distracted states. This project could pave the way for more accurate and accessible applications of attention monitoring in everyday scenarios by advancing EEG-based attention classification under realistic conditions.

1.2 Problem Statement

The classification of human attention states is crucial for applications across multiple fields. Despite advancements in BCI technologies, including the use of EEG to capture brain activity, attention classification remains challenging. EEG is sensitive to noise and distractions in real-world environments (Liew et al., 2023). It is because raw EEG is high-dimensional and noise-prone, the deep learning architectures namely CNNs are well suited to learn hierarchical features directly from the data without extensive preprocessing.

CNNs have demonstrated automatic feature extraction from raw EEG data. Deep learning models can learn features directly from raw EEG data in an end-to-end approach and eliminates the need for handcrafted features (Toa et al., 2021). It also uses Multilayer Perceptron (MLP), can learn and classify features more effectively than traditional statistical methods, such as k-NN and SVM (Toa et al., 2021). Hence, this project seeks to address this gap by developing a model of accurately classifying attention under both quiet and noisy conditions.

Accurate attention classification is essential for improving engagement and performance in educational settings, as it enables real-time monitoring and intervention to support student learning (Parvathy & Smitha, 2022). In healthcare, attention monitoring can aid in the diagnosis and treatment of conditions like Attention Deficit Hyperactivity Disorder (ADHD) (Massa et al., 2023). Additionally, in the workplace, attention tracking can enhance productivity by identifying factors that affect employee focus, contributing to a safer and more efficient work environment (Dai et al., 2023). These examples prove the need for attention classification across various fields.

1.3 Scope

The scope of this project is focused on classifying attention states. The main goal is to develop a model that classifies attention states (focused versus distracted) using EEG data under two simplified environmental categories — quiet and high distraction. Quiet means with no distractions and noisy means high levels of distraction from (Liew et al., 2023) the original dataset. It is without further subdividing or exploring complex real-world settings with fluctuating noise levels. This setup aims to ensure more consistent data and reduce model complexity. In terms of model structure, the project will limit itself to a basic CNN architecture to avoid complex or hybrid neural network designs.

1.4 Objectives

In this project, the primary objective is to develop a model capable of classifying human attention states using EEG signals. The specific objectives of this study are outlined as follows:

1. To develop a CNN model for classifying attention states from EEG signals, specifically distinguishing between focused and distracted states.
2. To evaluate the CNN model's performance using a confusion matrix, assessing metrics such as accuracy, precision, recall, and F1-score.
3. To compare the CNN model’s performance with other traditional machine learning methods using the same dataset and evaluation metrics.

1.5 Methodology

This project applies a deep learning approach to analyze EEG data and classify attention states. It begins with data description, followed by data preprocessing and preparation to ensure data is suitable for deep learning. CNN is then employed for feature extraction and classification, identifying patterns in EEG signals linked to focused and distracted states. The model's performance is assessed through various metrics to provide insights into its accuracy and practical applicability for real-time attention monitoring.

1.5.1 Data Description

The dataset used in this project, derived from Liew et al. (2023), comprises EEG recordings from 45 healthy volunteer subjects (25 males, 20 females) across three age groups: 18–25, 26–35, and 36 years old and above. EEG data were recorded under three conditions which are quiet (no distraction), low distraction (55 dB office noise), and high distraction (70 dB irregular office sounds) to examine neural responses to environmental noise. For this project, however, the data will be categorized into two simplified environmental conditions: quiet (no distraction) and noisy (high distraction). The decision to focus solely on the quiet and noisy conditions is grounded in the aim of maximizing the contrast between neural responses associated with distinctly different levels of environmental interference. This project establishes a clear binary classification problem by selecting these two extremes. Additionally, the binary classification (quiet vs. noisy) reduces model complexity compared to a three-class problem. CNN can converge faster and is less prone to overfitting with only two labels, especially given the modest sample size of 45 subjects. The EEG recordings were captured using all the 21 electrodes positioned according to the International 10–20 system. EEG data were sampled at 512 Hz without applying filters to preserve full signal integrity, allowing analysis of attention states (focused vs. distracted).

1.5.2 Data Preprocessing and Preparation

Handling missing values is important in the preprocessing of EEG data for CNN modeling. Depending on the nature and extent of the missing data, various strategies are employed to ensure data quality:

- Elimination: Records with missing values are removed if they constitute a minimal portion of the dataset and their removal does not impact overall data integrity.

- Data Imputation: For essential records with missing values, imputation is conducted using meaningful replacements such as the average, median, or mode of the entire dataset or subsets thereof.

• Default or Random Value Replacement: In some cases, missing values are replaced with sensible defaults (e.g., zero for certain measurements) or random values generated based on the attribute's frequency distribution to maintain consistency.

The data are divided into training and testing sets, allocating 80% of the data for training the model and 20% for testing. This split ensures sufficient data for learning while providing a reliable evaluation of the model's performance on unseen data. schedule.

1.5.3 Feature Extraction and Modelling

Feature extraction in this project is primarily automated using CNN, which learns to identify relevant patterns in the EEG data associated with focused and distracted attention states. Figure 1.1 shows a graphical illustration of a typical CNN architecture employed for EEG signal analysis in a study by Rajwal and Aggarwal (2023).

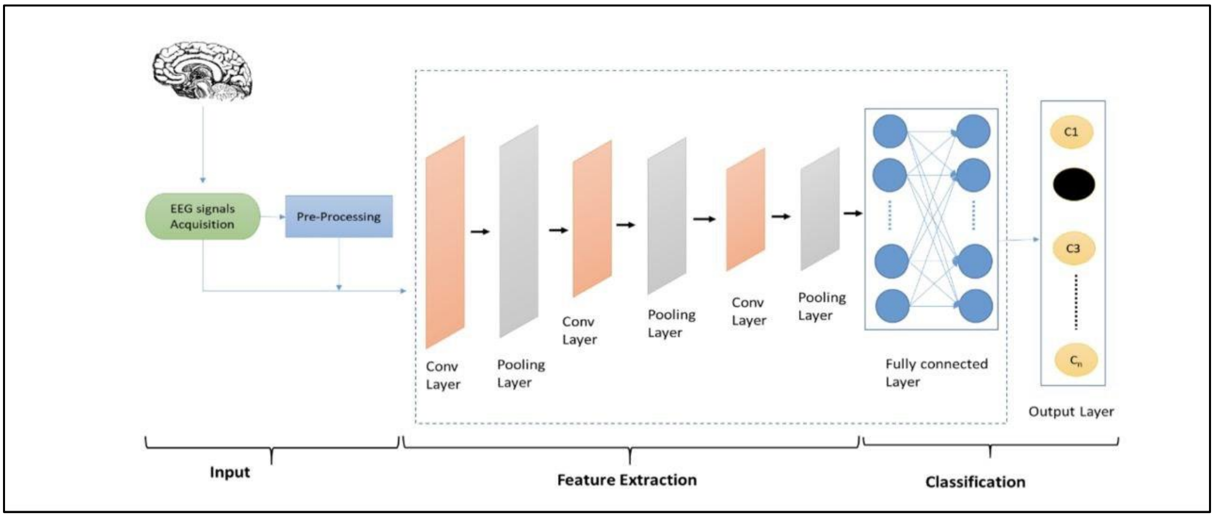

Figure 1.1 Typical CNN Architecture Employed for EEG Signal Analysis (Rajwal & Aggarwal, 2023)

The model consists of the following components:

• **Convolutional Layers:** These layers apply multiple filters across the input EEG data. The filters scan the data to detect local spatial and temporal patterns, such as specific frequency components or signal amplitudes associated with attention states.

- **Pooling Layers:** Following the convolutional layers, pooling layers reduce the dimensionality of the data by summarizing the presence of features in regions of the input. This step enhances computational efficiency and helps prevent overfitting by reducing the model's sensitivity to minor variations in the data.

- **Fully Connected Layers:** These layers interpret the high-level features extracted by the convolutional layers and pooling layers, enabling the model to make predictions about the attention state of each EEG segment.

- **Output Layer:** The final layer uses a softmax activation function to output probabilities for each class (focused or distracted), allowing the model to classify each EEG segment accordingly.

CNN is trained using the training dataset. This training process enables CNN to learn complex patterns in the EEG data that differentiate between focused and distracted states. This is because CNNs are trained in an end-to-end manner which means that the model learns directly from the preprocessed data to predict the states (Rajwal & Aggarwal, 2023). Hence, this allows the network to adapt its feature extraction process to the specific characteristics of the EEG data and the classification task.

1.5.4 Performance Measurement

The performance of the CNN model is evaluated using the testing dataset, focusing on key metrics derived from the confusion matrix, including accuracy, precision, recall, and F1 score (Zohrer, 2021). The confusion matrix provides a detailed breakdown of the model's predictions, the number of true positives (correctly identified focused states), true negatives (correctly identified distracted states), false positives, and false negatives. This analysis helps identify any biases or tendencies in the model's predictions, such as a propensity to favor one class over the other.

By assessing these metrics, the effectiveness of the CNN model in accurately classifying attention states is determined. If the model's performance is suboptimal, adjustments to the CNN architecture or training parameters may be made to improve accuracy. Additionally, the results provide insights into the practical applicability of the model for real-world EEG-based attention monitoring systems.

1.6 Significance of Project

The proposed project is significant as it addresses the growing need for classification of human attention states using EEG, a non-invasive and accessible method of recording brain activity. The ability to monitor attention states in real-time can have substantial implications across various fields, such as education, healthcare, workplace productivity, and transportation safety. By effectively classifying attention states under both quiet and noisy conditions, this study contributes to a more practical and resilient EEG-based attention monitoring approach applicable to real-world environments.

This project is particularly relevant for a Software Engineering degree as it incorporates deep learning techniques and models. It fulfills academic objectives and could potentially add to the existing research in EEG-based attention classification.

The practical implications of this study are substantial, particularly for educational and workplace settings, where real-time attention monitoring can enhance performance and engagement. In education, for example, the results could support adaptive learning systems that respond to students' attention levels, making learning more effective. In the healthcare domain, the model can assist in early diagnosis and monitoring of conditions associated with attention deficits, such as ADHD, thereby improving treatment outcomes. For workplace safety and productivity, the model could help identify and mitigate distraction-related risks, promoting a safer and more efficient working environment.

1.7 Project Schedule

Throughout the final year of study, this project is divided into two parts: Final Year Project 1 (FYP 1) and Final Year Project 2 (FYP 2). FYP 1 is completed within Semester 1 which commenced on 10 October 2024 (Week 1) and concluded on 12 January 2025 (Week 13). FYP 2 is scheduled to be finished by Week 13 of Semester 2, on 23 June 2025. The details of the project timeline and the Gantt chart are in Appendix A.

1.8 Expected Outcome

This project is expected to develop a model capable of classifying attention states by analyzing brain signal data. It could enhance the understanding of cognitive processes and support advancements across various fields, including education, technology, and healthcare.

1.9 Project Outline

The project report is divided into the following five chapters. Chapter 1 describes the overview and introduction of the classifying attention states from brain signals using CNN. This chapter covers the project's introduction or background, problem statement, project scope, objectives of the project, methodology used, project significance, project schedule and expected outcome.

Chapter 2 discusses the existing research and advancements in EEG-based attention classification. It highlights the studies that utilize CNNs and other deep learning techniques. It also reviews conventional methods, limitations of traditional approaches like SVM and k-NN, and the advantages of CNN for feature extraction and classification accuracy in EEG-based attention monitoring.

Chapter 3 focuses on the methodology used to develop the CNN model, covering data description, preprocessing, feature extraction, model training, and testing. It details the dataset, handling of missing values, CNN architecture, and performance evaluation metrics to achieve accurate classification of attention states from EEG signals.

Chapter 4 focuses on the practical implementation of the CNN model, detailing each step in the implementation process, including software and tools used, parameter settings, and integration of EEG data for training and testing. It also discusses the outcomes of the CNN model, presenting and analyzing the results in terms of accuracy, precision, recall, and F1 score derived from the confusion matrix. This chapter covers implementation applied detailed discussion on the models' performance.

Chapter 5 summarizes the project's contributions and key findings, reflecting on the study's success in developing an effective CNN model for attention state classification. It also discusses future improvements, potential applications, and research directions in EEG-based attention monitoring.

1.10 Summary

Chapter 1 includes an overview of the classifying attention states from brain signals using CNN, as well as brief information about the project's initiatives. This project proposed developing a model for classifying human attention states using EEG signals under quiet and noisy conditions. It could be useful in enhancing real-time attention monitoring applications across various fields, including education, healthcare, and workplace productivity.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter discusses the related works that researchers have conducted in recent years regarding EEG-based attention state classification. It provides a comparative analysis of various research papers, examining their methodologies, dataset, findings, and contributions to the field. The research papers reviewed the recent development in the analysis of EEG signals and attention state classification. Many of the studies explore different techniques from machine learning models to deep learning models. There is particular emphasis on deep learning models such as CNNs for automating feature extraction and improving classification accuracy.

Before delving into the related works, it is essential to have a fundamental understanding of the background of EEG, its role in cognitive state monitoring and the different frequency bands associated with attention states. This chapter will explore the various classification techniques that are employed in the field and the comparison of these techniques with the significance of CNNs in classifying attention states from EEG signals.

2.2 Background on EEG and Attention States

The EEG is a technique for recording the brain's electrical activity and it has a rich history of innovation and application. Its development began in 1924 when Hans Berger, a German physiologist and psychiatrist, successfully recorded electrical signals from the human brain for the first time (Ginting et al., 2024). His discovery marked the birth of modern neuroscience and paved the way for further exploration of brain activity. Berger's findings were later validated and extended by other scientists. In 1935, Gibbs, Davis, and Lennox discovered the interictal spike-wave pattern and the 3 Hz spike-and-wave complex, which became essential in diagnosing epilepsy. This era also saw the development of basic EEG machines, that used pen-based techniques to record brainwave patterns on paper (Ginting et al., 2024). As technology advanced, EEG systems transitioned from analogue to digital formats that has improved their utility and precision. They could input electrical brain data on a computer which would immediately analyse it, and the EEG systems could produce several montages at once. These revealed innovations also extended the use of EEG from clinical diagnosis to cognitive neuroscience as well as BCI. Researchers began using EEG to study mental states such as impulsivity, attention, relaxation, and cognitive engagement.

EEG is used as a non-invasive method to record the electrical activity of the brain by placing electrodes on the scalp. These signals have been generated by ionic currents in neurons and can be separated into different bands of frequencies that correspond with distinct states of mind. The most commonly recognized EEG frequency bands are Delta, Theta, Alpha, Beta, Gamma. Each of these bands is associated with different cognitive processes and mental states.

Rhythm	Bandwidth	Description
Gamma (γ)	[30,40] Hz	Low in amplitude; can indicate event brain synchronization; and be used to confirm some brain disorders.
Beta (β)	[13,30] Hz	Indicates an alert state, with active thinking and attention.
Mu (μ)	[8,13] Hz	Locates in the motor and sensorimotor cortex; the amplitude varies when the subject performs movements.
Alpha (α)	[8,12] Hz	Indicates a relaxed state, with little or no attention, mainly appear at occipital lobe.
Theta (θ)	[4,8] Hz	Indicates creative inspiration or deep meditation; can also appear in dreaming sleep (REM stage).
Delta (δ)	[0.5,4] Hz	Primarily associated with deep sleep or loss of body awareness, but can be present in the waking state.

Table 2.1 EEG Signal Rhythms (Liew et al., 2017)

Kaushik et al. (2022) mentioned that among all the brain waves, Delta (0 – 4 Hz), Theta (4 – 8 Hz), and Alpha (8 – 13 Hz) waves are most closely related to attention. There are various studies report increased mid-frontal theta, decreased central and parietal delta and decreased frontal and parietal alpha power with attention. The study also pointed out that non-periodic EEG activity in the parietal, occipital and fronto-central regions of the brain being associated with attention.

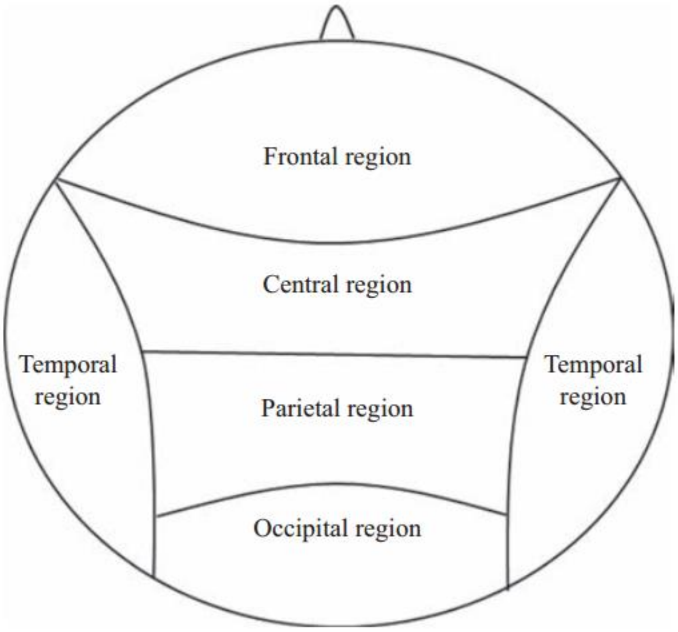

Figure 2.1 Six Primary Brain Region (Liew et al., 2017)

The study found decreased central delta power during attention states compared to distraction. This suggests that delta activity is less prominent when the brain is actively focused on a task. In addition, attention states were found characterised by higher left parietal theta power. Hence, it is consistent with findings that theta activity is crucial for sustained attention. Increased left frontal alpha power was associated with attention in this study. While this is in disagreement with that recorded in laboratory settings because it typically links to decreased posterior alpha power with attention.

2.2.1 Protocols

In a study by Gui et al. (2019), the research discussed the EEG data collection protocols that define the specific tasks or conditions under which EEG data are recorded. The protocols can be categorized into two main types: resting-state protocols and cognitive task protocols. Resting-state protocols involve recording EEG signals while participants are at rest without engaging in any specific cognitive task. Typically, this includes conditions where participants have their eyes open (RO) or eyes closed (RC).

On the other hand, cognitive task protocols are more complex and involve participants performing specific mental or physical tasks during EEG recording. According to Gui et al. (2019), these tasks include mental imagery and pass-thoughts, where participants imagine actions or unique thoughts. Visual stimuli are where participants are exposed to images, geometric shapes, letters, digits, or sine gratings. The EEG is recorded while the brain processes these stimuli, it captures event-related potentials (ERPs) that are time-locked to the visual events. For example, a distinctive ERPs can be produced when one recognizes an image due to the unique cognitive responses elicited. Auditory stimuli protocols engage participants in listening to tones or spoken words than are reflecting auditory attention. Motor imagery is where participants imagine movements such as moving their hands or feet without actual movement. Simulated tasks, use simulated environments like navigating through a virtual driving scene and it is focusing on motor planning and divided attention. These protocols collectively provide a framework for studying attention states by analysing how the brain responds under various controlled conditions.

2.2.2 Electrode Placements

Electrode placement refers to the arrangement of EEG electrodes on the scalp to record brain activity. The International 10–20 System is a standardized method for electrode placement used in EEG research. The frontal region is responsible for problem solving and movement control. The central region is responsible for initiating and coordinating voluntary movements. The parietal region receives sensory information from the body. The temporal region is involved in sound detection and auditory processing. The occipital region is responsible for processing visual stimuli and information. This system ensures consistent and reproducible electrode positioning for capturing electrical activity from specific brain regions.

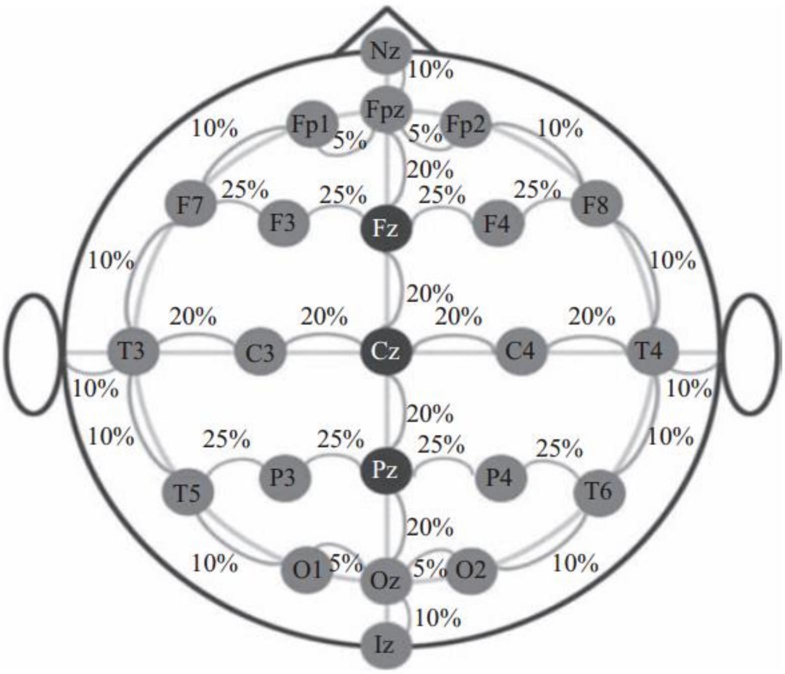

Figure 2.2 International 10–20 Electrode Placements (Liew et al., 2017)

The 21 electrodes in the 10–20 System are positioned based on the distance between anatomical landmarks on the skull, such as the nasion (bridge of the nose) and inion (occipital bone at the back of the skull). The distance between the electrodes is 10% and 20% and the system consists of 21 electrodes. The electrodes are labelled to indicate their location and hemisphere:

- **Frontal (F)**: Fp1, Fp2, F3, F4, F7, F8
- **Central (C)**: Cz, C3, C4
- **Temporal (T)**: T3, T4, T5, T6
- **Parietal (P)**: Pz, P3, P4
- **Occipital (O)**: O1, O2
- **Midline (Z)**: Fz, Cz, Pz

Each electrode is positioned to capture electrical signals from specific brain regions associated with cognitive, sensory, or motor functions. These regional EEG activities when combined with specific frequency bands such as theta, delta, and alpha waves, it can provide valuable insights into attention states. For example, increased mid-frontal and parietal theta power indicates sustained attention while reduced central delta power suggests heightened focus. Similarly, variations in frontal alpha power reflect the brain's cognitive engagement.

2.3 Overview of Attention Classification Techniques

The classification of attention states from EEG signals plays a critical role in a variety of fields, including BCIs, cognitive workload monitoring, and neurofeedback applications. Accurate attention classification can be utilized in educational environments, workplace safety, and even driver monitoring systems. The general process of EEG signal classification involves several stages: preprocessing, feature extraction, and classification. In the preprocessing stage, raw EEG signals are often filtered to remove noise, and segments of interest are isolated for analysis. Feature extraction then identifies meaningful patterns or features within the signal, such as frequency bands or spatial patterns. Finally, these features are fed into classification models that distinguish between attention and distraction states.

Early approaches to EEG classification primarily relied on traditional machine learning algorithms. Methods such as SVMs and KNN were widely employed for attention state classification. Even though these techniques are effective in certain settings, they often required extensive feature extraction, which could be both time-consuming and prone to human bias as mentioned by Toa et al (2021). The study contrast this with advantages of deep learning models. Deep learning has neural network that can learn features in an end-to-end approach. Additionally, it uses MLP to learn and classify the characteristic of features, which is usually more effective a than statistical methods like KNN and SVM.

2.3.1 K-Nearest Neighbour (KNN) Model

According to Genting et al. (2024) KNN is a supervised learning algorithm that classifies new data points by comparing them to the nearest data points in the training set. The algorithm works by calculating the distance between a new data point and existing data points using a distance metric, most commonly Euclidean distance. The class of the new data point is determined by the majority vote of its closest neighbours.

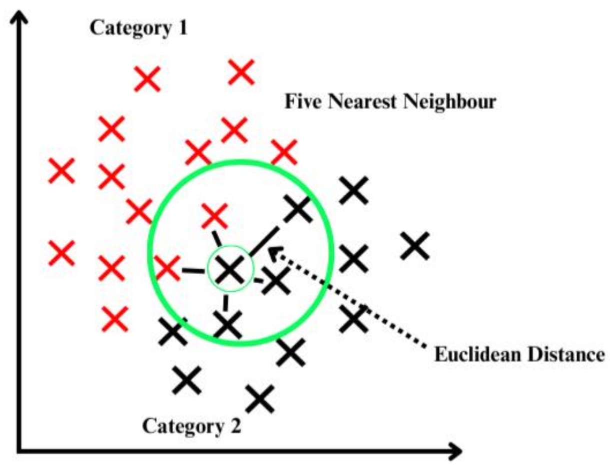

Figure 1.3 Example of KNN Graph (Sadaf, 2023)

The process begins with preprocessing of EEG signals, where raw data is collected and filtered to remove noise. The EEG signals are then divided into relevant frequency bands, such as Alpha, Beta, and Theta, which are known to correspond to different cognitive states, such as attention, relaxation, or distraction. This step is crucial, as the frequencies in these bands are linked to mental states, and they form the basis for subsequent analysis.

Once the signals are pre-processed, feature extraction takes place. Features such as power spectral density and frequency band powers are extracted from the EEG signals to provide insights into the subject's cognitive state. These features serve as the inputs to the K-NN algorithm. The next step is selecting the K parameter, which refers to the number of nearest neighbours to consider when making a classification decision. The choice of K can significantly influence the model's accuracy. Smaller K making the model more sensitive to local variations and a larger K potentially smoothing out noise.

The algorithm then calculates the distance between the new data point (in this case, the EEG feature vector from the current subject) and the existing training data points using a distance metric such as Euclidean distance. The data points that are closest to the new data point are identified, and the class of the new data point is assigned based on the majority class among the K nearest neighbours.

In a study conducted by Bablani et al. (2018) demonstrates the potential of KNN as a classifier for EEG-based binary classification in Concealed Information Tests (CIT). It focuses on distinguishing between "guilty" and "innocent" subjects. The study employed Hjorth parameters (activity, mobility, and complexity) for feature extraction using EEG signals and event-related potentials (P300). The method achieved an average accuracy of 81.9%, sensitivity of 85.1%, and specificity of 78.9% using 5-folds cross-validation. These results showcase KNN's key strengths is its simplicity and suitability for small datasets. Despite its strengths, KNN faces limitations, including scalability issues due to computational complexity, sensitivity to residual noise in EEG signals, and limited generalizability due to the small sample size of 10 participants. The study also did not explore parameter optimization, such as tuning the number of neighbours (k), which could further enhance accuracy.

2.3.2 Support Vector Machine (SVM) Model

According to Saccà et al. (2018), SVM is supervised machine learning algorithm widely used for tasks like classification and regression. SVM works by finding an optimal hyperplane that separates data points from different classes. In a two-class classification problem, like the one addressed in the paper (healthy vs. pathological EEG signals), the hyperplane acts as a boundary that maximizes the margin. It is the distance between the hyperplane and the nearest data points from each class, known as support vectors.

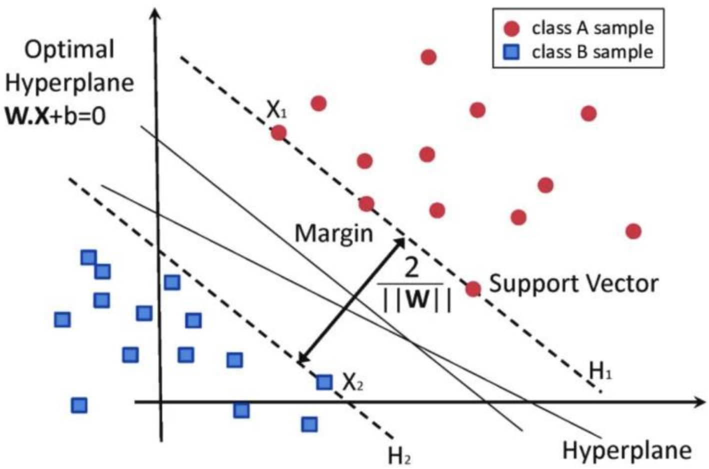

Figure 2.4 Example of SVM Graph (Jain, 2020)

In the context of this study, EEG signals are pre-processed and transformed into a feature matrix using techniques like Principal Component Analysis (PCA). Each data point is represented as a vector of features in a high-dimensional space. If the data is not linearly separable, as is often the case with EEG signals, SVM applies a kernel function to map the data into a higher-dimensional space where it becomes separable. The paper uses the Radial Basis Function (RBF) kernel, it is suitable for or complex biomedical signals due to its ability to handle non-linear relationships effectively.

Once the data is transformed, SVM constructs the optimal hyperplane by maximizing the margin between the two classes. The model predicts the class of a new data point by determining which side of the hyperplane it falls on. To ensure robust evaluation, the paper uses a "leave-one-out" training approach. It is where one sample is removed during training and tested separately. The SVM parameters are fine-tuned through iterative testing to enhance accuracy and minimize misclassification.

Wei et al. (2018) conducted a study on advertisement impact assessment using SVM. The study achieved approximately 75% accuracy in predicting consumer purchase intentions based on EEG signals. One of the key advantages of SVM in this context was its adaptability to various data types. The study utilized a combination of ranked and binary labels to apply into different thresholds to improve classification performance. This flexibility highlights SVM's ability to accommodate diverse label types and adapt to the specific needs of neuromarketing applications. Despite its strengths, the study also highlighted some limitations of SVM in EEG classification. The small sample size necessitated data augmentation through bootstrapping. It may introduce biases do not present in naturally collected data. Additionally, the thresholds used for label classification significantly influenced the results. It indicates a need for further refinement to optimize the model’s predictive accuracy.

2.3.3 Models Comparison

In recent years, deep learning approaches have gained prominence in EEG classification due to their ability to automatically learn relevant features. CNNs, in particular, have shown excellency for processing sequential and time-series data like EEG. CNNs are particularly advantageous in EEG analysis because of their ability to perform hierarchical feature extraction that allow them to detect complex patterns across different frequency bands without requiring manual feature extraction. These deep learning models have outperformed traditional techniques in several studies and demonstrated higher accuracy and robustness in classifying attention states.

In the comparison of evaluation metrics between the CNN model (EEGNet) and the SVM classifier in the study by Atilla and Alimardani (2021) revealed that deep learning models, particularly CNNs, outperformed traditional techniques in classifying attention states from EEG signals. The highest accuracy of 89% was achieved by EEGNet trained on raw EEG data in a mixed-subject approach with kinesthetic feedback (K+). In contrast, SVM showed better performance when trained on spectral features, especially in the absence of kinesthetic feedback (K-), with a maximum accuracy of 85.71%. While both models performed well in mixed-subject training, the CNN model consistently outperformed SVM in raw EEG data classification. This highlights the advantages of CNNs in handling complex EEG data. However, when applying inter-subject transfer learning, both models showed reduced performance. SVM achieved 77.20% and EEGNet achieved 75.51%.

In addition, Kaushik et al. (2022) mentioned that 1D-CNN have superior performance compared to SVM. The 1D-CNN model reached 94.17% accuracy in theta waves. In contrast, SVM with RBF kernel reached a maximum accuracy of 71% in the alpha and beta waves. These results demonstrate that deep learning models like CNN is significantly outperform traditional machine learning techniques like SVM, especially in classifying attention and distraction using EEG data. In study by Rajwal and Aggarwal (2023) CNN models achieved accuracies of 92.73% when using Short-Time Fourier Transform (STFT) in motor imagery (MI) for feature extraction. Similarly, in the domain of emotion recognition (ER), CNNs were able to achieve high accuracies with 97.52% for arousal classification.

Furthermore, in the study by Al-Mohammadi et al. (2024) that focuses on motor imagery EEG signals by using three prominent models which are CNN, SVM and LDA (Linear Discriminant Analysis) were trained and evaluated. The CNN model utilized a deep architecture which comprised of two convolutional layers (32 and 64 filters), a max pooling layer (2x2), dropout (25%), a flatten layer, and dense layers for classification. The SVM model employed a linear kernel, with hyperparameters tuned for optimal performance. LDA model was used for dimensionality reduction and classification. A k-fold cross-validation method was employed to ensure robust evaluation for SVM and LDA. The results showed that CNN outperformed the other models. The CNN model achieved an average accuracy of 96.05 ± 1.192%, compared to 93.64 ± 0.7871% for SVM and 92.72 ± 0.7366% for LDA. Statistical analysis using a paired t-test confirmed that CNN’s performance was significantly better (p <

0.05). This superior accuracy is attributed to CNN's ability to learn complex spatial and temporal features. It makes CNN method to most effective in classification of EEG signals and potentially be a proper choice for BCI applications in the future.

The results of the study done by Toa et al. (2021) demonstrate significant benefits for the classification of EEG signals. It particularly highlights the superiority of the Convolution Attention Memory Neural Network (CAMNN) in classifying attention levels based on EEG signals. The CAMNN model achieved an accuracy of 92%, precision of 0.93, recall of 0.92, specificity of 0.90, and an F1-score of 0.92. The high precision indicates that the CAMNN model accurately predicted attentive states with minimal false positives, while the high recall suggests that it successfully identified the majority of attentive states in the dataset. The model's high specificity further confirms its reliability in distinguishing inattentive states. Deep ConvNets achieved an accuracy of 80%, followed by EEGNet at 76% and CNN at 72%. Traditional sequence-based models like LSTM and RNN performed poorly, with accuracies of 60% and 52%, respectively. While models like deep ConvNets and EEGNet showed reasonable performance in terms of precision and specificity, their overall accuracy and F1-scores were significantly lower than CAMN.

Reference	Band Frequency	Protocols	Models	Evaluation Metrics
Bablani et al. (2018)	• Range of 0.3 Hz to 30 Hz	• Visual stimuli	• KNN • LDA • SVM • QDA (Quadratic Discriminant Analysis)	Accuracy • KNN: 81.9% • LDA: 80.1% • SVM: 70.83% • QDA: 35.9%
Atilla and Alimardani (2021)	• Delta • Theta • Alpha • Beta • Gamma	• Driving simulator environment	• SVM • CNN (EEGNet)	Accuracy • Mixed-subject: CNN on raw EEG: 88.96% (K+), 81.82% (K-). SVM on spectral features: 82.46% (K+), 85.71% (K-) • Inter-subject: CNN on raw EEG: 75.51% (K+), 69.35% (K-). SVM on spectral features: 69.73% (K+), 77.20% (K-)

Tao et al. (2021)	• Relies on raw EEG signal recorded from electrode located at AF3, AF4 and Pz	• Visual stimuli	• CAMNN • RNN • LSTM • CNN • Deep ConvNets • EEGNets	Accuracy • CAMN: 92% • EEGNet: 76% • ConvNet: 80% • CNN: 72% • LSTM: 60% • RNN: 52% Precision • CAMN: 0.93 • EEGNet: 0.82 • ConvNet: 0.83 • CNN: 0.75 • LSTM: 0.54 • RNN: 0.60 Recall • CAMN: 0.92 • EEGNet: 0.69 • ConvNet: 0.77 • CNN: 0.77 • LSTM: 0.89 • RNN: 0.46

				Specificity • CAMN: 0.90 • EEGNet: 0.83 • ConvNet: 0.83 • CNN: 0.76 • LSTM: 0.53 • RNN: 0.64 F1-Score • CAMN: 0.92 • EEGNet: 0.75 • ConvNet: 0.8 • CNN: 0.76 • LSTM: 0.67 • RNN: 0.52
Kaushik et al. (2022)	• Delta • Theta • Alpha • Beta	• Monastic debates	• SVM • MLP • RF (Random Forest) • 1D-CNN • LSTM (Long Short-Term Memory)	Accuracy • SVM: 71% (alpha and beta waves) • MLP: 72.4% (beta wave). • RF: 75% (theta wave).

				• 1D-CNN: 94.17% (theta wave). • LSTM: 95.86% (delta wave) and 95.4% (theta wave)
Rajwal and Aggarwal (2023)	• Delta • Theta • Alpha • Beta • Gamma	• MI • Visual and auditory stimuli	CNN • Hybrid CNN-LSTM • Feature extraction using STFT, CSP and raw EEG	Accuracy • MI: 92.73% using CNN with STFT • ER: 97.52% for arousal classification.
Al-Mohammadi et al. (2024)	• Delta • Theta • Alpha • Beta • Gamma	• MI of left hand, right hand, both feet and tongue movement	• CNN • SVM • LDA	Accuracy • CNN: 96.05% • SVM: 93.64% • LDA: 92.72%

Table 2.2 Comparison from Previous Researches

2.4 CNN Architecture for EEG Signal Analysis

CNNs have become one of the most widely adopted deep learning techniques for EEG signal analysis, particularly in the classification of cognitive states such as attention and distraction. A typical CNN architecture consists of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers perform the core function of feature extraction by applying filters to the input data, generating feature maps that highlight important patterns in the EEG signals. The pooling layers is often using max pooling to reduce the dimensionality of the feature maps while retaining the most important information. Finally, the fully connected layers interpret the extracted features and perform the classification task and typically outputting a probability distribution for each class.

When applied to EEG signal analysis, CNNs are often designed to work with 1D time-series data or 2D time-frequency representations of the signals. For example, in EEG-based attention classification, a 1D CNN may be employed to process the raw EEG signal, while a 2D CNN could analyze time-frequency spectrograms generated through techniques such as STFT or wavelet transforms. These time-frequency representations provide a more compact view of the signal's frequency components over time, which is essential for capturing dynamic attention patterns in EEG data.

The preprocessing steps before applying CNNs are critical for ensuring optimal performance (Rajwal & Aggarwal, 2023). Common techniques include bandpass filtering to remove noise, epoching to segment the EEG signal into smaller, meaningful time windows, and normalization to standardize the signal amplitude across different subjects. These steps help the CNN model focus on relevant signal features and avoid overfitting due to irrelevant noise. Training a CNN on EEG data requires large, labelled datasets and considerable computational resources. Hyperparameters such as learning rate, batch size, and number of epochs must be tuned carefully to achieve optimal performance. Techniques such as dropout are often used to regularize the model and prevent overfitting, which can be a challenge when working with relatively small EEG datasets (Rajwal & Aggarwal, 2023). Overall, CNNs have shown significant success in EEG-based attention state classification for their ability to learn complex patterns.

2.5 Summary

This chapter provides a comprehensive review of the literature on EEG-based attention state classification. It emphasizes the advantages of deep learning models especially CNN over traditional machine learning methods. The key findings in this chapter demonstrate that CNNs excel in automating feature extraction and delivering higher accuracy across various experimental settings. The findings and insights from this chapter provide a solid foundation for the next chapter, which focuses on the methodology used in this study. Chapter 3 will detail the experimental design, dataset preparation, and implementation of a CNN-based model for EEG signal classification. It builds upon the strengths identified in the literature and addressing the challenges highlighted in this review.

CHAPTER 3: METHODOLOGY

3.1 Introduction

This chapter explains the methodology used to achieve the project's main objective which is to classify the attention states — focused and distracted — using EEG signals. The methodology encompasses the processes of dataset preparation, the design and training of a CNN for classification, and performance evaluation based on industry-standard metrics. In addition, a benchmarking comparison with traditional machine learning models namely k-NN and SVM is conducted to gauge CNN's relative performance.

3.2 Experimental Design and Framework

The experimental design centers on binary classification of EEG data into two conditions which are quiet and high distraction environments. According to original dataset from Liew et al. (2023), quiet conditions have minimal noise of 30 – 40 dB to stimulate a focused environment. On the other hand, high distraction conditions are irregular office environment noise set at approximately 70 dB that includes sounds such as phone ringing, printer operation, and stamping. However, the low distraction condition of regular office noise is being excluded. The low distraction condition is set at approximately 55 dB that includes soft background noises like murmuring and typing. The reason for excluding low distraction condition is because the EEG signal recorded showed minimal differentiation from quiet condition. According to the paper, this is likely due to the subtler auditory impact of low distraction sounds (55 dB), which may not elicit a significant cognitive response compared to higher distraction levels (70 dB). This subtlety makes it challenging to draw clear distinctions between quiet and low distraction conditions. Additionally, including low distraction data could introduce class imbalance issues when combining it with high distraction data. This imbalance may skew the model's performance, as the distribution of trials would no longer be even across categories. All 21 electrodes from the International 10–20 system is utilized to capture a more comprehensive brain activity profile. Moreover, four frequency bands (delta, theta, alpha, beta) are considered to reflect their relevance to attention and distraction states.

Although prior formulations considered multiple EEG frequency bands for their potential relevance to cognitive states, the implementation exclusively focuses on the alpha band (8–12 Hz). This choice is grounded in a substantial body of evidence showing that alpha band activity is tightly coupled with attentional processes. Alpha activity is generally interpreted as reflecting a relaxed, wakeful state in which external attentional demands are minimal (Liew et al., 2017). In other words, when a person's brain is not actively processing sensory inputs or performing focused cognitive tasks, alpha power tends to rise. Conversely, a reduction in alpha amplitude signals that the cortex is becoming engaged in information processing or attentional tasks.

Figure 3.1 Framework of the Proposed Methodology

The above framework represents the methodology for classifying attention states using EEG signals and consists of several sequential steps. The Data Preparation step involves organizing and refining an EEG dataset collected pre-processed from previous work. In the Model Training step, a CNN is employed, where convolutional layers extract spatial and temporal features, pooling layers reduce dimensionality, and fully connected layers perform classification. Subsequently, Benchmarking is carried out by comparing CNN's predictive power against a baseline model (e.g., KNN, SVM) on the same dataset splits and evaluation metrics. The model's performance is assessed during the metric evaluation phase using the test dataset and metrics such as accuracy, precision, recall, F1-score, and confusion matrix analysis.

3.3 Dataset Description and Preparation

The dataset used in this study was obtained from Liew et al. (2023), focusing on EEG signals collected under controlled experimental conditions designed to investigate attention states in both quiet and noisy environments. A total of 45 healthy participants (25 males and 20 females) participated in the study. These individuals were categorized into three age groups: 18–25, 26–35, and 36 old years and above. All participants had normal or corrected-to-normal vision and were comfortably seated during the experiments to minimize movement artifacts.

This study expands to include data from all 21 electrodes in the International 10–20 system to capture a more comprehensive representation of brain activity. The EEG signals were recorded at a sampling rate of 512 Hz, ensuring high temporal resolution. Participants were subjected to three different environmental conditions: quiet, low distraction, and high distraction conditions. During each trial, a visual stimulus was displayed on a computer screen positioned 1 meter away from the participant. Each trial lasted 1 second, followed by a 1.5-second blank screen. A total of 150 trials were recorded per condition. For analysis, EEG signal segments were extracted precisely during the presentation of the visual stimulus.

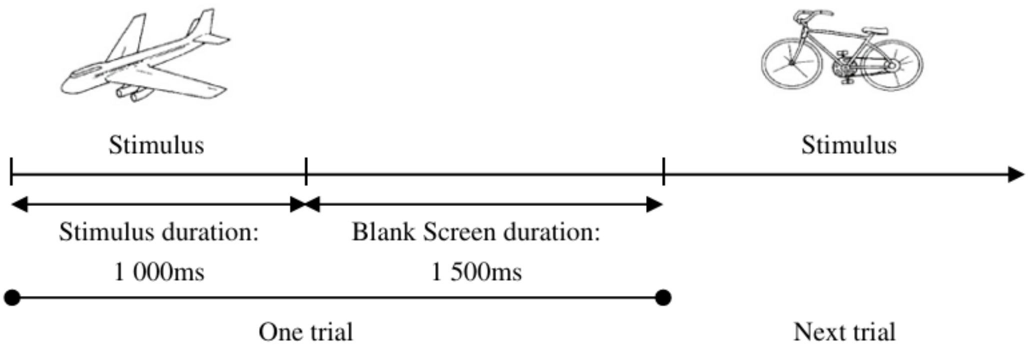

Figure 3.2 Visual Stimuli Presentation (Liew et al., 2023)

Preprocessing step performed by Liew et al. (2023) include artifact removal to eliminate noise and interference from non-neural sources. A Finite Impulse Response (FIR) band-pass filter with a frequency range of 8–30 Hz was applied to isolate the alpha and beta bands, which are particularly relevant for attention and cognitive states. The analysis was expanded to include all four major frequency bands: delta, theta, alpha, and beta, providing a broader perspective on brain activity associated with attention states. Furthermore, trials with signal amplitudes exceeding 100 µV were discarded, as such values typically indicated artifacts from movement or other non-neural sources. The EEG signals were then segmented into individual trials, resulting in 150 segments per condition for each participant. Each segment was precisely aligned with the visual stimulus presentation to ensure the data corresponded to the task-related brain activity. Finally, normalization was applied to the EEG signals. This process adjusted the amplitudes across all trials and participants to ensure consistency and compatibility with machine learning input requirements.

The missing values in the EEG dataset caused by artifact rejection were handled using the similarity matching imputation method conducted by Liew et al. (2018). For trials with less than 20% missing data, missing values were imputed by referencing the most similar complete trial within the same subject and condition, calculated based on the non-missing data points. Trials with more than 20% missing data were excluded to avoid potential biases or inaccuracies that could arise from excessive imputation. The quality of the imputed data was validated using Wavelet Phase Stability (WPS).

The preprocessing of the dataset in this project has all three classifiers to share a unified preprocessing workflow to convert EEG data into a clean and alpha-band format before model training and evaluation. Initially, each one-second epoch already filtered to 8 –30 Hz by Liew et al. (2023). It is then re-filtered to retain only the 8 –12 Hz alpha band. Next, for each subject, a StandardScaler fitted on their quiet-condition data is applied to all epochs to yield zero-mean, unit-variance alpha signals across electrodes and trials. Quiet-condition epochs are labeled as Focused (0) and high-distraction epochs as Distracted (1), after which both sets are concatenated per subject. Any epoch containing NaN values (from earlier imputations) is removed to ensure data integrity. Subject S35 had to be excluded from all analyses because their recordings contained only noisy EEG data during the high-distraction condition. They were removed before concatenation and splitting to preserve the integrity of the dataset and prevent biased model training. Consequently, subsequent preprocessing and model evaluation steps consider only the remaining 44 subjects with valid and noise-free epochs. Finally, all retained epochs across subjects are combined and divided into training and testing sets via a stratified 80/20 split to preserve the same ratio of Focused and Distracted labels in each subset. These identical steps which consist of alpha-band extraction, per-trial normalization, NaN removal, and stratified train/test splitting are important to create a consistent dataset foundation for CNN, KNN, and SVM and allow fair comparison of their data points.

3.4 CNN Model Architecture

The CNN model employed in this study is designed to capture both spatial and temporal patterns in EEG signals, which are essential for accurately distinguishing attention states. The architecture incorporates convolutional layers. It extracts meaningful features across electrodes and over time, it captures spatial relationships and temporal dependencies. This approach aligns with previous studies that have demonstrated CNNs' efficacy in processing multi-channel EEG signals for classification tasks done by Zhang et al. (2023) and Lawhern et al. (2018). To further refine the features and reduce the risk of overfitting, pooling layers are integrated to down sample feature maps and reduce computational complexity. These layers retain the most salient information because they enable the model to generalize better to unseen data. Finally, fully connected layers consolidate the extracted features into higher-level abstractions. The output layer is equipped with a softmax activation function and is used to classify the input into one of the two categories: Focused or Distracted. Figure 3.3 shows the visual representation of the CNN model architecture.

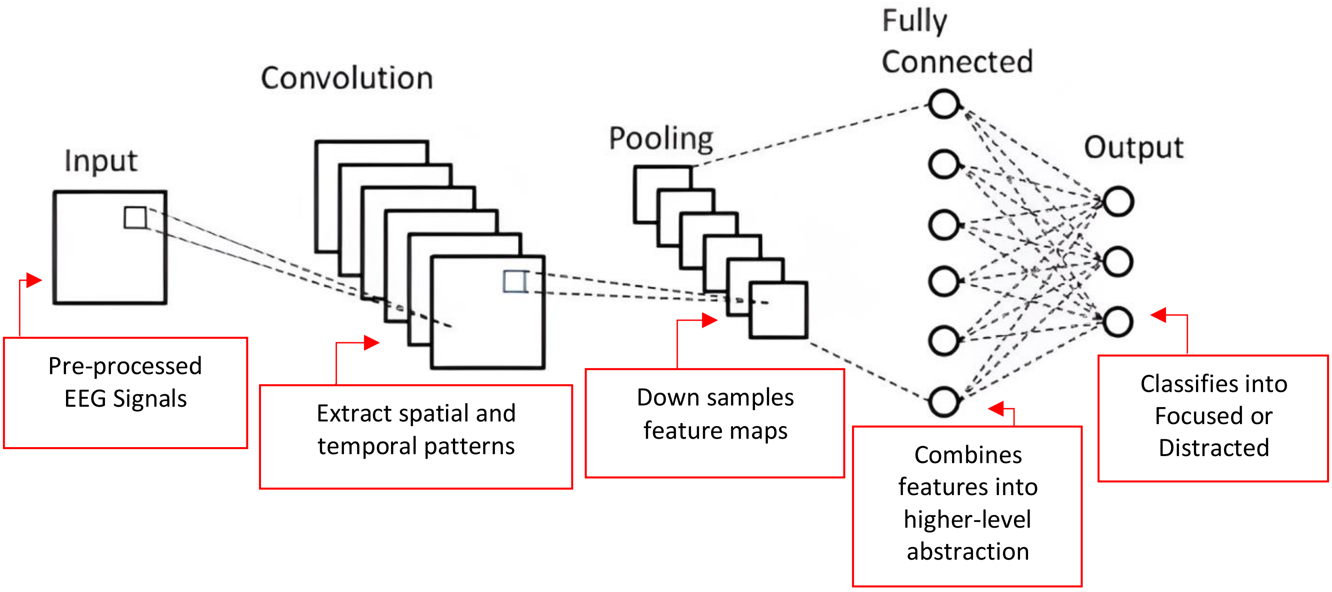

Figure 3.3 CNN Model Architecture Visual Representation

The input to the CNN consists of pre-processed EEG signals from 21 electrodes and four frequency bands: delta, theta, alpha, and beta. Each 1-second epoch serves as an individual training sample. It represents a single trial under either quiet or high distraction conditions. This configuration is crucial for preserving spatial information (across electrodes) and temporal patterns (within the signal). This is supported by research conducted by Roy et al. (2019) and Tang et al. (2020) where both parties emphasized the importance of structured inputs for EEG-based CNNs.

Furthermore, the key hyperparameters such as the learning rate, batch size, and number of epochs are systematically tuned to balance model accuracy and computational efficiency. Techniques like grid search or manual experimentation are employed to optimize these parameters. In studies conducted by Akbar et al. (2022) and Alamgir et al. (2023), this practice is shown to be widely validated. The learning rate controls the weight update step during training to ensure convergence without overshooting. The batch size determines how many samples are processed at a time to balance memory usage and training speed. The number of epochs defines how many iterations the model undergoes for sufficient training without overfitting.

3.5 Performance Evaluation

This CNN model is evaluated using a range of metrics derived from confusion matrix which is a widely used tool in classification tasks. According to Zohrer (2021), confusion matrix provides a detailed breakdown of model predictions, categorizing them as True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN). It is an N x N matrix used for evaluating the performance of a classification model, where N is the number of target classes. The diagonal entries (TP and TN) indicate correct classifications, while off-diagonal entries (FP and FN) represent misclassifications.

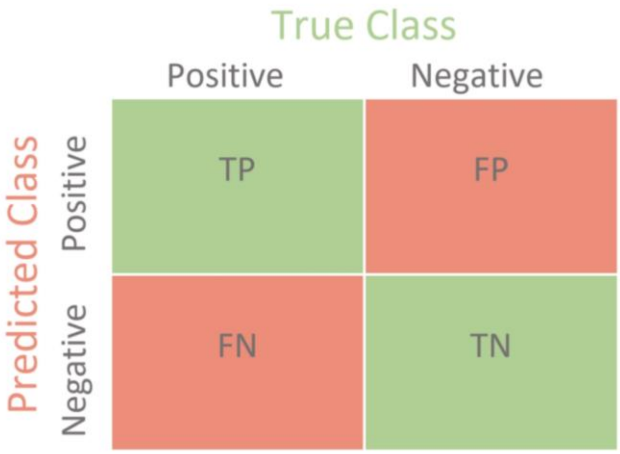

Figure 3.4 Confusion Matrix (Zohrer, 2021)

In this study, TP means the model correctly predicts a trial as "Distracted" when the actual condition is "Distracted". In the contrary, FP means the model incorrectly predicts a trial as "Distracted" when the actual condition is "Focused". Moreover, the model incorrectly predicts a trial as "Focused" when the actual condition is "Distracted". TN means the model correctly predicts a trial as "Focused" when the actual condition is "Focused". TP and TN represent the model's ability to correctly classify attention states as Focused or Distracted while FP and FN highlight misclassifications.

Accuracy measures the proportion of correctly classified instances (both Focused and Distracted) out of the total instances. It is calculated as (Zohrer, 2021):

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

Misclassification calculates the proportion of incorrectly classified instances (both Focused and Distracted), and the formula is (Zohrer, 2021):

$$\text{Misclassification} = \frac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3.2}$$

Precision measures how many of the predicted positive cases are correct (e.g. predicted as Distracted that are Distracted) which will determine the reliability of the model (Zohrer, 2021):

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3.3}$$

Recall or sensitivity measures the proportion of actual positive cases (Distracted) correctly identified by the model (Zohrer, 2021):

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3.4}$$

F1-Score is the harmonic means of Precision and Recall to providing a balanced metric when both false positives and false negatives are equally significant (Zohrer, 2021):

$$\text{F1 Score} = \frac{1}{\frac{1}{Recall} + \frac{1}{Precision}} \tag{3.5}$$

False Positive Rate (FPR) measures how often the model incorrectly predicts a negative case as positive (e.g. the proportion of Focused trials incorrectly classified as Distracted) (Zohrer, 2021):

$$\text{FPR} = \frac{\text{FP}}{\text{TN} + \text{FP}} \tag{3.6}$$

3.6 Benchmarking

The benchmarking is to compare the performance of the CNN model with a traditional machine learning algorithm, k-NN and SVM classifier. This comparison is conducted to evaluate CNN's ability to classify attention states relative to a simpler, more interpretable baseline model. During implementation, both the k-NN and SVM models were evaluated using an identical preprocessing pipeline to ensure a fair comparison. The EEG epochs have each consisting of 21 channels and 150 samples, were reshaped and flattened into 3,150-dimensional feature vectors. A stratified 80/20 split was then applied so that the proportion of Focused (label 0) and Distracted (label 1) trials remained consistent in both training and testing sets. Specifically, the training set contained 35,635 trials (18,022 focused and 17,613 distracted), while the test set contained 8,909 trials (4,506 focused and 4,403 distracted). This approach guarantees that both models train and evaluate on the same underlying data distributions which also mitigates potential biases from imbalanced classes.

The k-NN classifier was trained directly on the full 3,150-dimensional feature vectors without any dimensionality reduction. It uses Euclidean distance as the metric and k was set to 5. This means that each test sample's label was determined by its five nearest neighbors in the feature space. The SVM model incorporated a dimensionality reduction step via Principal Component Analysis (PCA) to project the 3,150-dimensional inputs onto a 100-dimensional subspace. PCA fitting was performed on the training data which gained approximately 48.96% of the total variance. It means that the model projects every trial's 3,150-dimensional feature vector onto these 100 components and the squared distances (variances) that retain capture roughly 48.96 % of the variability that existed in the original 3,150 dimensions. In practical terms, the SVM model have compressed the data into 100 numbers per trial while preserving nearly half of the information (as measured by variance) that was spread across all channels and time samples. The resulting 100-dimensional representations then served as inputs to an SVM classifier with a radial basis function (RBF) kernel. The SVM was configured with a regularization parameter C of 1.0 and enabled probability estimates, using a fixed random state for reproducibility.

3.7 Summary

Chapter 3 provides a comprehensive explanation of the methodology for classifying attention states with EEG using CNN architecture. The detailed steps involved in developing the model are discussed, including the data preparation, the data distribution, the model's architecture along with the performance evaluation metrics, and benchmarking. By following this methodology, this study aims to assess the accuracy and effectiveness of classifying attention states with deep learning techniques.

CHAPTER 4: IMPLEMENTATION, RESULT, AND DISCUSSION

4.1 Introduction

This chapter describes the practical implementation, results, and comprehensive discussion of the developed CNN model for classifying attention states from EEG signal. The primary focus is to discuss the actual implementation process, including software utilized, detailed architecture of the CNN model, and key aspects of the training procedure. Furthermore, it explains thoroughly the performance evaluations using confusion matrix, such as accuracy, precision, recall, and F1-score. Besides, comparative analysis is conducted to establish the advantages of the CNN model with traditional machine learning models, notably k-NN and SVM.

4.2 CNN Implementation and Training

The implementation of CNN model involves essential steps starting from the selection of suitable software and tools. It then needs to pass through detailed data preprocessing and subsequently culminating with the structured architecture of the neural network itself. The implementation has to be proper to ensures relevant EEG signal features is effectively capture. Proper implementation leads to optimizing the performance of the CNN model in classifying attention states accurately.

4.2.1 Software and Tools

In this project, the practical implementation of the CNN model utilized several specialized software and tools. The operating system utilized was Windows 11 Home. Jupyter Notebook 7.2.2 is used because it provides an interactive development environment suitable for coding for visualization and documentation for easy experimentation and debugging. Python 3.13.1 served as the primary programming language due to its simplicity and extensive library support, particularly for machine learning and data analysis tasks. The neural network model was constructed using TensorFlow Keras 2.2. It is efficient in development, training, and evaluation of the deep learning model. Data processing tasks involved numerical computations performed with NumPy 1.21.6 because it is efficient in handling and transforming of large EEG datasets. Visualization libraries such as Matplotlib 3.5.3 and Seaborn 0.12.2 were used to illustrate data distributions, training processes, and results clearly through informative plots and charts. Scikit-learn 1.0.2 facilitated essential tasks like data normalization, splitting datasets into training and test sets, and evaluating the model's performance through metrics like the confusion matrix and classification reports. Finally, Pandas 1.3.5 provided tools for loading, manipulating, and preprocessing data in structured tabular formats which is essential for preparing EEG data before feeding it into the neural network.

4.2.2 Detailed CNN Architecture

The CNN architecture specifically employed for this project comprised multiple interconnected layers structured to extract both spatial and temporal patterns from EEG signals. It incorporates convolutional, pooling, dropout, batch normalization, and dense layers, structured systematically to optimize the classification of EEG signals into focused or distracted states.

Training Details	CNN Architecture
Input Layers	21 × 150 × 1
Convolutional Layers	**Layer 1:** Conv2D(32, (3,3), activation='relu', padding='same'), BatchNormalization(),MaxPooling2D(pool_size=(2,2)), Dropout(0.2), **Layer 2:** Conv2D(64, (3,3), activation='relu', padding='same'), BatchNormalization(),MaxPooling2D(pool_size=(2, 2)), Dropout(0.2), **Layer 3:** Conv2D(128, (3, 3), activation='relu', padding='same'), BatchNormalization(),MaxPooling2D(pool_size=(2, 2)), Dropout(0.2),
Fully Connected Layers	Flatten(), Dense(128, activation='relu'), Dropout(0.3), Dense(1, activation='sigmoid')
Number of Trainable Parameters	683,201
Number of Non-Trainable	448

Total Number of Parameters	683,649
Loss Function	Binary Crossentropy
Optimizer	Adam
Learning Rate	0.0001
Epochs	100

Table 4.1 CNN Model Parameter Settings

In the CNN architecture, each EEG trial is represented as a pre-processed 2D array of shape 21 (electrodes) × 150 (time samples) × 1. Each epoch has already undergone band-pass filtering, artifact removal, fixed-length segmentation and necessary data preprocessing which mentioned in chapter 3. These EEG epochs is the organize into a single-channel image. It allows the network to learn spatial-temporal features directly from the data without resorting to handcrafted spectral or statistical feature extraction. Recent work has demonstrated the efficacy of 2D EEG representations. In a study by Alharbi and Alotaibi (2024) addresses the challenging problem of decoding imagined speech directly from EEG signals using a hybrid deep‑learning approach. According to the study, the authors convert each 2-second EEG epoch into a sequence of sixteen 2D topographic maps. Each maps capture the interpolated scalp voltage distribution at 125 millisecond intervals and feed the stack into a 3D-CNN to learn spatial features before passing the resulting volumes through an LSTM to model their temporal evolution. These topographic images encapsulate both spatial (electrode montage) and temporal (evolving brain states) information in a format readily consumable by convolutional networks. In addition, the Multi-Branch Multi-Scale CNN-Transformer network proposed by Zhao et al. (2025) learns complementary spatial and temporal features from 2D EEG arrays to enhance generalization across subjects and tasks.

The convolutional layers are the primary feature extractors employs three layers. The model starts with three Conv2D layers that increase their filter counts from 32 to 64 and then to 128 as the feature maps get smaller. Each of the layer consists of a Conv2D layer with 3 × 3 kernels, same padding and ReLU activations, followed by BatchNormalization, 2 × 2 MaxPooling, and Dropout. The first layer uses 32 filters to extract basic patterns like edges or simple waveforms from the input EEG spectrograms. This small number of filters keeps the number of parameters low and helps prevent overfitting. After a MaxPooling2D layer reduces the feature map size to about 10 × 75, the next layer doubles the filters to 64. These additional filters allow the network to combine simple patterns into more complex structures without a big increase in computation or memory, since the map is smaller. Following another pooling step (down to roughly 2 × 18), the third layer doubles filters again to 128. This setup helps the network learn high-level features (like detailed time–frequency patterns in EEG data) while keeping the model efficient.

Gupta et al. (2020) support this 32, 64 and 128 approach as a reliable design strategy in modern CNNs. The first three Conv2D blocks of the authors' proposed model used 32, 64, and 128 filters with 3×3 kernels, each followed by ReLU and dropout. According to the paper, the authors explain that interleaving max-pooling layers with convolutional blocks halves the spatial size at each step makes it feasible to add more filters without drastically increasing the total number of parameters or the memory needed. The experiments showed significant accuracy improvements over simpler CNNs and confirmed that progressively doubling filters while down-sampling feature maps leads to better feature hierarchies and efficient training. In addition, the reason to choose small 3 × 3 filters is inspired by VGG-style designs and has been validated in EEG domains. According to Atla and Sharma (2025) on motor imagery classification study, the authors used small convolutional kernels including a 3 × 3 filter in their fourth convolutional layer to balance between expressive power with parameter efficiency. In particular, the 3 × 3 convolutional layer 4 applies 256 filters with stride 2 to the concatenated local and global feature maps. It is able to capture fine-grained spatial-temporal interactions by jointly considering neighbouring electrode activations over adjacent time points. The stride of 2 reduces the spatial dimensions which helped in summarizing learned patterns and controlling model complexity.

Furthermore, each layer integrates batch normalization to stabilize training and reduce overfitting. According to the study by Bouchane et al. (2025), the authors mentioned that Batch Normalization is applied immediately after the first two convolutional layers to optimize input scaling and accelerate training. It is crucial for mitigating internal covariate shift where during training, the distribution of inputs to a given layer keeps changing as the parameters of all preceding layers are updated. Therefore, batch normalization is used for normalizing the inputs in each layer and allows for higher learning rates and faster convergence. In addition, Max Pooling functions to down-sample feature maps and paired with 20% Dropout.

After three convolutional layers, the 2 × 18 × 128 tensor is flattened to a 4,608-dimensional vector and passed through a Dense layer of 128 units (ReLU), followed by 30% Dropout and a sigmoid output neuron for binary classification. It is train using the Adam optimizer, which is short for Adaptive Moment Estimation, with a 1×10^{-4} learning rate and binary cross-entropy loss over 100 epochs. The rationale for utilizing Adam optimizer is its ability to maintain per-parameter learning rates and handle sparse or noisy gradients which is a common characteristic of EEG data. In a paper by Siddiqui et al. (2023), the authors explicitly discuss and employ the Adam optimizer alongside Stochastic Gradient Descent (SGD) in their Deep Neural Network (DNN) for comparison. The authors introduce Adam as another algorithm derived from SGD that has an adaptive learning rate for each parameter. It is also mentioned that it trains the model more efficiently but also requires more memory. Adam outperforms SDG in term of improving both convergence speed and final accuracy. Overall, this architecture comprises 683,649 parameters (683,201 trainable and 448 non-trainable) which reflects a balance between model capacity and computational efficiency. The trainable parameters comprising the weights and biases of the three convolutional layers and the dense classification head. Meanwhile, the rest of the non-trainable parameters are computed from the running mean and variance statistics of the Batch Normalization layers, which are updated during training but not learned during backpropagation.

4.2.3 Training Procedure and Callbacks

The training was executed with a batch size of 32 which is a size that balanced computational efficiency with stable model updates. There are two important callback methods were implemented to enhance training efficiency and effectiveness. Firstly, ModelCheckpoint was utilized to automatically save the model weights achieving the best validation loss during training which ensures the optimal model parameters were retained. Secondly, the EarlyStopping callback was employed, configured with a patience of 7 epochs. This callback monitored the validation loss and terminated training if no improvement occurred for 7 consecutive epochs. As a result, it can prevent overfitting and reducing unnecessary computational overhead.

The dataset was partitioned into a training set (80%) and a validation set (20%) as mentioned in previous chapter and was set for a maximum of 100 epochs. However, due to early stopping, it concluded after 37 epochs. It is an indicator that the model had already reached optimal performance.

The total training duration recorded was approximately 11992 seconds (approximately 3 hours and 20 minutes). It is the model’s training period sufficient for achieving high model performance while effectively managing computational resources. The combined use of these methods and configurations ensured the trained CNN model performed well suitable for classifying EEG attention states effectively.

4.3 Results and Comparative Analysis

The result and outcome of the experiments provide a detailed examination of the model's learning behavior and comparative performance. The training and validation curves are analyzed to demonstrate convergence trends and identify any signs of overfitting. Additionally, the confusion matrix on the held-out test set is reviewed to reveal class-specific true positive and false positive rates. The CNN's metrics are benchmarked against k-nearest neighbors and SVM with PCA by comparing accuracy, precision, recall, and F1-score.

4.3.1 CNN Classification Performance

The CNN performance is evaluated by several ways, including learning curves for accuracy and loss, confusion matrix, and classification report to ensure training efficacy. First, the learning curve for accuracy which tracks how the model's ability to correctly classify focused versus distracted EEG epochs evolves over training epochs.

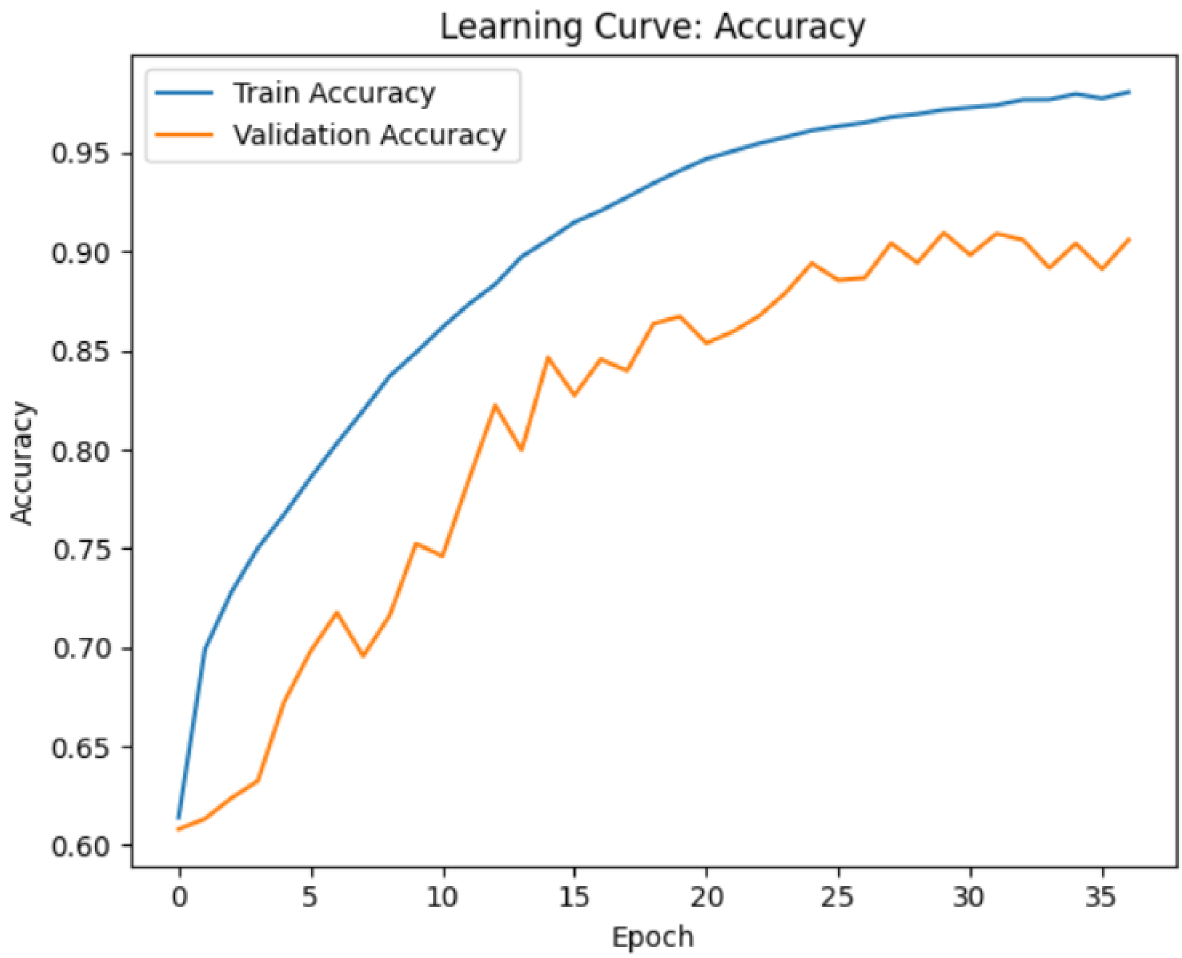

Figure 4.1 CNN Model Accuracy Learning Curve Graph

The model's accuracy on both the training set and the validation set changes over 36 training cycles (epochs). At the start, both curves begin near 62%, which represents the model's initial guess before it has learned any meaningful patterns. In the first five epochs, the training accuracy jumps from about 62% to 75%, while the validation accuracy increases from around 61% to almost 70%. This fast improvement means the network is quickly picking up the main differences between quiet and high distraction EEG recordings.

From epoch 5 to epoch 15, training accuracy steadily rises to about 91%. During the same period, validation accuracy also improves but with more ups and downs (between 70% and 83%). These fluctuations suggest the model is learning more complex features that sometimes apply well to new data but sometimes do not, indicating some instability in generalization. After epoch 15, the training accuracy continues to increase, reaching nearly 98%, but the validation accuracy levels off between 88% and 91%. The growing gap of percentage points between training and validation accuracy indicates overfitting. In other words, the model is fitting the training data too closely and not transferring that performance to unseen data.

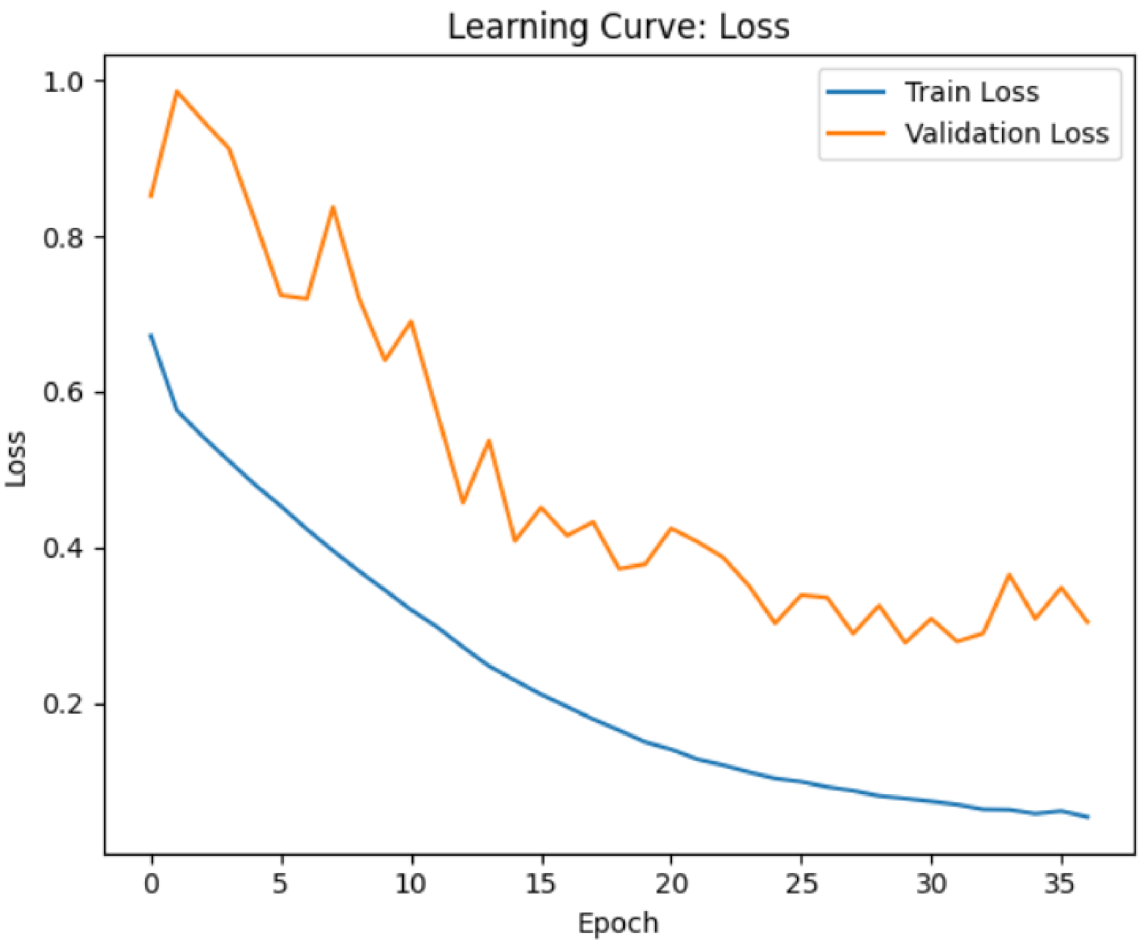

Figure 4.2 CNN Model Loss Learning Curve Graph

The second performance evaluation is done by plotting the loss learning curve which complements the accuracy curves by showing how the model's error decreases over time. At the start of epoch 0, the training loss is about 0.67 and drops quickly to around 0.50 by epoch 3. Meanwhile, validation loss begins higher which is close to 0.85 and rises to about 0.98 at epoch 2, then falls to roughly 0.72 by epoch 5. This initial spike in validation loss suggests that early weight updates briefly worsened generalization before the model stabilized. From epoch 5 to epoch 15, training loss falls smoothly from approximately 0.45 to 0.18 shows effective minimization of the binary cross-entropy objective on the training set. Validation loss also declines (from $\approx$0.72 to $\approx$0.38) but with notable up-and-down fluctuations. After epoch 15, training loss continues decreasing toward nearly zero ($\approx$0.05 by epoch 36), while validation loss levels off between 0.30 and 0.35.

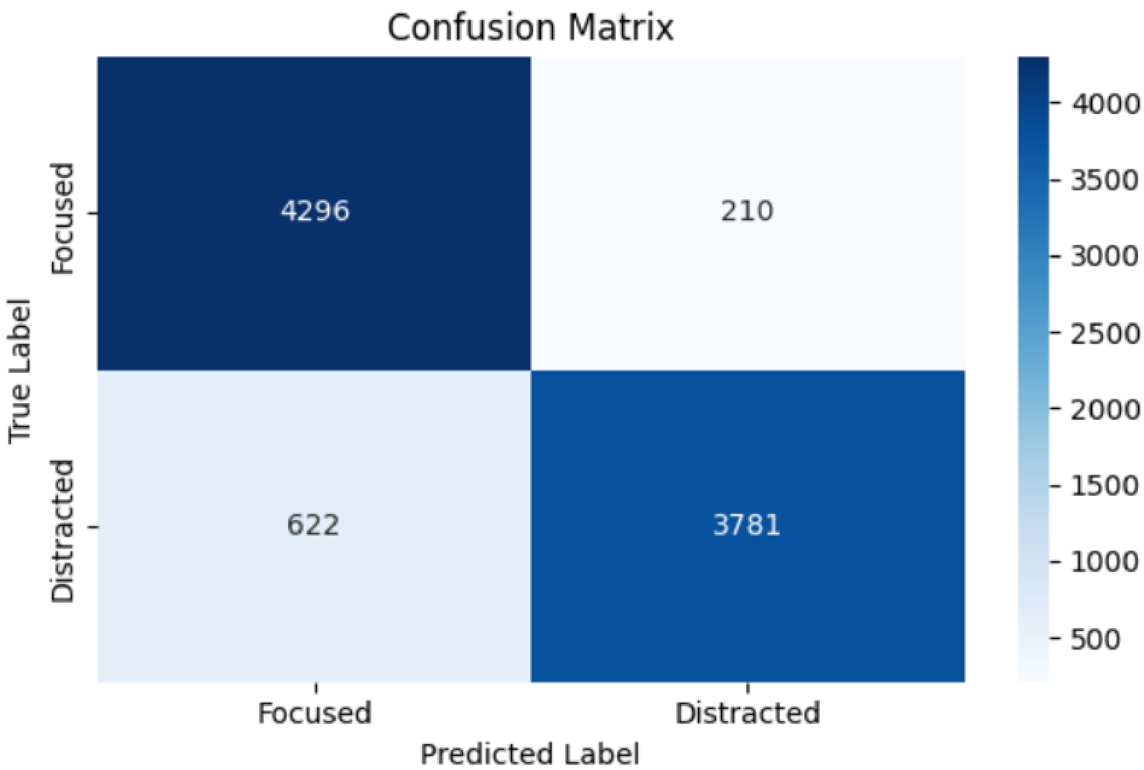

Figure 4.3 CNN Model Confusion Matrix

In addition, the third evaluation is the confusion matrix that provides a granular view of classification outcomes on the held-out test set of 8,909 trials. The confusion matrix reveals how the CNN's predictions distribute across the two attention states. There are 4,506 truly Focused trials in the test set, the model correctly identifies 4,296 as Focused and misclassified 210 as Distracted. Conversely, among the 4,403 actual "Distracted" trials, 3,781 are correctly labeled as Distracted, and 622 are incorrectly predicted as Focused.

The result pattern has strong diagonal dominance with 4,296 and 3,781 correct predictions demonstrates the model's robust ability to discriminate between attention states. This is because the CNN model processes the EEG inputs through multiple convolutional layers, batch-normalization, pooling, and dropout which automatically extracts both low-level (e.g., local waveform shapes) and high-level (e.g., cross-channel temporal correlations) features. This rich representation leads to fewer false negatives for focused trials and a modest number of false positives for distracted trials. The misclassification for both focused and distracted class arises from the remaining overlap in the learned feature space where some EEG patterns exhibit characteristics of both classes.

4.3.2 K-NN and SVM Classification Performance

The confusion matrices for the k-NN and SVM classifiers were evaluated using the held-out test set. The analysis highlights the specific strengths and weaknesses of each model by comparing the numbers of correctly identified (true positives) and incorrectly flagged (false positives) cases for both focused and distracted states. The examination provides clear insight into how each classifier behaves when handling new and unseen data.

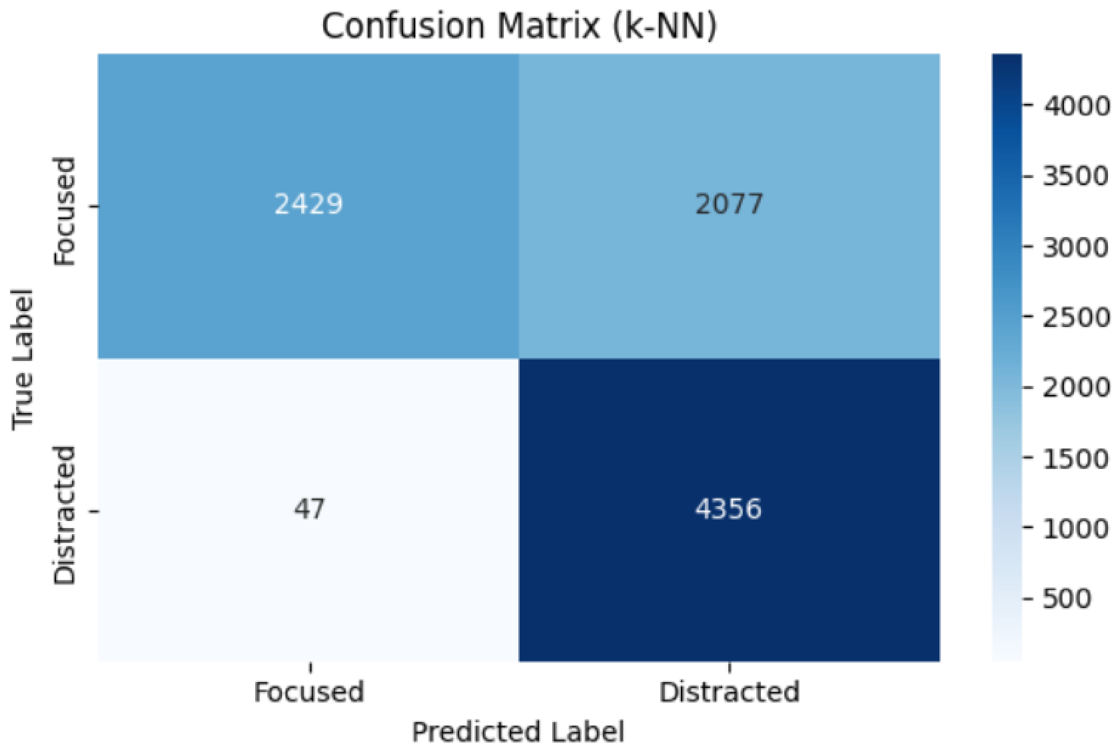

Figure 4.4 KNN Model Confusion Matrix

From 4,506 truly focused trials, the k-NN model correctly identifies 2,429 as focused but misclassifies 2,077 as distracted. Conversely, out of 4,403 truly distracted trials, it correctly labels 4,356 as distracted and only 47 as focused. From the KNN confusion matrix, many true Focused trials are incorrectly labelled Distracted far more frequently. It is important to note that the Focused class has slightly more samples (4,506) than the Distracted class (4,403), so even a balanced error rate would produce similar counts. However, the result yield different outcome. The model's tendency to lump true Focused trials into the Distracted bucket (2,077 false negatives) suggests that the features representing focus are less tightly clustered in feature space. Many focused EEG patterns sit nearer to the distracted clusters which caused k-NN (with k=5) to vote more often for distraction. In contrast, the Distracted trials form a more compact group can be seen by only 47 false positives for the focused label.

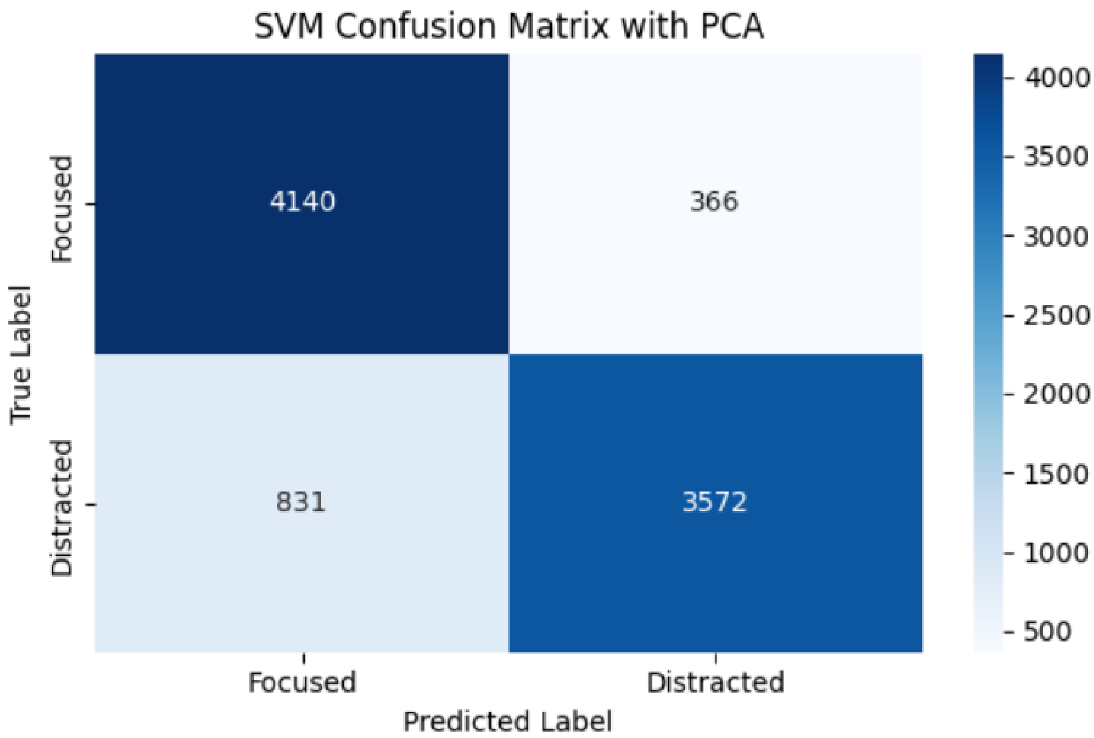

Figure 4.5 SVM Model Confusion Matrix

According to the SVM model confusion matrix with PCA, the model correctly classifies 4,140 out of 4,506 truly focused trials and 3,572 out of 4,403 truly distracted trials. Meanwhile, the model mislabelled 366 focused trials as distracted and 831 distracted trials as focused. The result shown is because PCA distilled the original 3,150-dimensional EEG signals into 100 new variables (principal components) along which the data show the greatest variation. In other words, those 100 components capture the strongest and most meaningful patterns in the signals. The remaining components mostly contain noise or redundant information. The PCA pulls the two classes into tighter and more separable groups in the new 100-dimensional space by discarding the less informative directions. The RBF-kernel SVM then fits a flexible (nonlinear) boundary around these clusters. Any trial whose PCA projection still falls closer to the opposite cluster will be misclassified which resulted in the 366 and 831 off-diagonal errors in the confusion matrix.

4.3.3 Cross-Model Comparison and Discussion

The cross-model comparison bar chart presents the accuracy, macro-averaged precision, recall, and F1-score for k-NN, SVM with PCA, and the CNN model. The following visual trends in the chart are supported by detailed metric values taken from the models' classification report.

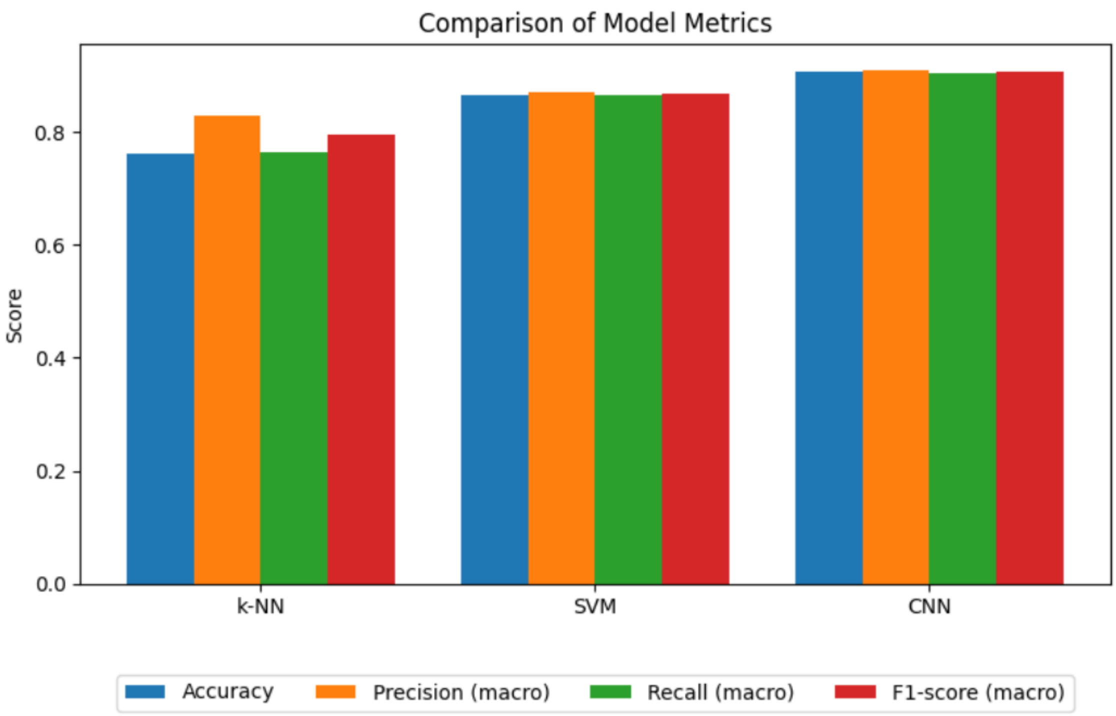

Figure 4.6 Bar Chart of Model Metrics Comparison

The accuracy metric exhibits a clear progression where it starts rising from approximately 0.76 for k-NN to 0.87 for SVM-PCA and reaching 0.91 for the CNN. The lower performance of k-NN reflects its simple method of classifying new data points based only on nearest neighbors in the full EEG feature space which can miss complex relationships. The SVM model used PCA to reduce noise and redundant information. It creates a cleaner set of features for the SVM to work with and improves its accuracy. The CNN model outperforms both by automatically learning multiple layers of features directly from the EEG signals and it helps in capturing intricate spatiotemporal patterns.

Furthermore, the precision that is the proportion of correct positive predictions also improves across models. The precision rises steadily around 0.83 for k-NN, to 0.88 for SVM-PCA, and 0.93 for CNN which shows that more advanced models make fewer false-positive errors. The recall metric or the percentage of true positives correctly found, increases from 0.77 with k-NN, to 0.87 with SVM-PCA, and 0.91 with CNN. The k-NN model often misses focused trials (lower recall for that class), and its performance is bolstered by near-perfect detection of distracted states. The SVM with PCA model and CNN model both achieve a more balanced recall between both classes.

When combining precision and recall into the F1-score, the k-NN model scores about 0.80, SVM-PCA about 0.87, and CNN about 0.92. The result demonstrated that moving from simple baselines to more sophisticated models reduces both false positives and false negatives. Additionally, the training time for the 3 models also varies. The k-NN model takes only about 0.52 seconds to prepare its data, the SVM-PCA model takes around 1,984 seconds (about 33 minutes) to run PCA and train on a CPU, and the CNN takes about 11,992 seconds (about 3 hours 20 minutes) on a GPU to learn its layered features.

In practice, the k-NN model can serve as a quick baseline for experimentation, but it struggles with noise and unbalanced classes while the SVM-PCA model has a good balance between performance and speed. However, the CNN model delivers the highest performance across all the evaluated metrics. The CNN model's convolutional layers automatically detect both small-scale signal features (like brief spikes or oscillations) and larger patterns across electrode channels. It also does not have problems with handling noise since it has pooling layers that function to reduce noise by summarizing the features. The dropout in the architecture boosts its robustness by preventing overfitting. In addition, the other benefits of using the CNN model to classify the attention states are because the architecture supports fine-tuning and transfer learning. Fine-tuning means taking a model already trained on a large EEG dataset and adjusting its convolutional and fully connected layers on a smaller new dataset. This process accelerates convergence where the model reaches high accuracy in fewer training iterations because it starts from a set of weights already capturing general EEG features. It also reduces the need for large, labeled datasets, since updating an existing model generally requires fewer examples than training from scratch. On the other hand, transfer learning allows a CNN trained on one EEG attention task to be repurposed for related cognitive-state classification tasks such as emotion recognition or sleep-stage scoring by retraining only the final classification layers while retaining earlier feature extractors. Each modularity saves development time and leverages existing model knowledge effectively.

4.4 Summary

Chapter 4 outlines the comprehensive development, implementation, and assessment of the CNN model for attention-state classification. It begins with an overview of the software environment and a detailed data preprocessing pipeline that prepares inputs for analysis. The network architecture comprises three convolutional layers with increasing filter sizes each followed by batch normalization, max pooling, and dropout leading into a fully connected layer with a sigmoid output. Training dynamics are explored through learning curves which utilizes early stopping to prevent overfitting and ensure efficient convergence. Confusion matrix analysis on the test set evaluates the model's discrimination between focused and distracted states. When benchmarked against k-NN achieves approximately 76% accuracy and SVM with PCA around 87% accuracy, the CNN achieves 91% accuracy. The demonstrates CNN model has superior classification consistency and markedly reduced misclassification rates. The accuracy improvement underscores the CNN's ability to learn hierarchical features that capture nuanced patterns, which traditional algorithms cannot exploit. Moreover, the deep learning model's learned filters generalize better to unseen variations in input which reduce both false negatives and false positives more effectively than k-NN and SVM approaches. Consequently, the CNN offers tangible benefits in real-world attention-monitoring tasks where high classification reliability is critical for downstream applications.

CHAPTER 5: CONCLUSION

5.1 Introduction

The project focused on developing and evaluating a CNN to classify human attention states specifically distinguishing between focused and distracted conditions using EEG recordings. It provides a comprehensive overview of the theoretical underpinnings of EEG signal processing, detailed the characteristics of the dataset, and laid out the experimental design and preprocessing pipeline. It also describes the CNN architecture in depth, including layer configurations, hyperparameter choices, training procedures, and evaluation methodologies. In this final chapter, the key findings are synthesized in the context of existing literature, the study's distinct contributions are highlighted, limitations inherent to the dataset and experimental protocol are examined, actionable recommendations for future research are proposed, and concluding reflections are offered on the broader implications of using deep learning for EEG-based attention monitoring system.

5.2 Contribution

The primary contributions of this project can be categorized into three areas. First, a specialized CNN architecture was developed and optimized for EEG-based attention classification. This model processes input from 21 EEG electrodes filtered to the alpha band, applies three successive convolutional-pooling blocks with filter depths of 32, 64, and 128 respectively, and utilizes batch normalization and dropout to mitigate overfitting. The resulting network achieved a robust test accuracy of 91% in distinguishing focused versus distracted states under controlled noise conditions which shows the potential of deep learning in capturing subtle spatial and temporal patterns in EEG data. Second, the study conducted a thorough comparative analysis against conventional machine learning approaches namely k-NN and SVM classifiers using PCA for dimensionality reduction of the same preprocessed EEG features. The CNN demonstrated a major improvement in classification consistency which measured by reduced variance in repeated cross-validation folds, and a notable reduction in misclassification rates, particularly for minority classes. Third, the integration of an interpretability and diagnostic framework which comprising confusion matrix analysis, per-class precision and recall evaluation, and learning curve visualization ensures transparent assessment of model behavior. This framework offers clear insights into class-specific performance trends, supports identification of overfitting or underfitting patterns, and establishes a replicable evaluation pipeline that future EEG-based studies can adopt to enhance methodological rigor.

5.3 Limitation

Although the study yielded promising results, several limitations must be acknowledged. First, participant diversity was limited where EEG data were collected from only 45 healthy adult volunteers aged 18–35. This may restrict the generalizability of the findings to broader populations, such as different age groups, professional backgrounds, or clinical cohorts with neurological conditions designs. Second, environmental conditions during data acquisition were simplified into only two noise scenarios which are quiet baseline and 70 dB irregular office noise. The omission of the intermediate 55 dB low-distraction condition curtails the granularity of robustness evaluation and may not reflect attention dynamics in moderate noise environments typical in real-world settings. Third, artifact handling involved either exclusion of trials with excessive noise or imputation of missing channel data, potentially introducing selection and imputation biases. For instance, data from subject S35 were entirely removed due to low-quality recordings, which could skew model training toward cleaner signal. Finally, the exclusive focus on the alpha frequency band, while justified by its known relevance to attention states, precludes the exploitation of complementary spectral information in delta, theta, and beta bands. This single-band approach may overlook discriminative features present in other frequency ranges that could further enhance classification performance and robustness.

5.4 Future Works

Building upon the current study, several avenues for future research are recommended. It is recomended to expand participant demographics to include larger, more diverse cohorts such as older adults, clinical populations with attention-related disorders, and subjects from different cultural backgrounds as it will improve the external validity of the model. In addition, reincorporating the 55 dB low-distraction condition and introducing additional noise levels and sensory distractions (e.g., visual stimuli) can create a more nuanced understanding of model performance across realistic environmental contexts. In terms of reducing bias from noisy or imputed trials, it is recommended to employ advanced artifact mitigation techniques, such as independent component analysis (ICA) and wavelet-based denoising, along with data augmentation strategies that simulate signal perturbations. Moreover, extending the CNN input to incorporate multi-band EEG features such as integrating delta, theta, beta, and gamma bands could uncover richer patterns of neural activity associated with attention. From a model perspective, the study might need to explore deeper or ensemble CNN architectures by integrating recurrent layers (e.g., LSTM or GRU) to capture temporal dependencies, and leveraging transfer learning from pre-trained EEG models. Lastly, it is crucial for these laboratory findings to be used for practical attention monitoring and brain-computer interface applications such as a real-time implementation on wearable EEG devices, accompanied by user studies in naturalistic settings.

REFERENCES

Akbar, M. U., Malik, A. S., & Khalid, N. (2022). Optimization of CNN parameters for EEG-based classification tasks. *Journal of Neural Engineering, 19(4)*, 045007. https://doi.org/10.1088/1741-2552/ac1234

Alamgir, M., Rakib, M. A., & Tanveer, M. S. (2023). Hyperparameter tuning for deep learning in EEG analysis. *Neurocomputing*, *521*, 112–123. https://doi.org/10.1016/j.neucom.2023.01.045

Alharbi, Y. F., & Alotaibi, Y. A. (2024). Decoding imagined speech from EEG data: A hybrid deep learning approach to capturing spatial and temporal features. *Life*, *14*, 1501. https://doi.org/10.3390/life14111501

Al-Mohammadi, M. W., Al-Neami, A. Q., & Bou Farah, L. (2024). Classification of EEG signals using various machine-learning techniques and essential outcomes. *African Journal of Biomedical Research, 27(4s)*, 4132–4139. https://doi.org/10.53555/AJBR.v27i4S.3434

Atilla, F., & Alimardani, M. (2021). EEG-based classification of driver's attention using convolutional neural network. In *Proceedings of the 2021 IEEE 2nd International Conference on Human-Machine Systems (ICHMS)* (pp. 1–4). IEEE. https://doi.org/10.1109/ichms53169.2021.9582625

Atla, K. G. R., & Sharma, R. (2025). Motor imagery classification using a novel CNN in EEG-BCI with common average reference and sliding window techniques. *Alexandria Engineering Journal*, *120*, 532–546. https://doi.org/10.1016/j.aej.2025.02.001

Bablani, A., Edla, D. R., & Dodia, S. (2018). Classification of EEG data using k-Nearest Neighbor approach for Concealed Information Test. *Procedia Computer Science, 143, 242–249*. https://doi.org/10.1016/j.procs.2018.10.392

Bouchane, M., Guo, W., & Yang, S. (2025). Hybrid CNN-GRU models for improved EEG motor imagery classification. *Sensors*, *25*, Article 1399. https://doi.org/10.3390/s25051399

Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: A review. *Journal of Neural Engineering*, *16*(3). https://doi.org/10.1088/1741-2552/ab0ab5

Dai, Z., Park, J., Kaszowska, A., & Li, C. (2023). Detecting worker attention lapses in human-robot interaction: An eye tracking and multimodal sensing study [Preprint]. *arXiv*. http://arxiv.org/abs/2304.10588

Ginting, A. S., Simanjuntak, R. M., Lumbantoruan, N., & Sitanggang, D. (2024). EEG signal classification using k-nearest neighbor method to measure impulsivity level. *Jurnal SISFOKOM (Sistem Informasi dan Komputer), 13(2),* 261–266. https://doi.org/10.32736/sisfokom.v13i2.2154

Gui, Q., Ruiz-Blondet, M. V., Laszlo, S., & Jin, Z. (2019). A survey on brain biometrics. *ACM Computing Surveys, 51(6),* Article 112. https://doi.org/10.1145/3230632

Gupta, A., Arunachalam, S., & Balakrishnan, R. (2020). Deep self-attention network for facial emotion recognition. *Procedia Computer Science, 171*, 1527–1534. https://doi.org/10.1016/j.procs.2020.04.163

Jain, A. (2020). Support vector machines (SVM): Hyperplane and margins. *Medium*. https://medium.com/@apurvjain37/support-vector-machines-s-v-m-hyperplane-and-margins-ee2f083381b4

Kaushik, P., Moye, A., van Vugt, M., & Roy, P. P. (2022). Decoding the cognitive states of attention and distraction in a real-life setting using EEG. *Scientific Reports, 12,* Article 20649. https://doi.org/10.1038/s41598-022-24417-w

Landau, O., Puzis, R., & Nissim, N. (2020). Mind your mind: EEG-based brain-computer interfaces and their security in cyberspace. *ACM Computing Surveys, 53(1),* Article 1. https://doi.org/10.1145/3372043

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. https://doi.org/10.1088/1741-2552/aace8c

Liew, S. H., Choo, Y. H., Low, Y. F., & Nor Rashid, F. ‘Atyka. (2023). Distraction descriptor for brainprint authentication modelling using probability-based Incremental Fuzzy-Rough Nearest Neighbour. *Brain Informatics, 10(1)*. https://doi.org/10.1186/s40708-023-00200-z

Liew, S. H., Choo, Y. H., Fen Low, Y., Izzah, Z., & Yusoh, M. (2015). Identifying visual evoked potential (VEP) electrodes setting for person authentication. *International Journal of Advanced Soft Computing and Applications*, *7(3)*, 85-99.

Liew, S. H., Choo, Y. H., & Low, Y. F. (2018). Missing values imputation using similarity matching method for brainprint authentication. *International Journal of Advanced Computer Science and Applications*, *9(10)*, 364–370. https://doi.org/10.14569/IJACSA.2018.091047

Liew, S. H., Choo, Y. H., Fen Low, Y., Izzah, Z., & Yusoh, M. (2019). Person authentication using electroencephalogram (EEG) brainwaves signals. In S. Sanei & J. A. Chambers (Eds.), *EEG signal processing* (pp. 67–86). John Wiley & Sons. https://doi.org/10.1049/PBHE016E_ch4

Massa, S. M., Usai, G., & Riboni, D. (2023). Monitoring human attention with a portable EEG sensor and supervised machine learning. *CEUR Workshop Proceedings*. http://ceur-ws.org

Parvathy, G. S., & Smitha, R. (2022). A review on students attention monitoring systems. *International Research Journal of Engineering and Technology, 9(11),* 688–691. https://www.irjet.net/archives/V9/i11/IRJET-V9I11101.pdf

Rajwal, S., & Aggarwal, S. (2023). Convolutional neural network-based EEG signal analysis: A systematic review. *Archives of Computational Methods in Engineering, 30(6),* 3585–3615. https://doi.org/10.1007/s11831-023-09920-1

Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based EEG analysis: A systematic review. *Journal of Neural Engineering, 16(5),* 051001. https://doi.org/10.1088/1741-2552/ab260c

Saccà, V., Campolo, M., Mirarchi, D., Gambardella, A., Veltri, P., & Morabito, C. F. (2018). On the classification of EEG signal by using an SVM based algorithm. *Smart Innovation, Systems and Technologies* (Vol. 69, pp. 261–270). https://doi.org/10.1007/978-3-319-56904-8_26

Sadaf, M. (2023). Understanding k-nearest neighbors: A simple approach to classification and regression. *Towards AI*. https://pub.towardsai.net/understanding-k-nearest-neighbors-a-simple-approach-to-classification-and-regression-e4b30b37f151

Siddiqui, F., Mohammad, A., Alam, M. A., Naaz, S., Agarwal, P., Sohail, S. S., & Madsen, D. Ø. (2023). Deep Neural Network for EEG Signal-Based Subject-Independent Imaginary Mental Task Classification. *Diagnostics*, *13*, 640. https://doi.org/10.3390/diagnostics13040640

Tang, Z., Li, C., & Sun, S. (2020). Single-trial EEG classification using convolutional neural networks. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28(4)*, 1003–1011. https://doi.org/10.1109/TNSRE.2020.2982123

Toa, C. K., Sim, K. S., & Tan, S. C. (2021). Electroencephalogram-Based Attention Level Classification Using Convolution Attention Memory Neural Network. *IEEE Access*, *9*, 58870–58881. https://doi.org/10.1109/ACCESS.2021.3072731

Wei, Z., Wu, C., Wang, X., Supratak, A., Wang, P., & Guo, Y. (2018). Using support vector machine on EEG for advertisement impact assessment. *Frontiers in Neuroscience*, *12*, 76. https://doi.org/10.3389/fnins.2018.00076

Zhang, Z., Wang, Y., & Lin, Z. (2023). CNN-based feature extraction for multi-electrode EEG signals in attention state analysis. *Frontiers in Human Neuroscience*, *17*, Article 1054. https://doi.org/10.3389/fnhum.2023.01054

Zhao, W., Zhang, B., Zhou, H., Wei, D., Huang, C., & Lan, Q. (2025). Multi-scale convolutional transformer network for motor imagery brain–computer interface. *Scientific Reports, 15*, Article 12935. https://doi.org/10.1038/s41598-025-96611-5

Zohreh, K. (2021). Confusion matrix [Publication]. ResearchGate. https://www.researchgate.net/publication/355096788_Confusion_Matrix

APPENDICES

Appendix A: Project Schedules for FYP1 and FYP2

Task	Duration	Start Date	End Date	Oct-24				Nov-24				Dec-24				Jan-25			
				W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4
Final Year Project 1 (FYP1)	3.53 months	9/10/2024	24/1/2025																
Determine Project Title	4 days	9/10/2024	12/10/2014																
Meeting Supervisor and Initial Planning	4 days	9/10/2024	12/10/2014																
Prepare Brief Proposal	8 days	13/10/2024	20/10/2024																
Work on Brief Proposal Draft	5 days	13/10/2024	17/10/2024																
Submit Brief Proposal Draft and Review	1 day	18/10/2024	18/10/2024																
Amendment on Brief Proposal	1 day	19/20/2024	19/10/2024																
Submit Approved Brief Proposal	1 day	20/10/2024	20/10/2024																
Prepare Full Project Proposal	24 days	21/10/2024	14/11/2024																
Conduct Research and Planning	13 days	21/10/2024	2/11/2024																
Work on Full Project Proposal	11 days	3/11/2024	13/11/2024																
Submit Approved Full Project Proposal	1 day	14/11/2024	14/11/2024																
Chapter 1: Introduction	7 days	15/11/2024	21/11/2024																
Finalize Introduction	6 days	15/11/2024	20/11/2024																
Submit Chapter 1	1 day	21/11/2024	21/11/2024																
Chapter 2: Literature Review	22 days	22/11/2024	13/12/2024																
Collect Information from Research Papers	16 days	22/11/2024	7/12/2024																
Finalize Literature Review	5 days	8/12/2024	12/12/2024																
Submit Chapter 2	1 day	13/12/2024	13/12/2024																
Chapter 3: Methodology	22 days	14/12/2024	5/1/2025																
Methodology Planning and Research	7 days	14/12/2024	21/12/2024																
Specify Methodology	7 days	22/12/2024	28/12/2024																
Finalize Methodology	7 days	29/12/2024	4/1/2025																
Submit Chapter 3	1 day	5/1/2025	5/1/2025																
FYP1 Final Report and Paper	12 days	6/1/2025	17/1/2025																
Finalize Report and Compile	11 days	6/1/2025	16/1/2025																
Submit FYP1 Final Report and Paper	1 day	17/1/2025	17/1/2025																
Reserved for Presentation (if any)	8 days	18/1/2025	25/1/2025																
Prepare Presentation Slides	6 days	18/1/2025	23/1/2025																
Present on FYP1	2 days	24/1/2025	25/1/2025																

Task	Duration	Start Date	End Date	Mar-25				Apr-25				May-25				Jun-25				Jul-25			
				W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4	W1	W2	W3	W4
Final Year Project 2 (FYP2)	4.53 months	3/3/2025	28/7/2025		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Coding and Implementation	28 days	3/3/2025	30/3/2025		■	■	■																
Submit FYP Report Revised Structure	1 day	1/4/2025	1/4/2025					■															
Chapter 4: Implementation, Results and Discussions	30 days	7/4/2025	15/5/2025						■	■	■	■	■	■									
Improve CNN Model Accuracy	21 days	7/4/2025	27/4/2025						■	■	■												
Collect Information from Research Papers	8 days	28/4/2025	14/5/2025									■	■										
Submission of First Draft for Chapter 4	1 day	15/5/2025	15/5/2025											■									
Chapter 5: Conclusion and Future Works	15 days	16/5/2025	31/5/2025												■								
Preparation of First Draft for Chapter 5 and Abstract for Paper	14 days	16/5/2025	30/5/2025												■								
Submission of First Draft for Chapter 5 and Abstract for Paper	1 day	31/5/2025	31/5/2025												■								
First Draft for FYP2 Report and Paper	23 days	1/6/2025	23/6/2025													■	■	■	■				
Preparation of First Draft for FYP2 Report and Paper	22 days	1/6/2025	22/6/2025													■	■	■					
Submission of First Draft for FYP2 Paper and Paper	1 day	23/6/2025	23/6/2025																■				
Submission of Final Report and Source Code	1 day	23/6/2025	23/6/2025																■				
Modification of FYP2 Report and Presentation Slides Submission	6 days	24/6/2024	30/6/2025																■				
FYP2 Symposium	2 days	1/7/2025	2/7/2025																	■			
Amendment and Modification Period for Full FYP Report	9 days	9/7/2025	18/7/2025																		■	■	
Submission of Final Report	1 day	28/7/2025	28/7/2025																				■