Heaven’s Light is Our Guide

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Rajshahi University of Engineering & Technology, Bangladesh

Exploring EEG Features for Emotion Recognition: An In-Depth Evaluation with Machine Learning and Deep Learning Approaches

Author
Tushar Das
Roll No. 1803108
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology

Supervised by
S. M. Mahedy Hasan
Assistant Professor
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology

ACKNOWLEDGEMENT

I would like to begin by expressing my deepest gratitude to the Almighty for bestowing His grace and benefits, which have endowed me with the perseverance and zeal necessary to complete my thesis successfully.

I extend my heartfelt thanks to my esteemed supervisor, **S. M. Mahedy Hasan**, Assistant Professor, Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi. His profound knowledge, expert guidance, and continuous encouragement have shaped this work. His mentorship has enhanced my technical skills and instilled in me a sense of confidence and determination.

I am also grateful to respected **Prof. Dr. Md. Ali Hossain**, Head of the Department of Computer Science & Engineering and all the respective teachers of Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi for their valuable suggestions and inspirations from time to time.

My heartfelt appreciation also goes to my parents, family, friends, and well-wishers, whose unwavering support and encouragement have been my pillars of strength. Their belief in my abilities and constant encouragement has kept me motivated and determined to achieve my goals.

Lastly, I express my gratitude to all those who have directly or indirectly contributed to this work. Their support, guidance, and encouragement have been invaluable, and I am truly grateful for their presence in my life.

October 9, 2024 Tushar Das

RUET, Rajshahi

Heaven’s Light is Our Guide

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Rajshahi University of Engineering & Technology, Bangladesh

CERTIFICATE

This is to certify that this thesis report entitled ***“Exploring EEG Features for Emotion Recognition: An In-Depth Evaluation with Machine Learning and Deep Learning Approaches”*** *submitted by* ***Tushar Das, Roll:1803108*** *in partial fulfillment of the requirement for the award of the degree of Bachelor of Science in Department of Computer Science & Engineering of Rajshahi University of Engineering & Technology, Bangladesh is a record of the candidate own work carried out by him under my supervision. This thesis has not been submitted for the award of any other degree.*

Supervisor

S. M. Mahedy Hasan
Assistant Professor
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204

External Examiner

Md. Azmain Yakin Srizon
Assistant Professor
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204

ABSTRACT

Emotion, the intricate interplay of physiological and cognitive processes, is a cornerstone of human experience, influencing behavior, decision-making, and social interaction. Accurately identifying emotional states greatly improves relationships and experiences, facilitating effective communication and decision-making, all of which highlight the critical role that understanding emotions plays in several areas. EEG signals have surfaced as a potentially effective method for discerning emotional states, providing valuable observations into the neural processes that underlie emotional reactions. Integrating deep learning and machine learning techniques with EEG has enabled the development of sophisticated models to discern emotions from brain signals. This research aims to evaluate the efficacy of different feature extraction techniques and classification models in precisely classifying human emotions through EEG data. The study incorporates various methods of feature extraction, including Statistical features, Power Spectral Density, and Differential Entropy with a variety of classification models, including Convolutional Neural Network (CNN), Support Vector Machine (SVM), and Decision Tree (DT), K-Nearest Neighbors (KNN), and Random Forest (RF) architectures. By conducting thorough assessments on SEED and DEAP datasets, this research illustrates that DE features exhibit superior performance compared to alternative extraction techniques. It is worth mentioning that DE features improve the accuracy of RF by 84% and SVM by 83.33% on the SEED dataset. Additionally, a CNN model that utilizes DE features achieves an accuracy of 89%. A hybrid CNN-BiLSTM model further improves accuracy, achieving an outstanding 93% accuracy on the SEED dataset. The hybrid model exhibits promising capabilities in enhancing precision and generalization, thus constituting the primary contribution of this study. This research enhances the development of emotion recognition technology and promotes a more profound comprehension of the interaction between humans and computers.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Introduction

Emotion, being an intricate psychological state, is expressed through a range of physical behaviors and physiological activities. In recent years, there has been a significant focus on recognizing emotions using affective information from various sources, including voices, physiological signals, videos, and text [4]. This endeavor primarily utilizes statistical machine-learning techniques to accurately detect and analyze users' emotional states in various scenarios, whether in real-time or offline. There has been a growing focus on the importance of physiological cues in emotion recognition due to recent advancements in affective computing. Although individuals may intentionally conceal their true feelings, computational psychophysiology provides a promising alternative to facial expressions and speech-based methods. Through analyzing brain activities, specifically using Electroencephalography (EEG), researchers strive to understand the inner workings of emotional cognition and accurately determine emotional states [5]. Considering the significant influence of the brain on emotions, utilizing EEG-based approaches is a valuable addition to conventional methods. Many HCI systems in use today struggle to understand and respond to human emotions, which limits their effectiveness. Although various methods exist for emotion recognition, some may be influenced by subjectivity and susceptible to manipulation. On the other hand, utilizing physiological signals like EEG for emotion identification provides a higher level of objectivity and reliability.

1.2 Overview

Brain activity monitoring is significantly facilitated by EEG, which was developed by Hans Berger in 1929 [6]. EEG provides insights into the functional and physiological alterations of the brain through the recording of electrical potentials produced by neuronal activity. The EEG signal provides Exhaustive psychophysiological information, reflecting neurons' collective electrical activity in particular brain regions. Significantly, EEG investigations have aided in identifying brain regions linked to emotional processing and the comprehension of the neural correlates of emotions. Ideological and interpersonal communication are significantly influenced by emotions. Improving the intelligence of Human-Computer Interaction (HCI) systems requires the precise comprehension and identification of emotions. By capitalizing on the brain's electrical activity to deduce emotional states, emotion recognition based on EEG provides a dependable and unbiased method. EEG provides a direct window into the neural mechanisms that underlie emotions by circumventing potential biases associated with behavior and speech-based methods. [2]

The potential applications of EEG-based emotion recognition are0 vast. For instance, developing emotion-aware driver assistance systems for cars is widely acknowledged as a promising approach to improve driving safety. Within the realm of neurology, the examination of emotions triggered by certain stimuli and the corresponding neural activities can provide valuable insights into the diagnosis of affective disorders like PTSD (post-traumatic stress disorder) and depression. Psychological studies have uncovered a fascinating phenomenon observed in individuals with depression. Their attention is often drawn more towards negative or dysphoric stimuli rather than positive ones. In addition, the identified emotion can be used to support different types of therapy for emotional disorders, such as robot-assisted therapy and music-assisted therapy. Emotion recognition, also known as sentiment analysis, has long been a vibrant area of research in Information Retrieval (IR). The identified emotional states can be utilized to fulfill various information retrieval requirements, such as automatically labeling multimedia content or enhancing user profiles to enhance the relevance of recommended multimedia content. In the field of leisure and entertainment, such as computer gaming, researchers aimed to identify the emotional states of gamers to adapt better to the game's difficulty level, punishment, and encouragement. Researchers have confirmed the impact of emotional states on memory in the realm of Virtual Reality (VR) applications, such as VR in education. It has been found that a positive mood can have advantageous effects on spatial learning. Therefore, it is important to acknowledge and consider the role of emotions when learning in a virtual reality environment. [4]

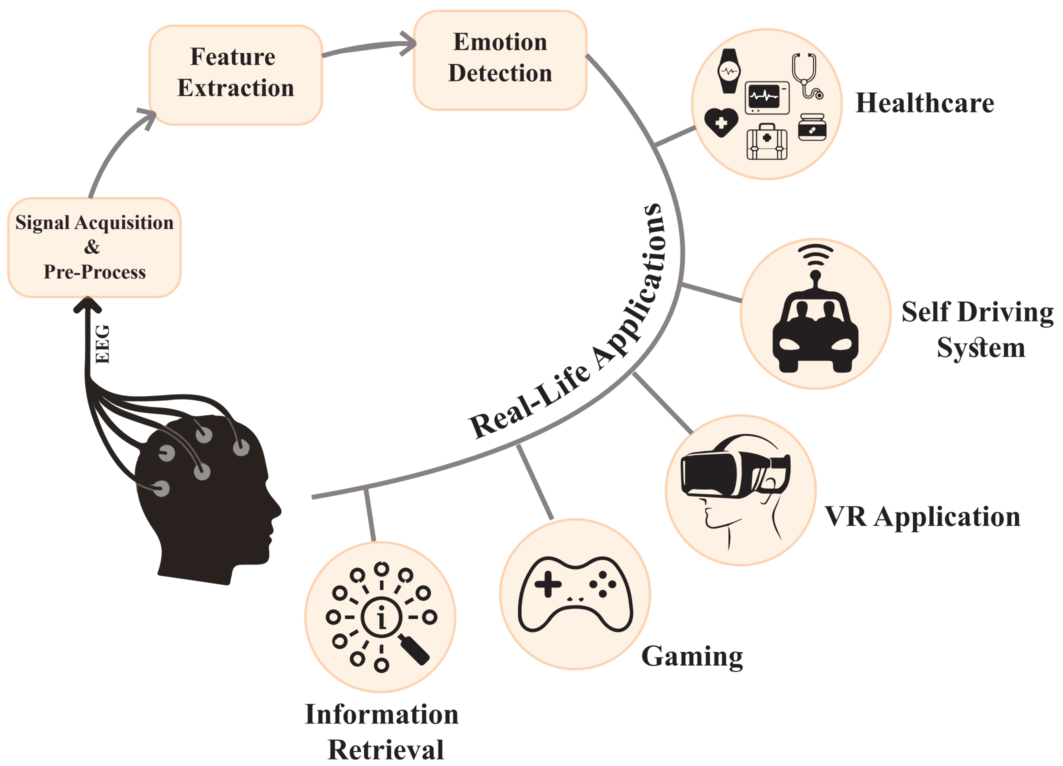


Figure 1.1: Potential applications of EEG-based emotion recognition

1.3 Motivation

The rapid advancement of technology in today's digital era has completely transformed how humans interact with computers. However, conventional approaches to interaction frequently need to consider a vital element of human communication - emotions. Given the constraints of current methods, it is becoming increasingly important to create advanced HCI systems that can comprehend and react to human emotions. Utilizing physiological signals such as EEG holds great potential for connecting humans and machines, enabling more empathetic and interactive experiences.

- The growing use of electronic devices and online activities calls for HCI systems to comprehend and react to human emotions to improve user experiences.
- Conventional approaches that rely solely on behavior and speech may not effectively

capture authentic emotional states, highlighting the significance of utilizing physiological signals such as EEG for enhanced emotion recognition.

- Utilizing EEG-based emotion recognition has the potential to establish a connection between humans and machines, allowing for more empathetic and personalized interactions that cater to individuals' emotional requirements.

- Adopting EEG-based emotion recognition is propelling the exciting field of affective computing research forward. This breakthrough technology has far-reaching implications across diverse healthcare, education, and entertainment domains.

- By integrating EEG-based emotion recognition into HCI systems, machines can better understand users' emotions, leading to more meaningful interactions and increased user satisfaction.

1.4 Objective of the Thesis

Exploring emotions' intricate nature and importance in human interactions, the study sets out to accomplish multiple goals to advance the field of EEG-based emotion recognition.

- Gain a deeper understanding of EEG patterns linked to various emotional states to enhance the accuracy and reliability of emotion recognition.

- Investigate and improve feature extraction methods to effectively capture subtle details in EEG signals that reflect different emotional states.

- Develop strategies to minimize the impact of individual differences in EEG-based emotion recognition, ensuring the reliability and applicability of models.

- Explore the potential of real-time emotion recognition using EEG data to enable interactive systems and interventions.

- The impact of EEG-based emotion recognition on user experience and engagement in different contexts will inform design decisions and system improvements.

1.5 Challenges

Although there are promising prospects for EEG-based emotion recognition, several challenges still need to be addressed to unlock its full potential.

- Understanding the intricate connections between EEG patterns and emotional states to extract meaningful and easily understandable features for emotion recognition algorithms.
- Accounting for individual differences in brain anatomy, physiology, and emotional expression is crucial to address the variability that can arise and complicate model generalization.
- Enhancing emotion recognition systems by incorporating EEG data with facial expressions, voice, and physiological signals leads to improved robustness and accuracy.
- Addressing the practical challenges of implementing EEG-based emotion recognition systems in real-world environments, including scalability, accessibility, and user acceptance.
- Understanding and respecting cultural diversity in emotional expression and perception is crucial for successfully developing and implementing EEG-based emotion recognition systems.

1.6 Thesis Organization

The report is organized into 6 chapters, including this chapter: ***Introduction*** where all the related topics are discussed, which are needed for understanding the research work. The outline of the remaining chapters is as follows:

Chapter 2

Topic - Background

This chapter presents an extensive overview of emotion recognition based on EEG, encompassing its foundational principles and historical progression. It establishes the groundwork for comprehending the forthcoming chapters and the importance of the research subject.

Chapter 3

Topic - Literature Review

This chapter conducts a review of existing literature related to EEG-based emotion recognition. By synthesizing existing knowledge, this chapter informs the research approach and methodology adopted in this study.

Chapter 4

Topic - Research Workflow & Implementation

This chapter presents the research workflow and methodology employed in this study. It discusses the dataset, data preprocessing techniques, feature extraction methods, and the design and implementation of machine learning and deep learning models for emotion recognition.

Chapter 5

Topic - Result & Performance Analysis

This chapter provides a detailed analysis of the experimental results and performance of the proposed methodology. It presents the evaluation matrices used to assess the effectiveness of the models. Additionally, it compares the performance of the proposed approach with existing methods and discusses the implications of the findings.

Chapter 6

Topic - Conclusion & Future Works

This concluding chapter summarizes the key findings of the study and discusses their implications. It highlights the contributions of the research and provides recommendations for future research directions.

1.7 Conclusion

In conclusion, EEG-based emotion recognition offers a captivating avenue for comprehending and deciphering human emotions. By capturing the brain's electrical activity, we can surpass the constraints of conventional emotion recognition techniques and open up possibilities for more advanced and compassionate human-machine interactions. With this thesis, our goal is to make a meaningful contribution to the field of affective computing research and its applications across different domains.

Chapter 2

Background

2.1 Introduction

Understanding human emotions is crucial in artificial intelligence, neuroscience, and psychology. Emotions affect behavior, thought, and well-being. Over time, scientists have developed discrete and continuous-dimensional models to measure and understand emotions. These models help categorize and understand emotions, advancing affective computing and emotion recognition systems. The brain controls emotion perception and regulation. The brain controls emotional reactions, sensory inputs, and responses through its complex architecture and systemic composition. By measuring brain electrical activity, electroencephalogram (EEG) data reveal cognitive and emotional processes. Deep learning and machine learning have improved emotion identification, making EEG data analysis more accurate and efficient. This chapter provides background for EEG-based emotion identification and lays the groundwork for techniques and applications.

2.2 Human Emotion

2.2.1 Understanding Emotion

Emotion is a complex psychological state that reflects human awareness and responds to external stimuli. It covers many experiences, from brief sensations to deep emotions. Emotion is crucial in human cognition, impacting how we make decisions, interact with others, and experience overall happiness. Studies indicate that people's emotional states significantly influence their decision-making and judgments, emphasizing the powerful role of emotions in shaping behavior and cognition. [7]

2.2.2 Models of Emotion

Quantifying emotional state labels within behavioral or physiological data is crucial for statistical analysis in psychology and intelligent computing-based emotion identification systems. It is important to capture and quantify the emotional state of target samples accurately. All quantitative emotion approaches fall into two types.

- **Discrete type of emotion quantification model:** The discrete model of emotion classifies emotional experiences into clear and recognizable states. Building upon Darwin's theory of evolution [8], esteemed scholars like Ekman [9] and Plutchik [10] have put forth models that outline a limited range of fundamental emotional states. These models propose that emotions are universally present and instinctive, crucial in human adaptation and survival. Anger, fear, joy, and sadness are fundamental to human emotional experiences. By classifying emotions into distinct states, researchers strive to create models that can effectively classify and interpret emotional expressions.

- **Continuous Dimensional Type of Emotion Quantification Model:** Unlike discrete models, the dimensional approach to quantifying emotions represents emotional experiences along continuous axes. The Valence-Arousal bipolar emotional quadrant system, developed by Russell [1], is a well-known dimensional model that presents emotions regarding their valence (positive or negative) and arousal (level of activation). This model offers a deeper comprehension of emotional states, considering the range of intensity and intricacy. Continuous dimensional models provide a versatile framework for understanding the complex nature of human emotions, allowing researchers to investigate the ever-changing relationship between various emotional aspects. Approaches like the Self-Assessment Manikins (SAM) scale-based method offer a way to measure emotional experiences across various dimensions, making conducting thorough evaluations of affective states easier.

The Valence axis covers a range of emotions, from positive to negative values, reflecting the varying levels of happiness and sadness experienced by individuals. Similarly, when the Arousal axis has positive values, it indicates an activated state that reflects excitement. On the other hand, negative values on this axis signify an unactivated state, reflecting calmness. Furthermore, incorporating extra dimensions can achieve a more comprehensive measurement of emotions beyond the standard axes (Figure 2.1). The Dominance dimension, for example, measures the level of control an individual has within the emotional process. When individuals are influenced by external factors, such as unexpected or frightening situations, their emotional state often shows reduced dominance. On the other hand, if the person controls their surroundings, their emotional state is more likely to show a greater sense of power.

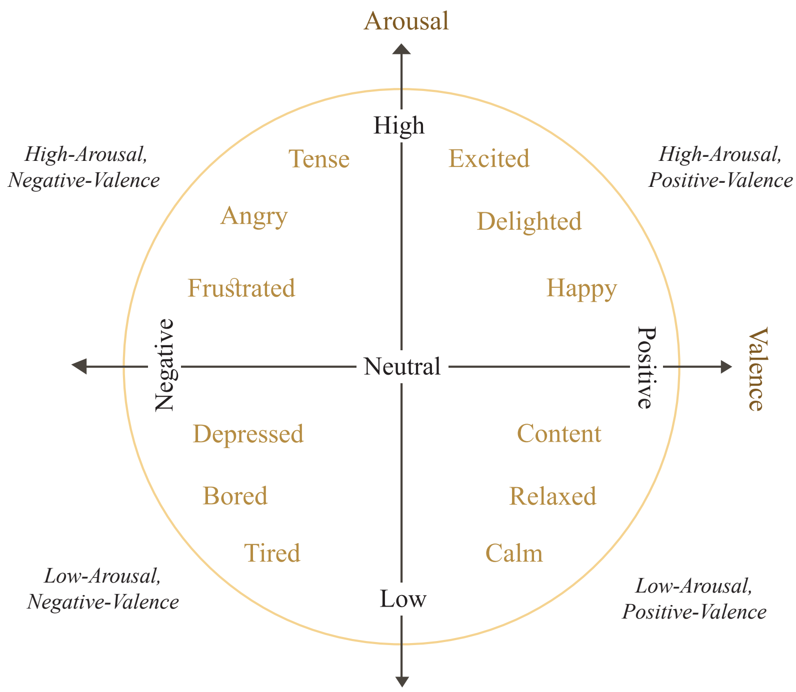


Figure 2.1: Valence-Arousal bipolar coordinate system proposed by Russell [1]

Assigning discrete emotional states to specific coordinates within the continuous dimensional space enables direct correspondence. For example, sadness is usually found in the quadrant with low Arousal, low Dominance, and low Valence. At the same time, happiness is located in the quadrant with high Arousal, high Dominance, and high Valence. The Self-Assessment Manikins (SAM) scale-based approach is commonly used for evaluating continuous emotional states. This approach utilizes visual representations, like manikins, in a questionnaire to evaluate the level of Valence and arousal. The SAM questionnaire typically uses a discrete scale that ranges from 1 to 9 (Figure 2.2).

When developing a recognition system based on the continuous type of emotion quantification model, researchers can choose between classification or regression modeling techniques.

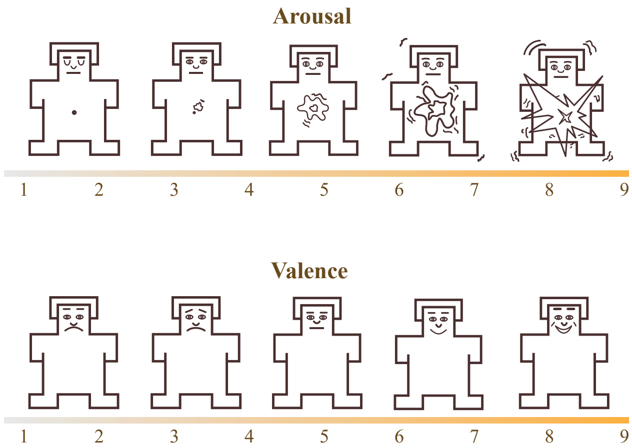


Figure 2.2: The Self-Assessment Manikins (SAM) scale

Regression modeling aims to create a model using data from samples with continuous emotional ratings. This allows for accurate prediction of emotional ratings for unknown samples. However, in classification modeling, emotional ratings are categorized into different levels, and samples are assigned specific class labels based on their emotional ratings. Classification algorithms are trained on labeled samples to determine the emotional classes of unknown samples.

2.3 The Brain

The brain is an incredibly intricate organ, one of the largest in the human body, consisting of over 100 billion nerves. The brain is the control center for all bodily actions and reactions, constantly processing sensory information and coordinating responses to maintain essential bodily functions necessary for human survival, including sensation, movement, and cognition.

When considering the human brain's functionality, it can be categorized into three primary sections. First and foremost, the cerebrum, commonly called the large brain, is responsible for overseeing advanced cognitive functions like language and reasoning. Additionally, the brainstem is crucial in overseeing fundamental physiological functions such as visual and auditory processes. Lastly, the cerebellum, the little brain, is crucial in coordinating movements and controlling motor functions. The brain is divided into different lobes [11], each with its own set of functions:

- **The frontal lobes**, located at the front of the brain, oversee problem-solving, reasoning, and aspects of speech.

- **The parietal lobes**, situated centrally, oversee movement and handwriting and process sensory information regarding touch, pressure, and temperature.
- **The temporal lobes**, located on the sides of the brain, are responsible for processing auditory information, storing memories, and influencing emotional reactions.
- Located at the back of the brain, **The occipital lobes** play a crucial role in processing visual information and impact memory and perception.

The brain comprises three main components: the cerebrum, cerebellum, and brainstem. The cerebrum comprises the cerebral cortex, brain nucleus, and limbic system. The cerebral cortex is particularly important for cognitive and emotional functions. The cerebral cortex is divided into left and right hemispheres, each containing distinct areas responsible for different functions. These functions include planning, sensory integration, and auditory and visual stimuli processing. The brain's complex structure and organization allow it to carry out various crucial functions necessary for human survival, including fundamental physiological processes and intricate cognitive abilities.

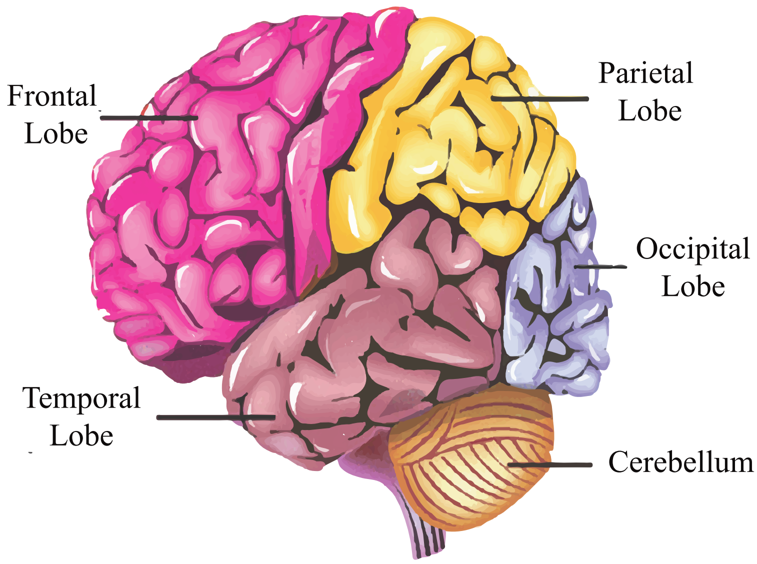


Figure 2.3: Physiological structure of the cerebral cortex [2]

2.4 EEG Signals

2.4.1 Basics of EEG

In 1929, Hans Berger made a groundbreaking discovery by recording the first human electroencephalogram (EEG). This began a field that completely transformed our understanding of brain activity [6]. EEG is a method used to measure potential variations between different areas of the scalp, which can give us valuable information about the electrical activity happening in the brain. Originally sparked by Richard Caton's research on animal brain activity in the 19th century, EEG rapidly progressed with further confirmation from electrophysiologists and neurophysiologists. The EEG signal, generated by the combined activity of neurons in certain brain regions, offers valuable psychophysiological insights useful for medical diagnosis and neuroengineering applications. For example, EEG signal analysis has proven valuable in clinical medicine for detecting diseases and developing Brain-Computer Interface (BCI) systems. These systems can empower disabled individuals by allowing them to control devices using their brain activity. [2]

EEG signals can be classified into five distinct frequency bands (Figure 2.4) associated with specific brain states and activities. Here are some of the frequency bands:

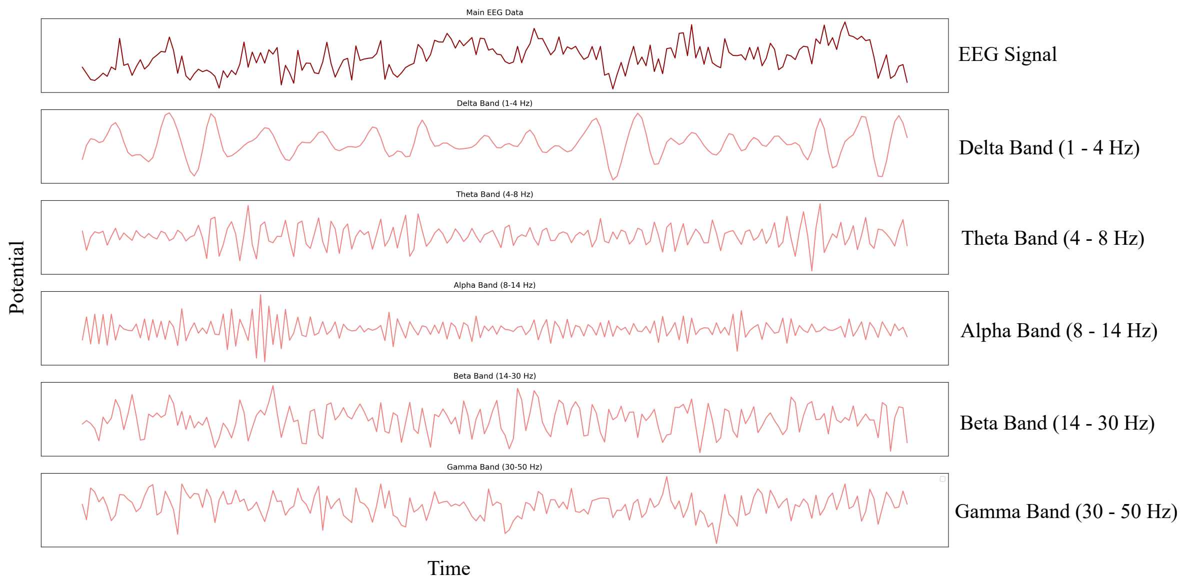


Figure 2.4: The five frequency bands of EEG signal

- **Delta (1–4 Hz):** Predominantly observed in the frontal cortex during unconsciousness or deep sleep, indicating lack of oxygen or anesthesia.

- **Theta (4–8 Hz):** Mainly detected in the parietal and temporal lobes, indicating relaxation and working memory load.

- **Alpha (8–14 Hz):** Prevalent in the occipital and parietal lobes during rest with closed eyes, disappearing in response to external stimuli or mental engagement.

- **Beta (14–30 Hz):** This is typically observed in the frontal lobe during active mental states, indicating high cognitive activity and focus.

- **Gamma (>30 Hz):** Associated with cognitive tasks, sensory processing, attention, and information integration in the brainstem, with low amplitudes.

Every frequency band offers researchers and clinicians a unique perspective on distinct facets of brain function, thereby supplying them with invaluable knowledge of neural processes. A fundamental comprehension of electroencephalogram (EEG) signals and their frequency bands is critical to deciphering cerebral activity and diagnosing a wide range of neurological disorders. With the ability to detect abnormalities in brain function and monitor sleep patterns, EEG continues to be a versatile instrument in neuroscience and clinical practice. [12, 6, 13]

2.4.2 EEG Signals in Emotion Recognition

Understanding the physiological causes of emotions is crucial to training computers to recognize and understand them. Emotions may be expressed through words, facial expressions, voice intonation, and nervous system responses. In contrast to facial expressions and voice tone, physiological signals are more accurate markers of emotional states since they are beyond the user's control. Recording from the scalp using conductive media and metal electrodes, EEG may capture brain alternating-type electrical activity. It helps understand human emotions since it can detect affective state changes. Existing research implies that emotional stimuli induce stronger low-frequency EEG responses than high-frequency. Positive emotions evoke less reaction than negative ones. The brain's midline power spectrum of Beta, Alpha, and Theta waves changes significantly whether happy, sad, or afraid. The EEG midline power spectrum is useful for categorizing emotions.

As the cerebral cortex regulates higher emotional and cognitive activities, EEG emotion detection aims to identify brain areas significantly connected to emotions. After cerebral cortex electrode placement categorization, emotions are classified by extracting EEG characteristics from each electrode group. Ranking electrodes by significance helps feature selection algorithms identify emotion-processing brain areas. EEG studies of functional brain connectivity have linked brain areas to emotions. The left frontal areas of the brain are activated by happiness but decreased by worry. Positive emotions are associated with frontal midline theta band power, but negative emotions are not. These findings show how neurophysiology links emotional states to EEG signal characteristics and how to recognize emotions from EEG data.

2.4.3 EEG signal acquisition

EEG has gained significant recognition as a reliable technique for capturing the brain's electrical signals. State-of-the-art equipment includes electrodes, a data storage unit, an amplifier, and a display unit. EEG signal acquisition can be done using two main methods: invasive and non-invasive. Some methods require the insertion of electrodes into the skull, allowing access to the brain's cortex. Although this method provides excellent signal quality and intensity, its operation can be quite demanding due to the surgical procedures' complexity. Alternatively, non-invasive methods involve the application of electrodes to the subject's scalp, which makes them more convenient to use and the prevailing approach in modern Brain-Computer Interface (BCI) research. Efficient acquisition of noninvasive EEG signals can be achieved through the use of affordable wearable EEG headsets and helmets with electrodes distributed across the scalp. These devices are easily accessible and widely utilized in research. The number and placement of electrodes used in EEG experiments vary depending on the specific research goals. However, the widely used International 10–20 electrode placement system is commonly utilized, with electrode counts varying from 6 to 62.
According to the research findings [14], electrodes related to emotions are mainly spread out in the frontal, parietal, occipital, and temporal lobes, as well as the central area of the brain. These regions are associated with the physiological processes involved in creating emotions. Abbreviations are used to represent specific brain regions. The left hemisphere is indicated by odd-numbered suffixes, while the right hemisphere is indicated by even-numbered suffixes. The abbreviations for the brain regions are as follows: FP (front polar), F (frontal), P (parietal), and O (occipital), T (Temporal). (Figure 2.5)
Modifying the electrodes' arrangement can substantially impact the feature dimension extracted from EEG signals, making experiments easier to conduct and reducing computational complexity. This simplified approach makes conducting experiments on emotion recognition easier and improves their practicality.

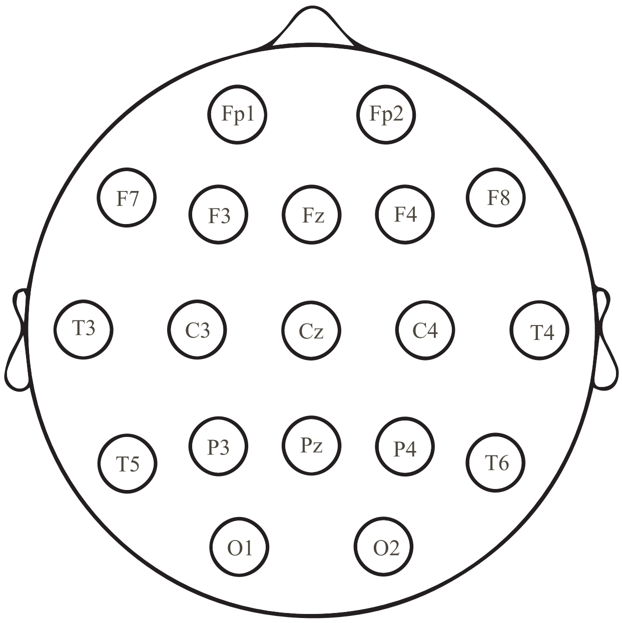


Figure 2.5: Electrodes for EEG recording in different lobes

2.4.4 EEG Signal Preprocessing

Preprocessing of EEG signals is crucial to ensure the removal of noise and improvement of weak EEG signals. These signals can easily be affected by various sources of interference, including electrodes and body movements. It is important to know that EEG recordings can be affected by various artifacts, such as electromyogram (EMG) signals originating from eye blinks and neck muscles. There are also concerns regarding motion artifacts caused by cable movement and electrode displacement. Various filters in the frequency domain are often employed to narrow the bandwidth of EEG signals and minimize artifacts. These filters include high-pass, low-pass, Butterworth, and notch filters. However, it is important to exercise caution when eliminating artifacts, as they may hold significant insights into emotional states.

Some commonly used preprocessing methods are independent component analysis (ICA) [15], principal component analysis (PCA) [16], common average reference (CAR) [17], and common spatial patterns (CSP) [18]. These methods employ blind source analysis to eliminate noise from multi-channel recordings and identify spatial filters to detect signals associated with muscular motions. To account for variations in individuals' physiological signals and minimize the influence of previous stimuli on emotional states, the baseline EEG features are subtracted from the post-stimulation features and then scaled to a range of. Emotion recognition research frequently centers around classifying emotions, specifically distinguishing between different arousal levels or positive and negative emotions. These emotional labels are typically derived from subjective rating data. [2]

2.4.5 EEG Feature Extraction

Once the preprocessing is complete, the subsequent task involves extracting features. This step is of utmost importance as it helps capture the information that reflects an individual's emotional state in EEG-based emotion recognition. Various methods can be used to gain valuable insights into emotional states, including analyses in the time, frequency, time-frequency domains, and nonlinear approaches. Research often incorporates features from the time-frequency domain (35%), followed by the frequency domain (27%) and the time domain (20%). Some researchers may opt to use raw data without features (11%) in their studies, as this approach can benefit deep learning algorithms. By doing so, they can ensure the preservation of important information. Nonlinear features also have some application (7%). [2]

Time domain analyses

EEG research has relied extensively on time domain studies since most EEG collection equipment operates in this domain. Time domain analysis uses several methods. These include histogram analysis, event-related potential (ERP), Hjorth parameters (activity, mobility, and complexity) [19], and Higuchi's fractal dimensions (FD) [20], which quantify self-similarity and complexity. Statistical variables including mean, standard deviation, variance, skewness, kurtosis, and geometric aspects underpin time domain analysis. These characteristics provide valuable EEG signal insights while reducing information loss. Key features commonly analyzed in the time domain include:
Mean: Represents the average signal value over a given period.

$$\mu = \frac{1}{N} \sum f(x) \tag{2.1}$$

where:

- μ: Mean
- $f(x)$: Signal values
- N: Number of samples in the signal

Standard Deviation: Measures the dispersion or variability of signal values around the mean.

$$\sigma = \sqrt{\frac{1}{N}\sum(f(x)-\mu)^2} \tag{2.2}$$

where:

- σ: Standard deviation

Variance: Indicates the average squared deviation of signal values from the mean.

$$\sigma^2 = \frac{1}{N}\sum(f(x)-\mu)^2 \tag{2.3}$$

where:

- σ^2: Variance

Kurtosis: Describes the peakedness or flatness of the signal distribution relative to a normal distribution.

$$\kappa = \frac{1}{N}\sum\frac{(f(x)-\mu)^4}{\sigma^4} \tag{2.4}$$

where:

- κ: Kurtosis

These features provide valuable insights into the characteristics and dynamics of EEG signals, facilitating the understanding and interpretation of brain activity.

Frequency domain analyses

Frequency-domain analysis outperforms time-domain analyses for automated emotion recognition in EEG recordings. To extract characteristics, these techniques transform time-domain EEG data into frequency-domain signals. Power Spectral Density (PSD), logarithm energy spectrum, HOS, and differential entropy are common frequency-domain analysis approaches. EEG signals are split into frequency subbands for analysis, and characteristics are retrieved from each. Short EEG segments are often analyzed for frequency using the fast Fourier transform (FFT). [21]

Differential Entropy: Differential entropy (DE) is the logarithm energy spectrum in a particular frequency band for a fixed-length EEG sequence. Like the entropy for assessing the complexity of continuous random variables, DE can be represented as:

$$H(X) = -\int_X f(x)\log(f(x))\,dx \tag{2.5}$$

DE, recognized for its efficacy in quantifying signal complexity, has successfully analyzed non-stationary and non-linear signals, such as EEG. Rooted in the extension of Shannon entropy, DE is employed to assess the intricacy of a continuous random variable. The introduction of DE to EEG-based emotion recognition was first introduced in [12].

If a random variable follows a Gaussian distribution $N(\mu, \sigma^2)$, the calculation of its differential entropy can be expressed using the following formulation:

$$\begin{aligned} DE &= -\int_{-\infty}^{+\infty} p(x)\log(p(x))dx \\ &= -\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\right) dx \\ &= \frac{1}{2}\left(2\pi e\sigma^2\right) \end{aligned} \tag{2.6}$$

Power Spectral Density: The Power Spectral Density (PSD) is computed using Welch's Method [22], which involves a two-step process. Firstly, each windowed segment's Discrete Fourier Transform (DFT) is computed. For each segment, the DFT $X_k(f)$ is calculated using the formula:

$$X_k(f) = \sum_{n=0}^{M-1} x_k(n) e^{-j2\pi fn/M} \tag{2.7}$$

Here, $x_k(n)$ represents the k-th windowed segment, M is the segment length, and $e^{-j2\pi fn/M}$ denotes the DFT basis function. Then, the squared magnitudes of these DFTs across all segments are averaged to obtain the estimated power spectral density $P_{xx}(f)$ at frequency f. This is achieved using the formula:

$$P_{xx}(f) = \frac{1}{K}\sum_k |X_k(f)|^2 \tag{2.8}$$

Here, K is the total number of segments, and $|X_k(f)|^2$ represents the squared magnitude of the DFT of the k-th segment. The PSD can be effectively computed using Welch's Method by following these steps.

Time-frequency domain analyses

Recent studies have revealed that electroencephalogram (EEG) signals exhibit non-stationary properties. This questions the linearity and quasi-stationarity assumptions underpinning conventional time and frequency-domain analysis techniques. Methods that depend on characteristics such as amplitude, duration, variance, and autocorrelation are inadequate for analyzing non-stationary signals like adult EEG signals. In response, alternative methodologies like time-frequency domain analysis techniques have surfaced. These techniques combine data from both domains, thereby facilitating analysis at the local level. These techniques demonstrate exceptional performance in capturing time-varying and non-stationary signals, which is essential for characterizing various emotional states. Out of these various techniques, the Wavelet transform is the most prevalent method utilized in time-frequency analysis. The Short-time Fourier transform (STFT), Hilbert Huang transform (HHT), and wavelet packet transform (WPT) are additional crucial techniques. [4]

An emotion recognition model requires EEG data collection, preprocessing, feature extraction, feature selection or reduction, and emotion categorization. Selecting the right classifier is crucial to proper emotion categorization. A mathematical function, the classifier, predicts the correct class of an unknown observation in a validation dataset. Different classification approaches have been used to categorize emotional EEG data in affective computing. Various classifiers are used, from support vector machines, decision trees, and linear discriminant analysis to recurrent neural networks and long short-term memory models. The right classifier greatly affects emotion recognition accuracy.

2.5 Overview of Machine Learning

Recently, machine learning (ML) has become an important tool for studying emotion detection, giving a range of methods for identifying EEG-based emotional states. Emotion recognition, sometimes described as a regression or classification issue, depends on the emotional model—dimensional or categorical. Dimensional models display emotions as continuous values along an axis, unlike categorical representations identifying emotions. Most current methods focus on classifying Russell's 2D emotion model [1] regions or categorical emotions. Using the Scopus database as a source of information, machine learning techniques, which classify emotional states using Brain-Computer Interface (BCI) data from EEG, have become more popular. Machine learning, an AI application, is vital to BCI data processing because it distinguishes distinct brain activity patterns. Supervised and unsupervised machine learning are the main categories. Unsupervised learning optimizes classifier parameters using input data and a cost function, whereas supervised learning uses training data. SVM, K-NN, DT, and RF are common machine-learning models used in EEG signal categorization for emotion identification. Each model has distinct advantages, which will be briefly outlined here.

2.5.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a powerful technique in supervised machine learning, capable of effectively addressing a wide range of regression and classification problems, regardless of their linearity. SVM has been widely used in various fields including face detection, disease diagnosis, and text recognition. SVM aims to identify hyperplanes that effectively separate distinct groups of data, referred to as support vectors, by maximizing the boundaries of separation according to specified labels or classes. Through the utilization of kernel functions, SVM has the capability to transform input data into feature spaces of higher dimensions. This allows for linear separation in situations where direct feasibility is impossible. Optimizing parameters, such as choosing appropriate kernel functions, plays a vital role in SVM-based classification, directly affecting the model's performance. Although SVM requires significant computational resources, it is widely favored for its intuitive and theoretically sound approach, making it a popular option for various classification tasks. [23]

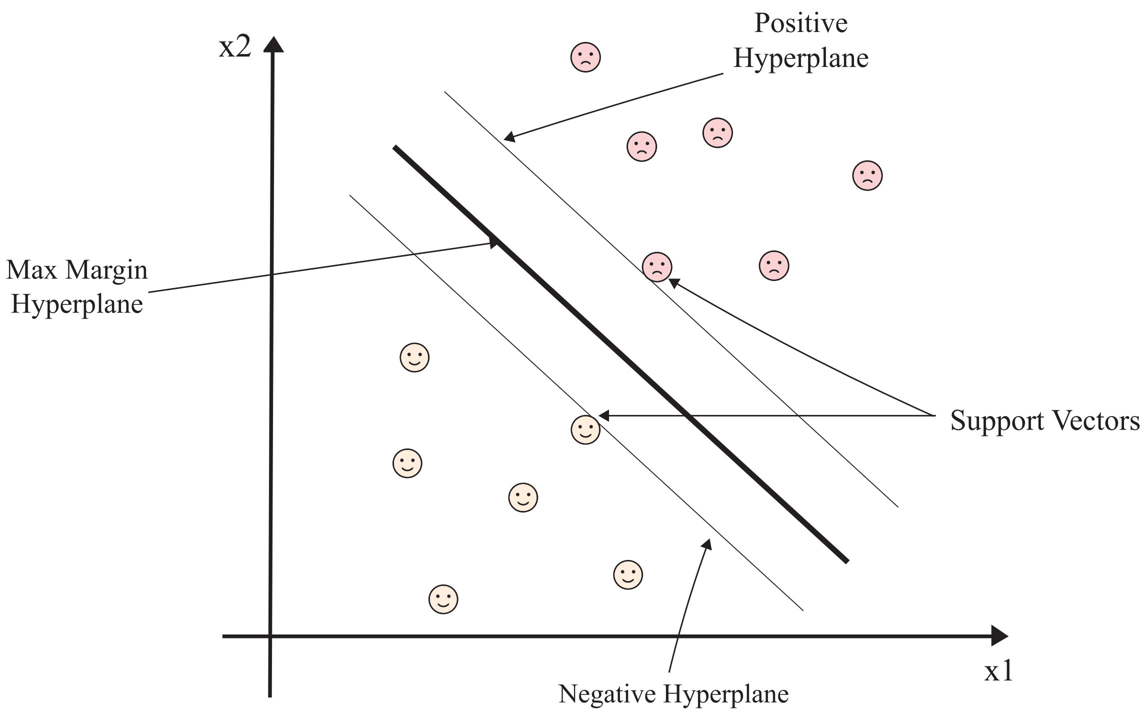


Figure 2.6: Support Vector Machine classifier

2.5.2 Decision Tree (DT)

Decision trees (DT) are incredibly flexible machine-learning tools that can be used for various tasks, including regression and classification. They work by repeatedly dividing the dataset into smaller groups using criteria that optimize the separation of data, resulting in a hierarchical structure. One commonly used criterion aims to maximize entropy reduction with each split. The decision tree assigns a class label to each leaf node. In contrast, the nonterminal nodes, such as the root and internal nodes, have attribute testing conditions to differentiate records with different characteristics. Decision trees can be seen as collections of rules, where higher nodes have a stronger impact on the overall accuracy of the samples. The ID3 algorithm underwent significant advancements and transformed into the C4.5 method. This evolution laid the groundwork for the development of novel supervised classification algorithms. It is worth mentioning that the J48 algorithm is a Java-based adaptation of C4.5, which has made it more accessible and user-friendly. Salvatore Ruggieri presented EC4.5, which provides decision trees similar to C4.5 but with notable enhancements in performance. Decision trees are highly regarded in machine learning for their clear and logical structure, allowing them to be used in various industries and applications. The advantage of interpretability is especially beneficial, leading to their extensive use in various applications. [24]

2.5.3 Random Forest (RF)

Random forest (RF) is a powerful ensemble learning technique that combines decision trees to tackle classification and regression tasks. Based on the bagging algorithm concept, RF can effectively handle large datasets by utilizing a subset of features when creating decision trees. Its impressive accuracy and efficiency have made it a widely favored classification technique. In RF, the final result is determined by the collective decision of all the decision trees in the ensemble. The operational framework of RF comprises various stages: First, sets for training purposes are randomly selected to match the size of the sample set. Next, a decision tree is constructed using each training set. Afterwards, a random selection is made from a group of attributes, and the most optimal attribute is then used to divide the nodes within this selected group. Afterwards, the predictions are derived from each decision tree, and a vote is made for each predicted outcome. Ultimately, the final decision is determined by choosing the outcome that receives the highest number of votes. [25]

2.5.4 K-Nearest Neighbor (KNN)

The k-nearest neighbor (KNN) algorithm is a powerful machine-learning tool for regression and classification tasks. Over time, KNN has evolved and is now extensively used in different fields including text recognition, emotion recognition, and face recognition. The core concept of K-NN is centered on evaluating the similarity between training and test data points, usually done using a distance function. When determining the class of a new test instance, KNN finds the training example that closely resembles the object and assigns it the corresponding class label. The algorithm's accuracy tends to increase as the number of nearest neighbors (k) increases. Initially, the KNN algorithm chooses a set of training samples and determines the number of neighbors (k). Afterwards, it computes the Euclidean distance between the test and training samples to determine the closest k samples in the training set for every new test instance. Ultimately, the most frequent class value among the k-nearest training samples determines the test instance's class. It is worth mentioning that as datasets grow larger, the computational complexity of KNN also increases substantially. As a result, K-NN is better suited for datasets with a moderate number of samples. [26]

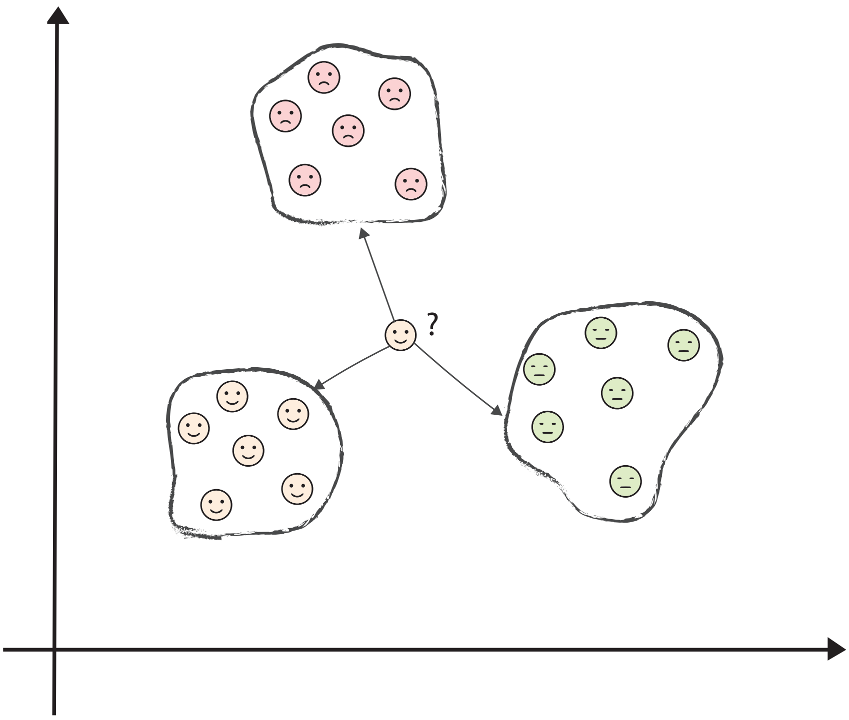


Figure 2.7: K-Nearest Neighbor classifier

2.6 Overview of Deep Learning

Machine learning and artificial intelligence include deep learning (DL), which learns from data. Deep learning (DL) is popular because it performs well in many classification and regression problems. Deep and superficial feature extraction is used to find a smaller yet relevant collection of characteristics. Manually extracting and choosing features is hard and domain-specific, unlike shallow features, which rely on algorithms and need plenty of labeled data. Traditional feature engineering approaches may struggle to extract complicated time series patterns. Thus, hierarchical feature representations are learned independently using deep learning. A 'deep' artificial neural network (ANN) has seven or more layers. DL began in the 1980s when Hinton and Salakhutdinov made substantial contributions to feature extraction in 2006. Deep neural networks extract relevant features via nonlinear transformations, eliminating feature space reconstruction and allowing direct processing of raw input. Deep learning (DL) algorithms, including Long Short-Term Memory networks (LSTMs), autoencoders, Deep Belief Networks (DBNs), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNNs), have had a significant impact in many areas. DL models have shown promise in emotion identification using EEG signal categorization, demonstrating their ability to abstract high-level data and adapt. [27]

2.6.1 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) have become fundamental components in the fields of computer vision and image processing owing to their exceptional capability of learning hierarchical representations of visual data automatically. CNN can accurately discern complex patterns and structures in images by effectively capturing spatial hierarchies of features through convolutional filters and pooling operations. Each layer of these networks is tasked with extracting progressively more intricate features from the input data. To generate feature maps, convolutional layers employ learnable filters to convolve input images. In contrast, pooling layers perform downsampling on the feature maps to preserve crucial information while decreasing computational complexity. By subjecting CNN to non-linear activation functions and repetitive application of these layers, they can acquire the ability to identify complex visual patterns and objects.

Alternating convolutional and pooling layers constitute the architecture of CNN, which are then followed by fully connected layers to facilitate classification or regression operations. Local features are extracted by the convolutional layers and subsequently refined and aggregated in subsequent layers to generate high-level representations of the input data. Support vector machines (CNN) have exhibited exceptional efficacy across various image-centric assignments, such as segmentation, object detection, and classification. Their achievement can be attributed to their capability of learning features autonomously from unprocessed data, which obviates the necessity for human feature engineering and facilitates comprehensive learning spanning from pixel-level inputs to overarching predictions. In general, CNN constitutes a potent instrument within the repertoire of machine learning algorithms, finding utility in various fields such as time series analysis and natural language processing, in addition to computer vision. [28]

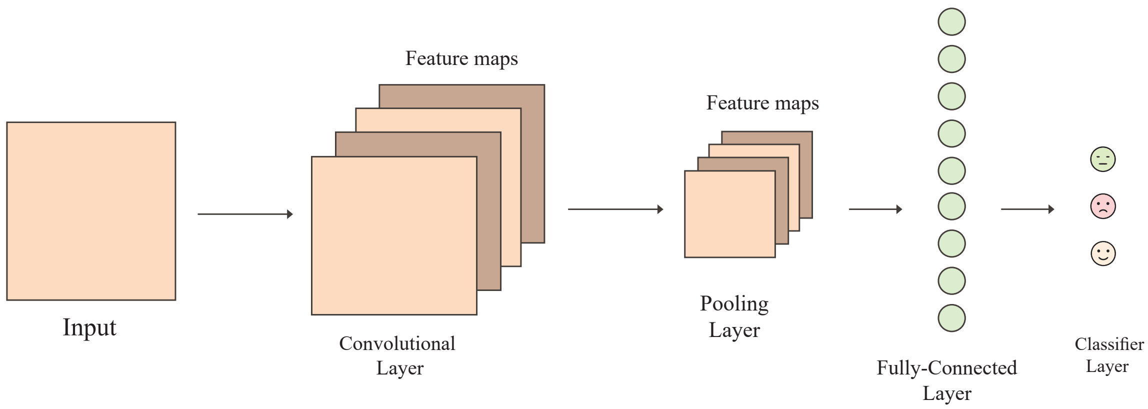


Figure 2.8: A simple Convolutional Neural Network (CNN) and its main layers

2.6.2 Bidirectional Long Short-Term Memory (BiLSTM)

Bidirectional Long Short-Term Memory (BiLSTM) networks are a type of recurrent neural network (RNN) architecture that has gained significant attention in sequence modeling tasks due to their ability to capture long-term dependencies in sequential data [29]. Unlike traditional RNNs, which process input sequences in a sequential manner from left to right, BiLSTMs process sequences in both forward and backward directions simultaneously. This bidirectional processing allows BiLSTMs to capture context information from past and future inputs, enabling them to better understand and model temporal dependencies in the data.

The basic principle behind BiLSTMs is the use of two hidden layers for each time step: one layer processes the input sequence from left to right, while the other processes it from right to left. The outputs of these two layers are then combined to produce the final output for each time step. This bidirectional approach helps overcome the vanishing gradient problem often encountered in traditional RNNs, allowing BiLSTMs to capture long-range dependencies more effectively.

The working method of BiLSTMs involves three main components: input, output, and forget gates. These gates control the flow of information within the network by regulating the input, output, and retention of information at each time step. The input gate determines how much new information is added to the cell state, the forget gate controls which information is discarded from the cell state, and the output gate determines the output based on the current cell state.

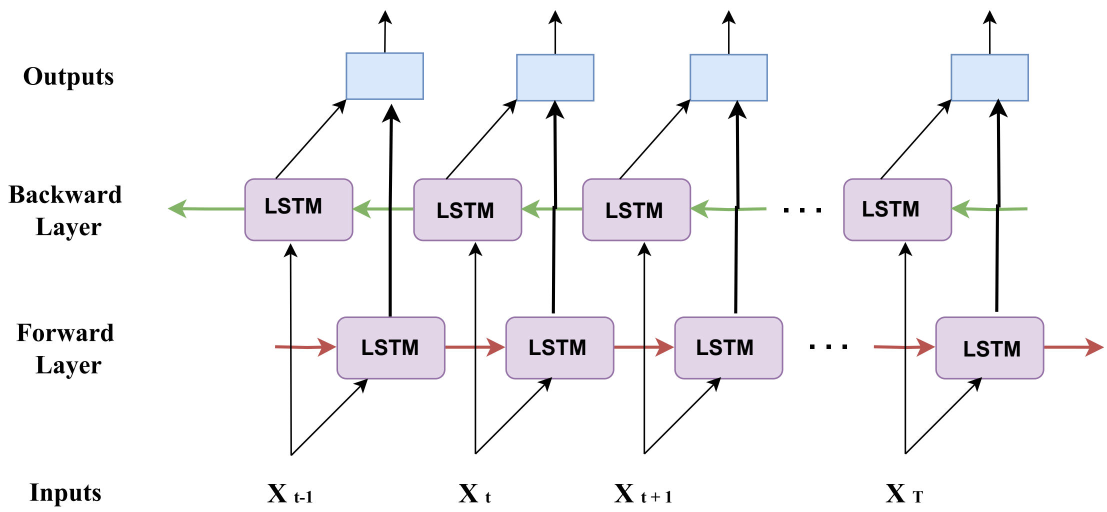


Figure 2.9: Architecture of a Bidirectional Long Short-Term Memory (BiLSTM) network [3]

BiLSTMs have been extensively implemented in machine translation, named entity recognition, and sentiment analysis, among other natural language processing tasks. Furthermore, they have demonstrated potential in tasks involving the prediction of time series, recognition of speech, and identification of emotions. The utilization of the bidirectional processing capabilities of BiLSTMs has enabled scientists to attain cutting-edge performance across a variety of sequential data analysis tasks. Additionally, BiLSTMs exhibit considerable promise in emotion recognition based on EEG [30, 31]. EEG signals possess an intrinsic sequential nature, wherein every data point corresponds to a distinct time increment. BiLSTMs are capable of capturing temporal dependencies and patterns in both forward and backward directions by bidirectionally processing EEG signals; this enables an improved understanding of the underlying emotional states.

2.7 Conclusion

This chapter has provided an overview of the fundamental components that comprise EEG-based emotion recognition. An investigation was conducted into signal preprocessing, encompassing noise elimination and signal amplification, and diverse methodologies for feature extraction spanning multiple domains. Furthermore, the discourse used machine learning and deep learning algorithms to categorize affective states according to EEG data. A comprehensive grasp of these foundational principles is deemed indispensable for subsequent investigations into emotion recognition based on EEG and its progressions.

Chapter 3

Literature Review

3.1 Introduction

The potential applications of emotion recognition from EEG signals have elevated it to a prominent research domain, encompassing affective computing, human-computer interaction, and mental health monitoring. Recent research has concentrated on classifying emotions from EEG signals using machine learning and deep learning techniques in an effort to improve the precision and robustness of emotion recognition systems. This chapter provides an analysis of various recent contributions in the field, emphasizing the research methodologies, algorithms, and performance metrics that were employed.

3.2 Related Works

Recent studies have prominently focused on detecting emotions in EEG signals, often utilizing statistical characteristics derived from EEG coupled with ML for mental state classification. Early research, exemplified by Duan et al. [32], proposed differential entropy for emotion recognition using EEG signals. Alhalaseh and Alasasfeh [5] devised an emotion recognition system based on machine learning and EEG signals. They combined feature extraction techniques to analyze EEG signals, including entropy and Higuchi's fractal dimension, with EMD/IMF and VMD. The system highlighted its potential for mental healthcare management while achieving a high accuracy rate of 95.20%. Using the SEED dataset, Lu [33] concentrated on automatic emotion decoding from EEG signals. The research examined various feature extraction techniques, including Hjorth mobility features, and assessed different machine learning approaches, concluding that a DNN classifier yielded the highest accuracy.

Mohutsiwa and Jamisola [34] conducted a comprehensive investigation into EEG-based human emotion classification, focusing on four emotions – amusement, disgust, sadness, and fear – elicited by custom video clips. Their study introduced an innovative approach combining various computational techniques to enhance emotion classification accuracy. Initially, Independent Component Analysis (ICA) was employed for artifact removal, ensuring the cleanliness and reliability of EEG data. Subsequently, feature extraction utilized band power analysis, Hjorth parameters, and their combination to capture diverse signal properties comprehensively. Additionally, Neighborhood Component Analysis (NCA) and Minimum Redundancy Maximum Relevance (mRMR) techniques were applied for feature selection, identifying the most informative features for emotion classification. Evaluation of six machine learning models revealed the LSTM RNN model as the top performer, achieving an impressive accuracy rate of 99.1%. The study highlighted the significance of preprocessing techniques and feature selection strategies in improving classification accuracy, with future research aimed at further validating and refining the proposed methodology.

Moshfeghi et al. [35] investigated the potential of EEG signals in the context of emotion classification. They utilized EEG data from controlled environments that had been collected previously and preprocessed to eliminate artifacts. The processed data was utilized to extract ten features, which were then combined in four unique ways to generate four separate datasets. WEKA software was utilized to prepare each dataset for classification. The classification task was completed using a support vector machine (SVM), which achieved an accuracy of 54% when categorizing three distinct emotional states and 74% on average when categorizing binary emotional states. Notwithstanding the encouraging outcomes, the scope remains for enhancement. Augmenting the number of subjects in the dataset may result in improved accuracy. Furthermore, techniques such as peak detection and thresholding could be investigated for EEG data processing. Another potentially intriguing outcome is separating EEG signals into distinct frequency regions, such as Alpha and Beta.

Saranya et al. [36] explored EEG-based emotion identification using the SEED dataset, aiming to understand and classify emotions such as happiness, sadness, fear, and neutrality. Emotions are fundamental to human experience, influencing decisions and interactions. EEG signals offer a direct representation of these emotions, but their classification poses challenges due to the instability of brain impulses. To address this, the researchers employed sampling and classification techniques, alongside thorough pre-processing to remove noise. Their proposed algorithm focused on accurately discerning between the four emotional states by analyzing EEG patterns unique to each emotion. Through their work, Saranya et al. contributed to advancing EEG-based emotion recognition, paving the way for more empathetic human-computer interactions. They achieved an average accuracy of 54% for three distinct emotional states and 74% for binary emotional states.

Various methodologies and algorithms are utilized in the studies under discussion to classify emotions from EEG signals. Emotion classification employs machine learning algorithms such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Support Vector Machines (SVM), and k-nearest Neighbours (k-NN). Notwithstanding the notable precisions attained in certain investigations, constraints such as the utilization of restricted datasets, absence of external validation, and failure to explicitly address constraints persist, potentially undermining the dependability and applicability of the results.

Deep learning uses neural network designs and machine learning to categorize EEG data for emotion recognition. Researchers use signal rhythm decomposition, entropy feature extraction, and CNN/LSTM integration. Emotion classification is improved via ensemble learning. Capsule Networks (CapsNets) show potential for emotion identification utilizing EEG signals' frequency domain, spatial characteristics, and frequency band properties. Deep learning methods increase emotion recognition from EEG signals regardless of dataset parameters or performance measurements, demonstrating its promise in affective computing and human-computer interaction.

Kumar and Molinas [37] designed deep neural networks and machine learning algorithms to examine the recognition of human emotions from EEG signals. Utilizing the SEED and DEAP datasets, they demonstrated a technique for automatic emotion recognition from EEG signals in their study. After decomposing the signals into rhythms, entropy features were extracted from EEG signals for classification using Multi-Layer Perceptron (MLP) and Convolutional Neural Network (CNN) models. The superior performance of the CNN model compared to the MLP model on the SEED dataset showcases the emotion recognition capabilities of CNNs.

In their study, Iyer et al. [38] introduced an ensemble learning methodology that integrates Long-Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models to iden-

tify human emotions from EEG recordings. Their method obtained a high degree of precision on the SEED dataset, suggesting bright prospects for emotion classification. However, the DEAP dataset exhibited a comparatively reduced level of accuracy. Combining CNN and LSTM models in an ensemble approach constitutes a novelty.
Joshi et al. [31] introduced an innovative approach to emotion recognition using EEG, specifically emphasizing subject-independent classification. An ensemble strategy was implemented, incorporating DEAP, SEED, and IDEA datasets characterized by unique technical aspects to fortify the emotion recognition system's resilience. Deep recurrent neural network (D-RNN) models were trained using derived features, including power spectral density (PSD), Hjorth parameters, differential entropy (DE), and linear formulation of differential entropy (LF-DE). In particular, bidirectional long short-term memory (BiLSTM) networks was trained using these features. The accuracy of the proposed method exhibited substantial enhancements; specifically, it increased by 8.2% when evaluated on SEED subsequent to its training on DEAP, and by 1.5% when the procedure was reversed. Notwithstanding these progressions, the research acknowledges certain constraints, including the requirement for vast datasets and possible prejudices stemming from cultural and technical variances among datasets. Nevertheless, the research acknowledges certain constraints, including the requirement for large datasets and possible prejudices arising from cultural and technical variances among datasets. These factors could have an impact on the model's ability to generalize and its efficacy in practical situations.

Guo et al. [39] introduced a framework for ensemble learning-based convolutional neural network feature integration in EEG-based emotion classification. They produced competitive results in emotion classification tasks by investigating feature extraction from synchronization likelihood matrices and correlation coefficient matrices of EEG signals. The proposed CNN-EL method showcased promising accuracy, with an accuracy of approximately 72.34%. Tripathi et al. [40] conducted DEAP data emotion classification experiments using deep and convolutional neural networks. Valence and Arousal categorization accuracy was superior to earlier studies. The top Deep Neural Network (DNN) model classified Valence into 2 classes with 75.78% accuracy. The study's uniqueness resides in neural networks' capacity to classify emotions using EEG signals, although dataset characteristics and generalizability were not addressed.
Chao et al. [41] introduced the Capsule Network (CapsNet) framework as a means of emotion recognition utilizing multiband EEG signals. A Multiband Feature Matrix (MFM) was con-

structed by combining frequency domain, spatial characteristics, and frequency band characteristics of EEG signals. CapsNet was then employed to perform emotion recognition. The superior performance of the CapsNet model in identifying emotions compared to prevalent models such as CNN, RDF, SVM, and k-NN demonstrates CapsNet's efficacy in utilizing EEG signal characteristics for emotion recognition.

Rahman et al. [42] presented non-linear features and ensemble learning for EEG-based emotion analysis. The DEAP and AMIGOS datasets compared classifiers and features for predicting six fundamental emotions. Higuchi Fractal Dimension, Sample Entropy, and Permutation Entropy were derived from EEG data. The suggested technique outperformed previous methods with 89.38% accuracy using thirty-two electrodes and 88.62% utilizing adaptive electrodes. Ensemble non-linear feature-based emotion detection approaches outperform others in the article.

Luo et al. [43] used spiking neural networks (SNN) to classify EEG emotions. Three feature extraction approaches were examined using DEAP and SEED datasets. The variance method was the most accurate. The SNN classified emotion states with 78% accuracy for valence, 74% for arousal, 80% for dominance, and 86.27% for liking. SNN successfully processed time series data for EEG-based emotion classifications in the article.

Sidharth et al. [44] developed a transfer learning-based emotion identification approach for EEG employing MPC, MSC, and DE characteristics. The research outperformed chance in subject-dependent (93.1%) and subject-independent (71.6%) classifications. The research discusses MPC and MSC's potential in neuroprosthetics and EEG-based emotion detection.

Although recent research on emotion recognition based on EEG has yielded encouraging outcomes, a number of limitations remain. These concerns encompass the possibility of overfitting and a dearth of generalizability resulting from the reliance on restricted datasets without external validation. Deciding which models should be utilized for real-world applications presents difficulties, given that personalized models might not be feasible. The reliability and consistency of the results are impacted by the variability observed in preparatory techniques and feature selection methods. Additionally, the absence of universally accepted performance metrics and evaluation methodologies poses a challenge in accurately comparing diverse approaches. In order to progress the discipline and enhance the dependability and strength of emotion recognition systems based on EEG, it will be imperative to surmount these constraints.

3.3 Conclusion

The literature review here highlights the increasing attention and progress in emotion recognition from EEG signals through the implementation of machine learning and deep learning methodologies. Scholars have investigated a range of approaches, such as ensemble learning, feature extraction, and neural network architectures, in an effort to enhance the precision and dependability of emotion classification systems. Several studies have exhibited encouraging outcomes; however, obstacles such as dataset limitations and models' generalizability still need to be resolved. Subsequent investigations ought to concentrate on rectifying these obstacles and augmenting the efficacy of emotion recognition systems based on EEG for pragmatic implementations.

Chapter 4

Research Workflow & Implementation

4.1 Introduction

Emotion recognition through EEG signals stands at the forefront of understanding human affective states, a field critical for applications ranging from mental health diagnostics to human-computer interaction. Profound progress in interpreting the intricacies of human emotions has resulted from the integration of machine learning and deep learning methodologies. The development of models with robust generalization capabilities is made possible by extracting meaningful features from EEG data, which is crucial to these trends. A thorough exposition of the research methodology is provided in this chapter, with particular emphasis on the methods employed for feature extraction, model development, and evaluation. With the ability to precisely categorize emotional states, the objective is to develop emotion recognition systems that are both dependable and efficient.

4.2 Proposed Workflow

The preliminary phase of the proposed research methodology, as delineated in Figure 4.1, entails the utilization of EEG recording devices to obtain EEG data from the participants. The data is then segmented into 1-second epochs to facilitate comprehensive analysis. Considering each epoch as a unique temporal window of EEG activity makes it possible to analyze brainwave patterns linked to different emotional states. After segmenting the data, characteristics are extracted from the EEG epochs. The feature extraction process involved computing Power Spectral Density (PSD), Differential Entropy (DE), and statistical features such as mean, variance, standard deviation, and kurtosis. Furthermore, frequency regions are extracted, encompassing alpha, beta, delta, theta, and gamma waves.

After the feature extraction process, four machine learning classifiers were trained: Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbours (KNN), and Random Forest (RF). The classifiers were trained to categorize EEG data into three emotional states: positive, neutral, or negative for the SEED dataset and valence and arousal levels for the DEAP dataset. Furthermore, a Convolutional Neural Network (CNN) architecture was constructed, alongside the machine learning classifiers. The CNN model used differential entropy (DE) as a feature and underwent training to categorize emotions directly based on the EEG data. In addition, a hybrid CNN-BiLSTM model was created to investigate the possibility of merging convolutional and recurrent neural networks. This model used a combination of Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) layers. Its purpose was to categorize emotional states by analyzing the EEG data.

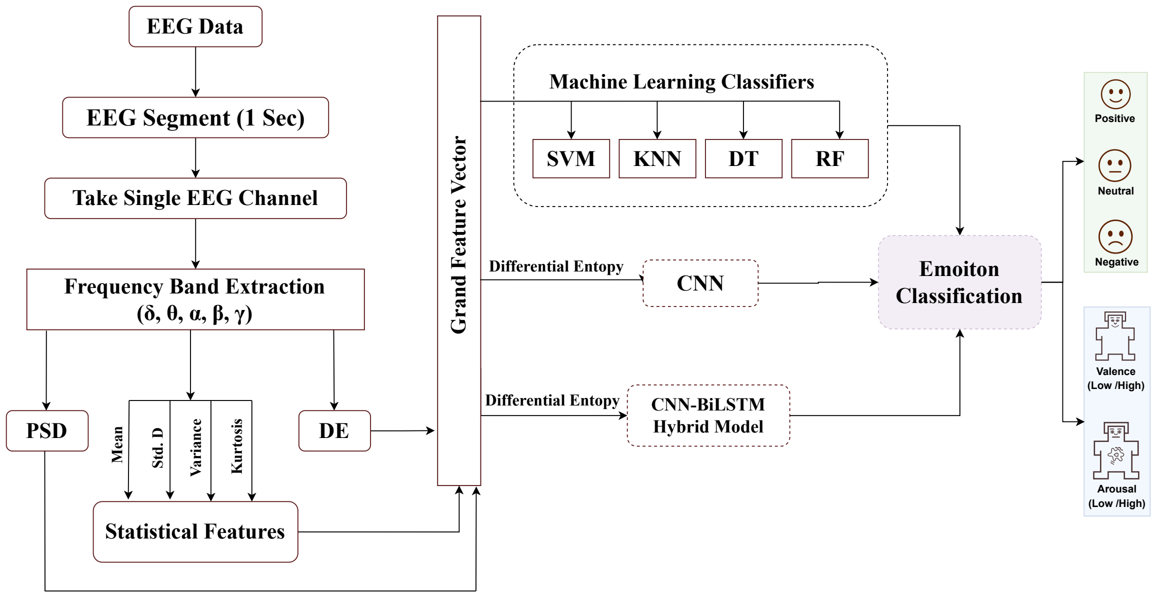


Figure 4.1: The proposed research workflow

The performance of each classifier is assessed using metrics including accuracy, precision, recall, and F1-score, after the training phase. Differential entropy (DE) is employed as a criterion in the model selection process. The model that demonstrates the most optimal performance in terms of DE is designated as the optimal classifier for classifying emotional states using EEG. An analysis of the outcomes is conducted to understand the efficacy of various features and classifiers in precisely classifying EEG data according to affective states. This analysis contributes to the development of robust emotion recognition systems using EEG data, providing valuable insights into selecting features and classifiers for accurate emotional state classification.

4.3 Dataset Description

4.3.1 SEED (SJTU Emotion EEG Dataset)

The SEED dataset [45], organized by the BCMI [46] laboratory under the leadership of Prof. Bao-Liang [47] Lu and Prof. Wei-Long Zheng [48], serves as a valuable resource for studying emotion and vigilance using EEG signals. The experimental design of SEED is highly detailed, with stimuli meticulously chosen to elicit precise emotional reactions. 15 Chinese film clips were selected, each designed to evoke positive, neutral, or negative emotions. The clips were carefully chosen to ensure that they could be understood without any additional explanation and that they would elicit a specific emotional response. These edited clips are about 4 minutes long and are designed to be coherent and emotionally impactful. The experimental procedure involves conducting 15 trials per session. Each trial is preceded by a 5-second hint, followed by periods for self-assessment and rest. (Figure 4.2)

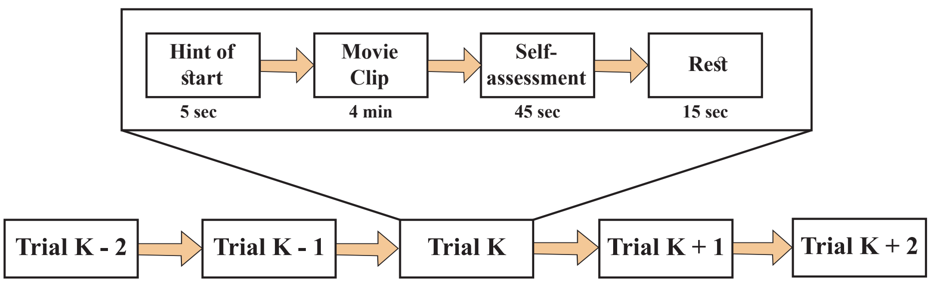


Figure 4.2: Experimental protocol for movie clip presentation of the SEED dataset

SEED's data collection process involves capturing both EEG signals and eye movements. By using the 62-channel (Figure 4.3) ESI NeuroScan System [49] and SMI eye-tracking glasses, the dataset captures a wide range of neurophysiological responses. A group of fifteen individuals from China, including males and females, participated in the experiments. Each participant repeated the experiment three times over about one week.

The key characteristics of the SEED dataset are summarized in Table 4.1

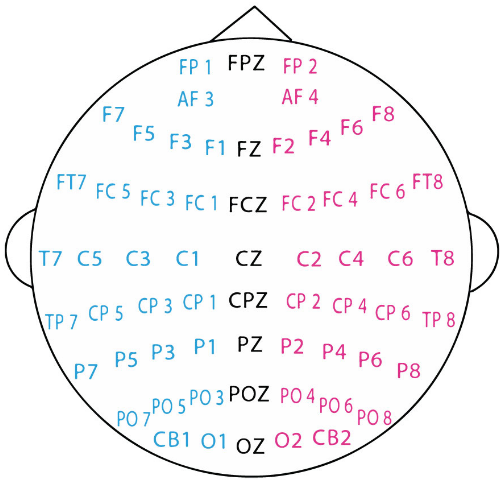


Figure 4.3: 62-channel electrode placement in The ESI NeuroScan System - SEED dataset

Table 4.1: SEED Dataset Characteristics

Attribute	**Value**
Subject	15
Duration of EEG	4 min
Sessions	3
Trials per Session	15
Total Trials	675
Number of Channels	62
Sample Frequency	1000 Hz
Down-Sampled Frequency	200 Hz
Bandpass Frequency	0-75 Hz
Labels	Positive, Neutral, Negative

4.3.2 DEAP Dataset

The DEAP dataset [50, 51] offers a valuable resource for investigating human emotional states using EEG, physiological signals, and video recordings. A group of thirty-two individuals participated in a study where they viewed 40 one-minute segments of music videos. Their brain activity (measured by electroencephalogram or EEG) and peripheral physiological signals were monitored and recorded throughout the experiment. Each participant provided ratings for the videos, evaluating their emotional impact, personal preference, perceived control, and level of familiarity. In addition, videos of the participants' faces were recorded from the front. The dataset is a valuable resource for researchers in the field, as it allows for exploring affective states using multimodal data.

The DEAP dataset was thoroughly collected, incorporating various modalities. Participants observed a subset of 40 one-minute music video excerpts while their EEG and physiological signals were recorded. Additionally, videos of participants' frontal faces were recorded. The stimuli selection process involved utilizing effective tags from the last.fm website, video highlight detection, and an online assessment tool. Participants were asked to rate their arousal levels, valence, like/dislike, dominance, and familiarity using SAM mannequins on a 9-point scale. (See in Chapter 2, Figure 2.2) This publicly available dataset encourages researchers to utilize it to test affective state estimation methods. The key characteristics of the SEED dataset are summarized in Table 4.2

Table 4.2: DEAP Dataset Characteristics

Attribute	**Value**
Subject	32
Duration of EEG	1 min
Session	1
Trials	40
Channel no	32
Sample Frequency	512 Hz
Down-Sampled Frequency	128 Hz
Bandpass Filter	4-45 Hz
Label	Arousal, Valence and Dominance

4.4 Dataset Preprocessing & Feature Extraction

In the preprocessing phase of the SEED dataset, only epochs corresponding to subjects' elicited emotions were selected. Raw EEG data underwent downsampling to 200 Hz, and recordings significantly affected by EMG and EOG contamination were manually eliminated. The identification of blink artifacts was facilitated through EOG recordings. A bandpass filter, ranging from 0.3 Hz to 50 Hz, was applied to rectify anomalies and filter noise. EEG segments aligning with the duration of each movie were extracted, and each channel was subdivided into non-overlapping 1-second epochs.

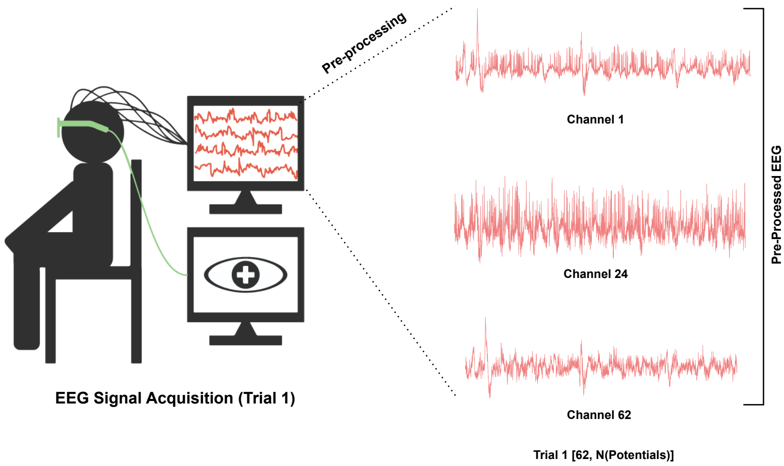


Figure 4.4: EEG acquisition & preprocess

The preprocessing steps for the DEAP dataset mainly focused on formatting and manipulating the data to ensure it was ready for analysis. The dataset was made available in downsampled and preprocessed versions, in both Matlab (.mat) and Python/Numpy (.dat) formats. After converting the data format, various preprocessing steps were implemented to guarantee the reliability and practicality of the EEG data. It is worth mentioning that the data was downsampled to a frequency of 128 Hz to minimize the computational burden. The dataset systematically removed EOG artifacts, ensuring noise elimination and signal integrity preservation by applying a bandpass frequency filter ranging from 4.0 to 45.0 Hz. In addition, the data was normalized by averaging to a common reference and rearranging the EEG channels based on a predefined layout. In addition, the data was divided into 60-second trials, where a 3-second pre-trial baseline was eliminated to concentrate on the significant neural activity during each experimental trial. More details about the preprocessing steps can be found in [45, 50].

A crucial step in EEG-based emotion recognition systems, feature extraction attempts to effectively map EEG segments to their corresponding emotional states by deriving salient features from the data. Three categories of features—Power Spectral Density (PSD), Differential Entropy (DE), and statistical features (Mean, Variance, Standard Deviation, Kurtosis)—were extracted to evaluate their performance in classifying affective states in this study. A distinct understanding of the neural activity that underlies emotional responses can be obtained from each feature type. Figure 4.5 shows the feature extraction flowchart for the SEED dataset.

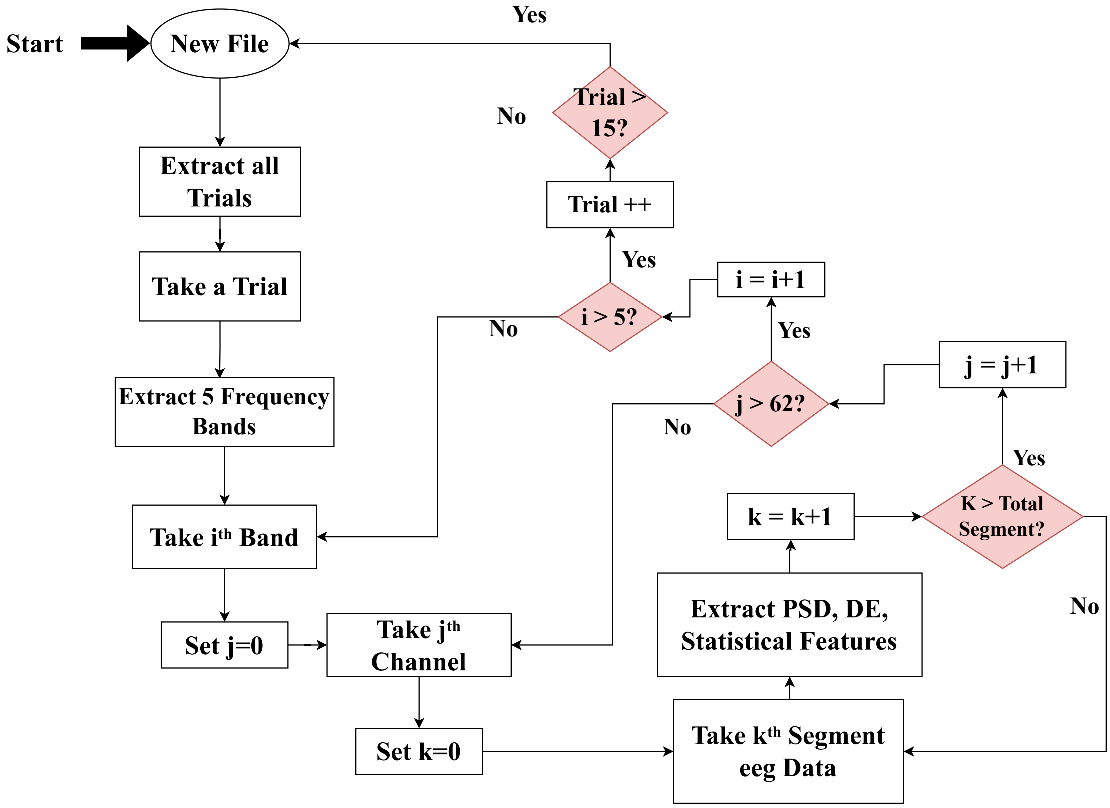


Figure 4.5: Feature extraction flowchart for the SEED dataset

1. **Segmentation**: The EEG data is segmented into n segments, and frequency bands are extracted.

2. **Channel Looping**: For each EEG channel:

    - **Segment Looping**: It iterates through each segment & frequency bands of the data.

    - **Feature Extraction**: For each segment:

        - **Power Spectral Density (PSD)**: Calculating the PSD of the segment. (See Section 2.4.5, Chapter 2)

– **Differential Entropy (DE)**: Computing the DE of the segment. (See Section 2.4.5, Chapter 2)

– **Statistical Features**: Calculating statistical properties such as mean, variance, standard deviation, and kurtosis of the segment. (See Section 2.4.5, Chapter 2)

3. **Aggregate Features**: The features extracted from each segment are aggregated to form feature vectors representing the entire EEG data. (Figure 4.6)

The emotion classifications in the SEED dataset are classified into three distinct categories: positive, neutral, and negative, each represented by the values -1, 0, and 1, respectively. DEAP dataset participants self-evaluated their arousal, dominance, and valence after each experiment. Participants might choose positions below or between 1 and 9 to establish a continuous scoring system. Thus, every participant gave quantitative values from one to nine. Thresives produce class IDs. Having more categories helps better describe feelings but makes emotion detection harder. The prior study used label-processing systems based on 1–5 and 5–9 self-assessment ratings. Binary classification tasks for high arousal (HA), low arousal (LA), high valence (HV), and low valence (LV) were used to recognize emotions [41]. The emotion identification task was split into valence and arousal binary classification tasks. Individuals with ratings below 5 were labeled “low”; “high” was the default. There were four identifiers: HA, LA, HV, and LV.

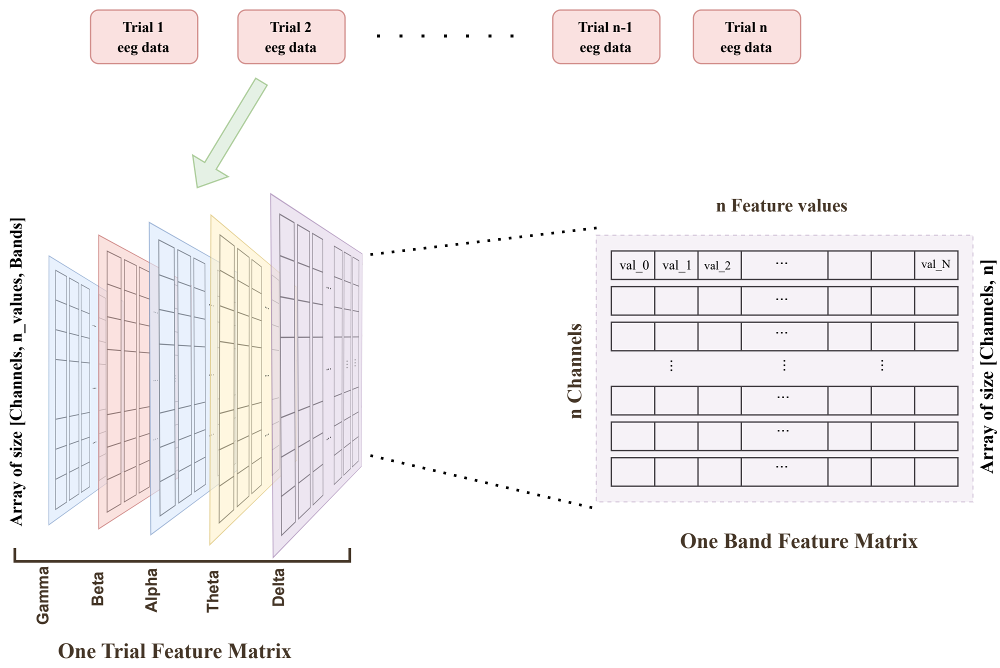


Figure 4.6: Structure of extracted feature matrix

After extracting the features, a moving average filter was applied to smooth the feature values [2, 5, 4]. Figure 4.7 illustrates the EEG data from the SEED dataset (channel 3, trial 1) after the feature extraction and smoothing process.

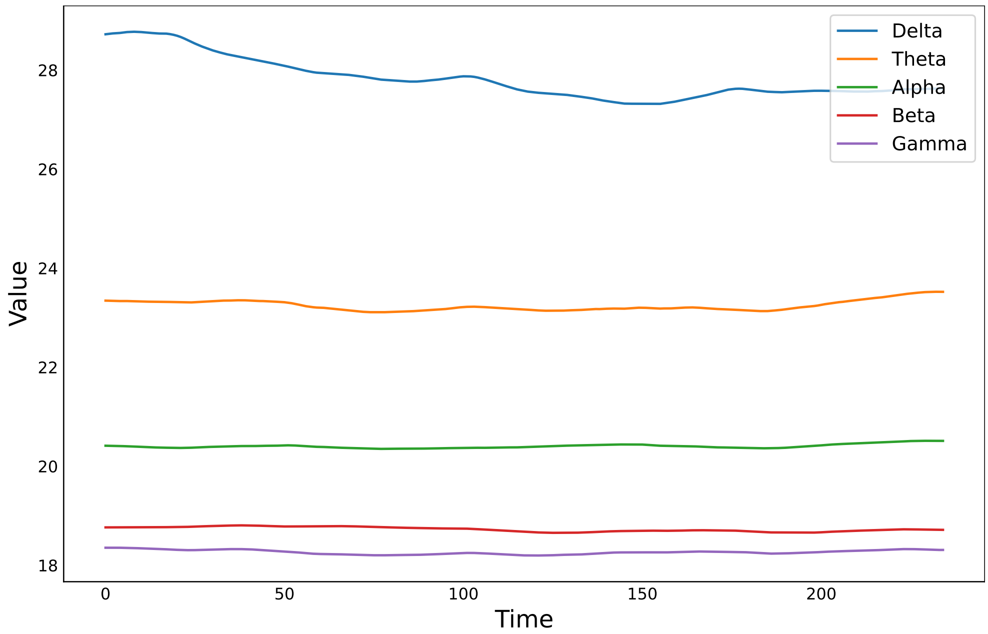


Figure 4.7: SEED EEG data (channel 3, trial 1) after feature extraction and smoothing

The Table 4.3 provides the feature map sizes for both the SEED and DEAP datasets. Each dataset contains three types of features: Differential Entropy (DE), Power Spectral Density (PSD), and Statistical features. The shape represents the dimensions of the feature maps as (trials, channels, feature values, and frequency bands).

Table 4.3: Feature Map Sizes for SEED and DEAP Datasets

Dataset	**Feature Type**	**Shape**
SEED	DE	(675, 62, 265, 5)
	PSD	(675, 62, 184, 5)
	Statistical	(675, 62, 736, 5)
DEAP	DE	(1280, 32, 42, 4)
	PSD	(1280, 32, 42, 4)
	Statistical	(1280, 32, 168, 4)

4.5 Implementations

4.5.1 Machine Learning-Based Classification

Machine learning models offer a versatile approach to classifying EEG data by identifying and utilizing intrinsic patterns that enable the prediction of emotional states. These models can identify relationships between characteristics and emotional states by employing statistical measures and features extracted from EEG signals. A variety of algorithms were trained and evaluated in this study, including Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbours (KNN), and Random Forest (RF). Significant findings were obtained concerning the effectiveness of machine learning classifiers in categorizing EEG data as affective states through the implementation of this methodology. Furthermore, the research assessed the efficacy of these classifiers by utilizing statistical measures (mean, variance, standard deviation, kurtosis), Differential Entropy (DE), and Power Spectral Density (PSD) in order to classify emotions from the SEED and DEAP datasets.

4.5.2 CNN-Based Model

Convolutional Neural Networks (CNNs) have shown considerable promise in analyzing EEG data for emotion recognition tasks. Their ability to automatically learn spatial and temporal patterns from raw data makes them well-suited. [38] The Convolutional Neural Network (CNN) illustrated in Figure 4.8 is designed to classify EEG signals based on affective states, utilizing the SEED dataset. The CNN model utilizes Differential Entropy (DE) features as input, which were selected due to their capability of capturing the temporal dynamics of EEG signals. DE offers vital insights into the underlying neural activity by quantifying the signal's uncertainty or randomness. By integrating DE features, CNN can leverage this temporal information more efficiently to differentiate between various emotional states.

The CNN architecture begins with three convolutional layers, each equipped with 64 filters of size 5x5 and 3x3, to extract spatial features from the EEG data. The convolutional layers effectively capture patterns at various levels of granularity, which is crucial for distinguishing minute details within the EEG signals. The following blocks incorporate convolutional layers with 128, 256, and 512 filters, thereby augmenting the model's capacity to acquire hierarchical data representations. The network can efficiently capture spatial hierarchies by selecting convolutional filter sizes of 5x5 and 3x3. CNN can learn fine-grained and coarse-grained features by employing larger filters to capture more global patterns and concentrate on local features.

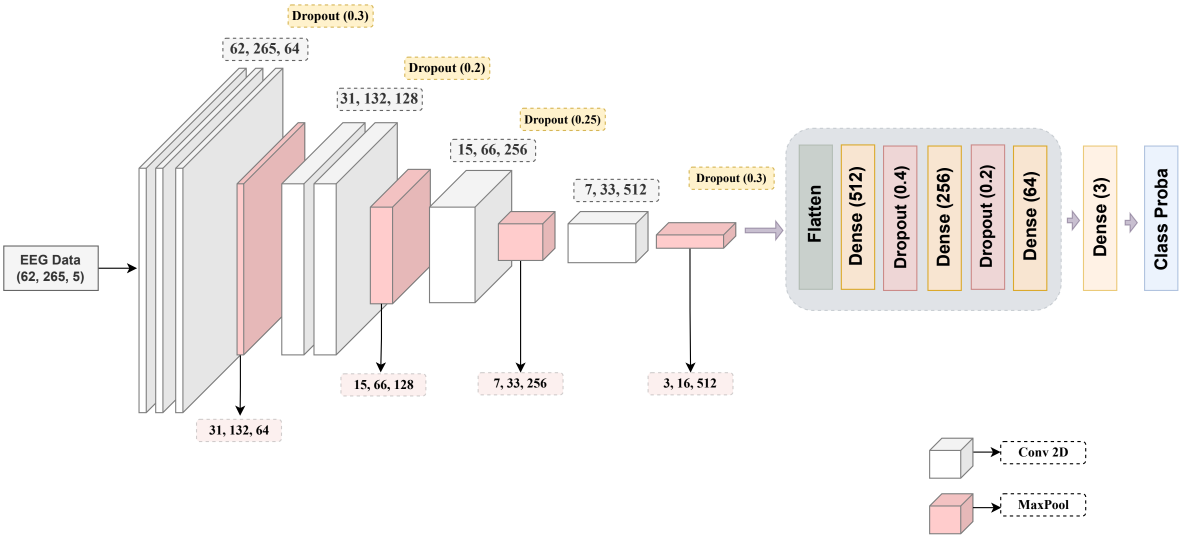


Figure 4.8: CNN-based model architecture employed in the SEED Dataset

After max-pooling layers, dropout layers with rates of 0.3, 0.2, and 0.25 are implemented to prevent overfitting by deactivating neurons randomly during training. By employing max-pooling layers with a pool size of 2x2, salient information is maintained while the spatial dimensions of the feature maps are reduced in subsequent convolutional blocks. Following dense layers comprising 512, 256, and 64 neurons, the network concludes with a softmax output layer comprising three neurons that symbolize the following affective states: positive, neutral, and negative. Multiclass classification is facilitated by the softmax activation, which distributes the emotion categories according to a probability distribution.

A comparable CNN architecture classified EEG signals by valence and arousal for the DEAP dataset. The first convolutional block has three layers with 32, 32, and 32 filters and 5x5, 4x4, and 3x3 kernels to capture spatial characteristics. After each convolutional layer, 2x2 max-pooling and 0.3 dropout layers prevented overfitting. Each block reduced the number of filters until the penultimate block had 128. The final block has dense layers with 512, 256, and 64 neurons and a softmax output layer with valence (Low or High) and arousal (Low or High) neurons. EEG hierarchical characteristics were retrieved using this CNN architecture. The design remained stable across EEG datasets despite input form and parameters, proving its adaptability.

After the convolutional blocks, fully connected dense layers incorporate high-level features learned by convolutional layers, transforming spatial hierarchies into a flattened feature vector.

The ReLU activation function is selected due to its backpropagation efficacy and non-linearity, which effectively addresses the issue of vanishing gradients. Implementing dropout regularisation following each layer mitigates the risk of overfitting and improves the model's generalization ability. The convolution of these layers strategically enables the CNN to efficiently derive significant hierarchical features from the input EEG data, resulting in enhanced efficacy in emotion recognition tasks.

4.5.3 CNN-BiLSTM Based Hybrid Model

Convolutional neural networks (CNNs) are exceptionally adept at discerning spatial patterns in data, especially in the case of time series and images, which present grid-like structures [28]. Convolutional neural networks (CNNs) have demonstrated significant effectiveness in EEG data analysis when obtaining hierarchical depictions of spatial attributes, including complex frequency and amplitude patterns distinguished among various brain regions. Using numerous convolutional layers, convolutional neural networks (CNNs) independently generate abstract representations of the input data as they increase in complexity.

Alternatively, Bidirectional Long Short-Term Memory (BiLSTM) networks encode temporal dependencies inherent to sequential data with extraordinary proficiency. EEG signals, which represent the complex patterns of electrical activity that occur gradually, are effectively deciphered by these networks in terms of temporal relationships. BiLSTM networks effectively discriminate minute details of transient oscillations and the fundamental long-term interdependencies inherent in EEG signals through the seamless integration of forward and backward processing of input sequences.

The CNN-BiLSTM hybrid model (Figure 4.9) utilized for the SEED dataset effectively extracts spatial and temporal features from EEG data for emotion recognition by synergistically combining convolutional and recurrent layers. Commencing with a convolutional block consisting of three layers, each containing 64 filters with a dimension of 5x5, the model incrementally progresses by utilizing two further convolutional layers containing 128 and 256 filters, respectively. After each convolutional layer, feature maps are downsampled using max-pooling operations, and dropout layers are incorporated to mitigate the tendency toward overfitting. The reshaped features are then inputted into a Bidirectional Long Short-Term Memory (BiLSTM) layer, which comprises numerous units distributed across different layers: two have 64 units, two have 128 units, one has 256 units, and 512 units. The LSTM's bidirectional characteristic permits the detection and representation of temporal dependencies in both the forward and backward orientations.

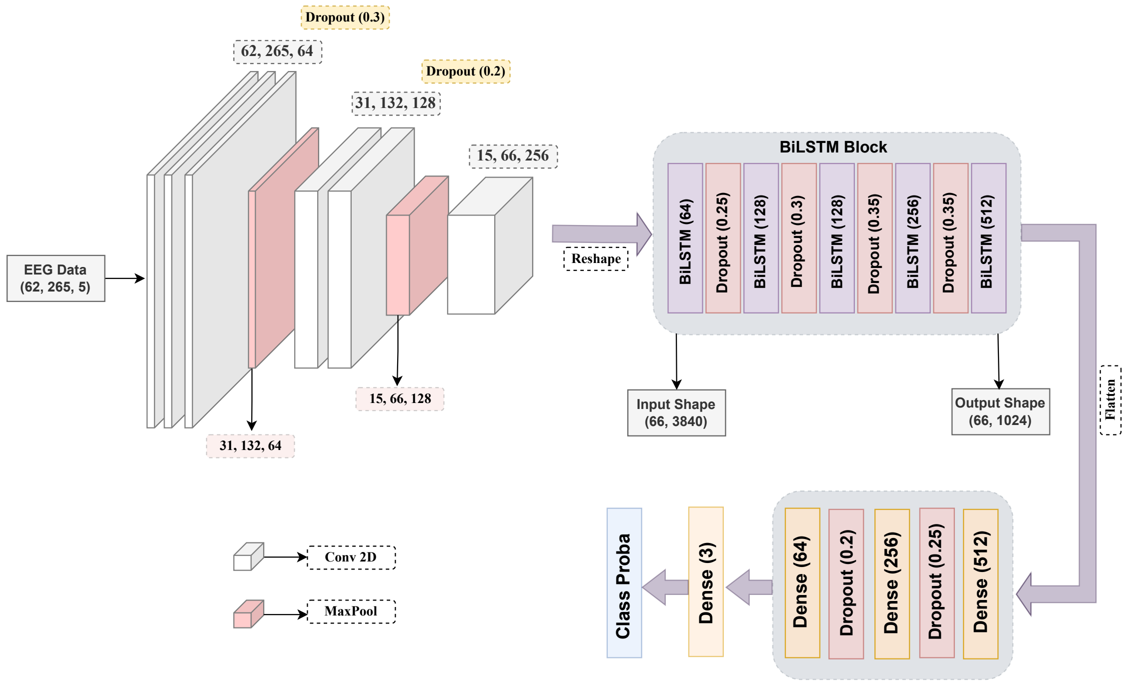


Figure 4.9: CNN-BiLSTM based model architecture employed in the SEED Dataset

Following the LSTM layers, the output is flattened and passed through dense layers with entirely connected connections, consisting of 512, 256, and 64 neurons in that order. Following this, a softmax output layer comprising three neurons symbolizes the following emotional states: positive, neutral, and negative. Rectified Linear Unit (ReLU) activation functions are implemented to incorporate non-linearity in the dense layers, while dropout regularisation is utilized to prevent overfitting. The CNN-BiLSTM hybrid model effectively incorporates temporal and spatial information extracted from EEG signals, thereby providing a resilient structure for classifying emotions in the SEED dataset.

Minor modifications are made to the CNN-BiLSTM architecture of the DEAP dataset in comparison to the SEED dataset to account for variations in input shape and final dense unit. The model's predictive capability is encapsulated in the final dense layer, which provides discrete valence and arousal level classifications. The CNN-BiLSTM framework offers a resilient structure for emotion recognition tasks on the datasets by synthesizing spatial and temporal information extracted from the EEG data. The accuracy and generalizability of emotion recognition tasks can be significantly improved through the meticulous architecture design, which enables the identification of intricate patterns within EEG data.

4.5.4 Training, Testing & Validation Set

The data partitioning process for the SEED dataset was as follows: the initial 550 samples were designated as the training set, samples 550 to 599 comprised the validation set, and samples 600 to 674 comprised the final testing set. The identical division was implemented for the corresponding designations, ensuring uniformity throughout the dataset. Similarly, the DEAP dataset was partitioned into three sets: training, validation, and testing. The initial 1080 samples were included in the training set, whereas samples 1080 to 1179 were set aside for validation. The last samples comprised the concluding testing batch, ranging in value from 1180 to 1280. Labels were partitioned once more in conjunction with the data. The research process was conducted methodically, always adhering to this approach. This ensured that the models underwent training on representative data, and additionally, distinct datasets were made available for testing and ultimate evaluation.

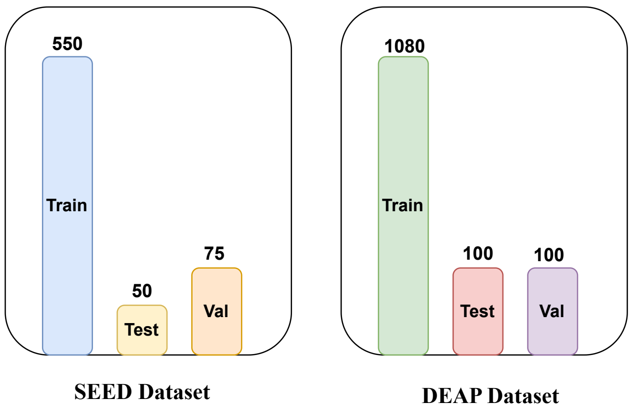


Figure 4.10: Train, test & validation sizes of SEED & DEAP dataset

4.5.5 Hyperparameter Tuning with GridSearchCV

Hyperparameter tuning is crucial in achieving optimal performance for machine learning models. A systematically investigated investigation was conducted using GridSearchCV [52] to identify the most effective combinations of hyperparameters for all the models, including CNN, machine learning algorithms, and the CNN-BiLSTM hybrid model. To investigate the effect that hyperparameter variation within predetermined ranges had on model performance, these parameters were selected with consideration. For the CNN, adjustments were made to learning rates, dropout rates, kernel sizes, and the number of filters. For the machine learning models, hyperparameter-tuning was performed on regularisation, kernel type, number of estimators, maximum depth, and number of neighbors. The CNN-BiLSTM hybrid model was configured with the following hyperparameters: learning rates, dropout rates, kernel sizes, dense layer units, and LSTM units. Each model was rigorously assessed by GridSearchCV using an extensive variety of hyperparameter combinations. To optimize performance, configurations were chosen that simultaneously enhanced accuracy and generalization. The outcomes of hyperparameter optimization yielded significant insights regarding the optimal configurations for every model, thereby improving their overall performance in tasks involving emotion recognition. Employing this methodical strategy for hyperparameter optimization enhanced the models' dependability and efficacy, rendering them more appropriate for practical implementations.

4.6 Conclusion

In conclusion, this chapter has provided an in-depth study methodology for developing EEG-based emotion recognition systems. Significant progress in categorizing human affective states has been achieved by implementing sophisticated machine learning and deep learning methodologies. By effectively deriving significant characteristics from EEG signals and constructing resilient models, it has become possible to develop systems that exhibit exceptional accuracy and generalization capabilities when applied to a wide range of datasets. Employing this methodological framework, not only facilitates the progression of affective computing as a discipline, but it also presents opportunities for the development of practical applications in domains including neuromarketing, artificial intelligence, and healthcare.

Chapter 5

Result & Performance Analysis

5.1 Introduction

EEG signal-based emotion recognition is an intricate and demanding field of study that has wide-ranging implications, including but not limited to healthcare and human-computer interaction. Machine learning methodologies have exhibited encouraging outcomes in recent times when applied to the analysis of EEG data for the purpose of identifying and categorizing emotional states. In this chapter, the efficacy of diverse deep learning architectures and machine learning models in the domain of emotion recognition is investigated by analyzing EEG data extracted from two datasets: SEED and DEAP. Statistical features, power spectral density (PSD), and differential entropy (DE) were employed in the SEED dataset to train machine learning models, including k-nearest neighbors (KNN), support vector machines (SVM), and random forests (RF). Furthermore, an evaluation was conducted on the efficacy of convolutional neural network (CNN) and CNN-BiLSTM models in identifying emotions based on EEG signals. Likewise, to compare performance across datasets, the identical models and feature sets were applied to the DEAP dataset.

5.2 Experimental Setup

In both the SEED and DEAP datasets, grid search cross-validation was utilized to refine the hyperparameters of the machine learning models. All three feature sets—Differential Entropy (DE), Power Spectral Density (PSD), and Statistical Features—were subjected to this tuning procedure. To determine the optimal configuration, different tuning ranges were chosen for each hyperparameter. Various hyperparameter combinations were employed to train and assess the models to determine which configurations produced the best results. By implementing this methodology, it was guaranteed that the models would be optimized to attain optimal levels of precision and generalizability when it came to assignments involving the classification of emotions. The optimal parameters acquired for each classifier and feature set via the tuning procedure are displayed in the Table 5.1.

Table 5.1: Optimized Hyperparameters for Machine Learning Models for Each Feature

Dataset	**Feature Set**	**Model**	**Best Parameters**
SEED	DE	RF	n_estimators=6, max_features=sqrt
		SVM	kernel=linear, C=10
		DT	criterion=entropy, min_samples_split=2
		k-NN	n_neighbors=15, metric=euclidean
	PSD	RF	n_estimators=200, max_depth=None
		SVM	kernel=linear, C=10
		DT	criterion=gini, min_samples_split=2
		k-NN	n_neighbors=15, metric=euclidean
	Statistical	RF	n_estimators=200, max_features=sqrt
		SVM	kernel=linear, C=100
		DT	criterion=gini, min_samples_split=2
		k-NN	n_neighbors=3, weights=distance
DEAP	DE	RF	n_estimators=400, max_depth=10
		SVM	kernel=linear, C=0.5
		DT	criterion=gini, min_samples_split=5
		k-NN	n_neighbors=11, weights=distance
	PSD	RF	n_estimators=200, max_depth=40
		SVM	kernel=rbf, C=100
		DT	criterion=entropy, min_samples_split=2
		k-NN	n_neighbors=11, weights=distance
	Statistical	RF	n_estimators=500, max_depth=None
		SVM	kernel=linear, C=10
		DT	criterion=entropy, min_samples_split=5
		k-NN	n_neighbors=3, weights=uniform

The effectiveness of Convolutional Neural Network (CNN) and CNN-BiLSTM models was assessed in the investigation of enhanced emotion recognition. The Differential Entropy (DE) feature was utilized from the SEED and DEAP datasets. The CNN model, compiled with the Adam optimizer and a learning rate of 1×10^{-4} for SEED and 3×10^{-5} for DEAP, underwent 100 epochs of training with a batch size of 32, using categorical cross-entropy loss. Similarly, the CNN-BiLSTM model, compiled with the Adam optimizer and a learning rate of 1×10^{-4}, underwent the same training regimen. By meticulously configuring these models and fine-tuning their parameters, efforts were made to develop robust architectures for accurate emotion recognition from EEG data, contributing to the advancement of emotion recognition technology.

Table 5.2: Experimental Setup for CNN and CNN-BiLSTM Models for Differential Entropy Feature

Model	Dataset	Optimizer	Learning Rate	Loss Function	Metrics	Epochs	Batch Size
CNN	SEED	Adam	4×10^{-5}	Categorical Crossentropy	Accuracy, RMSE	100	64
CNN	DEAP	Adam	3×10^{-5}	Categorical Crossentropy	Accuracy, RMSE	100	32
CNN-BiLSTM	SEED & DEAP	Adam	3×10^{-4}	Categorical Crossentropy	Accuracy, RMSE	100	32

5.3 Evaluation Metrics

When evaluating models, a variety of metrics are commonly used, including the confusion matrix, Precision, Recall, Accuracy, and F1-score. Before discussing these metrics, it is crucial to acknowledge the importance of EEG emotion recognition and the fundamental significance of machine learning, particularly deep learning, in feature extraction and model generalization. This paper explores the role of these evaluation metrics in assessing the effectiveness of EEG-based emotion recognition models.

5.3.1 Confusion Matrix

The confusion matrix is a vital tool for evaluating binary classification models, providing a concise overview of the model's predictions compared to the actual data labels. It consists of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). From these values, important performance metrics such as Precision, Recall, Accuracy, and F1-score can be computed, enabling a comprehensive assessment of the model's ability to distinguish between classes.

5.3.2 Accuracy

Accuracy is crucial for assessing machine learning algorithms' performance, particularly in classification tasks. It measures the proportion of correct predictions compared to the total number of predictions. While accuracy offers a basic performance measure, it may not provide a complete picture, especially with imbalanced datasets or varying false positives and negative costs. Therefore, cross-validation and hold-out testing are essential to evaluate the model's generalization capability and avoid overfitting.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.1}$$

5.3.3 Precision

Precision evaluates a binary classification model's ability to identify positive instances among all instances predicted as positive correctly. A high precision score indicates fewer false positive predictions and greater reliability in identifying positive instances correctly. Conversely, a low precision score suggests more false positive predictions and less reliability in distinguishing between positive and negative instances.

$$Precision = \frac{TP}{TP + FP} \tag{5.2}$$

5.3.4 Recall

Recall, also known as sensitivity, measures the model's ability to identify positive instances out of all actual positive instances correctly. A high recall value indicates proficiency in capturing positive instances from the dataset accurately. Conversely, a low recall value suggests missing positive instances and less accuracy in identifying all positive cases.

$$Recall = \frac{TP}{TP + FN} \tag{5.3}$$

5.3.5 F1-Score

The F1-score balances precision and recall, comprehensively evaluating a binary classification model's performance. When precision and recall have similar values, the F1-score is close to 1, indicating a balanced model performance. However, if precision and recall differ significantly, the F1-score will be lower, indicating the model's inability to balance the two metrics.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5.4}$$

In conclusion, the evaluation metrics discussed in this chapter are crucial for assessing the effectiveness of EEG-based emotion recognition models. Understanding these metrics enables researchers and practitioners to make informed decisions about model optimization and improvement strategies, leading to more accurate and reliable emotion recognition systems.

5.4 Performance Evaluation on SEED Dataset

5.4.1 Evaluating Machine Learning Models

EEG-based emotion recognition involves training machine learning models to analyze brain signals to comprehend mental states and emotional responses. These models depend on extracting pertinent features from EEG data, which capture latent patterns that indicate various emotional states. The aforementioned attributes may span from rudimentary statistical metrics to more intricate depictions of signal properties, including frequency content and entropy. In emotion recognition on EEG data, the efficacy of various feature sets—Differential Entropy (DE), Power Spectral Density (PSD), and Statistical features—is critical. Within this particular framework, Differential Entropy (DE) emerges as a highly promising characteristic for emotion recognition on EEG signals, effectively capturing the subtleties of emotional expression. The classification outcomes for machine learning models trained on the SEED dataset utilizing distinct feature sets—namely Differential Entropy (DE), Power Spectral Density (PSD), and Statistical features—are displayed in the table 5.3.

Compared to models employing PSD and Statistical features, those with Differential Entropy (DE) features consistently attained superior precision, recall, and F1-scores. From the options considered, Random Forest (RF) achieved the highest precision (0.84), recall (0.84), and F1-

Table 5.3: Classification Report for Machine Learning Models on SEED Dataset

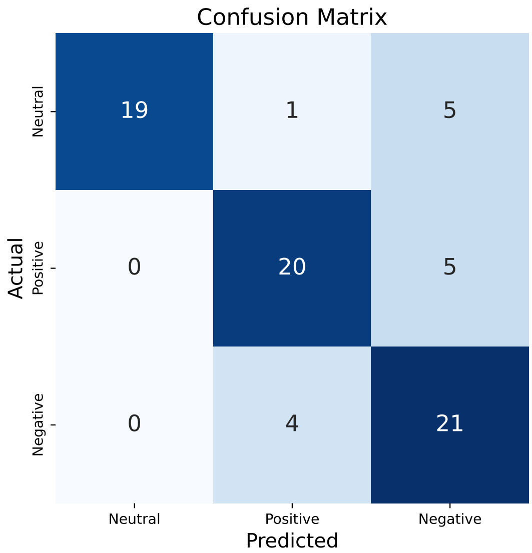
Feature	Model	Avg Precision	Avg Recall	Avg F1-score	Accuracy (%)
DE	DT	0.83	0.80	0.80	80.00
	KNN	0.75	0.73	0.74	73.33
	RF	0.84	0.84	0.84	84.00
	SVM	0.84	0.83	0.83	83.33
PSD	DT	0.45	0.45	0.45	45.00
	KNN	0.52	0.51	0.50	51.33
	RF	0.65	0.65	0.65	65.33
	SVM	0.45	0.40	0.37	40.00
Statistical	DT	0.32	0.32	0.32	32.00
	KNN	0.49	0.44	0.45	44.00
	RF	0.55	0.52	0.53	52.00
	SVM	0.49	0.44	0.42	44.00

score (0.84), all of which suggest that it exhibited a strong capability in the classification of affective states. Achieving precision, recall, and F1-scores ranging from 0.80 to 0.83, the Decision Tree (DT) and Support Vector Machine (SVM) models also performed admirably with DE features.

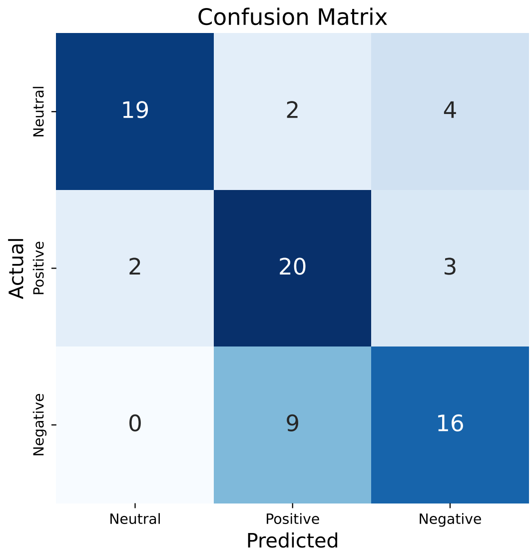
The decision tree, K-nearest Neighbours (KNN), random forest, and support vector machine (SVM) models exhibited diminished precision, recall, and F1-scores compared to models trained on DE-based features when exposed to PSD features. Among models based on PSD, Random Forest attained the highest accuracy of 65%. This finding highlights the difficulties of accurately capturing affective patterns using exclusively PSD features. Analogous to DE-based models, the efficacy of models trained on statistical features was comparatively subpar. With an accuracy of 32%, the Decision Tree model demonstrated the inadequacies of emotion recognition based solely on statistical features.

Differential Entropy emerges as the most effective feature for emotion recognition on the SEED dataset. DE-based models consistently outperformed PSD and Statistical feature-based models, showcasing their ability to capture subtle emotional cues in EEG signals. These results highlight the importance of feature selection in EEG-based emotion recognition, emphasizing the significance of Differential Entropy in accurately classifying emotional states from EEG data.

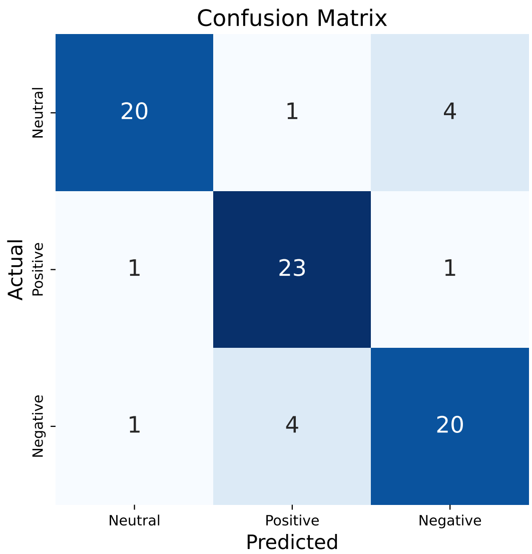
The confusion matrix comprehensively illustrates the classification performance of DE-based models (Figure 5.1). The decision tree models accurately categorized twenty-one instances as negative, nineteen as neutral, and twenty as positive emotions. While the performance of the K-Nearest Neighbours (KNN) algorithm was marginally inferior, it still achieved satisfactory results, accurately classifying 19 neutrals, 20 positives, and 16 negatives. Random Forest (RF) models attained similar accuracy levels, accurately classifying 20 Neutrals, 23 Positives, and 20 Negatives. Support Vector Machine (SVM) models exhibited reliable performance with the correct classification of twenty neutrals, twenty-one positives, and twenty-one negatives.



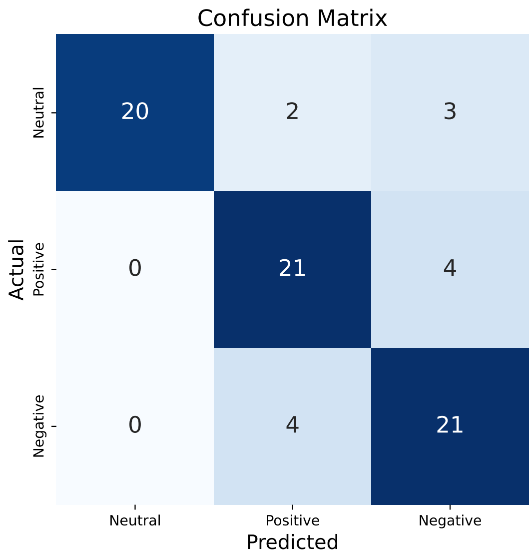
(a) Decision Tree



(b) KNN



(c) Random Forest



(d) SVM

Figure 5.1: Confusion matrices for ML classifiers on Differential Entropy - SEED dataset

5.4.2 Evaluating the CNN Model

The CNN model, trained using Differential Entropy (DE) features and the SEED dataset exhibits remarkable efficacy in classifying affective states. The model demonstrates an overall accuracy of 89%, with Positive and Neutral classes attaining a precision of 100% and Negative classes 88%. This demonstrates its capacity to differentiate among various emotional states effectively. The CNN model accurately classifies 21 out of 25 neutral samples, all 25 positive samples, and 21 out of 25 negative samples, as the confusion matrix indicates (Figure 5.2). This further underscores the model's efficacy in emotion recognition when employing DE features.

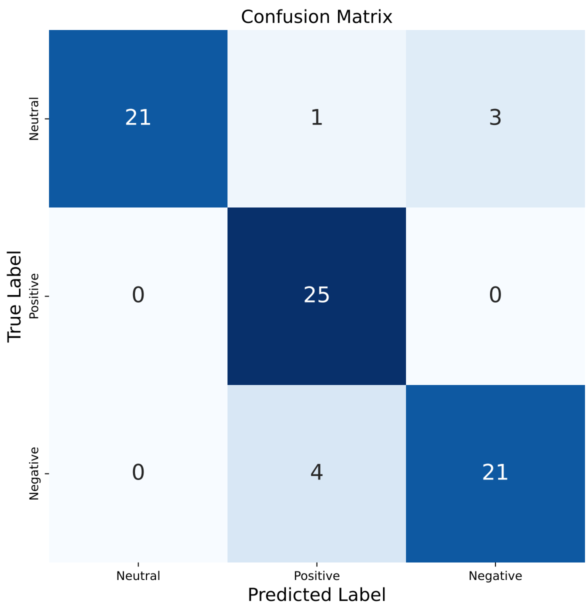


Figure 5.2: Confusion matrix for CNN model on Differential Entropy - SEED dataset

Table 5.4: Classification Report for CNN Model with DE Features on SEED Dataset

Class	**Precision**	**Recall**	**F1-Score**
Neutral	1.00	0.84	0.91
Positive	0.83	1.00	0.91
Negative	0.88	0.84	0.86
Accuracy		0.89	

The CNN model outperforms alternative machine learning classifiers regarding F1-score, accuracy, precision, and recall. As indicated by their superior performance, CNNs are ideally adapted for capturing complex patterns in EEG data when employing DE features for emotion recognition on the SEED dataset. The balanced performance of the CNN model is demonstrated by its high precision, recall, and F1-score across all emotional states. The accuracy-loss curve (Figure 5.3) showcases a consistent decrease in the loss function over the training epochs, indicating that the model is effectively learning from the data. This optimizes the accuracy of emotion recognition by reducing misclassifications. In contrast, the CNN model incorporating DE features demonstrates superior performance in emotion recognition tasks based on EEG data to precisely classify emotional states.

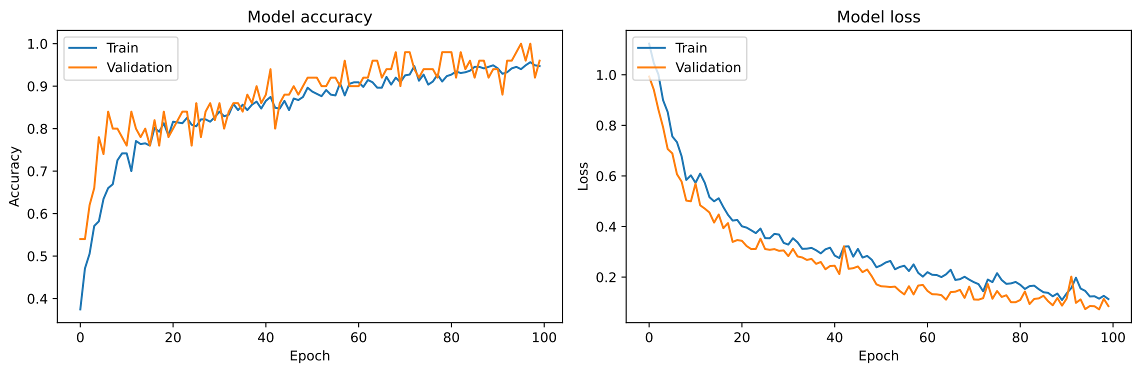


Figure 5.3: Accuracy-loss curve for CNN model on Differential Entropy - SEED dataset

5.4.3 Evaluating the CNN-BiLSTM Model

The CNN-BiLSTM model demonstrates superior performance in emotion recognition tasks based on EEG data, particularly when using DE features on the SEED dataset. The CNN-BiLSTM model outperforms alternative machine learning classifiers by attaining high precision, recall, and F1-score across all emotional states, with an accuracy of 93%. This exemplifies the efficacy of convolutional and recurrent neural network combinations in capturing temporal and spatial dependencies within EEG signals.

A marginal enhancement in precision and additional performance indicators is discernible in the CNN-BiLSTM model relative to the CNN model. By integrating temporal information via the bidirectional LSTM layers, the CNN-BiLSTM model improves its capability to discern sequential patterns in EEG signals, thereby bolstering its performance in emotion recognition tasks. Furthermore, the AUC scores associated with each emotional state indicate the exceptional capacity to differentiate among various emotional states. The model's efficacy is demonstrated by AUC values of 1.00 for neutron and positive and 0.98 for negative states.

Table 5.5: Classification Report for CNN-BiLSTM Model with DE Feature on SEED Dataset

Class	**Precision**	**Recall**	**F1-Score**
Neutral	1.00	0.88	0.94
Positive	0.93	1.00	0.96
Negative	0.88	0.92	0.90
Accuracy		0.93	

The efficacy of the CNN-BiLSTM model in capturing both immediate and extended dependencies within the EEG data is the primary factor contributing to its success. The capture of intricate patterns found in EEG signals is possible by utilizing the memory retention properties of LSTMs and the hierarchical representation learning capabilities of CNNs. When applied to DE features, which symbolize the entropy of the EEG signals, the CNN-BiLSTM model can proficiently acquire knowledge of the temporal fluctuations in entropy, thereby facilitating the classification of emotions with greater precision.

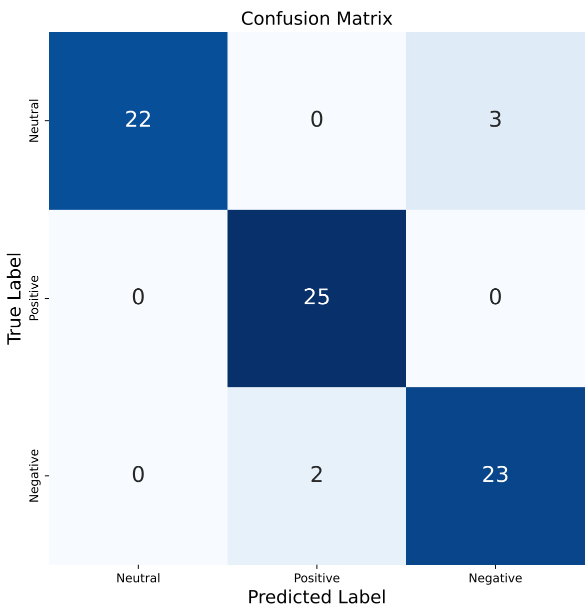


Figure 5.4: Confusion matrix for CNN-BiLSTM model on Differential Entropy - SEED dataset

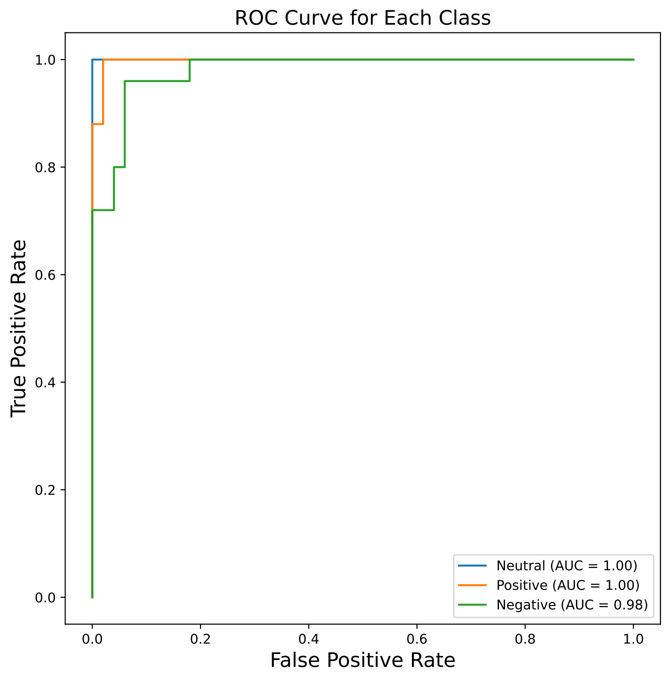


Figure 5.5: ROC-AUC curve for CNN-BiLSTM hybrid model on Differential Entropy - SEED dataset

5.5 Performance Evaluation on DEAP Dataset

5.5.1 Evaluating Machine Learning Models

This section assesses the performance of machine learning models on the DEAP dataset, which comprises low and high classes for the valence and arousal dimensions of emotion recognition, respectively. A variety of classifiers for machine learning were trained utilizing three distinct feature sets: statistical features, power spectral density (PSD), and differential entropy (DE). DE features exhibited superior predictive performance in valence and arousal on a consistent basis, surpassing both PSD and Statistical features. The average accuracy of DE features for valence prediction was 55.5%, surpassing that of PSD features (52.5%) and Statistical features (55%). Likewise, in arousal prediction, DE features achieved an average accuracy of 58%, surpassing PSD features by 57.5% and Statistical features by 51.5%.

These findings demonstrate that DE features can capture the subtleties of emotional states in EEG data. Their efficacy in overcoming various machine learning classifiers' performance challenges underscores their resilience and pertinence in the realm of emotion recognition. Consistent with prior research on the SEED dataset, DE features demonstrated superior accuracy (80–89%) compared to PSD and Statistical features. Based on the results obtained, DE features are demonstrated to be the most effective method for precisely discerning emotional states from EEG data. This implies that DE features may have wide-ranging applications in emotion recognition. (Table 5.6)

Table 5.6: Accuracy Table for Machine Learning Classifiers on DEAP Dataset

Feature	**Model**	**Accuracy (Valence)**	**Accuracy (Arousal)**	**Avg Accuracy %**
DE	DT	0.60	0.58	59%
	RF	0.59	0.58	58.5%
	KNN	0.55	0.56	55.5%
	SVM	0.52	0.59	55.5%
PSD	DT	0.51	0.54	52.5%
	RF	0.52	0.63	57.5%
	KNN	0.54	0.61	57.5%
	SVM	0.51	0.63	57%
Statistical	DT	0.56	0.54	55%
	RF	0.57	0.61	59%
	KNN	0.55	0.55	55%
	SVM	0.46	0.57	51.5%

5.5.2 Evaluating the CNN Model

The accuracy of the CNN model for valence and arousal levels was comparatively low when applied to the DEAP dataset containing DE features (Table 5.7). The model demonstrated a precision, recall, and F1-score of 0.55 on average for valence classification, with an accuracy of 55%. Comparatively, the model achieved an average precision, recall, and F1-score of 0.49 and an accuracy of 49% in the arousal classification domain. As indicated by these results, the CNN model trained using the DE features encountered difficulties in accurately differentiating between various affective states in the DEAP dataset. In contrast to its exemplary performance on the SEED dataset, the CNN model encountered more formidable obstacles when handling the DEAP dataset. Multiple factors contribute to the CNN model's inferior performance on the DEAP dataset compared to the SEED dataset. Initial concerns may arise when attempting to generalize the learned patterns from the potential disparity in emotional state distribution between the DEAP and SEED datasets. Furthermore, it is possible that the model's delineation of distinct emotional states would be hindered by variations in the complexity of emotional features captured by DE across datasets. Differences in the consistency and quality of EEG signals between the two datasets may also affect the model's efficacy.

Table 5.7: Classification Report for CNN Model on DEAP Dataset with DE Features

Emotion Level	**Avg Precision**	**Avg Recall**	**Avg F1-Score**	**Accuracy (%)**
Valence	0.55	0.55	0.55	55
Arousal	0.49	0.49	0.49	49

5.5.3 Evaluating the CNN-BiLSTM Model

The hybrid model comprising CNN and BiLSTM exhibits promise in improving accuracy compared to CNN models operating independently on both the DEAP and SEED datasets. Nevertheless, the lower accuracy obtained by DEAP compared to SEED is a noteworthy observation. The combination of CNN and BiLSTM architectures nevertheless exhibits potential for enhancing overall performance. This implies that incorporating temporal information furnished by BiLSTM layers can efficiently capture intricate patterns in EEG data, thereby facilitating the recognition of emotions with greater precision.

Even though DEAP has a lower accuracy, it is critical to remember that the feature extraction method is critical. The potential cause for the reduced precision identified in DEAP may be the difficulties encountered during the efficient extraction of Differential Entropy (DE) from the dataset. The efficacy of the DEAP dataset could be substantially improved by developing a generalized method for extracting DE customized to its specific attributes. This demonstrates the significance of feature extraction techniques, particularly in complex datasets like DEAP, in enhancing the efficacy of machine learning models.

In general, the results of this study highlight the importance of integrating Convolutional Neural Network (CNN) and BiLSTM architectures in the context of emotion recognition. Additionally, they show how additional advancements could be achieved by implementing optimized feature extraction techniques.

Table 5.8: Classification Report for CNN-BiLSTM Model on DEAP Dataset with DE Features

Emotion Level	**Avg Precision**	**Avg Recall**	**Avg F1-Score**	**Accuracy (%)**
Valence	0.59	0.59	0.59	59
Arousal	0.51	0.51	0.51	51

5.6 Related Works to the Proposed Study

The works listed in Table 5.9 are pertinent to the proposed EEG signal emotion recognition research. Subject-wise accuracy ranging from 86% to 97.29% was achieved on the SEED and DEAP datasets using differential entropy (DE) features in conjunction with multi-layer perceptron (MLP) and convolutional neural network (CNN) models. Various EEG features were explored, and their effectiveness was evaluated with machine learning and deep learning models in this research, aligning with this approach. Wavelet transform features were employed along with support vector machine (SVM), random forest (RF), and k-nearest neighbors (KNN) models, resulting in respective accuracies of 63.33%, 60.23%, and 58.54% on the SEED-IV dataset. Statistical features were utilized in conjunction with deep neural network (DNN) and CNN models, resulting in accuracies of 73.13% and 73.36% on the DEAP dataset. Additionally, an accuracy of 57.5% on the DEAP dataset was achieved by combining wavelet transform features with a KNN model. An accuracy of 81.4% on the DEAP dataset was achieved by combining energy-based and statistical features with a CNN model. These studies offer valuable insights into the efficacy of various machine learning models and feature extraction techniques utilized in the domain of emotion recognition from EEG signals, providing a valuable context for research in this area.

Table 5.9: Related Works to the Proposed Study

Authors	Dataset	Feature	Model	Accuracy
Kumar and Molinas [37]	SEED, DEAP	DE	MLP, CNN	Subject-wise accuracy range from 86% to 97.29%
Putra et al. [53]	DEAP	Wavelet transform	KNN	57.5%
Song et al. [54]	SEED, DREAMER	DE	DGCNN	90.40%
Tripathi et al. [55]	DEAP	Statistical and energy-based features	CNN	81.4%
Li et al. [56]	SEED, SEED - IV	DE	DAN	83.81%
Saranya et al. [36]	SEED-IV	Wavelet transform	SVM, RF, KNN, ANN	63.33%, 60.23%, 58.54%, 73.8%

5.7 Conclusion

This research investigation examined the effectiveness of various deep learning architectures and machine learning models in identifying emotions based on EEG signals. Utilizing the SEED and DEAP datasets revealed several significant discoveries. In the analysis of the SEED dataset, it was consistently observed that models trained using power spectral density (PSD) and statistical features outperformed those trained using differential entropy (DE) features. The CNN-BiLSTM model exhibited exceptional precision in assessing temporal dependencies, which implies that it may possess the capability to recognize emotions effectively. A thorough evaluation of machine learning and deep learning techniques for emotion recognition based on EEG signals constitutes the contribution of this work. Gaining valuable insights into the merits and drawbacks of various methodologies is accomplished through the methodical comparison of performance across datasets. The investigation into a hybrid model that combines Bidirectional Long Short-Term Memory (BiLSTM) networks with Convolutional Neural Networks (CNN) is a noteworthy contribution. The sequential learning capability of BiLSTM networks is utilized to capture temporal dependencies, while CNNs are employed to extract spatial features from EEG signals in this hybrid architecture. Particularly in capturing spatial and temporal information from EEG signals, the CNN-BiLSTM model exhibited encouraging results. To capture temporal dependencies on the SEED dataset, this model demonstrated remarkable performance, surpassing alternative models. Furthermore, it showed promise in enhancing accuracy when applied to the DEAP dataset.

Chapter 6

Conclusion & Future Works

6.1 Introduction

This chapter summarises prospective future research directions and provides conclusions from the study's findings. Exploring feature extraction methods, machine learning models, and deep learning architectures, the study examined various techniques for emotion recognition utilizing EEG data. Furthermore, the article explores the constraints that were encountered throughout the research and suggests potential avenues for future investigation that hope to overcome these obstacles.

6.2 Summary

Emotion recognition using EEG data presents a complex yet promising avenue of research with wide-ranging applications across various domains. This work aimed to examine the efficacy of various techniques for extracting features and models for machine learning and deep learning architectures in properly categorizing emotional states based on EEG data. The generalization potential of the suggested approaches was assessed using two datasets, SEED and DEAP.

The process of feature extraction involves three primary techniques: Power Spectral Density (PSD), Differential Entropy (DE), and Statistical characteristics (mean, variance, standard deviation, kurtosis). Each approach is designed to capture distinct properties of EEG data. After extracting the features, four conventional machine learning classifiers (Support Vector Machine, Decision Tree, K-Nearest Neighbours, and Random Forest) were trained using the retrieved features. The results obtained from the SEED dataset demonstrated that models trained on DE features consistently achieved superior performance compared to models trained on PSD and Statistical features, as measured by accuracy, recall, and F1-score. The Random Forest model had superior accuracy, precision, recall, and F1-score performance compared to other models, indicating its strong capacity to categorize emotional states.

A Convolutional Neural Network (CNN) architecture was subsequently created utilizing the DE characteristics as input. The CNN model demonstrated exceptional effectiveness in categorizing emotional states on the SEED dataset, with an overall accuracy of 89%. Nevertheless, when the CNN model was tested on the DEAP dataset, it had difficulties, leading to decreased valence and arousal categorization accuracy scores. To tackle this issue, a novel hybrid CNN-BiLSTM model was introduced with the objective of capturing both temporal and spatial dependencies present in EEG data. The CNN-BiLSTM model demonstrated superior performance to the solo CNN model on both the SEED and DEAP datasets, highlighting its potential to enhance accuracy compared to conventional CNN architectures.

6.3 Conclusion

In conclusion, this research endeavor has yielded significant findings pertaining to identifying emotions from EEG data by implementing diverse machine learning architectures, feature extraction techniques, and models. After a thorough analysis, differential Entropy was determined to be the most effective feature; it consistently performed better in classification assignments than Power Spectral Density and Statistical features. Using DE features improved accuracy, precision, recall, and F1 scores by effectively capturing subtle patterns in EEG signals linked to emotional states. Significantly, the Random Forest model demonstrated outstanding performance, attaining an accuracy of 84% in SVM and 83.33% in RF.

Additionally, encouraging outcomes were observed with the application of DE features in the construction of a 2D Convolutional Neural Network model that achieved an accuracy of 89% on the SEED dataset. Nevertheless, implementing the 2D CNN model on the DEAP dataset revealed certain obstacles, which strongly suggest the necessity for additional iteration and applicability. A CNN-BiLSTM hybrid model was suggested to tackle this issue, which took advantage of the interdependencies between space and time in EEG signals. The hybrid model achieved an accuracy of 93% on the SEED dataset, surpassing the performance of the standalone CNN architecture.

6.4 Limitation

The remarkable 93% accuracy attained by the CNN-BiLSTM model on the SEED dataset serves as a clear indication of the efficacy of the methods that have been proposed. Nonetheless, the suboptimal performance observed on the DEAP dataset, specifically the CNN model's challenge to attain satisfactory outcomes, underscores the models' limitations in terms of generalizability. The interpretability of deep learning models, the possibility of information loss during feature extraction, and dataset specificity are some limitations of this work.

- The comprehension of deep learning models' predictions is frequently impeded by their operation as ”black boxes,” which complicates their interpretability. Trust may be eroded by this lack of transparency, particularly in critical applications like EEG-emotion recognition.

- In addition, information loss is a possibility during feature extraction due to the simplification of intricate EEG data by techniques such as DE and PSD. This simplification may negatively impact performance, as evidenced by the results obtained when applying models to the DEAP dataset.

- The practical applicability of models is constrained by their dependence on particular datasets for training and validation. Although successful on SEED, DEAP performance declined, indicating that the model was overfitting to SEED's characteristics.

6.5 Future Works

- **Investigate advanced DE feature extraction methods**: Autoencoders and deep learning-based approaches should be investigated as potential advanced methods for extracting more generalized DE features from EEG data.

- **Improve the preprocessing of EEG data**: Incorporate methodologies such as artifact elimination, noise mitigation, and precise filtration into EEG data preprocessing to augment the quality of input data intended for emotion recognition models.

- **Integrate sources of multimodal data**: For more robust emotion classification, incorporate supplementary data sources such as physiological signals or facial expressions to furnish more comprehensive contextual information.

- **Construction of more generalized models**: Construct models that exhibit strong generalizability across diverse datasets and conditions through the optimization of training procedures, model architecture, and feature extraction. This process necessitates thorough validation and refining of hyperparameters on various datasets.
- **Perform real-world validation investigations**: Conduct practical assessments of the proposed models' performance, usability, and efficacy in domains such as mental health monitoring and human-computer interaction by validating them in real-world scenarios.

REFERENCES

[1] J. A. Russell, "Affective space is bipolar.," *Journal of personality and social psychology*, vol. 37, no. 3, p. 345, 1979.

[2] E. H. Houssein, A. Hammad, and A. A. Ali, "Human emotion recognition from eeg-based brain–computer interface using machine learning: a comprehensive review," *Neural Computing and Applications*, vol. 34, no. 15, pp. 12527–12557, 2022.

[3] S. Nath, "Why is bilstm better than lstm?," April 2024. `https://medium.com/@souro400.nath/why-is-bilstm-better-than-lstm-a7eb0090c1e4`.

[4] X. Li, Y. Zhang, P. Tiwari, D. Song, B. Hu, M. Yang, Z. Zhao, N. Kumar, and P. Marttinen, "Eeg based emotion recognition: A tutorial and review," *ACM Computing Surveys*, vol. 55, no. 4, pp. 1–57, 2022.

[5] R. Alhalaseh and S. Alasasfeh, "Machine-learning-based emotion recognition system using eeg signals," *Computers*, vol. 9, no. 4, p. 95, 2020.

[6] H. Berger, "Über das elektroenkephalogramm des menschen," *Archiv für psychiatrie und nervenkrankheiten*, vol. 87, no. 1, pp. 527–570, 1929.

[7] M. IB, "Measures of emotion: a review," *Cognition and emotion*, vol. 23, no. 2, pp. 209–237, 2009.

[8] P. Ekman, "Darwin's contributions to our understanding of emotional expressions," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3449–3451, 2009.

[9] P. Ekman *et al.*, "Basic emotions," *Handbook of cognition and emotion*, vol. 98, no. 45-60, p. 16, 1999.

[10] R. Plutchik, *Emotions and life: Perspectives from psychology, biology, and evolution.* American Psychological Association, 2003.

[11] T. Nykopp, "Statistical modelling issues for the adaptive brain interface (m. sc diploma)," *Helsinki University of Technology, Department of Electrical and Communication Engineering*, 2001.

[12] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Transactions on autonomous mental development*, vol. 7, no. 3, pp. 162–175, 2015.

[13] E. P. Torres, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "Eeg-based bci emotion recognition: A survey," *Sensors*, vol. 20, no. 18, p. 5083, 2020.

[14] J. Zhang, Z. Yin, P. Chen, and S. Nichele, "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review," *Information Fusion*, vol. 59, pp. 103–126, 2020.

[15] Y. Wang and T.-P. Jung, "Improving brain–computer interfaces using independent component analysis," *Towards practical brain-computer interfaces: Bridging the gap from research to real-world applications*, pp. 67–83, 2013.

[16] P. Ofner and G. R. Müller-Putz, "Movement target decoding from eeg and the corresponding discriminative sources: A preliminary study," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1468–1471, IEEE, 2015.

[17] M. J. Alhaddad, M. Kamel, H. Malibary, K. Thabit, F. Dahlwi, and A. Hadi, "P300 speller efficiency with common average reference," in *Autonomous and Intelligent Systems: Third International Conference, AIS 2012, Aveiro, Portugal, June 25-27, 2012. Proceedings*, pp. 234–241, Springer, 2012.

[18] X. Li, H. Fan, H. Wang, and L. Wang, "Common spatial patterns combined with phase synchronization information for classification of eeg signals," *Biomedical Signal Processing and Control*, vol. 52, pp. 248–256, 2019.

[19] B. Hjorth, "Eeg analysis based on time domain properties," *Electroencephalography and clinical neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.

[20] Y. Liu and O. Sourina, "Eeg-based subject-dependent emotion recognition algorithm using fractal dimension," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 3166–3171, IEEE, 2014.

[21] M. Murugappan and S. Murugappan, "Human emotion recognition through short time electroencephalogram (eeg) signals using fast fourier transform (fft)," in *2013 IEEE 9th International Colloquium on Signal Processing and its Applications*, pp. 289–294, IEEE, 2013.

[22] O. M. Solomon Jr, "Psd computations using welch's method," *NASA STI/Recon Technical Report N*, vol. 92, p. 23584, 1991.

[23] Z. S. Torabi, M. H. Nadimi-Shahraki, and A. Nabiollahi, "Efficient support vector machines for spam detection: a survey," *International Journal of Computer Science and Information Security*, vol. 13, no. 1, p. 11, 2015.

[24] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE transactions on systems, man, and cybernetics*, vol. 21, no. 3, pp. 660–674, 1991.

[25] T. K. Ho, "Random decision forests," in *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1, pp. 278–282, IEEE, 1995.

[26] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.

[27] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, p. 420, 2021.

[28] M. Krichen, "Convolutional neural networks: A survey," *Computers*, vol. 12, no. 8, p. 151, 2023.

[29] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of lstm and bilstm in forecasting time series," in *2019 IEEE International conference on big data (Big Data)*, pp. 3285–3292, IEEE, 2019.

[30] F. Cui, R. Wang, W. Ding, Y. Chen, and L. Huang, "A novel de-cnn-bilstm multi-fusion model for eeg emotion recognition," *Mathematics*, vol. 10, no. 4, p. 582, 2022.

[31] V. M. Joshi, R. B. Ghongade, A. M. Joshi, and R. V. Kulkarni, "Deep bilstm neural network model for emotion detection using cross-dataset approach," *Biomedical Signal Processing and Control*, vol. 73, p. 103407, 2022.

[32] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for eeg-based emotion classification," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 81–84, IEEE, 2013.

[33] R. Lu, "Automatic emotion decoding from eeg signals using the seed dataset," 2023.

[34] L. O. Mohutsiwa and R. S. Jamisola, "Eeg-based human emotion classification using combined computational techniques for feature extraction and selection in six machine learning models," in *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 1095–1102, IEEE, 2021.

[35] M. Moshfeghi, J. P. Bartaula, and A. T. Bedasso, "Emotion recognition from eeg signals using machine learning," 2013.

[36] K. Saranya, D. Sudharson, N. Gokila, S. Lakshmipriya, and K. Suriyavathi, "Emotion recognition using eeg signal classification of seed dataset," in *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, pp. 1–6, IEEE, 2023.

[37] M. Kumar and M. Molinas, "Human emotion recognition from eeg signals: model evaluation in deap and seed datasets," in *Proceedings of the First Workshop on Artificial Intelligence for Human-Machine Interaction (AIxHMI 2022) co-located with the 21th International Conference of the Italian Association for Artificial Intelligence (AI* IA 2022), CEUR Workshop Proceedings, CEUR-WS. org*, 2022.

[38] A. Iyer, S. S. Das, R. Teotia, S. Maheshwari, and R. R. Sharma, "Cnn and lstm based ensemble learning for human emotion recognition using eeg recordings," *Multimedia Tools and Applications*, vol. 82, no. 4, pp. 4883–4896, 2023.

[39] K. Guo, H. Mei, X. Xie, and X. Xu, "A convolutional neural network feature fusion framework with ensemble learning for eeg-based emotion classification," in *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, vol. 1, pp. 1–4, IEEE, 2019.

[40] S. Tripathi, S. Acharya, R. Sharma, S. Mittal, and S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on deap data," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, pp. 4746–4752, 2017.

[41] H. Chao, L. Dong, Y. Liu, and B. Lu, "Emotion recognition from multiband eeg signals using capsnet," *Sensors*, vol. 19, no. 9, p. 2212, 2019.

[42] M. M. Rahman, A. K. Sarkar, M. A. Hossain, and M. A. Moni, "Eeg-based emotion analysis using non-linear features and ensemble learning approaches," *Expert Systems with Applications*, vol. 207, p. 118025, 2022.

[43] Y. Luo, Q. Fu, J. Xie, Y. Qin, G. Wu, J. Liu, F. Jiang, Y. Cao, and X. Ding, "Eeg-based emotion classification using spiking neural networks," *IEEE Access*, vol. 8, pp. 46007–46016, 2020.

[44] S. Sidharth, A. A. Samuel, H. Ranjana, J. T. Panachakel, *et al.*, "Emotion detection from eeg using transfer learning," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 1–4, IEEE, 2023.

[45] Brain and Cognition Machine Intelligence (BCMI) Lab, Shanghai Jiao Tong University, "SEED Dataset." `https://bcmi.sjtu.edu.cn/home/seed/seed.html`. Accessed: November 25, 2023.

[46] Shanghai Jiao Tong University, "Bioinformatics and Computational Medicine Lab." `https://bcmi.sjtu.edu.cn/`. Accessed: April 20, 2024.

[47] L. Zhang, "Dr. Liang Zhang's Homepage." `https://bcmi.sjtu.edu.cn/~blu/`. Accessed: April 20, 2024.

[48] W. Zheng, "Weilong Zheng's Homepage." `https://weilongzheng.github.io/`. Accessed: April 20, 2024.

[49] "Compumedics Neuroscan." `https://compumedicsneuroscan.com/`. Accessed: April 20, 2024.

[50] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.

[51] Queen Mary University of London, School of Electronic Engineering and Computer Science, "DEAP Dataset." `https://www.eecs.qmul.ac.uk/mmv/datasets/deap/index.html`. Accessed: April 20, 2024.

[52] scikit-learn contributors, "GridSearchCV." `https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html`, 2024. Accessed: April 22, 2024.

[53] A. E. Putra, C. Atmaji, and F. Ghaleb, "Eeg-based emotion classification using wavelet decomposition and k-nearest neighbor," in *2018 4th International Conference on Science and Technology (ICST)*, pp. 1–4, IEEE, 2018.

[54] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532–541, 2018.

[55] S. Tripathi, S. Acharya, R. Sharma, S. Mittal, and S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on deap data," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, pp. 4746–4752, 2017.

[56] H. Li, Y.-M. Jin, W.-L. Zheng, and B.-L. Lu, "Cross-subject emotion recognition using deep adaptation networks," in *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13–16, 2018, Proceedings, Part V 25*, pp. 403–413, Springer, 2018.

SIMILARITY INDEX

Thesis_Book_1803108.pdf

ORIGINALITY REPORT

26%
SIMILARITY INDEX

PRIMARY SOURCES

1	link.springer.com Internet	357 words — 2%
2	arxiv.org Internet	316 words — 1%
3	www.mdpi.com Internet	290 words — 1%
4	Essam H. Houssein, Asmaa Hammad, Abdelmgeid A. Ali. "Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review", Neural Computing and Applications, 2022 Crossref	167 words — 1%
5	Xiang Li, Yazhou Zhang, Prayag Tiwari, Dawei Song, Bin Hu, Meihong Yang, Zhigang Zhao, Neeraj Kumar, Pekka Marttinen. "EEG Based Emotion Recognition: A Tutorial and Review", ACM Computing Surveys, 2022 Crossref	131 words — 1%
6	export.arxiv.org Internet	85 words — < 1%
7	ceur-ws.org Internet	72 words — < 1%