A Hybrid Signal Processing Method for the Classification of Emotions from Physiological Signal: EEG with Deep Neural Network

A thesis submitted to the
UPES

For the award of

Doctor of Philosophy

In

Computer Science & Engg.
By
Amrendra Nath Tripathi

July 2025

Supervisor(s)
Dr.Tanupriya Choudhury
Dr. Hitesh Kumar Sharma

School of Computer Science
UPES
Dehradun-248007: Uttarakhand, INDIA

A Hybrid Signal Processing Method for the Classification of Emotions from Physiological Signal: EEG with Deep Neural Network

A thesis submitted to the
UPES
For the award of

Doctor of Philosophy
in
Computer Science & Engg.
By
Amrendra Nath Tripathi
(SAP ID 500066159)

July 2025

Supervisor
Dr. Tanupriya Choudhury
Professor
School of Computer Science,
UPES

Co-Supervisor
Dr. Hitesh Kumar Sharma
Professor
School of Computer Science,
UPES

School of Computer Science
UPES
Dehradun-248007: Uttarakhand; INDIA

DECLARATION

I declare that the abstract entitled "***A Hybrid Signal Processing Method for the Classification of Emotions from Physiological Signal: EEG with Deep Neural Network***" has been prepared by me under the guidance of Dr.Tanupriya Choudhury, Professor (Supervisor) and Dr. Hitesh Kumar Sharma, Professor (Co-Supervisor) School of Computer Science, UPES, Dehradun . No part of this Abstract has formed the basis for the award of any degree or fellowship previously.

Amrendra.

Amrendra Nath Tripathi
School of Computer Science
UPES,Dehradun-248007, Uttarakhand

UPES
UNIVERSITY OF TOMORROW

Great Place To Work® Certified FEB 2025 - FEB 2026 INDIA ™

Date: April 16,2025

CERTIFICATE

We certify that Amrendra Nath Tripathi has prepared his thesis entitled **"A Hybrid Signal Processing Method for the Classification of Emotions from Physiological Signal: EEG with Deep Neural Network** for the award of PhD degree of the UPES, under our guidance. He has carried out the work at the School of Computer Science, UPES.

15/4/2025

Dr.Tanupriya Choudhury (Supervisor)
Professor
School of Computer Science,
UPES,Dehradun -248007,
Uttarakhand, India

15/04/2025

Dr.Hitesh Kumar Sharma (Co-Supervisor)
Professor
School of Computer Science,
UPES,Dehradun -248007,
Uttarakhand, India

ABSTRACT

The brain is the human body's most complex and fascinating organ. Scientists and researchers are studying the relationship between action, emotion, memory, thought, consciousness, and the physical body. Rather than using MRI and Functional MRI for brain studies, Electroencephalography (EEG) signals are used to better understand the dynamicity of the brain. Electroencephalograph is a term used to describe the recording of electrical potentials around the scalp.

The voltage variations obtained from ionic current flows within the neurons of the brain can be determined with EEG. It is a noninvasive procedure (or) technique based on a neuroimaging method (or) technique in which voltage measurements are recorded via a series of electrodes kept on the scalp (or) within the cortex field (intracranial EEG). Hans Berger was the first person to record an EEG from a human brain in the 1920s. He was also the first to read the cyclic patterns of an EEG signal and to discover the relationship between various states of consciousness (e.g. sleep, attention processes, wake-fullness, etc.) It is referred to as EEG wave patterns, and it has a frequency range of 8 to 12 Hz for the alpha frequency.

EEG analysis by a physician is complex, time-consuming, and requires knowledge. As a result, an automated EEG analysis system must be built. Data collection, preprocessing, signal transformation, feature extraction, and manifestations are all part of the EEG signal processing process. EEG is a method of obtaining a recording of electrical activity in the brain by placing an electrode on the scalp. It is a rough approximation of the neuron's total electrical activity. The active potential is the cell's active property, with an amplitude range of up to 100mV. EEG recordings are obtained using the 10-20 International electrodes method, which is the most used system of electrode placement. Researchers are more likely to use the 10-20 method to compare their data for visualization. EEG is a method of recording electrical activity along the scalp caused by the firing of neurons in the brain.

The importance of biological time series analysis has long been recognized in the field of nonlinear analysis, and it is primarily useful for hidden dynamical features. For the

interpretation of EEG signals in patients with various cognitive behaviors, nonlinear analysis methods are used. The study considers data sets from various cognitive states or tasks.

Medical signals, known as vital signs, reflect the internal operations of the human body. Examining and analyzing these signals provides insights into the condition of essential organs such as the heart, lungs, and brain. Among these, EEG presents the greatest challenge in interpretation. Recorded at multiple locations on the head surface, EEG signals contain various frequency components with lower amplitude than other medical signals. A distinctive characteristic of EEG is the superimposition of signals from distant sources when measured at a single point, further complicating analysis. Due to these complexities, EEG remains one of the most difficult signals to evaluate. However, careful processing enables the development of applications designed to assist individuals with multiple limb amputations. Brain-Computer Interface (BCI) serves as a rehabilitative system, facilitating the restoration of partial limb movement. This study focuses on EEG signals to determine upper-limb movement. Signal characteristics, including frequency, amplitude, time-frequency distribution, and event-based features, form the basis of medical signal analysis. Various techniques enable this analysis, but due to the inherent nonlinearity of signals, artificial intelligence emerges as the most effective approach for classification.

EEG signals exhibit key characteristics such as frequency and event-related brain activity. Utilizing their time-frequency combined features as inputs for artificial intelligence enables effective signal analysis. Unlike other classification methods, artificial intelligence adapts to signal nonlinearity, offering a significant advantage. Identifying the origins of various brain activities is essential for BCI applications aimed at enhancing movement prediction. Variability in signal attributes across individuals presents a challenge in algorithm generalization, often leading to inconsistent results across different databases. Addressing this issue requires the development of algorithms that incorporate individual subject constraints, ensuring adaptability and improved performance.

The initial study focuses on the increasing interest in automatic human ER based on EEG signals, driven by the growing importance of HCI. Previous models have often overlooked the contextual information embedded within EEG signals. This research manuscript introduces a novel automated model designed to enhance ER using EEG signals. Initially, signals are sourced from the online DEAP, followed by denoising using EMD and VMD filters to eliminate artifacts and noise. Feature extraction utilizes 20 statistical features to extract discriminative information from the decomposed EEG signals. Lastly, the Long Short-Term Memory network (LSTM) is employed for human ER in terms of arousal or valence, with optimal hyper-parameters selected using the Improved Rat Swarm Optimization Algorithm (IRSOA). The resulting IRSOA-LSTM network demonstrates superior performance, achieving a mean accuracy of 84.89%, sensitivity of 86.95%, specificity of 86%, precision of 83.68%, and f1-score of 85.28% on the DEAP database, as evidenced in the results and discussion section. Simulation outcomes affirm the superiority of the proposed IRSOA-LSTM network over existing machine-learning models.

In the subsequent investigation, recent progress in emotion detection, particularly concerning arousal and valence, has demonstrated its utility in the domains of psychology and Human-Computer Interaction (HCI). Given the rapid evolution of computer applications, there is substantial interest among researchers in automatically classifying human emotions based on EEG. Nevertheless, prevailing methodologies have not sufficiently accounted for the contextual information embedded within EEG signals. To address this gap, this study employs an automated model to enhance EEG-based emotion recognition. Initially, EEG signals are sourced from the DEAP. Subsequently, data decomposition is conducted using both VMD and EMD filters. Following this step, the novel Modified Tunicate Swarm Optimization Algorithm (MTSOA) is employed for feature selection to enhance accuracy by identifying the most relevant features conducive to the emotion classification process. Additionally, optimal hyper-parameters are determined using the IRSOA. Finally, a Bi-Directional LSTM(Bi-LSTM) is deployed for emotion classification in the concluding phase to differentiate between arousal and valence. According to the results, the proposed method yields superior outcomes, achieving arousal accuracy of 84.90 % and valence accuracy of 87.3%.

The proposed method enhances emotion recognition by utilizing a hybridized CNN and LSTM architecture to capture spatial and temporal features for superior classification accuracy. Raw EEG data from the DEAP dataset is preprocessed using hybrid VMD and EMD technique to remove artifacts, followed by the extraction of statistical features. These features are classified using the hybrid CNN-LSTM model, further optimized with IRSOA, achieving notable arousal accuracy of 89.58% and valence accuracy of 92.29%. This method effectively demonstrates its capability to handle diverse human emotions, ensuring robust and efficient EEG-based emotion recognition.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my research supervisor(s), Dr. Tanupriya Choudhury, Professor, SCS, UPES and Dr. Hitesh Kumar Sharma, Professor SCS, UPES, for their continuous guidance, support, and encouragement throughout my Ph.D. journey. Their insightful feedback and valuable suggestions have been instrumental in shaping the direction and outcome of this research work.

I also take this opportunity to thank my former supervisor, Dr. Nilima Fulmare, for her early support and guidance during the initial phase of my doctoral research, which laid a strong foundation for the work that followed.

I am sincerely grateful to Dr. Ram Sharma, Hon'ble Vice Chancellor, UPES, for providing a visionary academic environment and institutional support that enabled me to carry out this research.

I am also thankful to the members of my doctoral committee, Dr. Vijaysekhar Chellaboina, Dean SCS, Dr. Pankaj Kumar, Associate Dean (R&D) and Dr. Sonali Vyas, Sr. Associate Professor, Research Co-Ordinator for their time, evaluation, and constructive comments at various stages of my research.

I extend my appreciation to the faculty and staff of the School of Computer Science**,** UPES for providing a conducive academic environment and access to necessary resources during my research.

Special thanks to my colleagues and friends at the research lab, whose camaraderie and support helped make the challenging times more manageable. Finally, I am deeply grateful to my family—especially my parents and spouse and my son for their unwavering love, patience, and motivation. Their belief in me has been my greatest strength.

Amrendra.

Amrendra Nath Tripathi

School of Computer Science

UPES, Dehradun-248007, Uttarakhand

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
BCI	Brain-Computer Interface
CNN	Convolutional Neural Network
DEAP	Dataset for Emotion Analysis using Physiological Signals
DNN	Deep Neural Network
ECG	Electrocardiogram
EEG	Electroencephalography
EMD	Empirical Mode Decomposition
ER	Emotion Recognition
GWO	Grey Wolf Optimizer
HCI	Human-Computer Interaction
IRSOA	Improved Rat Swarm Optimization
KNN	K-Nearest Neighbors
LSTM	Long Short-Term Memory
MRI	Magnetic Resonance Imaging
MTSOA	Modified Tunicate Swarm Optimization
PSO	Particle Swarm Optimization
RMSE	Root Mean Square Error
SVM	Support Vector Machine
VMD	Variational Mode Decomposition

CHAPTER 1

Introduction

Overview

The human brain is a most complex and wonderful organ of human body. Scientists and researchers are discovering the intricate relationships between action, emotion, memory, thought, consciousness, and how they interact with the physical body. To study the dynamic nature of the brain, EEG signals have become a substitute for MRI and Functional MRI in brain researches. Electrodes are applied to the scalp to measure the electrical activity produced by the brain, which is known as EEG.

Electroencephalography is a technique to measure voltage variations due to ionic currents within neurons in the brain. This non-invasive neuroimaging method records electrical activity through electrodes placed on scalp or, in case of intracranial EEG, within cortical field. The first human brain EEG signals were recorded in the 1920s by Hans Berger. He also found some cyclic patterns in the signals of EEG and disclosed how these patterns relate to diverse states of consciousness, like sleep or wakefulness, as well as to attention processes. Such cyclic signals are called wave EEG patterns; the most studied is the alpha frequency band: 8-12 Hz.

Analysing the EEG signals manually is such a complex, time-consuming job that needs specialized skill. Therefore, an automated analysis system for EEG is an important challenge. The overall EEG signal processing workflow could involve data collection, data preprocessing, signal transformation, feature extraction, and then interpreting the results. With EEG, the electrical activity of brain is recorded by placing electrodes on the scalp, giving an approximation of the collective electrical activity of neurons.

One of the features of active neuronal potentials is an amplitude up to 100 mV and is well captured in the EEG recording. Such recording is typically obtained with a widely recognized 10-20 electrode placement system throughout the world, widely accepted by researchers for data to be visualized and then compared. Thus, the EEG serves as a powerful technique for recording electrical activity over the scalp, which relates to the firing patterns of neurons in the brain.

The importance of biological time series analysis has long been recognized in the field of nonlinear analysis, and it is primarily useful for hidden dynamical features. For the interpretation of EEG signals in patients with various cognitive behaviours, nonlinear analysis methods are used. The study considers data sets from various cognitive states or tasks.

Medical signals, often referred to as vital signs, provide crucial insights into the internal functioning of the human body. By analyzing these signals, both the health and performance of vital organs such as heart, lungs, and brain are anlaysed. Among these, the EEG is particularly complex to understand. EEG signals are recorded from various points on the scalp and consist of multiple frequency components with lower amplitudes than most other medical signals. Additionally, EEG signals often include superimposed signals originating from distant regions of the brain, making their interpretation even more challenging.

Despite these complexities, careful analysis of EEG signals can lead to remarkable applications. For instance, a BCI system can be developed to assist individuals with multiple limb amputations. Such systems are rehabilitative, enabling partial restoration of limb movement by decoding brain activity.

The analysis of medical signals typically involves examining features like frequency, amplitude, and time-frequency combinations, among other characteristics. While various methods can be employed for signal analysis, the inherent nonlinearity of these signals makes artificial intelligence an ideal approach for accurate classification and interpretation. By leveraging AI, it is effective to analyze EEG signals and unlock their potential for innovative rehabilitative applications.

EEG signals are primarily characterized by their frequency and event-related components associated with brain activity. To analyze these signals effectively, their combined time-frequency attributes serve as features in artificial intelligence models. AI offers a significant advantage over traditional classification methods by accommodating the inherent nonlinearity of EEG signals. Thus, for BCI applications seeking enhanced movement prediction, it would be important to identify the sources of various behaviors in neural processes. Due to differences in signal characteristics from one subject to another, generalizing an algorithm among people is difficult. An algorithm that works well on a particular dataset may not even work well on another set. This problem can be overcome by designing algorithms based on individual-specific constraints and ensuring more accurate and person-based output.

1.1 The Brain's Functional Organization

The functional organization of the brain offers a simplified way through which its various functionalities could be understood. Each region is highly localized, giving room for precise specialization. For instance, the secondary motor areas are located anterior to the primary motor cortex in the frontal lobes. They involve preparing and initiating movement, just like the secondary visual lobe, which handles detailed visual information such as color and motion near the primary visual cortex.

The tertiary areas, or association cortices, process higher-order processes such as interpretation and memory. Even though these areas of function are well-defined, lacking a common reference template from which to compare becomes challenging when dealing with a large number of individuals. Differences in brain size between individuals could cause positional misalignments and the misinterpretation of functional areas.

1.2 Background of Electroencephalography (EEG)

EEG measures voltage variations brought on by the ionic currents passing through brain neurons to record the electrical activity along the scalp. With its long-standing presence in neuroscience, EEG remains an essential tool for studying and understanding brain activity. The human brain, being the ultimate control center of the body, plays a critical role in the nervous system in managing functions of all organs. Brain, composed of billions of interconnected neurons linked by axons and dendrites, works as a highly efficient network. Neurons are the basic data-processing units of the brain; they receive electrical signals from other neurons. An electrical discharge known as an action potential or nerve impulse is created when these signals converge and the total input surpasses a certain threshold. The subsequent neuron in the network receives this action potential as its input.

Fig. 1.1 illustrates the three main anatomical components of the brain: the brainstem, cerebellum, and cerebrum. Every one of these elements has a distinct and crucial purpose in brain function.

(1) Cerebrum: The cerebrum is a largest and most vital part of human brain, playing a central role in thought processes, breathing, emotions, and motor functions. Its outermost layer, known as the cerebral cortex, is crucial for advanced brain activities which is further divided into two hemispheres such as right and left connected by corpus callosum and separated by a deep fissure (Nolte, 2002).

Each hemisphere is further divided into four distinct lobes: frontal, temporal, parietal, and occipital.

- **Frontal lobe** is responsible for emotions, cognitive processes, executive functions, speech production, and voluntary movement.
- **Parietal lobe** governs recognition, sensory perception, spatial orientation, and movement coordination.
- **Occipital lobe** specializes in visual processing.
- T**emporal lobe** manages auditory processing, long-term memory, and language comprehension.

Together, these lobes enable the cerebrum to regulate complex and diverse brain functions critical to daily life.

(2) Cerebellum: The cerebellum, located at the back of the head, second-largest part of brain. It plays an important role in regulating muscle movements, maintaining balance, controlling coordination, and processing sensory perception.

(3) Brainstem, located at the base of brain where it connects the spinal cord to the cerebrum, serves as the body's primary control center. It regulates vital functions such as breathing, movement, appetite, and heartbeat, as well as sensory inputs like heat and pain.

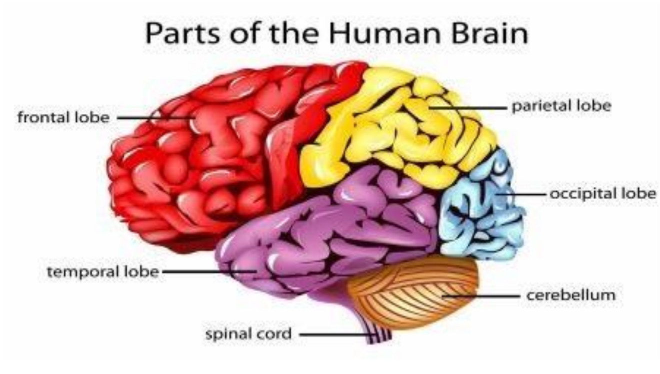


Figure 1. 1: Parts of the Human Brain

1.3 Electroencephalography

EEG is a method for detecting and recording electrical activity in brain using electrodes positioned on scalp. A device used for this is an electroencephalograph, which produces a visual representation known as an EEG. This record is valuable for studying and diagnosing brain disorders. Dr. Hans Berger introduced the EEG machine in 1929, marking a significant advancement in medical research. Multichannel EEG refers to a method that uses a greater number of electrodes to measure brain activity. While EEG is typically recorded using scalp electrodes, it can also be done directly on the brain's surface. EEG offers several advantages, including high temporal resolution, non-invasiveness, ease of use, and low cost. EEG behavior is categorized into three types based on measurement techniques: random, evoked potential, and single neuron. Spontaneous action, which involves continuous EEG measurement from scalp or brain, falls under this classification. When recorded from the scalp, EEG has an amplitude of about 1-100 µV, while surface brain recordings typically range from 1-2 mV. The duration of EEG recordings depends on the experimental protocol. Evoked potentials refer to EEG components that arise in response to stimuli (such as music or video), requiring multiple stimuli and signal averaging for accurate measurement. This study primarily focuses on spontaneous action, with EEG being measured from the scalp.

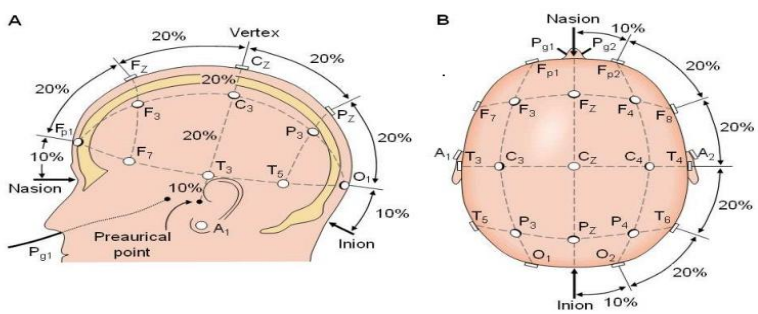


Figure 1. 2: The International 10-20 electrode positioning system, a) left and b) above the head

1.4 EEG Electrode System

For recording electric voltage through EEG, the International 10-20 electrode placement system is widely used. The numbers 10 and 20 represent the actual distances between adjacent electrodes, based on the overall front-back and right-left dimensions of the skull.

As illustrated in Fig. 1.2, these distances are measured from two reference points: the nasion (located between the eyebrows) and the inion (at the back of the head). Electrodes are placed on the scalp's surface, with each electrode labeled according to the brain region it corresponds to (F = frontal; FP = prefrontal; T = temporal; P = parietal; O = occipital; Z = midline electrode). Odd-numbered electrodes are placed on left side of brain, while even-numbered electrodes are on right side. Reference electrodes, A1 and A2, are typically attached to earlobes. EEG recordings can be made using anywhere from 2 to 256 channels, with different electrodes employed for different setups. For higher resolution, the 10-10 or 10-5 electrode placement systems are often used.

1.5 EEG Rhythms

As illustrated in Fig. 1.3, EEG rhythms consist of five types of waves, each with varying amplitudes and frequencies. The delta band, with a frequency range of 0 to 4 Hz, is associated with deep sleep. The theta band, ranging from 4 to 7 Hz, appears when a person is drowsy. The alpha band, found between 8 and 12 Hz, occurs when a person is relaxed, alert, and has their eyes closed. The beta band, with frequencies ranging from 12 to 30 Hz, reflects an active brain state and is linked to concentration, cognitive tasks, problem-solving, and stress. The gamma band, which is seldom observed, has frequencies above 30 Hz and is characterized by small amplitudes.

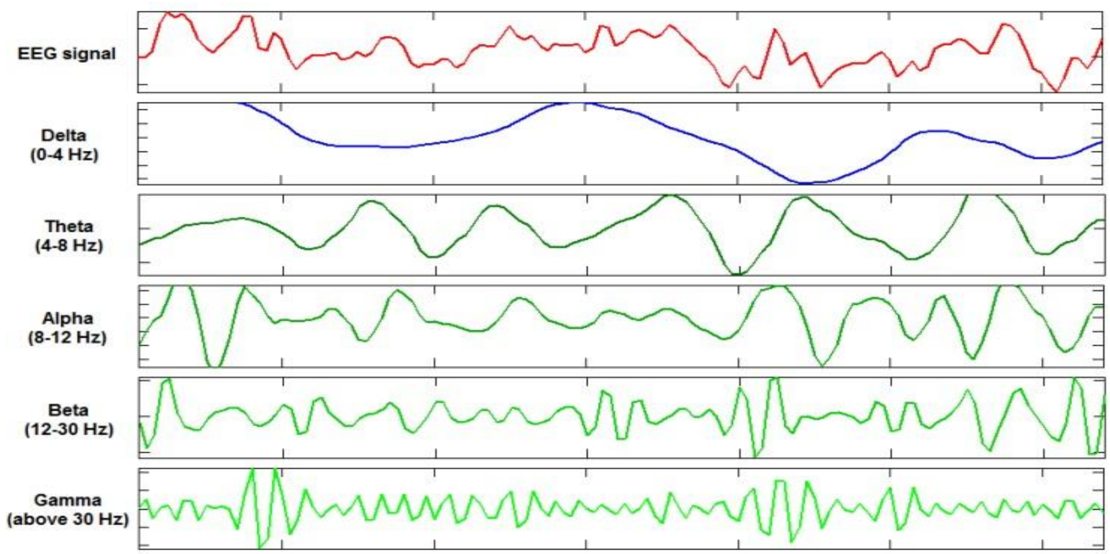


Figure 1. 3: EEG Frequency/Amplitude Bands

1.6. Background Of Emotion

Emotion plays a crucial role in human behavior and communication. It is a complex interplay of subjective and objective experiences, mediated by neural networks. Emotion also influences the brain's state, impacting various processes and serving as a key factor in human life. Given that EEG is a non-invasive tool for measuring brain activity with millisecond temporal resolution, physiological signals like EEG are frequently utilized in emotion research.

Emotions are characterized and described using both discrete and dimensional models. For instance, Ekman et al. (2013) proposed discrete categories for emotions. Fig. 1.4 illustrates Ekman's list of six universal emotions, Plutchik's emotion wheel and Russell's valence-arousal scale (Russell, 1980) are examples of three-dimensional emotion models. Due to its simplicity and universal applicability, the Russell model is widely used in emotion recognition research. Fig. 1.5 shows the Russell model. In this study, the Russell Valence-Arousal model is employed to recognize four emotions: happy, angry, sad, and relaxed.

Figure 1. 4: Ekman's six universal emotions

1.7. Emotion Representation

Emotions need to be identified and quantified for effective recognition. The first proposal for classifying basic emotions dates back several decades (Shu et al., 2018). However, psychologists have yet to reach a consensus on a definitive description. Emotions are typically modelled in two ways: one approach categorizes emotions into distinct groups, while the other uses various dimensions to describe them.

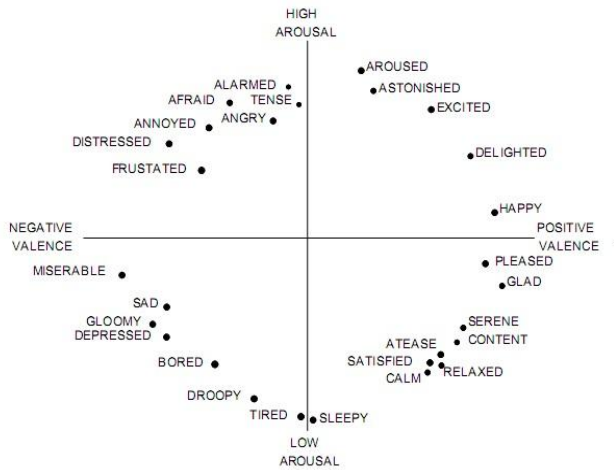


Figure 1. 5: Russell valence –Arousal Model

During emotion elicitation, subjects are presented with emotionally evocative materials designed to provoke specific emotions. In recent years, pictures, music, and films have been the most commonly used stimuli. Additionally, new methods, referred to as situational stimulation, have become increasingly popular. Video games and reminiscence are examples of such stimuli. One particularly promising approach gaining traction is affective virtual reality.

1.7.1. Discrete Emotion Models

Shu et al. (2018) states that emotions are distinct, measurable, and closely connected to physiological responses. He identified several characteristics associated with fundamental emotions, including: (I) emotions are innate; (II) individuals experience the same emotions in similar situations; (III) emotions are expressed in comparable ways across people; and (IV) physiological responses to the same emotions are consistent. He also suggested that other emotions arise as evolved responses or combinations of these fundamental emotions.

1.7.2. Multi-Dimensional Emotion Space Model

Subsequent psychological studies have explored the relationships between various emotions, such as the connections between dislike and liking or hatred and fury, as well as variations in emotional intensity. For instance, happiness can range from mild contentment to extreme joy. In response, psychologists have developed multifaceted models of emotion to better capture these complexities. Lang introduced a two-dimensional framework using valence and arousal to classify emotions. In this model, arousal represents the intensity of emotional experiences, ranging from passive (low) to active (high), while valence indicates the emotional quality, spanning from negative (unpleasant) to positive (pleasant). Different emotions can be mapped inside this 2D space, as shown in Fig. 1.6. For instance, rage has a high arousal and negative valence, whereas depression has a low arousal and negative valence.

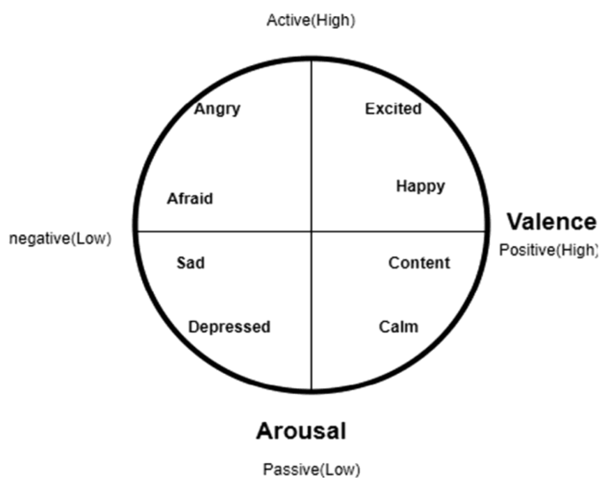


Figure 1. 6: Valence-Arousal model of emotions

Emotion Recognition (ER) plays a vital role in HCI, enabling machines to interpret human emotional states effectively. Advances in computer science, cognitive science, modern neuroscience, and psychology have all contributed to the creation of ER (Alhalaseh & Alasasfeh, 2020; Nandi et al., 2021). Computer scientists have developed automatic emergency response systems to improve HCI in a variety of fields, such as gaming, healthcare, industrial applications, and the military. In general, there are two primary categories of human emotion identification methods: (i) physiological signals, including electromyography (EMG), electrocardiography (ECG), and EEG; and (ii) non-physiological signals, like body language, facial expressions, and voice (Wang & Guo, 2020; Tan et al., 2021). Among these, EEG stands out as the most widely used physiological signal for ER, as it provides an accurate representation of human mental states and offers real-time brain activity monitoring (Tuncer et al., 2021; Yang et al., 2021). Advancements in EEG electrode technology have significantly increased the adoption of

EEG-based ER systems (Bălan et al., 2019). Emotion classification primarily relies on two approaches: Discrete Basic Emotion Description (DBED) method and the dimensional approach (Islam et al., 2021; Hu et al., 2021). The DBED method categorizes emotions into six fundamentals (An et al., 2021; Liu & Fu, 2021; Song et al., 2021). In contrast, the dimensional approach represents emotions along continuous scales, such as arousal and valence, to capture variations in emotional states (Shen et al., 2021; Aslan, 2022). Recently, DL models have gained substantial attention in ER due to their scalability and reliability (Liu et al., 2020; Koelstra et al., 2011). This study intends to address non-stationary nature of EEG signals and boost classification performance by adopting a unique optimization-driven DL model. The following is a summary of this manuscript's main contributions:

The raw EEG signals are sourced from DEAP database. To ensure the signals are clean and suitable for processing, Integrated EMD and VMD filters are applied to remove noise and artifacts from collected EEG data.

Among 20 statistical features extracted from the decomposed EEG signals are Hjorth Activity, Mobility and Complexity, Sample and Shannon Entropy, standard deviation, variance, Mean Curve Length, Normalized First Difference, Mean Teager Energy, Auto-Regressive Model, Zero Cross Rate, Band Power for delta, theta, alpha, beta, and gamma frequencies. Additional features include the Ratio of Band Power (alpha/beta), Minimum, and Maximum. Statistical feature extraction offers several advantages, such as reducing the risk of overfitting, enhancing accuracy, improving data visualization, and accelerating training process of the LSTM model.

Human emotions are classified using LSTM networks based on arousal and valence dimensions. IRSOA technique optimizes the selection of LSTM hyperparameters, thereby improving both processing time and computational efficiency. To evaluate the effectiveness of the IRSOA-LSTM network, various performance metrics are employed.

1.8 Deep Learning Methods for Emotion Recognition

Emotions are an essential part of human communication, influencing decision-making and behavior. The ability to recognize emotions using EEG signals has gained attention, particularly in Brain-Computer Interface (BCI) systems. EEG-based emotion recognition is crucial for improving human–machine interactions and has applications in healthcare, cognitive research, and HCI.

Despite significant advancements, emotion recognition still faces challenges, particularly in achieving high accuracy. To enhance recognition performance, Deep Learning (DL) models have been widely adopted. Several studies have explored different methods, including feature extraction, feature selection, and classification techniques, to improve the efficiency of emotion recognition systems.

1.8.1. Purpose Of Preprocessing In Emotion Recognition

Preprocessing is a fundamental step in EEG-based emotion recognition, ensuring that raw EEG signals are cleaned, structured, and optimized for feature extraction and classification. Since EEG signals are highly susceptible to noise, artifacts, and interference, effective preprocessing enhances the accuracy and reliability of emotion recognition models.

1. Filtering EEG Signals:

EEG signals often contain noise from external sources, muscle movements, and electrical interference, requiring filtering techniques for noise elimination. Bandpass filtering (0.5–45 Hz) removes frequencies outside the relevant EEG range, preserving brainwave activity across delta, theta, alpha, beta, and gamma bands. Notch filtering (50/60 Hz) eliminates powerline interference commonly present in EEG recordings due to electrical noise. Independent Component Analysis (ICA) separates EEG signals into independent sources, enabling the removal of artifacts such as eye blinks, muscle movements, and heartbeats, ensuring cleaner and more reliable EEG data for analysis.

2. Signal Decomposition

EEG signals are often non-stationary, with frequency content changing over time, requiring decomposition methods to extract meaningful components. VMD and EMD break EEG signals into multiple Intrinsic Mode Functions (IMFs), isolating emotion-relevant patterns, while Wavelet Transform (WT) analyzes signals across different frequency scales, capturing both time and frequency characteristics essential for emotion recognition.

3. Standardization and Normalization

EEG signals vary in amplitude across individuals and recording sessions, requiring standardization for consistency. Z-score normalization centers the data around zero, reducing bias from varying amplitudes, while Min-Max scaling rescales data within a fixed range (e.g., [0,1]), preventing domination by high-amplitude signals and improving model performance.

1.8.2 Feature Extraction in EEG-Based Emotion Recognition

Feature extraction is essential for converting raw EEG signals into meaningful representations that is processed by Machine Learning (ML) and DL models. Several types of features are commonly extracted in EEG-based emotion recognition to capture meaningful patterns from brain signals. Statistical features, such as mean, standard deviation, and entropy, help characterize EEG signal variations. Wavelet features utilize WT to analyze frequency-based changes in EEG data. The Hurst exponent measures long-term dependencies in brain activity, providing insights into cognitive and emotional states. Additionally, frequency band analysis decomposes EEG signals into theta, alpha, beta, and gamma bands, each associated with different emotional and cognitive processes. These extracted features serve as inputs for classifiers, enabling effective differentiation between emotional states.

The extracted features serve as inputs to ML/DL models, helping distinguish between different emotional states. Without effective feature extraction, raw EEG signals contain too much noise and redundant information, making it difficult for classifiers to learn meaningful patterns. By selecting the most relevant features, emotion recognition systems can achieve higher accuracy and better generalization across different datasets.

1.8.3. Classification Based on Emotions

DL models play a crucial role in EEG-based emotion recognition by automatically extracting meaningful patterns from EEG signals. Unlike traditional ML approaches, which rely on handcrafted features, DL models learn spatial and temporal representations directly from raw or pre-processed EEG data, enhancing classification performance.

Convolutional Neural Network (CNN)

CNNs are widely used for feature learning and classification in EEG-based emotion recognition due to their ability to capture spatial and temporal patterns. EEG signals are represented as 2D feature maps or converted into spectrograms, enabling CNNs to extract both local and global patterns. Through spatial feature extraction, CNNs learn dependencies between EEG channels, capturing relationships across different brain regions. Temporal feature learning is achieved by applying convolutional filters across time-series EEG data, allowing the model to recognize fluctuations in brain activity associated with emotional states. The structured feature extraction process enhances the model’s ability to differentiate between various emotional states with high accuracy.

Long Short-Term Memory (LSTM) and Bi-Directional LSTM (Bi-LSTM)

Recurrent Neural Networks (RNNs) (Gilakjani & Al Osman, 2023) process sequential EEG data, with LSTM and Bi-LSTM playing a key role in emotion recognition. LSTM captures long-term dependencies in EEG sequences, mitigating the vanishing gradient problem seen in standard RNNs. Bi-LSTM enhances this by processing EEG signals in both forward and backward directions, improving context-awareness in recognizing emotional states. Stacked Bi-LSTM, a deeper architecture with multiple Bi-LSTM layers, enables the model to learn complex relationships in EEG signals across valence, arousal, and liking dimensions, leading to more accurate emotion classification.

Hybrid Deep Learning Models

Hybrid models integrate multiple deep learning architectures to enhance classification accuracy and robustness. ResNet-50 (Chao et al., 2019) based CNNs utilize skip connections to retain information across layers, enabling deeper hierarchical feature extraction while mitigating gradient vanishing issues in EEG-based emotion recognition. Ensemble learning methods further improve classification by combining predictions from multiple models; bagging reduces variance and enhances stability by aggregating classifiers, while boosting techniques like Extreme Gradient Boosting (XGBoost) (Pan et al., 2020; Khateeb et al., 2021) iteratively refine weak classifiers for better performance, and Adaptive Boosting (AdaBoost) focuses on difficult samples, improving the model's ability to distinguish similar emotional states.

Fine-tuning DL models for ER involves optimizing a well-defined objective function using specialized algorithms to enhance performance. The objective function serves as a guiding metric that evaluates model performance, while optimizers iteratively adjust the model parameters to achieve the best possible accuracy.

1. Objective Function in Emotion Recognition

The objective function is a mathematical function that measures the effectiveness of a deep learning model in predicting emotional states from input features. In ER, the objective function is often a loss function designed to minimize classification errors in valence-arousal detection.

Mathematically, the objective function $J(\theta)$ is defined in Eqn. (1.1).

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N} L(y_i, \hat{y}_i) \tag{1.1}$$

where, N is total number of training samples, y_i is the actual emotion label, $y^{\wedge}{}_i$ is the predicted label, $L(y_i, \hat{y}_i)$ represents the loss function. For classification-based valence-arousal prediction Mean squared Error (MSE) is used which is expressed in Eqn. (1.2).

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{1.2}$$

Once an objective function is defined, optimizers are used to adjust DL model parameters (such as weights and biases) to minimize the objective function. This process is represented in Eqn. (1.3).

$$\theta^* = \arg\min_{\theta} J(\theta) \tag{1.3}$$

where, θ^* represents the optimized parameters that yield the best performance. Different optimizers use various techniques to update the model parameters and minimize $J(\theta)$.

2. Optimization Algorithms for Emotion Recognition

(a) Particle Swarm Optimization (PSO)

PSO is a metaheuristic optimization technique inspired by swarm intelligence. It optimizes deep learning hyperparameters, such as learning rate, dropout rate, and the number of layers, by simulating the behaviour of particles searching for an optimal solution. PSO updates model parameters using Eqn. (1.4).

$$v_i^{t+1} = wv_i^t + c_1 r_1 (p_i - x_i) + c_{2r_2} (g - x_i) \tag{1.4}$$

where, v_i is velocity of a particle, $x_i^{t+1} = x_i^t + v_i^{t+1}$ is the position, p_i is a best position found by a particle, g_i is the best global position found, w, c_1, c_2 are hyperparameters (Li et al., 2022).

(b) Grey Wolf Optimizer (GWO)

GWO is inspired by the hunting behavior of wolves. It adjusts deep learning parameters by mimicking the social hierarchy of wolves, where alpha, beta, and delta wolves guide the search. Parameter updates are performed are given in Eqn. (1.5).

$$X(t+1) = X_a - A.D_a \tag{1.5}$$

where, $X(t+1)$ is the updated parameter, X_a represents the best solution (alpha wolf), A and D_a are control vectors guiding the optimization (Chen, 2023).

FFO is inspired by firefly behavior in nature, where fireflies move toward brighter individuals. It is used for feature selection and parameter tuning in deep learning models. The update equation is expressed in Eqn. (1.6).

$$X_i^{t+1} = X_i^t + \beta e^{-\gamma r^2}(X_i - X_j) + \alpha \in \tag{1.6}$$

where, X_i, X_j are firefly positions, β is attractiveness, absorption coefficient as γ, $\alpha \in$ is a randomization factor (Abgeena & Garg, 2023).

(c) Rat Swarm Optimization (RSO)

RSO is a bio-inspired metaheuristic optimization algorithm that mimics the behavior of rats searching for food. It is effective for hyperparameter tuning in DL. RSO models rats as agents searching in a solution space, where the best solution (i.e., the lowest loss) corresponds to the best food source. The movement of rats is influenced by a leader-following mechanism which is given in Eqn. (1.7).

$$X_i^{t+1} = X_i^t + A.(X_L - X_i) + B.(X_R - X_i) + C.(X_G - X_i) \tag{1.7}$$

where, X_i^{t+1} is the position of the i^{th} rat in iteration t, X_L is the position of the leading rat, X_R is a randomly selected rat, X_G is the global best solution, A, B, C are adaptive coefficients controlling movement.

(d) Tunicate Swarm Optimization (TSO)

TSO is inspired by tunicate swarm behaviour, where tunicates use jet propulsion and swarming to move towards optimal food sources. This optimization technique is highly effective in fine-tuning deep learning models. TSO algorithm updates the position of tunicates based on Eqn. (1.8).

$$X_i^{t+1} = X_i^t + S.(X_B - X_i) + W.(X_G - X_i) \tag{1.8}$$

where, X_i^t is the position of the i^{th} tunicate in iteration t, X_B is the best tunicate, X_G is the global best solution, S and W are adaptive coefficients representing tunicate swarming behaviour.

However, PSO, GWO, FFO, RSO, and TSO each face challenges limiting their optimization performance. PSO struggles with premature convergence, parameter sensitivity, and scalability issues (Yin et al., 2023). GWO suffers from local optima stagnation, slow convergence, and weak local search capability (Zhang et al., 2023). FFO's computational complexity increases with population size and is highly sensitive to the light absorption coefficient (Wen & Zhou, 2024). Among the discussed optimization algorithms, RSO and TSO are selected due to their superior adaptability in balancing exploration and exploitation, making them more suitable for complex optimization tasks. However, to overcome their limitations, such as susceptibility to local optima and sensitivity to initial parameters (Abdulla et al., 2024; Guodong et al., 2024; Si et al., 2024), an improved version of RSO and TSO is developed using an adaptive mechanism. This enhancement dynamically adjusts key parameters, ensuring better convergence stability, improved global search efficiency, and enhanced performance in high-dimensional optimization problems.

1.9. DEAP Database

The DEAP multimodal dataset is used to study human affective states. It contains peripheral physiological and EEG signal data collected from 32 participants while they viewed 40 one-minute music video segments. Participants rated each video based on like/dislike arousal, superiority, valence and familiarity. Additionally, 22 of 32 individuals had frontal face video recordings (Koelstra et al., 2012).

- The dataset consists of 40 electrodes—32 for recording EEG signals and 8 for other physiological signals.
- There are 32 .mat files, each corresponding to one subject, along with a metadata file containing the participants' ratings.
- The data dimensions for a single subject are 40×40×8064, where the first dimension represents the video clips, the second dimension corresponds to the electrode

positions, and the third dimension shows the voltage readings. Graphical representation of DEAP dataset is given in Fig. 1.7.

Characteristics of the DEAP Dataset for Emotion Recognition

- Sequential EEG Data: The dataset consists of 40×40×8064 dimensions per subject, where the last dimension corresponds to time-series EEG signals. This makes it highly suitable for models like LSTM and BiLSTM, with improved optimizations which specialize in learning long-term dependencies from sequential data.
- High-Dimensional Features: The dataset contains 40 electrode channels, requiring models that can extract spatial and temporal patterns, making CNN+LSTM an ideal choice for hybrid feature extraction.
- Affective Labels: ER in DEAP is based on valence, arousal, liking, and dominance scores, which require sophisticated models capable of capturing non-linear relationships between brain signals and emotional states.

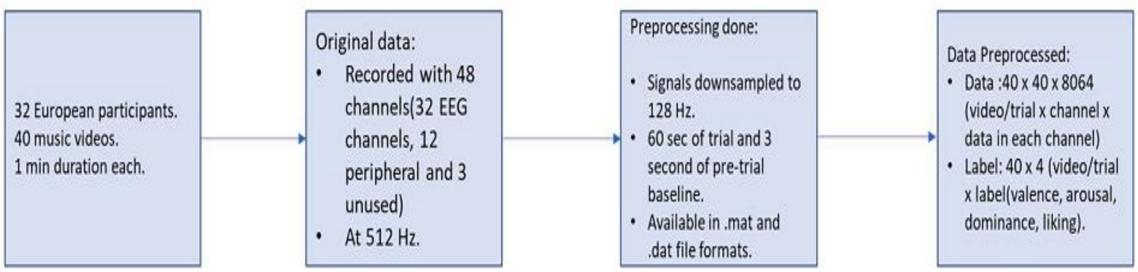


Figure 1. 7: Overview of DEAP dataset

Table 1. 1: Limitations of traditional classifiers on the DEAP Dataset

Model	Limitations faced by traditional classifiers using DEAP Dataset
SVM	Requires handcrafted features and struggles with high-dimensional EEG data, leading to suboptimal generalization.
k-NN	High computational cost due to the large dataset size, and it lacks the ability to learn temporal features from EEG signals.
Random Forest	Cannot capture sequential dependencies in EEG signals, as it treats features independently without considering time-series relationships.
Traditional RNN	Suffers from the vanishing gradient problem, making it difficult to retain long-term dependencies in EEG sequences, which are crucial for emotion recognition.

The DEAP dataset consists of high-dimensional, sequential EEG data, making it challenging for traditional classifiers to effectively process and classify emotions which is given in Table 1.1

1.10. Objectives

The main objective of this research is to design and implement advanced signal processing and DL techniques to effectively extract and classify emotional states from EEG data in the DEAP dataset. The goal is to improve accuracy and robustness in emotion recognition by leveraging optimized feature extraction, selection, and classification strategies.

1. **Novel Multidimensional EEG Signal Processing Method**:

DEAP dataset, requiring effective signal decomposition techniques to remove noise and extract meaningful emotional features. A robust method is developed using a hybrid approach of EMD and VMD to eliminate artifacts and enhance relevant EEG components. The extracted features are classified using an LSTM network, which captures long-term dependencies in EEG sequences to improve the accuracy of ER.

2. **Binary Classifier for Emotion Classes:**

A hybrid optimization algorithm, MTSOA, is developed to select the most discriminative features from the extracted EEG data, reducing computational complexity and improving model efficiency. A Bi-LSTM model is designed for classification, leveraging its bi-directional processing capability to capture EEG patterns in both forward and backward directions. To further enhance performance, IRSOA is applied for hyperparameter tuning, optimizing the model's learning rate, dropout rate, and hidden layer configurations, leading to improved accuracy and robustness.

3. **Hybrid CNN-LSTM for Arousal and Valence Detection:**

The dataset includes EEG signals associated with emotional arousal and valence**,** necessitating models that effectively capture both spatial and temporal dependencies. To address this, a hybrid CNN-LSTM is developed, where CNN extracts spatial features by identifying relationships between different electrode positions, while LSTM captures temporal dependencies, learning the emotional states evolve over time. To further enhance performance, the model is refined using IRSOA**,** which optimizes feature selection and hyperparameters, ensuring high-precision detection of emotional states like arousal and valence with improved reliability.

1.11 Organization of Thesis

The thesis is structured into six chapters, each centered on the design and performance evaluation of a hybrid DL method for recognizing various classes of human emotions. The overall organization of a thesis is depicted in Fig. 1.8.

Chapter 1 lays the foundation for this research by providing an in-depth overview of background and concepts related to ER and EEG. It begins by discussing the functional organization of the brain and the role of EEG in understanding neural activity. The chapter introduces the basics of EEG, including its electrode systems and rhythms, which are crucial for capturing brain signals. Additionally, it explores the concept of emotions, their representation, and the models used to study them. Two prominent approaches to emotion representation are highlighted: discrete emotion models and the multi-dimensional emotion space model. This chapter serves as a comprehensive introduction to the core topics underpinning the research.

Chapter 2 reviews the existing research papers on EEG-based ER, summarizing various feature extraction techniques and DL approaches. It highlights the applications of EEG as a motivation for the proposed study and outlines the problem statement and objectives. A detailed description of the DEAP dataset, central to this research, is provided. The chapter concludes with a comparative analysis of existing models, emphasizing their limitations and setting the stage for the proposed framework.

Chapter 3 presents the hybrid DL methodology for ER. It details the preprocessing steps applied to the EEG data, feature extraction techniques, and the architecture of the hybrid model. The chapter explains the integration of DL components and it outlines the training and validation strategies employed to optimize the model's performance.

Chapter 4 presents the proposed methodology for enhancing EEG-based ER using MTSOA and RSOA. It begins by addressing the limitations of existing methods in managing context information in EEG signals and introduces automated approach for improved feature selection and hyperparameter tuning.

Chapter 5 presents the implementation and evaluation of the proposed hybrid IRSOA-CNN+LSTM framework for EEG-based emotion recognition. It begins by addressing the limitations of prior models in capturing the context information of EEG signals and introduces a novel automated approach for improving the recognition process. The chapter concludes with an evaluation metrics used to assess the effectiveness of the proposed approach.

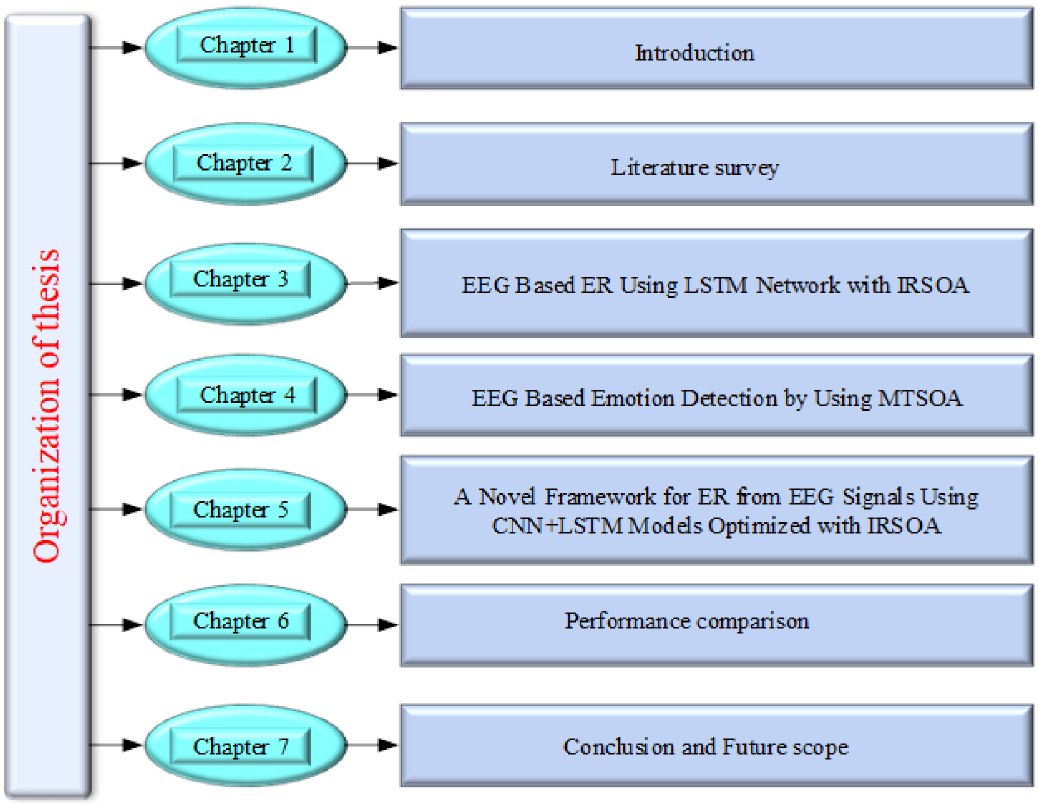


Figure 1. 8: Overview of thesis

Chapter 6 presents a comprehensive comparative analysis of three proposed hybrid methods for EEG-based emotion recognition, focusing on their performance in classifying emotions into arousal and valence classes. The chapter aims to evaluate strengths and limitations of each method, providing insights into their effectiveness in handling EEG signals for emotion detection.

Chapter 7 concludes the research by summarizing the key findings and contributions made in field of EEG-based ER. The chapter then discusses the limitations encountered during the research, such as the dependency on the DEAP dataset. Finally, the future scope of the research is outlined, suggesting potential directions for further improvement.

CHAPTER 2

Literature Review

2.1 Review of Emotion Recognition Through EEG Signal

The complexity and nonlinear nature of EEG signals have driven extensive research on automated analysis techniques. Artificial intelligence and machine learning have been widely applied to enhance EEG signal interpretation, particularly in BCI applications. Researchers have explored various feature extraction methods, including time-frequency analysis, to improve classification accuracy. However, subject-specific variations pose challenges, highlighting the need for personalized algorithms.

Beyond neurological diagnosis, EEG has gained prominence in monitoring tasks such as emotion state recognition and identity verification. Since emotions originate from brain activity, EEG provides a direct measure, making it a powerful tool for emotion recognition. Researchers extract key features from EEG data, and studies have identified gaps in existing approaches while proposing improvements. This chapter reviews advancements in EEG-based classification, with a focus on emotion recognition, emerging challenges, and potential research directions.

(a) The model should be clear and include the fewest possible EEG channels.

(b) To improve model efficiency and minimize noise, highly effective filtering techniques need to be proposed, as real-time data often contains significant artifacts.

(c) Various feature extraction techniques and ML classifiers were applied to analyze different datasets, with the integration of ensemble methods and DL architectures further enhancing model accuracy.

(d) Exploring non-linear and functional connectivity aspects can provide valuable insights into the dynamics of EEG data, opening up new possibilities for diverse applications.

Many excellent experiments have already been conducted using EEG signals in field of signal processing.

1. Tan et al. (2020) presented SNN-based approach for multimodal ER, leveraging NeuCube's evolving architecture to classify emotional valence. Their model effectively handled spatio-temporal data, demonstrating competitive accuracy

without relying on EEG signals. Evaluations on the MAHNOB-HCI dataset achieved 73.15% accuracy using feature-level fusion. The study highlighted SNN's adaptability and efficiency, making it suitable for real-time applications compared to traditional DL methods.

2. Ullah et al. (2019) suggested an Ensemble-Learning Algorithm (ELA) to automatically select discriminative EEG channel subsets for emotion recognition. This approach not only enhances classification accuracy but also improves computational efficiency and reduces processing time. Experiments using the DEAP database showed that ELA model outperformed previous models. However, model's performance came at the cost of increased memory usage and longer processing times.
3. Using EEG signals, Krishna et al. (2019) developed the Tunable QWT (TQWT) method for categorizing human emotions. Several features, including mobility, root mean square, activity, crest, clearance, shape factor, average change in amplitude, log identifier, absolute sum and square root are extracted from the raw EEG signals in this study using TQWT method, which breaks them down into sub-bands. To categorize emotions as sad, fear, happy, or relaxed, these extracted feature vectors are subsequently fed into an Extreme Learning Machine (ELM). According to simulation studies, the suggested model fared better at classifying emotions than current methods. However, the framework's computational cost rose when more than one model was added.
4. Liu et al. (2020) used CNN, sparse based encoder, and Deep Neural Network (DNN) to create a novel hybrid DL model in EEG-based emotion categorization. A sparse autoencoder is used to encode and decode the discriminative deep features that the CNN first derives from the gathered databases. The DNN is then fed the processed data to classify the emotions. The efficacy of this hybrid DL model in emotion recognition was shown by experiments using the DEAP database. However, the model turned out to be costly for practical implementations.
5. Zeng et al. (2019) employed Sinc-Net, a model consisting of three deep convolutional layers, to classify emotions. However, implementation results revealed that this model outperformed traditional classifiers in terms of classification performance and

demonstrated faster convergence. However, the model required a large amount of training data, making the computational cost quite high.

6. Bajaj et al. (2018) presented an emotion-based BCI system using EEG signals for classifying four emotions by introducing Flexible Analytic WT(FAWT) for feature extraction, which decomposes EEG signals into sub-bands, capturing emotion-specific characteristics. The classification performance was evaluated using variants of K-Nearest-Neighbor (KNN) classifier, where weighted-KNN achieved the highest accuracy of 85.1%. The study demonstrated that FAWT-based approach outperformed existing four-class emotion recognition methods, highlighting its effectiveness for emotion classification in BCI applications. Additionally, Kim and Choi (2020) proposed an LSTM network with attention mechanisms for EEG-based emotion classification, emphasizing the need to address class-imbalance issues as a future improvement.
7. Chakladar and Chakraborty (2018) employed Higher-Order Statistics (HOS) combined with correlation subset selection in reduction of dimensionality and ER. Their approach successfully distinguished four types of emotions: harmony, anger, negative, and positive. However, the model, which utilized a kernel-based SVM, faced challenges in effectively handling non-linear data.
8. Yin et al. (2021) designed a method for ER, combining LSTM with Graph CNN (GCNN). The LSTM network was employed to capture the temporal relationships between extracted feature vectors, while GCNNs were utilized to extract graph-domain feature vectors. The emotion classification results were derived using a dense layer. Experiments on the DEAP database demonstrated that the proposed framework outperformed existing methods in terms of classification accuracy. However, the main challenges identified in the study were vanishing gradients and overfitting.
9. A lightweight human emotion identification system developed by Subasi et al. (2021) follows four key steps: pre-processing, feature extraction, dimensionality reduction, and emotion classification. The system begins by removing artifacts then, feature extraction and dimensionality reduction performed using TQWT alongside six different techniques. For emotion recognition, various ML classifiers were tested, with the rotating forest ensemble combined with an SVM classifier yielding the best

results. However, a major challenge with MSPCA was balancing dimensionality reduction with the risk of losing valuable information.

10. Salankar et al. (2021) implemented decomposition of signal using EMD, followed by use of second-order difference plots for feature extraction. To ensure statistical significance of extracted features, the Wilcoxon test was conducted, confirming a p-value below 0.05. For classification, both Multilayer Perceptron (MLP) and an SVM were employed for binary and multi-class classification. However, the MLP required careful hyperparameter tuning, including adjustments to the number of iterations, neurons, and hidden layers, as it was highly sensitive to feature scaling.

11. Gao et al. (2022) conducted extraction using PSD, sample and differential entropy which was then given into a SVM classifier for emotion classification. When tested on DEAP database, the suggested model outperformed other classifiers. However, despite its effectiveness, SVM was found to be unsuitable for multi-class emotion identification due to its inherent limitation of supporting only binary classification.

12. Sharma et al. (2020) integrated a DL model with higher-order statistics for automated emotion classification. Similarly, Yang et al. (2019) presented a multi-column CNN approach for classifying emotions using signals. While these DL methods demonstrated strong classification performance, they were computationally expensive, requiring high-end systems to achieve optimal results.

13. Mert and Akan (2018) enhanced emotion identification, specifically for arousal and valence, by combining Multivariate EMD(MEMD) with frequency-domain features such as Power Spectral Density (PSD), parameters of Hjorth, power ratio, entropy. This approach improved the efficiency of emotion classification.

14. Song et al. (2018) inaugurated a multi-channel EEG based ER using an iterative Dynamic GCNN (DGCNN) model. The extensive experimental results have established that DGCNN model provides even better recognition than other conventional techniques. However, DGCNN model has important downsides such as overfitting issues and slow convergence.

15. Chen et al. (2018) hybridized approximate entropy and EMD in extraction of features extraction and signal decomposition which were then fed into a hybrid model for classification integrating SVM and Deep Belief Networks (DBN). While this approach enhanced classification performance, the DBN was computationally expensive due to its high data requirements for achieving superior accuracy.

16. Chao et al. (2019) enhanced emotion classification by integrating multi-band feature matrices with Capsule Networks (CapsNet). While the proposed CapsNet model improved classification efficiency, its real-world implementation proved to be highly expensive.

17. Zhang et al. (2020) implemented DNNs with Kernel Machines (KM) in recognition of emotions. An exhaustive experimental analysis proved that the proposed model accomplished well in modelling the interactions within the multi-modal physiological EEG data for even better recognition performance

18. Pan et al. (2020) utilized wavelet packets, fuzzy samples and approximate entropy for feature extraction then fed into an ensemble machine classification method, combining SVM, ELM, and Decision Trees (DT) for emotion categorization.

19. Taran and Bajaj (2019) proposed an audio-video stimulus method used to set up EEG recordings of positive, anxious, sad, and relaxed emotions, and a two-stage filtering method for emotion EEG signal recognition is proposed. Initially, a correlation criterion is proposed for removing noisy IMFs from IMFs obtained by EMD on the raw EEG signal. Denoised EEG signal is reconstructed with improved characteristics of stationarity using noise-free IMFs. In the second level of decomposition, denoised EEG signal is decomposed further into modes VMD technique. The filtered modes are retained for reconstructing the denoised EEG signal within required frequency range. After this two-stage filtering process, non-linear features of the filtered EEG signals are given into a SVM for emotion classification.

20. Pandey and Seeja (2019) proposed an independent ER in which features derived from IMFs such as PSD peak value and first difference were used for ER. To extract IMFs from EEG signals, they employed two different techniques: EMD and VMD. The proposed system was evaluated using the benchmark DEAP dataset, where features extracted from top three IMFs using VMD outperformed those based on EMD. For subject-independent emotion recognition, DNNs demonstrated superior performance over SVM classifiers. Since the model was trained and tested using EEG data from diverse subjects, the proposed approach is a versatile emotion identification system, capable of recognizing emotions from an individual's EEG even if their data was not included in the training process.

21. Salankar et al. (2021) developed a highly effective classifier for accurately identifying emotions and their corresponding quadrants in EEG signals. The signals are first decomposed into oscillatory IMF using EMD as a preprocessing technique. Second-order difference plots (SODP) were then employed to calculate tendency measure of elliptical region. For both emotion classification across valence, arousal, superiority, and liking quadrants, MLP utilizes an SVM with two hidden layers. The models' performance is evaluated using statistical metrics such as accuracy, specificity, and sensitivity. The MLP outperformed the SVM, achieving a maximum classification accuracy of 99.99% with high and very low arousal space.

22. Fathalla and Alshehri (2020) reviewed and summarized the processes, techniques, and methods of emotion detection. The paper highlights EEG-based ER as an example of emotion detection approaches, outlining necessary steps beginning with acquisition of EEG signals during emotion elicitation process. It then describes extraction of features using various techniques like EMD and VMD. Finally, the paper discusses the use of different classifiers, including SVM and DNN, to categorize emotions.

23. Nawaz et al. (2020) employed the three-dimensional (arousal, valence, and dominance) model of emotion to determine feelings elicited by music videos. During EEG recording, participants were asked to watch a one-minute video, and the study aimed to identify which features most effectively differentiate between various emotions. The study examined several features about EEG signals, including wavelet energy, statistical features, fractal dimension, entropy, and power. To validate the best features, the researchers used two feature selection techniques: PCA and a relief-based algorithm. Various classifiers were employed to assess the efficacy of the selected features. PCA was found to provide enhanced performance. The system achieved overall classification accuracies of 73.62% for valence, 79.95% for arousal, and 79.60% for dominance which demonstrated that time-domain statistical features of EEG signals were effectively distinguished between various emotional states. Additionally, the three-dimensional emotion model helped differentiate similar emotions that might be mislabeled in a two-dimensional framework (e.g., fear vs. anger). These findings hold potential for developing real-time EEG-based emotion identification systems.

24. Zhuang et al. (2017) proposed an EMD technique for feature extraction and ER. EEG signals are automatically decomposed into IMF using EMD. The multidimensional features of each IMF include normalized energy, first difference of time series and phase. This approach was evaluated on a publicly available emotion-based database, and outcomes showed that these collected three features are effective in emotion detection. Upon examining role of each IMF, it was found that the high-frequency component, IMF1, plays a crucial role in identifying different emotional states. The EMD technique was also applied to analyze informative electrodes. Additionally, classification accuracy of suggested method was compared with other conventional techniques. Experimental evaluation on DEAP dataset depicted that proposed approach enhances precision of emotion recognition.

25. Bhardwaj et al. (2015) classified EEG signals into seven distinct emotions using ML methods like SVM and Linear Discriminant Analysis (LDA) and performance achieved by both algorithms was calculated and compared. With SVM and LDA, the system identified the seven emotions with an average overall accuracy of 72.13% and 67.50%, respectively. High precision opens up future applications in emotion detection, including music players, EEG-based music therapy, neuromarketing, and market research.

26. Doma and Pirouz (2020) and El Ayadi et al. (2011) conducted a comparative study of various ML approaches, including SVM, KNN, LDA, Logistic Regression, and DT, using iteration data from sensor channels of EEG. Each model was evaluated with and without PCA for reducing dimensionality. A key aspect of the study was the application of grid search hyperparameter tuning, implemented on the Spark cluster, which significantly reduced execution time. DEAP a multimodal dataset, was employed to analyze human affective states. Participants rated 40 one-minute musical samples and classifiers were trained for each of four classes using time-segmented, 15-second iteration of data. Hybridized PCA- SVM yielded a best result in 31th to 46th interval of segmentation, achieving an accuracy of 84.73% and a recall of

98.01%. The study found that a common classification model converged to higher performance for each time segment and binary training class, highlighting the need for different models to distinguish between various states of emotion.

27. Zangenehet al. (2018) proposed several emotional concepts and methods related to emotion recognition systems. They explored various topics, including different physiological measurements such as ECG, and the distinction between offline and online recognition systems. The discussion also covered emotional stimulus types and fundamental emotion models like the valence-arousal model and the discrete model. The primary focus of their work was on emotion recognition using EEG data, which gained significant popularity in recent years for ER applications.

28. Li et al. (2018) proposed a mode determination approach based on a similarity between input signals and a sum of mode functions using difference of average signal method to determine mode number for VMD technique. This approach was validated with real-time high-speed gear and bearing defect data which demonstrated that suggested method could accurately determine number of modes for VMD technique in rotating machinery applications, offering improved performance when compared to the traditional EMD process in frequency spectrum diagnosis.

29. To enhance emotion recognition performance, Liu et al. (2019) developed a hybrid approach to feature extraction in the EMD domain that combines the ability to extract suitable information from multiple components while eliminating redundant features with the best feature selection technique using Sequence Backward Selection (SBS). The method was assessed using the DEAP dataset, where KNN and SVM were employed to distinguish emotional states in the both states, respectively. The study compared three distinct EEG signal rhythms and three temporal windows of varying durations. The outcome indicated that the EEG signal with a one-second temporal-based window outperformed several state-of-the-art approaches, achieving 83.46% highest accuracy in the Valence state and 82.90% in the Arousal dimension. This technique shows promise for real-time ER in a multimodal framework for emotional communication-based human-robot interaction.

30. To improve subject-independent recognition performance, Cimtay et al. (2020) proposed using pre-trained conventional CNN architectures. They used raw EEG data, performing windowing, pre-adjustments, and normalization, in contrast to other methods which extracts spectral power information from signals. This method improves DNN's capacity to find unknown attributes by eliminating human features from training process, which lowers the possibility of overlooking hidden patterns in raw data. During emotion prediction, a filter was used to remove data across predetermined intervals to increase classification accuracy and reduce false detections. Using LUME dataset, this approach produced mean cross-subject accuracy of 78.34% for three emotion classes and 86.56% for two classes. Additionally, on DEAP, it obtained an average accuracy of 72.81%. The model trained on SEED dataset demonstrated an average prediction efficiency of 58.01% across all participants and emotion classes when evaluated on the DEAP dataset. The suggested method performs on par with or better than previous subject-independent EEG-ER techniques in the literature and has a lower complexity because it does not require feature extraction.

31. Atkinsonet al. (2016) developed a method which combines SVM emotion classifiers with statistically-based feature selection techniques. The foundation of their model for classification of emotion was based on valence/arousal dimensions.

32. Asghar et al. (2020) proposed an innovative signal processing system utilizing a DNN to extract relevant features from all channels of EEG relevant to emotion, specifically using VGG-16 model. To reduce computational costs and improve performance, the study introduces a Fast EMD (FEMD) model, which greatly reduces size of features for faster analysis. This approach transforms signal into a 2D wavelet spectrogram, capturing each subject's properties. A DNN-based emotional classification model is then applied to SJTU SEED and DEAP datasets. While the DEAP dataset is classified using Arousal/Valence dimensions, the SEED dataset is categorized into positive, negative, and neutral dimensions using Random Forest (RF), SVM, and KNN. The proposed model demonstrates superior performance over other advanced techniques in detecting human emotions on both datasets.

33. Daz et al. (2019) presented a study of EEG applied to emotions such as anxiety, tenderness, pleasure, and surprise, with a success rate of 63.36 percent on average.

34. To break down EEG signals with different channel characteristics into discrete IMFs, Pusarlaet al. (2020) used EMD. The first four IMFs were employed as feature vectors for emotion identification once their Sample Entropy (Samp En) was determined. To differentiate between three types of human emotions these feature vectors were input into a hybrid RF classifier and were tested on the publicly available SEED dataset, achieving a maximum accuracy of 88% and 96% when considering all 62 EEG channels.

35. In an experiment conducted by Awan et al. (2022), several high-dimensional features were extracted, leading to the development of a cross-subject emotion detection system called ST-SBSSVM. This system combines SVM with significance test and sequential backward selection (SBSS) to enhance ER. ST-SBSSVM was tested on SEED, DEAP, and SJTU datasets. Compared with traditional ER methods, ST-SBSSVM improved cross-subject detection accuracy by 13.54% on EAP dataset and 26.95% on SEED dataset, using multiple dimensional features. Recognition accuracy on DEAP dataset was comparable to SBS, but on the SEED dataset, ST-SBSSVM outperformed SBS by 6%. Additionally, ST-SBSSVM reduced program runtime by 97% on DEAP and 91% on SEED. The ER accuracy across all subjects was 72% on DEAP and 89% on SEED, demonstrating the effectiveness of the method compared to recent studies in this field.

36. Chao et al. (2019) presented a DL system which combines a capsule network (CapsNet) with a multiband feature matrix (MFM) to improve ER. The system integrates spatial frequency domain, and band features from multi-channel EEG signals to create the MFM. Using the MFM data, the CapsNet model categorizes emotional states. Experiments conducted on the DEAP dataset, which includes EEG, physiological, and visual information, show that developed method outperforms many existing models.Results highlight that the capsule network excels in mining and leveraging the three correlation qualities, while the three MFM features complement each other, enhancing the system's performance.

37. Jeong et al. (2001) conducted a study in which EEG and GSR data were collected from 28 subjects using commercially available equipment. Various stimuli were presented during the data collection, including two engaging video clips and one designed to induce boredom. Collected samples were named based on participants' self-reported levels of boredom, as indicated in a questionnaire. Using this data, the researchers trained 35 models with 18 different ML techniques and selected best three methods. After hyperparameter tuning, they validated final models 1,000 times with cross-validation to enhance robustness of results. The Multilayer Perceptron model emerged as the best performer, achieving an average accuracy of 80.98% (AUC: 0.771). Additionally, the study indicated a connection between boredom and combined properties of GSR and EEG signals. This finding may contribute to the development of accurate emotional computing systems and provide deeper insights into the physiological aspects of boredom. Table 2.1 depicts the comparative analysis of various existing methods.

Table 2. 1: Comparison of an approach based on an electroencephalogram (EEG) signal that uses a different feature extraction method and classifier to recognize human emotion states.

Literature	Methodology	Findings	Challenges and Considerations
Tan (2020)	SNN for human ER using skin conductance, EEG, facial expressions)	SNN achieved comparable classification accuracy but was computationally complex.	Computational complexity.
Ullah et al. (2019)	ELM for EEG channel subsets selection for ER	Improved computational time and classification accuracy demonstrated on the DEAP database.	Higher requirement of space and increased time for computation
Krishna et al. (2019)	TQWT technique for emotion classification using EEG signals	Four different types of emotions were correctly classified	Inclusion of multiple models increased computational complexity.

Liu et al. (2020)	Hybrid DL model integrating CNN, sparse autoencoder, and DNN for EEG-based emotion classification	Effective emotion classification demonstrated, but computationally costly.	High computational cost.
Zeng et al. (2019)	Sinc-Net for ER using DNN and convolutional layers	Quick convergence but required large training data and was computationally costly.	Enormous training data requirement.
Bajaj et al. (2018)	FAWT technique with KNN	Effective method for outlier removal but prone to overfitting.	Overfitting issue.
Kim and Choi (2020)	Attention based LSTM network for ER using EEG signals	Attention to class-imbalance problems required as future extension.	Class-imbalance concerns.
Chakladar and Chakraborty (2018)	Combining correlation-based subset selection with HO to reduce dimensionality in ER	Non-linear data handling required, suggesting the use of kernel based SVM.	Focus on non-linear data required.
Yin et al. (2021)	LSTM with GCNN integration for ER using DEAP database	Achieved better classification results but faced issues with overfitting and vanishing gradients.	Overfitting and vanishing gradients concerns.
Subasi et al. (2021)	Light-weighted ER framework using MSPCA, DWT, TQWT, and machine-learning classifiers	RF ensemble with SVM achieved superior classification	Trade-off between dimensionality reduction and information loss.

Salankar et al. (2021)	EMD followed by feature extraction with 2nd order difference plots, and classification with MLP and SVM	Validation with statistical measures, scaling sensitivity, and tuned with MLP.	Sensitivity to feature scaling and hyperparameter tuning.
Gao et al. (2022)	PSD based extraction of suitable features followed by classification with SVM	Effective for binary-classification but not suitable for multi-class ER.	Suitable for binary-classification.
Sharma et al. (2020)	Combination of DL model with HOS for automatic ER	High computational cost and requirement of high-end systems for significant performance.	High computational cost.
Mert and Akan (2019)	Integration of MEMD with frequency domain features for ER	Achieved arousal or valence classification but computationally costly with DBN model.	High computational cost with DBN model.
Song et al. (2018)	DGCNN	Superior recognition performance but faced issues with vanishing gradients and overfitting.	Challenges with gradient disappearance and model overfitting.
Chen et al. (2018)	SVM + DBN classification	Effective classification but computationally costly due to DBN model requirement.	Computational cost with DBN model.
Chao et al. (2019)	CapsNet	High computational expense.	High computational expense.
Zhang et al. (2020)	Integration of KM with DNNs	Effective capture of relations between physiological EEG	Computational intensity.

		signals but computationally intensive.	
Pan et al. (2020)	Feature extraction with wavelet packet, by ensemble ML method for classification with SVM, ELM, and decision tree for ER	Effective classification but computationally complex with multiple classifiers.	Computational complexity with multiple classifiers.

2.2 Feature Extraction

The preprocessing data for each electrode is analyzed, and relevant features are extracted in frequency, time, as well as time-frequency domains. Jenke et al. (2014) reviewed various feature extraction techniques for EEG-based ER. IMFs of EEG signal offer valuable time-frequency insights into signal. In developed method, the IMFs are calculated using both EMD and VMD techniques. IMFs represent the oscillatory components of a signal, which are decomposed into their distinct oscillatory parts.

Since EEG signals (Pandey & Seeja, 2019) are nonstationary with unpredictable frequency contents, time-frequency correlated features are of value for extracting accurate information. The use of IMF as a method to detect features is good. Techniques involved in calculating the IMF of the EEG signal include EMD and VMD. Since these methods decompose an EEG signal into its respective oscillatory elements, also known as IMFs. Two features of EEG signals are computed using first two components: PSD peak value and first difference of signals. While there are a variety of techniques for categorizing emotions, SVM and DNN are the most used in EEG-based ER.

2.3 Empirical Mode Decomposition

In terms of the signal g(t), which denotes EEG, all components of the IMF are measured through a technique termed shifting. Two requirements must be met by all IMF components: mean value of upper and lower envelopes, and each IMF component must have an equal number of zeroes crossing points or a difference of no more than one point is given in Fig. 2.1. The following are the measures involved in finding IMF:

1. Identify all the and minima, maxima of signal $f(t)$

2. Determine lower $L(t)$ and upper envelope $U(t)$ by integrating the maxima and minima respectively.

3. Calculate mean envelope $M(t)$ as an average of lower and upper envelopes, as shown in Eqn. (2.1)

$$M(t) = \frac{L(t)+U(t)}{2} \tag{2.1}$$

4. Calculate the detail signal as $d(t)$ given in Eqn. (2.2) and check whether it satisfies the two conditions required for IMFs.

$$d(t) = f(t) - M(t) \tag{2.2}$$

5. If the computed $d(t)$ satisfies IMF conditions, then first IMF component $IMF_1(t)$, is set to $d(t)$ If not, repeat steps 1 to 4 until $d(t)$ meets required conditions.

6. Next IMF component is derived by calculating residue $r(t)$ as shown in Eqn. (2.3), and considering $r(t)$ as new signal then repeat the steps.

$$r(t) = f(t) - IMF_1(t) \tag{2.3}$$

7. These steps are iterated until $r(t)$ becomes constant.

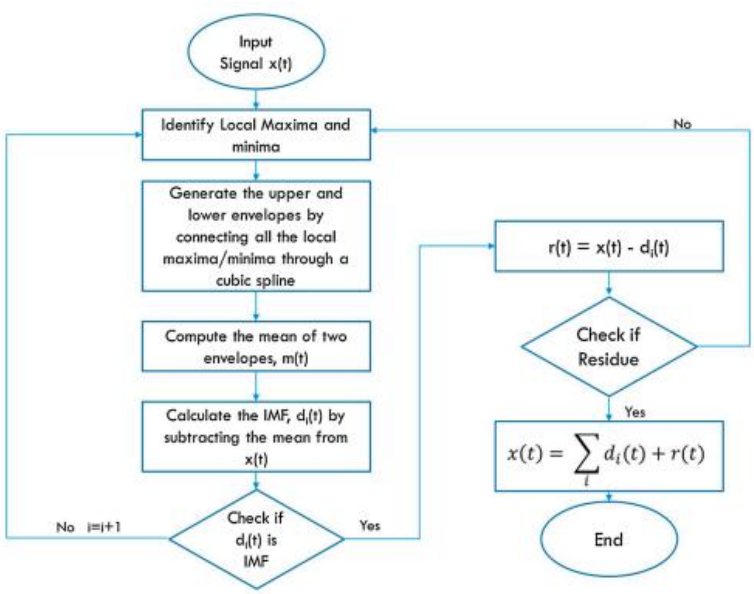


Figure 2. 1: Empirical Mode Decomposition Algorithm

2.4 Variational -Mode-Decomposition

VMD is an audio processing technique for breaking down a signal into a countable number of modes, each representing a different frequency component within a signal. Here's a simplified explanation of the VMD algorithm:

1. **Input Signal**: Start with a signal $x(t)$ that you want to decompose.
2. **Objective**: The goal is to find modes $u_k(t)$and associated frequencies ω_k that best represent the signal while promoting sparsity in the frequency domain.
3. **Algorithm Overview**:
   - **Step 1: Fidelity Term**: Minimize the difference between the original signal $x(t)$ and a sum of the modes $u_k(t)$. This ensures the reconstructed signal closely matches the original.
   - **Step 2: Regularization Term**: Encourage sparsity in the frequency domain by penalizing the complexity of the modes.
   - **Step 3: Balancing**: Adjust a parameter λ to balance fidelity and sparsity terms.
4. **Iterative Optimization**:
   - **Update Modes**: Adjust each mode to minimize the fidelity term and promote sparsity.
   - **Update Frequencies**: Adjust frequencies to further promote sparsity in the signal.
   - **Convergence**: Iterate mode and frequency updates until convergence criteria are met.
5. **Subproblems**:
   - **Mode Update**: Minimize the fidelity term and regularization term for each mode.
   - **Frequency Update**: Minimize both fidelity and regularization term for each frequency.
6. **Optimization Techniques**: Techniques like gradient descent or iterative thresholding are used to solve the subproblems.
7. **Convergence Criteria**: Stop iterating when a specified number of iterations is reached or when energy functional changes insignificantly.

The VMD algorithm effectively decomposes the input signal into its constituent modes while promoting sparsity in the frequency domain, making it useful for analyzing signals with non-stationary and oscillatory components. VMD is particularly useful for signals with non-stationary and oscillatory components, such as EEG signals, where it can help extract meaningful features for further analysis, such as emotion classification.

VMD is a decomposition in the time-frequency. It is designed to solve some of the problems in EMD, which are detailed below.

1. EMD lacks backward error correction and is an iterative algorithm
2. It cannot stand loud noises.

VMD follows concurrent technique rather than a recursive one to generate IMF from a signal. The signal is broken down into k IMFs using this procedure, which results in several modes uk with corresponding center frequencies ck. Total of these modes represents original signal. It does not generate residual noise and is less prone to noise than EMD.

2.5 Deep Neural Network

It is a type of ML model used for data classification and feature learning. They incorporate various optimization techniques and utilize four activation functions sigmoid, tanh, ReLU, and softmax for both binary and multiclass classification tasks. One of the key advantages is, its ability to process large and complex datasets with minimal computational resources. With the ability to modify the number of hidden layers and the neurons in each layer, a DNN normally has more than three layers, including input and output layers. In this study, DNN was chosen for emotion recognition due to its capacity to extract stronger, more meaningful features from the data.

2.6 Research Gaps

As per literature till time the work has been done on EMD or VMD individually for understanding the behavior of the signal. But nobody has attempted to create hybrid version of combination of step-by-step EMD/VMD. This combination can be beneficial in certain signal processing applications, particularly when dealing with complex and non-stationary signals like EEG. One approach could involve using EMD as a preprocessing step to decompose the signal into IMFs and then applying VMD on each IMF to further refine the decomposition. This can potentially enhance the effectiveness of mode separation and noise removal, leading to improved signal representation and analysis.

According to a literature survey, some research gaps can still be worked upon to improve the existing technology or system. These are:

- EEG-based ER studies can be distinguished by various factors, such as the technique used for emotion identifier, extracter of features, distinguishers, the number of participants, and the number of emotions classified. Despite extensive research on this topic, challenges and unresolved issues remain in EEG-based emotion recognition. For example, there are concerns regarding the accuracy of ER, number of electrodes utilized, and number of emotion categories identified.
- In ER from EEG, it is not generally agreed upon which features are most appropriate and only a few works exist.
- Available EEG Signal processing technique for emotion recognition analysis provide lesser accuracy. Various signal processing techniques for non-stationary signals, such as wavelet, multi-WT, and time-frequency distributions, were analyzed for complex signals. Among these, EMD is widely utilized for examining nonlinear and non-stationary signals.
- Feature extraction techniques and classifiers available for emotion classification need improvement (Dragomiretskiy & Zosso, 2014; Khosla et al., 2020).
- EMD is a versatile method that works independently of the signal's linearity and stationarity. Recently, an adaptive technique called VMD has been introduced for analyzing multi-component signals. VMD outperforms traditional methods like EMD and wavelet decompositions in separating multiple components of signals . However, a limitation of VMD is its direct preprocessing of highly non-stationary data, such as EEG and EMG signals. To address this, VMD requires an enhancement of the stationary properties of these signals before processing. To better handle non-stationary EEG signals, a combination of the strengths of both EMD and VMD is employed. EMD improves the stationary characteristics of EEG signals by eliminating noise, while VMD further processes the signal to extract the desired components. Combination of EMD and VMD feature extraction of EEG signal can detect emotion more accurately.

- Given that emotions originate in the brain and EEG measures brain activity, there is a significant connection between the two in identifying emotional states. Consequently, research focused on EEG-based emotion detection in humans has rapidly increased.

2.7 Problem Statement

Emotion recognition techniques span various domains, including text, speech, facial expressions, and neurological signals, offering valuable insights into individuals' emotional states. However, the challenge lies in effectively extracting discriminant features from electroencephalography (EEG) signals due to their dynamic nature. Recent advancements in emotion detection, particularly concerning arousal and valence, present significant opportunities for psychology and HCI. Despite the increasing interest in automated emotion classification using EEG, current methods often overlook essential contextual cues embedded within EEG signals. So, there is a need to develop an automated model aimed at improving EEG-based ER.

2.8 Proposed Approaches

EEG-based emotion recognition has gained significant attention for its ability to analyze brain activity and classify emotional states such as arousal and valence. However, traditional methods face challenges in feature extraction due to complexity of EEG signals, as well as noise and artifacts, making manual feature selection both error-prone and time-consuming. To overcome these limitations, this research proposes an automated hybrid DL framework optimized for accurate emotion recognition. The process begins with advanced preprocessing techniques, utilizing hybrid EMD and VMD to remove artifacts and enhance signal quality. EMD adaptively decomposes EEG signals into IMF, effectively capturing non-linear characteristics, while VMD mitigates mode mixing by precisely decomposing signals into predefined modes. Fig. 2.2 presents the work flow developed approaches.

Following preprocessing, statistical features, and higher-order moments are extracted to capture critical signal characteristics, including amplitude variation and distribution asymmetry. Feature selection is further refined using the MTSOA. Classification is performed through DL models, including LSTM for capturing temporal dependencies, Bi-LSTM for bidirectional processing, and a hybrid CNN+LSTM approach for spatial-temporal analysis. To optimize performance, the IRSOA fine-tunes the hyperparameters of these models, ensuring higher classification accuracy and robustness. This automated and optimized framework addresses the limitations of traditional methods, offering a reliable and efficient solution for EEG-based ER.

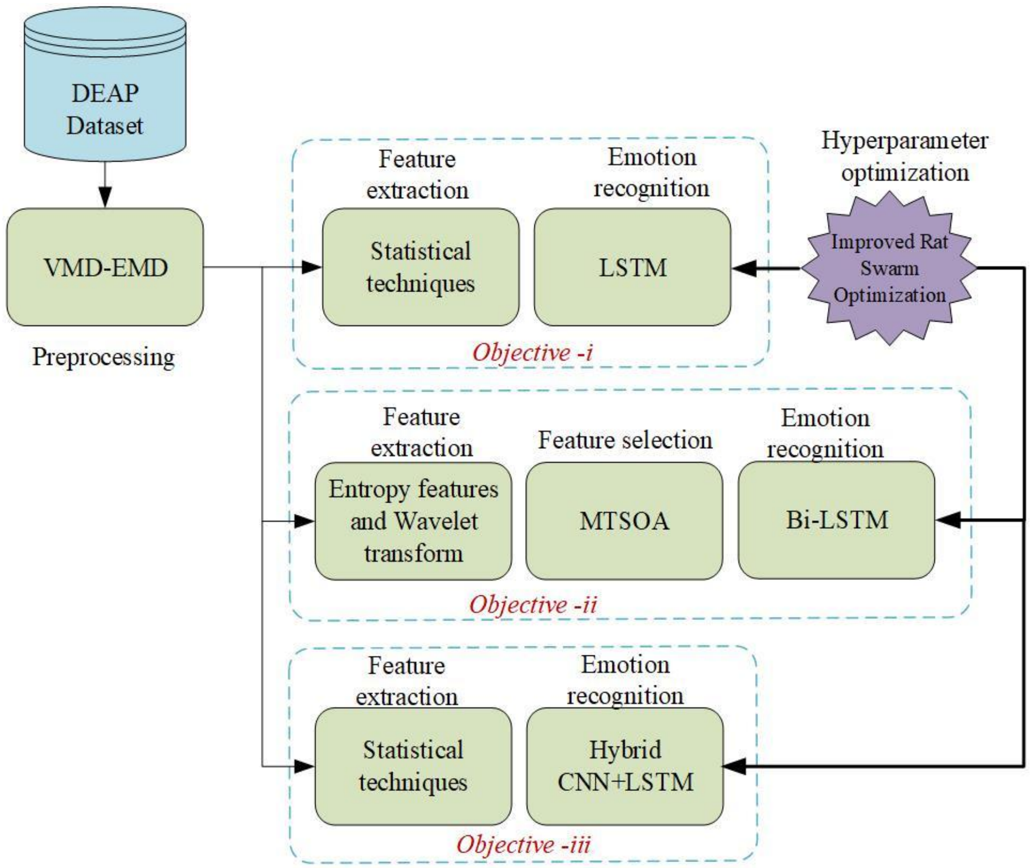


Figure 2. 2: Overall system model

2.9 Summary

This chapter reviews existing methods in EEG for ER, emphasizing key challenges and research gaps. Traditional approaches relying on handcrafted feature extraction are error-prone, time-intensive, and struggle with non-linear, non-stationary nature of EEG signals. Additionally, conventional classification methods often fail to capture the complex temporal and spatial patterns required for accurately identifying emotional states like arousal and valence. Preprocessing limitations further compound these issues, as standard noise removal techniques often leave artifacts, degrading signal quality and impacting feature extraction. These gaps highlight the need for automated hybrid DL frameworks fine-tuned with novel optimization techniques, these frameworks address existing limitations, enabling efficient feature extraction and accurate emotion classification.

CHAPTER 3

EEG Based Emotion Recognition Using Long Short Term Memory Network with Improved Rat Swarm Optimization Algorithm

Overview

Building on the challenges in EEG-based emotion recognition, this study proposes a DL framework integrating statistical feature extraction with an optimization algorithm for accurate classification. Unlike traditional methods, this approach efficiently handles EEG's non-linearity and noise. By enhancing feature selection and classification, the framework improves emotion recognition, specifically for arousal and valence.

3.1 Introduction

One of the most promising emerging research areas is automated human ER, which is commonly derived from EEG signals and has gained significant attention as HCI continues to expand. However, few techniques have concentrated on contextual details within EEG signals. In this research, an innovative approach to enhance ER from EEG signals is developed. In initial phase, EEG signals are collected through DEAP database. To reduce glitches and noise from raw EEG data, denoising is performed using EMD and VMD filters. Following this, 20 statistical features are extracted from the cleaned data to capture discriminative data from decomposed signals. Finally, an LSTM network is employed for emotion recognition, specifically targeting arousal and valence. Optimal hyperparameters for LSTM network are determined through IRSOA technique. As demonstrated in results and discussion, IRSOA based LSTM network achieves an efficiency of 84.89% on DEAP dataset. These results indicate that suggested IRSOA-LSTM network outperforms existing ML models.

ER plays a vital role in HCI by allowing machines to comprehend human emotional and mental states. The development of ER is closely tied to progress in fields such as computer field, cognitive science, modern neuroscience, and psychology. The inclusion of external emotion-state recognition by automated systems is expected to expand the scope of HCI, leading to applications in areas such as gaming, military, industry, and healthcare. Human-oriented emotion recognition approaches are divided into two types: (i) non-physiological, such as voice, body gestures, expressions in face, and (ii) physiological-based, such as EMG, ECG, and EEG. Among the various types of functional signals, EEG is most widely used for ER because it promptly measures brain cortical activity, which is closely linked to state of mind. Adoption of EEG-based ER have increased significantly due to advancements in electrode technologies. Emotions of human are typically classified into two main approaches: dimensional methods and DBED methods. DBED methods categorize emotions into six distinct types: Happiness, Anger, Sadness, Fear, Disgust, and Surprise. On the other hand, the dimensional method represents emotions along two continuous dimensions: arousal and valence. In recent years, several automated emotion recognition architectures have been proposed, with DL emerging as the most popular choice among researchers due to its reliable and scalable characteristics.

3.2 Contribution

This research work presents a novel optimization-based DL model for improving classification accuracy and resolving unstable properties of EEG signals. Major contributions are

- DEAP database is used to gather the input raw signals using EEG. The obtained signals are clean and prepared for processing and classification by integrating EMD with VMD filters to eliminate noise and glitches.
- To extract meaningful features from decomposed EEG signals, 20 analytical features were utilized. These statistical features offer several advantages, such as reducing the risk of overfitting, improving classification accuracy, enhancing data illustration, and accelerating the learning process of DL models like LSTM.
- A correlation analysis is conducted to study the relationships among these features. This analysis helps identify whether certain features are redundant (highly correlated) or complementary (weakly correlated). For reference, the correlation between seven representative features are analyzed: 'Hjorth Activity,' 'Shannon Entropy,' 'Variance,' 'Mean Curve Length,' 'Band Power Delta (BPD),' 'Band Power Theta (BPT),' and 'Band Power Alpha (BPA)’ shown in Fig. 3.1. Obtained results indicate that most of features exhibit moderate to weak correlations with each other. This weak correlation signifies that these features likely capture distinct

information about the EEG signals, which is beneficial for model performance, as it reduces redundancy and ensures a more diverse feature set.

- For example, BPD, BPT, and BPA related to frequency-specific brainwave activities had low correlation values with features such as 'Shannon Entropy' and 'Hjorth Activity,' which measure signal complexity and variability. Similarly, 'Variance' showed only a moderate correlation with 'Mean Curve Length,' indicating that both features provide unique insights into the data. The weak correlations confirm that these features collectively enrich the feature space, allowing the LSTM model to effectively learn temporal dependencies without being overwhelmed by overlapping or redundant information.

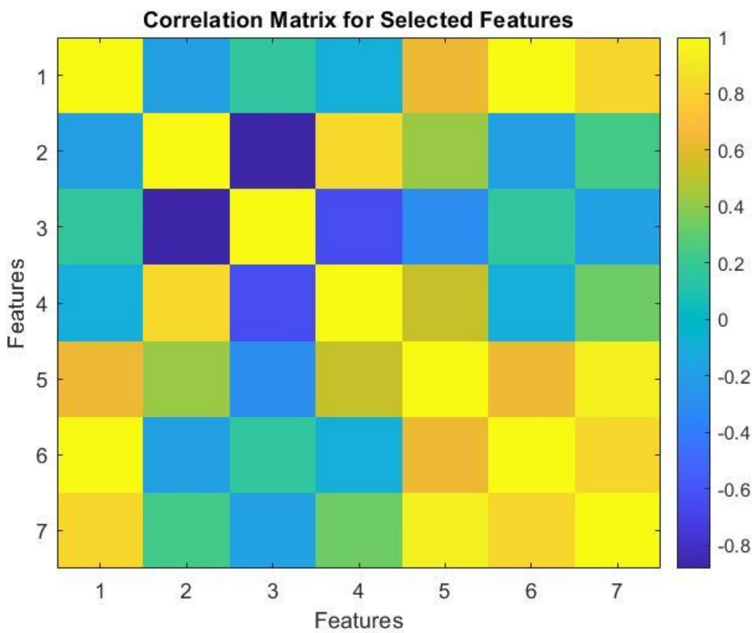


Figure 3. 1: Correlation analysis of different features

The application of LSTM in classifying human emotion arousal and valence is enhanced by the IRSOA technique, which optimizes the selection of LSTM hyperparameters, improving computational efficiency and operation duration. The effectiveness of suggested IRSOA-LSTM network is evaluated through various performance metrics.

3.3 Proposed Methods

In EEG-driven ER, suggested framework comprises four essential stages: signal acquisition, data preprocessing, extraction of features, and emotion classification.

The process begins with signal collection using the DEAP dataset, a widely recognized benchmark for emotion analysis that includes multimodal physiological signals recorded while participants viewed emotionally evocative videos. These EEG signals serve as the raw input for the framework.

Next, in the preprocessing phase, a hybrid approach combining EMD and VMD is employed to enhance signal quality. EMD adaptively decomposes the EEG signals into IMFs, effectively capturing non-linear and non-stationary components. However, EMD alone may suffer from mode mixing, where overlapping components appear within a single IMF. To overcome this, VMD is integrated into the preprocessing step, which decomposes the signals into predefined modes with high precision, ensuring comprehensive noise removal and artifact reduction. This hybrid EMD-VMD approach significantly enhances the reliability and integrity of the processed EEG signals.

Following preprocessing, the feature extraction phase focuses on deriving meaningful statistical features from the denoised signals. A total of 20 statistical features, including mean, variance, skewness, kurtosis, and higher-order moments, are computed to capture critical characteristics of the EEG signals. These features encapsulate essential information such as amplitude variations, frequency components, and distribution patterns, which play a crucial role in differentiating emotional states.

Finally, the emotion recognition phase employs a classification model based on the IRSOA-LSTM network. LSTM networks are utilized for their ability to capture temporal dependencies in sequential EEG data, ensuring accurate classification of emotional states like arousal and valence. To further enhance the model's performance, IRSOA is applied to enhance the hyperparameters of LSTM method. This optimization step improves the classification accuracy and robustness of the framework, enabling precise emotion recognition.

By integrating advanced preprocessing methods, robust feature extraction, and optimized DL techniques, this framework addresses the limitations of traditional methods and ensures reliable emotion classification. The block diagram representing this comprehensive framework is presented in Fig. 3.2.

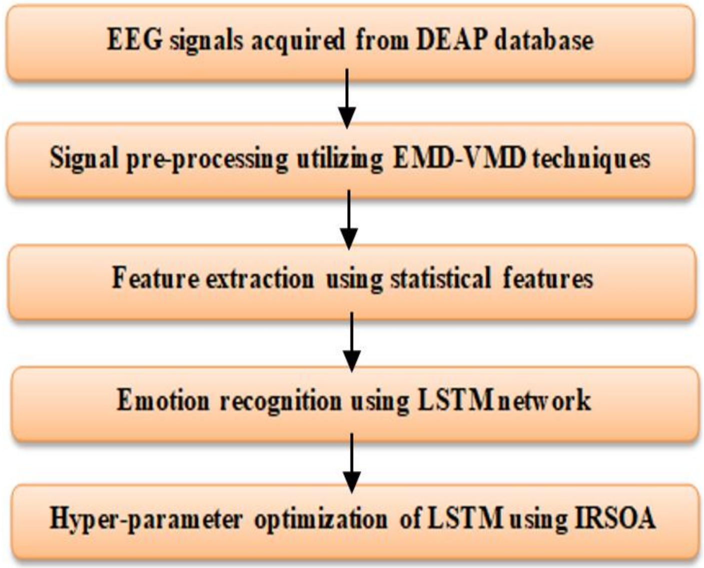


Figure 3. 2: Block diagram of the developed framework

3.3.1 Database Description

The initial step here involved the acquisition of EEG Signals from DEAP database. This dataset contained EEG recordings from 32 individuals, each viewing 40 music video clips, each lasting one minute. The real-time data was then pre-processed to a standardized format, down sampled to 128Hz, and further segmented into different trials. This was done to make it more suitable for ML applications. This research work used the existing preprocessing steps, which included the removal of EOG artifacts, along with bandpass filtering (4.0-45.0 Hz) and reordering of EEG channels to adhere to the Geneva configuration. This dataset was further pre-processed using EMD and VMD to handle non-linear and non-stationary features of EEG signals more effectively. EMD decomposed EEG signals into IMFs to separate the different signals into their sub-signals corresponding to different frequency bands. This would enhance the discriminative and differentiating potential of the data. These decomposition techniques allowed for more effective noise reduction and artifact removal, which would ensure that only relevant signal components were used for further subsequent analysis.

IRSOA-LSTM network's performance was evaluated using DEAP database, which contains physiological EEG data recorded from 32 participants. Each participant evaluated various videos based on dimensions of both classes. For 22 of 32 participants, the EEG data collection was accompanied by frontal facial video recordings. EEG signals in DEAP were sampled at 512 Hz from 32 electrodes placed at locations including O1, Fp1, AF4, FC2, CP5, FC1, Fp2, PO3, PO4, Fz, FC6, O2, CP1, CP2, Cz, T8, FC5, F4, C4, Oz, CP6, AF3, T7, P4, F8, P7, F3, P3, Pz, P8, C3, and F7. The EEG data comprises recordings from 32 subjects, with each channel associated with arousal and valence labels, and includes 40 videos. A summary of DEAP dataset, which includes physiological analyses and self-reported labels, is provided in Table 3.1. Fig. 3.3 illustrates some example signals from DEAP database.

Table 3. 1: Overview of DEAP Dataset

Category	Details
Physiological analyses	
Criteria based on rating	-Similarity: Discrete scale (1 to 5)
	- Other ratings: Continuous scale (1 - 9)
Measured Attributes	Valence/Arousal
Total number of Videos	40
Participants	32 subjects
Captured Data	- Facial video recordings (22)
	- 32-channel EEG signals
	- EEG sampling rate: 512 Hz
Online self-reported labels	
Annotation Scale	Discrete scale (1 to 9)
Evaluation Parameters	Valence/ Arousal
Ratings per each Video	14 - 16
Video preference method	- 60 videos elected manually
	- 60 videos chosen using affective tags
Video Duration	1 minute per video
Total Videos in Dataset	120

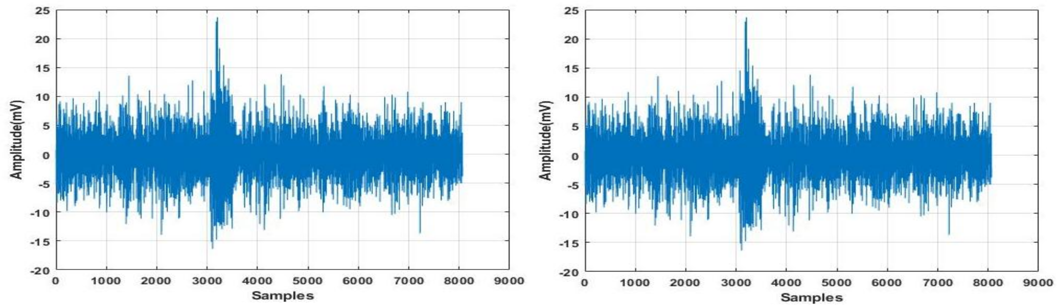


Figure 3. 3: Sample acquired EEG signals.

3.3.2 Data Preparation and Feature Engineering

Subsequent step in feature extraction aimed to identify meaningful features that could accurately represent the EEG signals and their corresponding emotional states, aiding in the interpretation of emotional state fluctuations. Some of the extracted statistical features included the mean, variance, skewness, and kurtosis of each decomposed signal component. These features provided valuable insights into the nature and dynamics of brain activity as it responds to different emotions. Additionally, frequency-domain features, such as PSD, were utilized to capture the oscillatory behaviour of signals across various frequency bands. The combination of all these features formed a robust feature space representation of the EEG signals, enhancing the model's ability to effectively distinguish between different states of emotion.

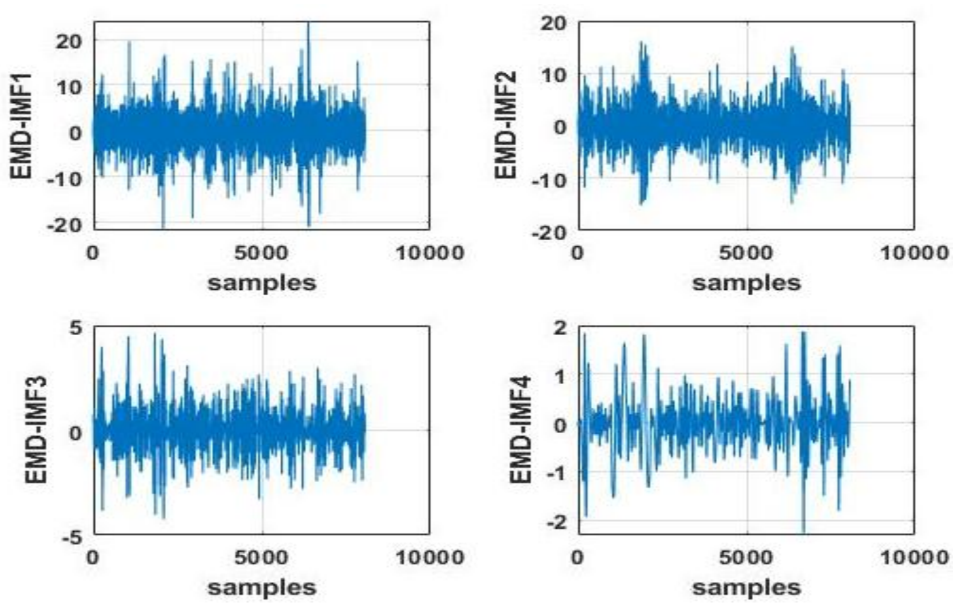


Figure 3. 4: Example of signal decomposition using EMD Method

The EMD/VMD techniques are employed to decompose non-stationary and complex EEG signals, removing unnecessary noises and glitches after acquiring signals from DEAP database.

EMD serves as the initial step, breaking down the EEG signals into multiple IMFs. Each IMF represents a sub-band signal, capturing both low and high-frequency components of the original signal. A filtering process then identifies high-frequency elements, enhancing the signal's sharpness while preserving its critical features. This smoothing filter ensures that the signal remains clear for classifier interpretation without losing essential information (Meng et al., 2019; Zheng & Pan, 2020). Fig. 3.3 illustrates an example of a decomposed signal using EMD.

On the other hand, VMD method decomposes EEG signals into distinct modes (sub-signals). Each VMD mode undergoes the Hilbert transform, enabling the assessment of frequency spans and the unilateral frequency spectrum for each mode (Kaur et al., 2021). The frequency bands of these modes are then combined with an exponential function, shifting them to a predefined central frequency (Taran & Bajaj, 2018).

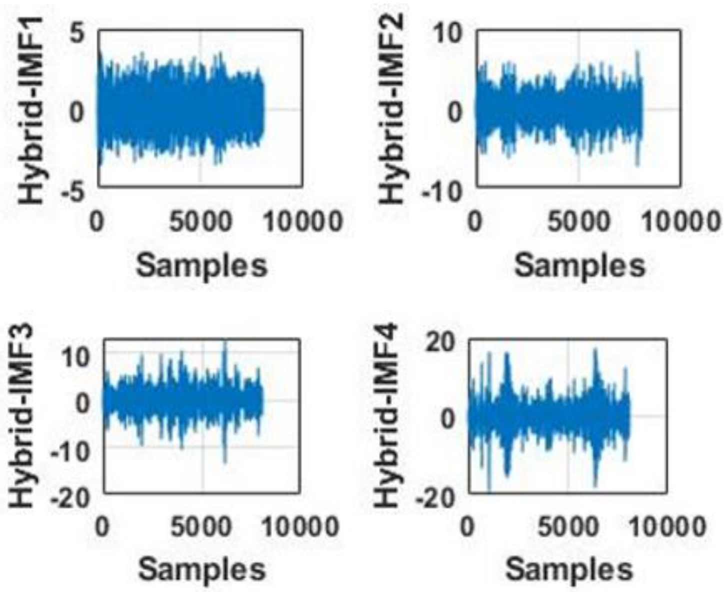


Figure 3. 5: Hybrid EMD-VMD-Based EEG signal decomposition

Gaussian smoothness of decoded signals is used to assess the bandwidth. The EMD-VMD hybrid filters effectively remove noise, artifacts, and external disturbances that may occur during signal recording. The decomposed signals produced by this hybrid method are shown in Fig. 3.4 and 3.5. Following signal decomposition, hybrid feature extraction is performed to generate feature vectors from cleaned signals. Hybrid extraction of feature process includes a comprehensive number of statistical features (Sharma et al., 2022). For 2D emotion classification, this hybrid method yields a feature of length=74 for both classes which improves classification accuracy, reduces the threat of overfitting, enhances data representation, accelerates learning process of LSTM network.

3.3.3 Emotion Classification

The core of proposed methodology lies in the design of a robust DL architecture, specifically tailored to learn complex patterns in EEG signals for accurate emotion recognition. Feature extracted with a length of 74, is provided as input to LSTM network for 2D-ER. LSTM units are particularly well-suited for learning temporal dependencies and long-range correlations in time-series data, making them ideal for EEG-based emotion recognition tasks. The layers of the LSTM network effectively capture the sequential nature of brain activity and its variations over time, which are crucial for understanding emotions. Additionally, fully connected dense layers with dropout regularization are integrated into the architecture. These layers learn high-level abstract features while minimizing the risk of overfitting, ensuring robust model performance. Compared with similar classification methods, LSTM network demonstrates superior capacity to maintain significant data and manage long-term dependencies through its self-feedback mechanism. This is facilitated by memory cells in LSTM, which consist of three gates, designed to store and process information effectively, even in scenarios involving long-term dependencies.

1. **Forget Gate**: It determines data from previous cell state c_{t-1} should be retained or removed. This decision is made using an activation function (sigmoid) σ, which outputs a value $f(t)$ between 0 and 1. If $f(t)=1$, the input data is fully passed into the model; if $f(t)=0$, it is discarded. Current input state y_t and previous hidden state h_{t-1} serve as inputs for this gate. This gate ensures that only relevant past information is retained.
2. **Input Gate**: which decides what new data to be stored in current cell state c_t and also uses a sigmoid activation function, which outputs $i(t)$ with values between 0 and 1.

This output is multiplied with the activation function $\tanh$, which generates $\hat{c}(t)$, a new candidate cell state. The resultant product is then added to the existing cell state, effectively updating c_t to store new and relevant information.

3. **Output Gate**: Output gate $O(t)$ determines the information that is passed to next cell and final output. The current cell state c_t is processed through a sigmoid activation function, and its output is multiplied with a $\tanh$ activation function. This regulates the range of values between -1 and 1, ensuring that relevant information is passed forward for accurate predictions.

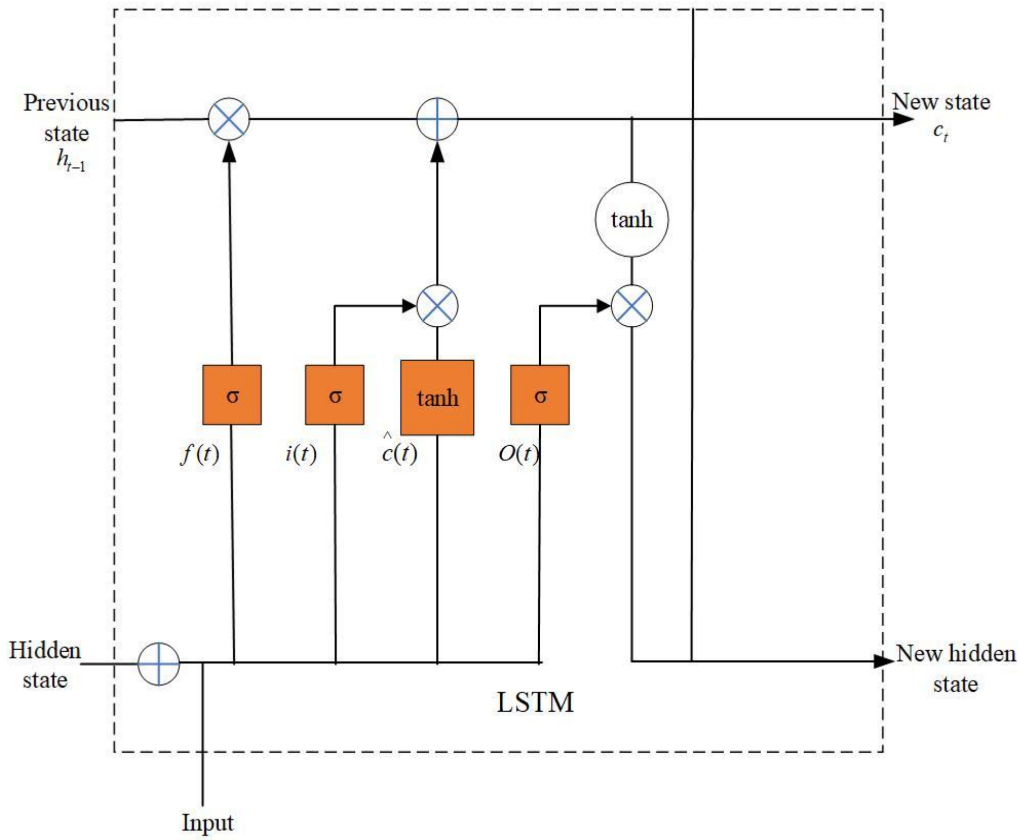


Figure 3. 6: Architecture of LSTM

These steps enable the LSTM network to retain important sequential information, making it particularly effective for emotion recognition tasks by accurately predicting emotional states based on past activations in EEG signals. The inclusion of fully connected dense layers further enhances the network's capability to generalize and classify emotions with high accuracy. Fig. 3.6 depicts the architecture of LSTM.

Thus, the proposed LSTM-based framework not only captures temporal dependencies in EEG signals but also ensures robust classification by leveraging its architecture's inherent advantages. This makes it a highly effective tool for emotion recognition, offering a significant improvement over conventional method. The following are the steps used in LSTM-based ER analysis.

The cell state $\hat{C}_t$ in an LSTM network is mathematically formulated in Eqn. (3.1). In this equation w_c represents weight matrix, b_c denotes bias term, C_t indicates the memory cell value, h_{t-1} signifies previous hidden state, $\tanh$ refers to hyperbolic tangent activation function. Additionally, x_t represents the input at time step t.

$$\hat{C}_t = \tanh(w_c \times [h_{t-1}, x_t] + b_c) \tag{3.1}$$

Input gate i_t regulates the information flow into the memory cell, as described in Eqn. (3.2). Here, σ represents sigmoid activation function. Similarly, forget gate $f(t)$ determines amount of past information retained in cell state and is mathematically expressed in Eqn. (3.3).

$$i_t = \sigma(w_i \times [h_{t-1}, x_t] + b_i) \tag{3.2}$$

$$f_t = \sigma(w_f \times [h_{t-1}, x_t] + b_f) \tag{3.3}$$

Previous cell state is denoted as C_{t-1}, and updated memory cell C_t is computed using Eqn. (3.4),

$$C_t = f_t \times C_{t-1} + i_t \times \hat{C}_t \tag{3.4}$$

where, symbol $\times$ represent dot product.

Output gate O_t is final outcome of LSTM and is computed based on cell state. This is represented in Eqn. (3.5). The hidden state h_t which serves as final result of LSTM block, is defined in Eqn. (3.6).

$$O(t) = \sigma(w_0 \times [h_{t-1}, x_t] + b_0) \tag{3.5}$$

$$h_t = O(t) \times \tanh(C_t) \tag{3.6}$$

Architecture of developed LSTM model, as shown in Fig. 3.6, is designed with two sequential LSTM layers that effectively capture temporal dependencies from the input EEG data. The first layer consists of 64 and second layer contains 32 hidden units, both utilizing Rectified Linear Unit (ReLU) activation functions to enhance the model's robustness and performance. The output layer comprises three neurons with a sigmoid activation function to handle the classification of emotions into two dimensions (e.g., arousal and valence). To optimize the model, IRSOA algorithm is employed, ensuring superior weight tuning and maximizing classification accuracy. Incorporating gate control and memory cell for long-term information retention, updating, and resetting, LSTMs achieved outstanding performance in two-dimensional emotion classification (Dhiman et al., 2021).

LSTM model was trained with optimized parameters to enhance efficiency. It was configured with a maximum of 100 training epochs, gradient decay constant ranging between [0,1], and primary rate of learning is adjusted between 0.001 and 0.01. Batch size was set to a minimum of 27, while gradient threshold was fixed at 1 to prevent instability. Execution environment was set to auto, and L2 regularization is applied with a value of 0.1 to mitigate overfitting.

Hyperparameter tuning was performed using IRSOA, demonstrating versatility and effectiveness in extracting meaningful patterns from EEG data for precise emotion classification.

3.3.4 Optimizing the Hyperparameters

RSOA is a hybrid heuristic approach inspired by attacking and hunting behaviors of rats in groups. It explores and exploits solutions within the search space, similar to other group-based metaheuristics. By modeling aggressive and following actions of rats, it optimizes hyperparameters effectively.

In population-based metaheuristics, an initial random solution is generated, where each rat's position represents a candidate solution in the search space. The objective function is evaluated iteratively, with adjustments made based on attacking and following behaviors. RSOA then optimizes LSTM model hyperparameters, enabling efficient two-dimensional emotion classification by maintaining a balance between exploration and exploitation for improved performance (Eslami et al., 2022).

1. Initialization of Rat Positions

Initial positions of rat population (s_i) represent candidate solutions in the search space. These positions are randomly initialized within the specified bounds for each hyperparameter is depicted in Eqn. (3.7).

$$s_i = s_{i-\min} + rand.(s_{i-\max} - s_{i-\min}), i = 1,2....N \tag{3.7}$$

where, $s_{i-\min}$ and $s_{i-\max}$ denote lower and upper bounds of i^{th} hyperparameter, $rand$ is a random number between [0, 1], N is number of agents (population size).

2. Attacking Procedure and Position Updates

The attacking procedure simulates how rats adjust their positions to find better solutions. The updated positions of the rats are computed using Eqn. (3.8).

$$\vec{P}_i(s+1) = \left| \vec{P}_r(s) - \vec{P} \right| \tag{3.8}$$

where, $\vec{P}_i(s+1)$ is an upgraded position of i^{th} rat, $\vec{P}_r(s)$ is a best solution (position of optimal rat) found, $\vec{P}$ is an intermediate position, calculated using Eqn. (3.9).

$$\vec{P} = A.\vec{P}_i(s) + C.(\vec{P}_r(s) - \vec{P}_i(s)) \tag{3.9}$$

where, $\vec{P}_i(s)$ is current position of i^{th} rat. A and C are parameters that balance exploration and exploitation, defined in Eqns. (3.10 and 3.11).

$$A = R - s.\left(\frac{R}{Iter_{\max}} \right), x = 1,2,3,...Iter_{\max} \tag{3.10}$$

$$C = 2.rand \tag{3.11}$$

where, R and $rand$ are random number between [1, 5] and [0, 2].

Each candidate solution (rat position) represents a combination of these hyperparameters. The performance of each candidate solution is evaluated using the objective function, which computes the accuracy or loss of the LSTM model during training.

4. Convergence and Optimal Solution

The algorithm iteratively updates the positions of the rats until convergence is achieved or maximum iteration is reached ($Iter_{\max} = 100$). Final solution corresponds to the hyperparameter combination that yields the best performance on the validation set.

5. Parameter Settings for RSOA

The assumed parameters for RSOA:

- Population size N : 30.
- Maximum iterations $Iter_{\max} = 100$.
- Search space dimensionality: 4
- Range for R : [1, 5].
- Range for $rand$: [0, 2].

The RSOA faces challenges like premature convergence, parameter sensitivity, high computational costs, and scalability issues in high-dimensional problems. Its stochastic nature also leads to inconsistent results. To overcome these limitations, the IRSOA is introduced, offering better convergence, scalability, and performance.

To improve DL model performance, IRSOA algorithm is employed for hyperparameter tuning. Inspired by the social behavior of rats in locating resources in complex environments, IRSOA incorporates adaptive learning rates and dynamic population sizes to efficiently explore the search space and avoid local minima. Critical parameters, including rate of learning, batch size, hidden layer count, and dropout rates, are optimized to maximize the model's accuracy for the emotion recognition task. To further improve its performance, an opposition-based learning concept was introduced, utilizing a random initialization technique to generate candidate solutions (as described in Equation 3.7) when prior knowledge is unavailable. The convergence speed and optimization effectiveness are influenced by distance between initial and best solutions. By improving quality of randomly generated solutions, IRSOA achieves superior performance.

Leveraging this technique, enhance method IRSOA, is introduced to improve speed of convergence and increase chances of identifying a global optimal solution compared to conventional RSOA. IRSOA estimates opposite positions of each solution using concept of opposite numbers. To explain the algorithm behind initializes of new population, it is essential to first understand the concept of opposite numbers.

Define the Search Space:

Define the hyperparameters to be tuned

Initialize Population:

Use random initialization to generate a population of candidate solutions (hyperparameter configurations). For each candidate, compute its opposite solution using the opposite number formula. If the learning rate of a candidate is 0.001, its opposite is calculated using Eqn. (3.12)

$$\textit{Opposite learning rate} = \max(LR) + \min(LR) - Current(LR) \tag{3.12}$$

where, LR is learning rate.

Evaluate Initial Population:

Train the LSTM model for each candidate and its opposite using a subset of the emotion dataset Evaluate their performance based on an objective function (e.g., validation accuracy or F1 score). Select the better solution between the candidate and its opposite.

Update Population:

- Identify the worst-performing candidate in the current population (highest error or lowest accuracy).
- Replace it with a new candidate generated using Eqn. (3.13).

$$x_{new} = \begin{cases} rand \times P_r(t)_1 & if\ \ rand_2 < 0.5 \\ x_{\max} - rand_3 \times (x_{\max} - x_{\min}) & if\ \ rand_2 > 0.5 \end{cases} \tag{3.13}$$

where, $P_r(t)$ is best solution ensuring the exploration of new hyperparameter configurations.

Repeat Iterations:

Continue the process of evaluating, updating, and replacing solutions until maximum number of iterations is reached or performance metric converges. Table 3.2 depicts the algorithm form IRSOA.

The implementation of IRSOA for tuning the hyperparameters of the LSTM model significantly improves classification accuracy.

Table 3. 2: Pseudocode of IRSOA for tuning LSTM

```
Define hyperparameters to tune (e.g., units, dropout, learning rate, batch size, optimizer)
Define algorithm parameters: Population size (N), Max iterations (Iter_max)
# Step 1: Initialization
For i = 1 to N do
   Initialize each rat's position x_i as random hyperparameter values
   Compute opposite position x̂_i based on the opposite number formula
   Train LSTM using x_i and x̂_i, evaluate fitness
   If f(x̂_i) < f(x_i) then
      Substitute x_i with x̂_i
   End if
End for
# Step 2: Main Optimization Loop
Initialize control parameters (A, C, R)
Identify the best set of hyperparameters p_r with the lowest fitness value
While T < Iter_max do
   For i = 1 to N do
      Parameter updation (A, C)
      Update hyperparameter set x_i using movement equations
      Train LSTM with the updated hyperparameter set and evaluate fitness
      If x_i goes beyond allowable hyperparameter ranges then
         Adjust x_i back within limits
      End if
   End for
   Replace the worst-performing set of hyperparameters with a new set (using Eqn. 3.13)
```

Update the best-performing hyperparameter set p_r

Increment counter iteration $t = t + 1$

End

Step 3: Return Result

Output is best hyperparameter set and its fitness value

3.4 Analysis and Interpretation of Experimental Findings

Simulations of IRSOA with LSTM network carried out on MATLAB R2020b, utilizing an Intel i7 processor with RAM of 16GB. Performance of IRSOA-LSTM network was assessed through multiple assessment criteria. Specificity measures the ability of IRSOA with LSTM network to correctly identify true negatives among all given categories, ensuring accurate differentiation between distinct classes. Sensitivity, on the other hand, evaluates ability to detect true positives for each category, highlighting effectiveness in capturing relevant patterns from input data.

F1-score is particularly crucial in scenarios involving binary classification for valence and arousal, as it represents mean value of precision and recall, ensuring a balanced evaluation of model advancement. Accuracy in EEG-based event-related paradigms directly reflects effectiveness of IRSOA-LSTM network in classifying emotional states. Precision, defined as ratio of correctly identified positive instances to total predicted positives, further strengthens assessment of classification reliability. Mathematical formulations of these evaluation metrics, crucial for comprehensive performance analysis, are detailed in Eqns. (3.14) to (3.18).

$$Accuracy = \frac{TN+TP}{TN+TP+FP+FN} \times 100 \tag{3.14}$$

$$Precision = \frac{TP}{TP+FP} \times 100 \tag{3.15}$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \tag{3.16}$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \tag{3.17}$$

$$F1-score = \frac{2TP}{2TP+FP+FN} \times 100 \tag{3.18}$$

where, FP is False Positive, TP denotes True Positive, False Negative denoted as FN, and True Negative is represented as TN.

3.4.1. Comprehensive Analysis of DEAP Dataset

Simulated experiments evaluated the proposed IRSOA-LSTM network using the DEAP database, with validation performed through five-fold cross-validation, where 80% for training and 20% for is used for testing. Performance comparison with traditional ML classifiers, including RF, DT, Multi-SVR (MSVR), and KNN, demonstrated that IRSOA-LSTM outperformed these models in emotion classification. Table 3.3 provides a detailed comparison, highlighting significant improvements in accuracy and reliability.

Experimental analysis focused on individual classification of valence and arousal, two critical dimensions in emotion recognition. In arousal classification, IRSOA-LSTM achieved an accuracy of 83.54% which indicate that the network effectively distinguished between different levels of arousal with high reliability. In valence classification, performance was even more remarkable, reaching 86.25% accuracy. These findings confirm that IRSOA-LSTM successfully minimized information loss and extracted meaningful patterns from EEG signals. Fig. 3.7 provides a graphical representation comparing IRSOA-LSTM with other classifiers.

Further comparisons presented in Table 3.4 emphasize that IRSOA-LSTM significantly outperformed other metaheuristic optimization methods, including GWO (Chen, 2023), PSO (Li et al., 2022), and FOA (Abgeena & Garg, 2023). Unlike these conventional approaches, IRSOA effectively balanced exploration and exploitation, leading to better hyperparameter tuning and improved classification outcomes. For arousal classification, IRSOA-LSTM consistently achieved 83.549% for accuracy, 86.99% and 84.15% for sensitivity and specificity, F1-score of 84.50%, and 82.10% precision. Similarly, in valence classification, results were 86.25% accuracy, 86.925% sensitivity, with a specificity of 87.93%, 86.11% for F1-score, and precision of 85.312%.

Beyond classification performance, IRSOA-LSTM demonstrated exceptional efficiency in training. Hyperparameter tuning using IRSOA significantly reduced computational cost, enabling faster convergence and optimal model performance. With a training time of just 34.5 seconds, IRSOA-LSTM required considerably less computational time than traditional classification techniques. This reduction in processing time highlights its effectiveness in handling large-scale EEG datasets without compromising accuracy.

Table 3. 3: Comparative analysis of IRSOA-LSTM Network performance

Metrics	Classes									
	Arousal					Valence				
	RF	KNN	DT	MSVM	LSTM	RF	KNN	DT	MSVM	LSTM
Accuracy (%)	80.11	81.55	50.12	68.54	83.549	82.71	79.17	50.10	82.79	86.25
Sensitivity (%)	84.57	75.69	51.651	68.74	86.99	80.21	82.86	49.27	79.07	86.925
Specificity (%)	83.89	78.98	53.18	68.07	84.15	79.30	82.95	51.52	80.78	87.93
F1-score (%)	83.67	82.40	50.28	68.29	84.50	81.96	82.32	49.11	79.981	86.11
Precision (%)	82.76	90.18	49.10	69.85	82.10	83.26	83.72	48.95	80.38	85.31

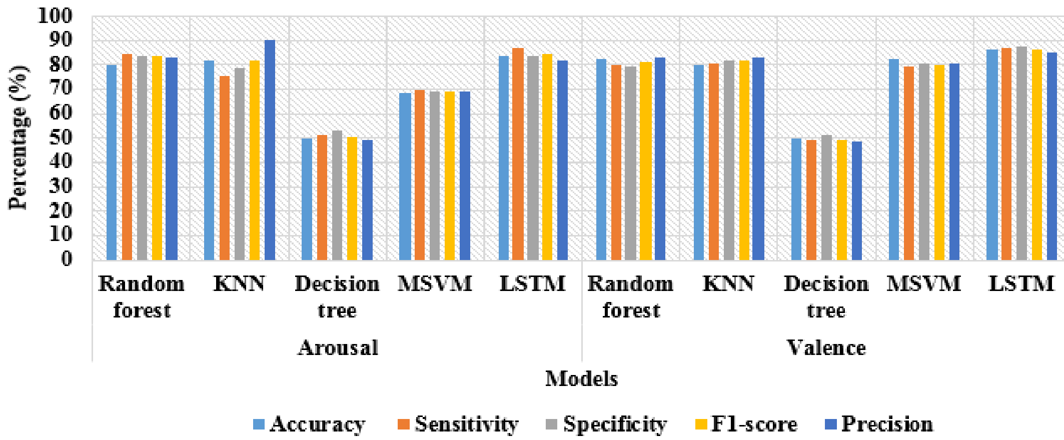


Figure 3. 7: Performance comparison of IRSOA-LSTM Network with alternative classifiers

Table 3. 4: Comparative analysis of IRSOA with LSTM Network and alternative optimization techniques

Metric (%)	Arousal				Valence			
	PSO	GWO	FOA	IRSOA	PSO	GWO	FOA	IRSOA
Accuracy	80.47	77.44	73.35	83.54	81.32	80.95	80.18	86.25
Sensitivity	82.38	72.43	73.86	86.9	82.58	77.77	84.60	86.93
Specificity	84.1	74.32	73.55	84.14	85.93	82.83	80.44	87.97
F1-score	81.28	73.32	73.12	84.45	82.92	79.56	83.51	86.11
Precision	81.12	73.13	70.47	82.05	84.38	79.35	81.43	85.311

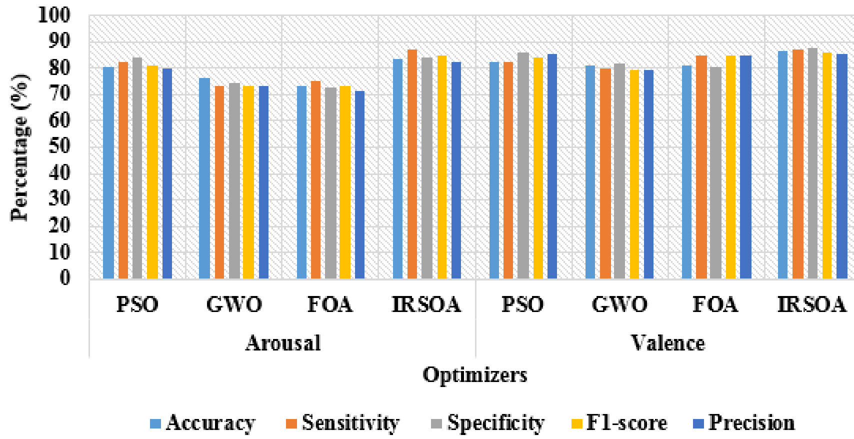


Figure 3. 8: Comparative performance analysis of IRSOA based LSTM and other optimization techniques

Table 3. 5: Classification performance of suggested IRSOA optimized by IRSOA model across four categories

Metric (%)	RF	KNN	Decision Tree	MSVM	LSTM(Proposed)
Accuracy	83.15	83.46	54.91	70.38	88.42
Sensitivity	84.99	79.72	52.92	71.46	89.21
Specificity	82.89	81.28	52.96	72.23	85.15
F1-score	84.05	83.45	55.15	69.81	88.65
Precision	85.19	81.42	52.75	72.89	89.45

For PSO, following parameters were set: dimension of 4, Larger iterations count of 100, size of population = 30, threshold value= 0.5, inertia weight = 0.9, social and cognitive factors of 2. GWO was configured with a threshold of 0.50, maximum iterations of 100, dimension = 4, and population size = 30. For FOA, the chosen parameters included threshold = 0.5, dimension = 4, alpha = 1, theta = 0.97, gamma = 1, population size = 30, beta value of one, and total count of iterations = 100. Graphical representation of LSTM enhanced by IRSOA alongside alternative hyperparameter optimizers is presented in Table 3.4 and Fig. 3.8.

Additionally, Fig. 3.9 and Table 3.5 provides results of suggested IRSOA-based LSTM network and other different classifiers across four categories: Low value of arousal and valence similarly high percentage of valence as well as arousal. Compared to classifiers in Table 3.5, IRSOA based LSTM network consistently outperforms with maximum values across all metrics. The suggested network's superior classification performance is attributed to many parameters in the LSTM, which do not require proper tuning which inherent flexibility in tuning parameters allows LSTM to achieve better results than competing models.

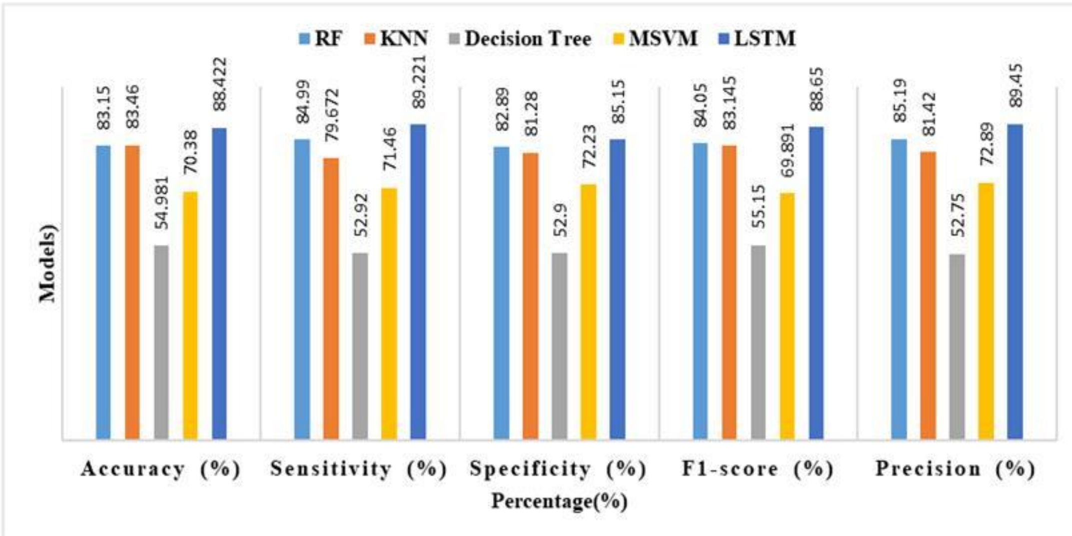


Figure 3. 9: Metrics evaluation of LSTM+IRSOA method

3.4.2. Comparative Investigation

Table 3.6 presents a comparative analysis of suggested IRSOA with LSTM network against existing models. In an EML for computing a distinctive EEG channel subset for ER. Their vast experimentation provided classification accuracies of 77.45% for valence and 70.20% for arousal. In employed a hybrid extraction approach combining PSD and GoogleNet to extract distinctive vectors which were then used for two-dimensional emotion classification with an SVM classifier, achieving 76.22% valence accuracy and 79.52% arousal accuracy.

In integrated Kernel matrices with DNN for ER. Their experimental findings indicated that the Kernel fusion of machines+ DNN model attained 63.50% valence accuracy and 64.45% arousal accuracy.

Table 3. 6: Comparative performance analysis of IRSOA-LSTM Network with different DL methods for Emotion classification

Classes	ML/DL Methods			
	ELM	CNN+SVM	Kernel Fusion + DNN	IRSOA+LSTM(Proposed)
Arousal	70.10	80.52	64.50	83.54
Valence	77.40	75.22	63.10	86.25

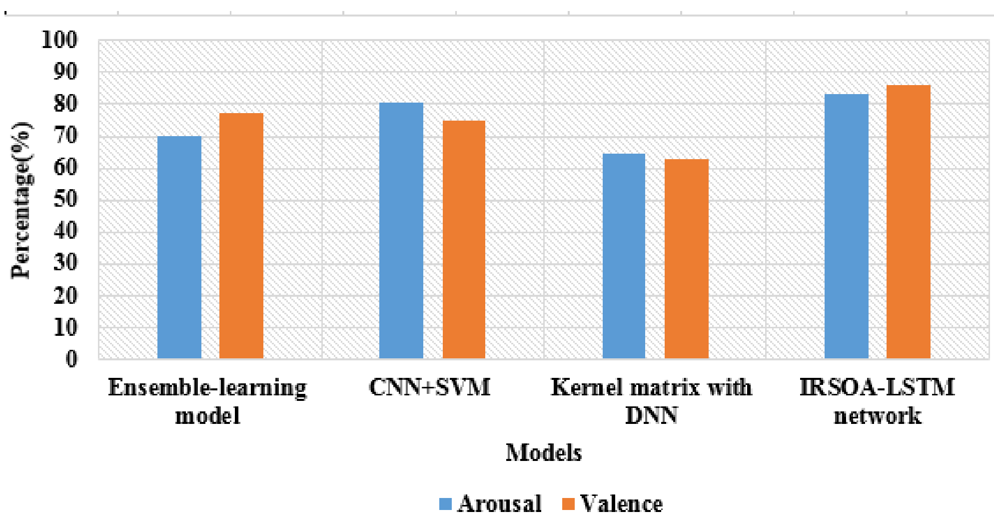


Figure 3. 10: Comparative graphical representation of results for IRSOA-LSTM network and conventional models.

In contrast, proposed IRSOA-LSTM model demonstrated superior performance in two-dimensional ER, achieving 86.25% valence accuracy and 83.54% arousal accuracy, as illustrated in Table 3.6 and Fig. 3.10.

3.5. Summary

This research involves designing a new IRSOA-LSTM model for effective classification of human emotions by EEG. Over the last few decades, it has proven to be one of the most difficult applications in human-gathered emotions since they are analyzed through EEG signals. this research is based on the online DEAP database, which helps in discovering and classifying human emotions. IRSOA-LSTM developed in this research includes data pre-processing, extraction of suitable features followed by classification, and optimization through IRSOA in which preprocessing of signals uses two distinct methods such as EMD hybridized with VMD, which is used to eliminate artifacts from raw signals so that best possible detail is retrieved. next is the statistical feature extraction area where around 20 techniques are used to obtain significant amounts of feature data. At last, refined information is taken as input for LSTM classification and definition of emotions such as arousal and valence followed by selection of optimal hyper-parameters by improved RSO. As depicted in computed segment, suggested IRSOA based LSTM network's performance is analysed by adopting different evaluation criteria. Thus, simulated result reveals that suggested IRSOA based LSTM network achieves an accuracy of 84.89, with a sensitivity score of 86.95%, specificity and f1-score of 86%, 85.28%, and precision of 83.68%. Additionally, IRSOA-LSTM network has also proven effective in processing time. future work will also cover the evaluation of developed IRSOA based LSTM on various other repository for emotion classification in e-learners. real-time alternative databases provide better insights regarding emotion regulation

CHAPTER 4

EEG Based Emotion Detection by Using Modified Tunicate Swarm Optimization Algorithm

Overview

The IRSOA-LSTM model effectively classifies emotions from EEG signals, but LSTM alone has limitations in capturing long-range dependencies due to its unidirectional processing. Additionally, relying solely on feature extraction without selection increases redundancy and computational complexity, potentially affecting classification accuracy. To address this, Bidirectional LSTM (Bi-LSTM) processes data in both forward and backward directions, preserving contextual information. Integrating feature selection alongside extraction further refines the input, enhancing efficiency and improving emotion recognition.

4.1. Introduction

In recent years, rapid advancements in computer programs have focused on automating the classification of human emotions from EEG signals. However, existing techniques struggle to incorporate context information effectively, limiting their accuracy. This research enhances EEG-based emotion recognition by introducing MTSOA, which improves context management in signal processing.

MTSOA enhances feature selection by efficiently navigating complex search spaces, while IRSOA optimizes hyperparameters for better exploration. These refinements lead to more robust emotion detection. Suggested Bi-LSTM optimized with IRSOA achieves an arousal accuracy of 84.9% and a valence accuracy of 87.3%.

Decision-making, mental health, and cognitive function are all impacted by emotions, which are vital to human existence (Iyer et al., 2023). They are extremely individualized and impacted by a wide range of elements, such as outside stimuli, experiences, memories, mood and personality (Hosseini et al., 2023). While negative emotions can weaken and decrease behavior, positive emotions can enhance it. In particular, severe negative emotions, such as sadness, can have profoundly harmful impacts on mental and physical health (Gupta et al., 2023). Accurately recognizing and comprehending emotions has enormous promise and real-world applications in many different fields (Zhang et al., 2022). Applications for emotion recognition are helpful in HCI, psychology, and healthcare. Psychology examines emotional states to better understand and treat mental health conditions. It enables medical professionals to keep an eye on patients' mental and physical well-being.

By adapting interfaces to emotional inputs, HCI enhances the user experience. All things considered, emotion identification promotes advancements across a range of domains by supporting more intuitive HCI, personalized therapies, and improved mental health care. Physiological signals provide unmatched benefits because of their spontaneity, even if behavioral signals like speech and facial emotions are frequently used for affective computing (Song et al., 2021). By identifying the electrical patterns produced by neurons, EEG signals describe the underlying activity of the brain. While elevated alpha and theta rhythms are linked to lower emotional arousal, greater beta and gamma frequencies are linked to increased emotional arousal. Emotional emotions can change brain activity, producing distinct EEG patterns.

By capturing the dynamic interaction of brain activity linked to different emotions, instrument for comprehending and defining emotional experiences. By examining these frequency variations, states of various emotions can be detected. Inter band and neural oscillations coupling, which involves connections among many frequency bands, may be studied thanks to its great temporal precision (Cui et al., 2023). Many affective computing research methodologies have redirected their attention to the analysis of physiological signals as a result of these capabilities (Hernández et al., 2023). More than any other aspect, emotions impact mental health, decision-making, and cognitive processes in human existence. They are incredibly individualized, influenced by mood, personality, experiences, memories, and stimuli. While negative emotions might reduce behavior, positive feelings can increase it. The most severe condition, like depression, causes serious harm to the body and mind (Gupta et al., 2023). ER and comprehension have enormous potential in many fields (Zhang et al., 2023).

ER finds applications in psychology, healthcare, and HCI. In mental study, studying emotional states aids in understanding and treating mental health disorders. In healthcare, it helps monitor patients' well-being and response based on emotions. In HCI, ER enhances user experience by enabling connections to adapt to emotional inputs. Overall, identification of emotions supports personalized treatments, improves health care, and fosters intuitive interactions in technology, driving advancements across multiple sectors.

Although emotional computing frequently uses behavioral signals, such as speech and facial emotions, physiological signals provide unparalleled insights because of their spontaneity (Song et al., 2021). Through the detection of neuronal electrical patterns, EEG signals record brain activity. Different EEG patterns indicate different emotional states. Increased alpha and theta rhythms correspond to lower arousal, while elevated beta and gamma frequencies indicate higher arousal.

EEG is a useful tool for analyzing emotional experiences since it can be used to analyze these frequency shifts and distinguish between different emotions. Its great temporal precision makes it possible to examine neural oscillations and cross-frequency interactions, exposing intricate relationships between various frequency bands (Cui et al., 2023). Because of these benefits, physiological signals for emotion analysis are currently given priority in many affective computing studies (Hernández et al., 2023; Baradaran et al., 2023; Baradaran et al., 2023; Luo et al., 2023; Samal & Hashmi, 2023).

A major challenge in emotion recognition lies in identifying the true context in which stimuli were presented during experiments, complicating annotation and interpretation. Ideally, emotions should be defined precisely, yet establishing authentic baseline remains difficult. Biased emotional ratings or self-reports from test subjects are considered the most reliable method for recognizing and attributing emotions in experimental settings (Pusarla et al., 2022; Ramzan & Dawn, 2023).

Although EEG-based emotion recognition systems have achieved measurable success, little effort has been made to artificially evoke emotional states to optimize test accuracy. Addressing limitations in older methodologies, recent approaches focus on selecting central nervous system signals minimally affected by interfering variables. By applying wavelet and statistical analyses, signals are processed to extract features that, when combined with classification techniques, enable moderate success in classifying three, four, and five emotional states. Despite EEG's merits, further experimentation is needed to refine recognition rates and deepen insights into mood mechanisms (Gao et al., 2022; Vargahan et al., 2023; Zhang et al., 2020; Zong et al., 2023).

4.2 Contribution

The key contributions of this research are as follows.

1. MTSOA framework is introduced to improve emotion-recognition accuracy by optimizing feature selection. By effectively identifying most relevant features, the approach minimizes redundant data, enhances classification performance, and ensures a more accurate representation of emotional states. The approach optimally and adaptively selects relevant features, improving the model's discriminatory capability. This refined representation strengthens emotion recognition by focusing on the most informative aspects of input data.
2. Signals collected from DEAP database are decomposed using hybridized EMD -VMD filters. These filters ensure noise-free signals before further processing and categorization.
3. To classify human emotional states as "valence and arousal," a Bi-LSTM network is designed to optimize computational efficiency and processing time for optimal hyperparameters.

4.3 Proposed Method

The primary goal of MTSOA is to enhance EEG signal accuracy in emotion-detection tasks. The process unfolds in several phases: data collection, preprocessing, feature extraction, feature selection, hyperparameter optimization, and classification. Initially, data acquisition is conducted using publicly available datasets. The next phase focuses on eliminating irrelevant or unsuitable features through preprocessing. A feature selection method then identifies relevant, non-redundant features. IRSOA is applied to optimize hyperparameters, ensuring optimal model performance. Finally, recognition through Bi-LSTM, enabling reliable emotion prediction. Fig. 4.1 presents the schematic representation of suggested approach.

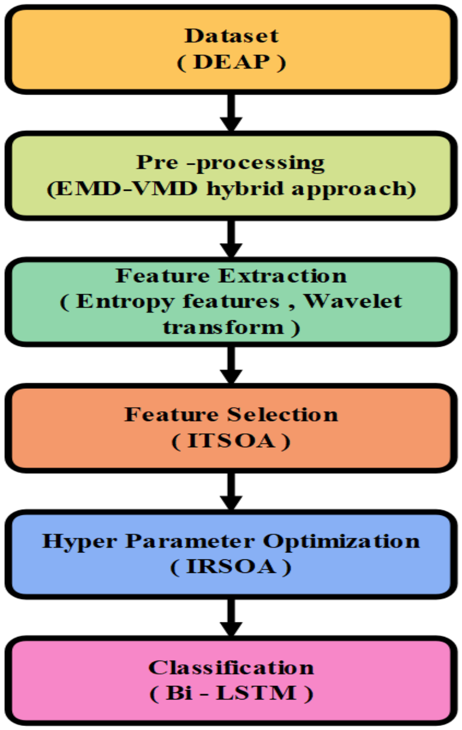


Figure 4. 1: Workflow of MTSOA-BiLSTM method

4.3.1 Data Augmentation

Raw EEG signal is collected from DEAP dataset (Salankar et al., 2021) and a detailed description of this dataset is provided below:

It contains EEG data from 40 channels recorded from 32 individuals. This study associates' data with two emotional variables: arousal and valence. It includes 40 EEG trials, each corresponding to a 60-second film, with scores assigned on a 0-9 scale. For analysis, selected EEG trials were taken from the second half of each recording.

Scores ranging from 0-5 indicate low levels, while those from 5-9 represent high levels. Emotional states in DEAP database are categorized into two groups and separated into four distinct classifications: High Arousal Low Valence (HALV) to frustration, High Arousal High Valence (HAHV) for joy, Low Arousal High Valence (LAHV) linked to leisure, and Low Arousal Low Valence (LALV) for sorrow. This categorization frames the study as a four-class classification problem. EEG signal representations are shown in Fig. 4.2.

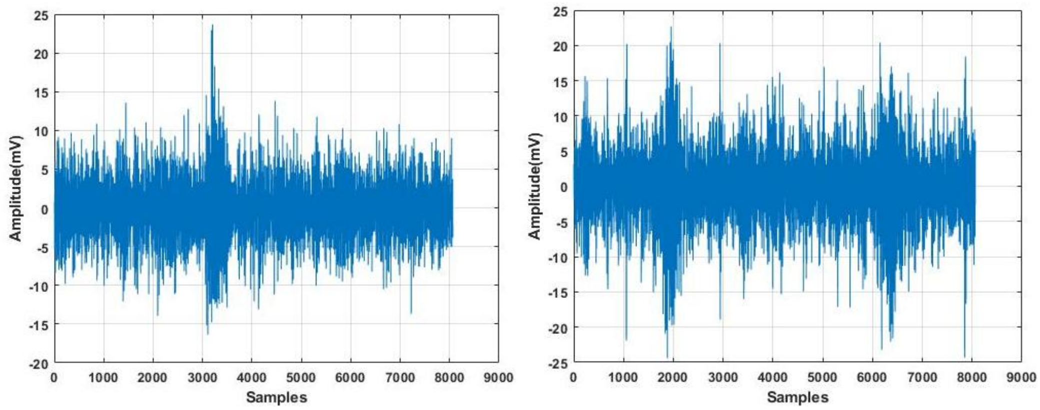


Figure 4. 2: Acquired sample signals

- **Data Preprocessing**

Once data is collected, preprocessing is applied to refine raw data without complications. EEG signals undergo artifact removal, noise filtering, and re-referencing before decomposition filters are applied. This process includes EMD (Salankar et al., 2021) and VMD (Xu et al., 2023), enhancing EEG data quality for more precise findings. These decomposition filters play a crucial role in feature extraction for emotion detection algorithms.

(i) Empirical Mode Decomposition

Signals retrieved from DEAP database undergo processing with EMD to accurately decompose unstable and non-stationary signals, and discarding unwanted noise and glitches. Process begins by separating recorded EEG signals into several IMFs using EMD method. These IMFs serve as sub-band components for further decomposition of the signal. Eqns. (4.1) and (4.2) define the process by establishing mean of both envelopes (upper, lower).

$$y(t) = \sum_{n=1}^{N} C_n(t) + r_N(t) \tag{4.1}$$

where $C_n(t)$ represents intrinsic mode obtained through time-series EEG signal $y(t)$ using EMD. Here, $y(t)$ is denoted as concatenation of decomposed IMFs along with residual component.

$$q(t) = \frac{u_h(t) + u_l(t)}{2} \tag{4.2}$$

where $u_h(t)$ is upper and $u_l(t)$ is lower envelope, $q(t)$ is mean envelope. Two essential criteria must be met during process of decomposition, ensuring that each iteration follows specific conditions. Fig. 4.3. shows a sample for EMD mechanism.

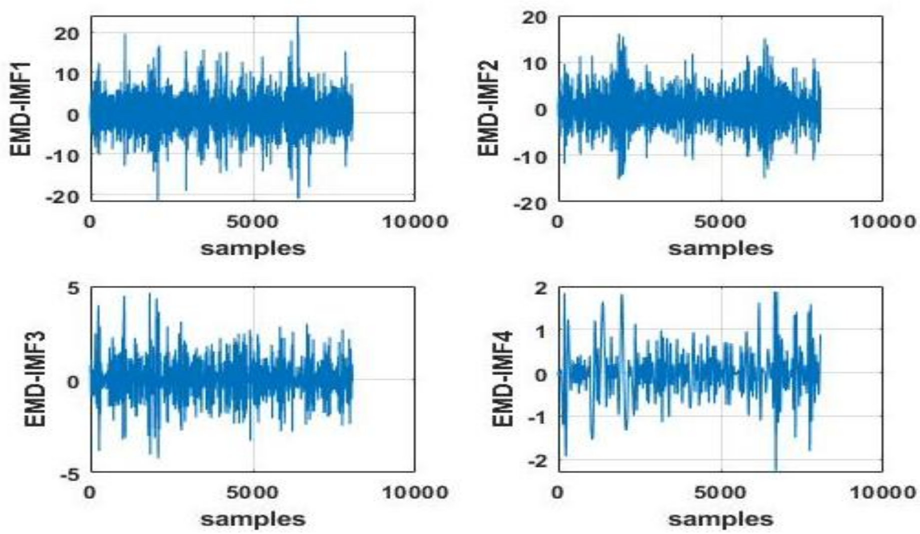


Figure 4. 3: EMD-Based Decomposition of EEG Signal

- **Variational Mode Decomposition**

VMD is an advanced signal processing technique that enables quasi-orthogonal adaptive signal decomposition. It works by iteratively solving a variational equation to generate IMFs with distinct center frequencies and bandwidths, making it highly effective for analyzing complex signals. Unlike EMD, which is prone to mode aliasing, VMD provides a more structured and adaptive approach to handling nonlinear and non-stationary signals. By optimizing frequency localization, VMD ensures better separation of signal components, making it a preferred choice for applications requiring precise feature extraction. The mathematical formulation of this variational constrained problem is detailed in Eqn. (4.3).

$$\{S_m\}^{\min},\{\xi_m\}\sum_m \left\|\partial_t(\gamma(t)+j/\pi t)*s_m(t)]e^{-j\xi_m t}\right\|_2^2 \tag{4.3}$$

where $\{S_m\}=\{s_1,...s_m\}$ represents IMFs extracted through VMD. m is total number of IMFs, and $\{\xi_m\}=\{\xi_1,...\xi_m\}$ represents centre frequency of mode. g is actual input signal, j is imaginary unit, ∂_t is a partial derivative with respect to t, $*$ denotes convolutional operation, and $\gamma(t)$ represents impulse response. The constraint in Eqn. (4.3) is incorporated using Lagrange multiplier $\mu(t)$ leading to the function in Eqn. (4.4):

$$(\{s_m\},\{\xi_m\},\mu) = \alpha\sum_m \left\|\partial_t[(\gamma(t)+j/\pi t)*s_m(t)]e^{-j\xi_m t}\right\|_2^2 + \left\|g(t)-\sum_m s_m(t)\right\|_2^2 + (\mu(t),\sum_m s_m(t)) \quad (4.4)$$

To find the optimal solution of Eqn. (4.4), an iterative approach is used, leading to updates for $s_m^{(n+1)}$, $\xi_m^{(n+1)}$, $\mu_m^{(n+1)}$.The Alternating Direction Method of Multipliers (ADMM) is applied to approximate the saddle point of Eqn. (4.5), where updated mode function $s_m^{(n+1)}$ is given by,

$$s_m^{(n+1)} = \arg\min\left\{\alpha\left\|\partial_t[(\gamma(t)+j/\pi t)*s_m(t)]e^{-j\xi_m t}\right\|_2^2 + \left\|g(t)-\sum_i s_i(t)+(\mu(t))/2\right\|_2^2 \quad (4.5)$$

where, $\xi_m = \xi_m^{(n+1)}, \sum_i s_i(t) = \sum_{i\neq m} s_i^{(n+1)}(t)$

Applying Fourier transform to Eqn. (4.6) provides frequency-domain representation of each IMF and its corresponding center frequency is given in Eqn. (4.7).

$$\hat{s}_m^{(n+1)}(w) = \frac{\hat{g}(w)-\sum_{i\neq m}\hat{s}(w)+\hat{\mu}(w)/2}{1+2\alpha(w-\xi_m)^2} \quad (4.6)$$

$$\xi_m^{(n+1)} = \frac{\int_0^\infty w\left|\hat{s}(w)\right|^2 dw}{\int_0^\infty \left|\hat{s}_m(w)\right|^2 dw} \quad (4.7)$$

In each iteration, $s_m^{(n+1)}$ is obtained by Wiener filtering $\hat{g}(w)$ after subtracting sum of all other components $\hat{g}(w)-\sum_{i\neq m}\hat{s}(w)$ ensuring that only the desired mode remains. The term $\xi_m^{(n+1)}$ represented as updated centre frequency, while $\hat{s}_m(w)$ corresponds to inverse Fourier transform of extracted component where real part of result defines the mode function $s_m(t)$.

Given the unstable and non-stationary nature of EEG signals, VMD provides an effective approach to minimizing instability. Hybrid form of sample decomposed signal is depicted in Fig. 4.4.

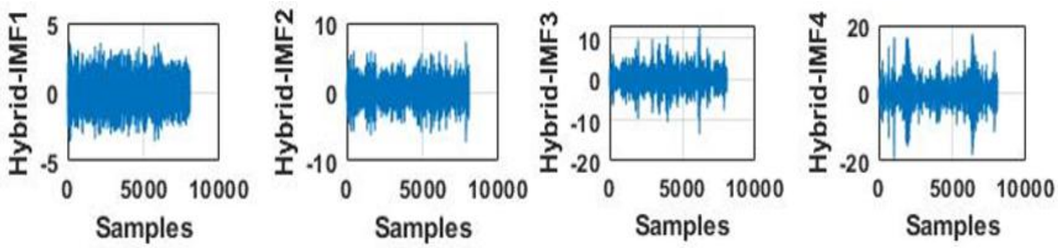


Figure 4. 4: Hybrid decmposed signal

4.3.2 Feature Extraction

After processing raw signal yields a decomposed signal, it serves as input for next stage. The extracted features aim at discovering and electing most informative and relevant attributes from original data so that these attributes can prove beneficial for analysis and learning algorithms in several applications such as ML and pattern recognition. The hybrid feature extraction entails a wide-ranging set of statistical methods for extracting features from processed signals (Al-Salman et al., 2019; Dhiman et al., 2021; Wang, 2020). When it comes to two-dimensional emotional predictions, this hybrid feature extraction gives rise to a considerably large feature length of 2676, which is then reduced to 1204 relevant features in relation to both classes. Working of this hybrid technique has several benefits, such as increased efficiency, minimal chances of overfitting, improved data illustration, and an accelerated learning process when deployed in conjunction with Bi-LSTM networks.

4.3.3 Optimal Selection

At this stage, preprocessed output is fed into the proposed feature selection process. TSOA excels in feature selection for emotion classification by adaptively exploring the feature space. It intelligently navigates feature subsets, maintaining equilibrium between discovery and utilization to identify key characteristics. This adaptive search strategy enhances the selection of distinctive features, improving model’s ability to capture strong emotional patterns. Due to its innovative approach, TSOA outperforms previous techniques, making the feature selection process more dynamic and optimizing it for emotion classification. The following is a summary of key aspects of feature selection methods:

Tunicate Swarm Optimization Algorithm

The newly employed technique, TSA, excels in finding better optimal solutions than other competing ones. It is effectively used for real-life engineering design problems. In the water, tunicates have two strategies; to discover food, Tunicates use swarm intelligence and to find optimal states for feeding in their environments.

- **Conflict-Avoidance Mechanism for Search Agents**

To reduce conflicts among tunicates, the position of a search agent is established by integrating the vectors X , Y and Z , as described in Eqns. (4.8)–(4.10):

$$X = \frac{Y}{Z} \tag{4.8}$$

$$Y = d_2 + d_3 - W \tag{4.9}$$

$$W = 2.d_1 \tag{4.10}$$

Here, Y represents gravitational influence, Z denotes social interaction forces, and W corresponds to oceanic water flow advection. The values d_1, d_2 and d_3 are random variables within a range [0,1] and estimation of Z is determined using Eqn. (4.11):

$$Z = [Q_{\min} + d_1.(Q_{\max} - Q_{\min})] \tag{4.11}$$

where $Q_{\min}$ and $Q_{\max}$ define minimum and maximum speeds for social interaction, set to 1 and 4, respectively.

- **Movement Towards Best Neighbour**

According to Eqn. (4.12), tunicates adjust their position toward their most optimal neighbour:

$$P_D = \left| W_S - \rho.P_L(s) \right| \tag{4.12}$$

Here, P_D represents movement direction of tunicate toward best food source, W_S signifies the food source's optimal position, $P_L(S)$ represents the tunicate's current location, and ρ is a random between [0,1].

- **Movement Towards the Optimal Search Agent**

Using Eqn. (4.13), the most effective position of a tunicate relative to the food source can be determined:

$$P_L(S) = \begin{cases} W_s + X.P_D, & if \quad \rho > 0.5 \\ W_S - X.P_{D,} & if \quad \rho < 0.5 \end{cases} \tag{4.13}$$

This approach ensures an efficient search mechanism by dynamically adjusting positions based on social interactions, gravitational forces, and environmental factors, ultimately leading to an optimized exploration process.

- **Swarm behaviour**

To model tunicates' behavior, the return between the first two optimal solutions must be assessed. Once established, the remaining search agents adjust their movements efficiently based on the position of the best-performing tunicate. The swarm characteristics can be estimated using Eqn. (4.14).

$$Q_L(s+1) = \frac{Q_L(s) + Q_L(s+1)}{2 + d_1} \tag{4.14}$$

where $Q_L(s+1)$ represents the position of a tunicate in the search space, and d_1 is a random variable within a range [0,1]. This formulation ensures a balanced adjustment of movement within the swarm, enhancing the efficiency of search process.

- **Modified form of Tunicate Swarm Optimization Algorithm**

For search solutions of the TSA globally and locally, MTSOA is developed in original $\vec{A}$, $\vec{G}$ and $\vec{M}$,so these vectors enable search agents in order to navigate through search space randomly while exploring efficiently without conflicts. Alterations in these vectors allow for improvement in exploitation and exploration phases. MTSOA has been chosen to perform selection of features in detection of different emotions. MTSOA is best suited to explore through complicated search spaces. Also, this system is supposed to optimize the feature subsets to improve model performance by selecting the EEG signal feature combinations with maximum relevance. This greatly enhances the flexibility and efficiency of emotion recognition systems. However, some variations, especially in high dimensions and complicated scenarios, may fall into local optima, thereby complicating global optimization. This paper proposes improving TSA's exploration ability as follows:

- **Improved exploration ability:**

Based on this research, the values of denominator and numerator were modified significantly at the beginning: high for denominator and low for numerator. Conventionally, as a result of this change, the search agents would take long steps to make the search space explore a regional area in a well-defined way. In this enhanced method, tunicates focus local search, therefore creating a balance is a must case. With the increase in number of iterations, values are lowered leading to their lowest levels. With these changes, the searching phase was also enlarged so that local exploration would even be more effective. Enhanced form vector A can be deduced mathematically as shown in the subsequent Eqn. (4.15).

$$\vec{A} = \frac{2-4(t/t_{max})}{4-3(t/t_{max})} \tag{4.15}$$

where, t and t_{max} signifying the value of the current iteration concerning the number of iterations, the suggested output of feature selection has been fed into hyper parameter optimization to better accuracy in the emotion recognition briefly detailed in the next stage.

4.3.4 Hyper Parameter Optimization

After the aforementioned feature selection stage, a hyperparameter optimization was applied via the RSOA towards accurate emotion recognition. Owing to its capability of efficiently exploring parametric space, the RSOA is utilized in emotion recognition for hyperparameter selection. The adaptive nature of RSOA and its ability to locate global optima further enhance the performance of any model, thereby ensuring that the hyperparameters are finely tuned for emotion recognition tasks, culminating in high accuracy and robustness. Hyperparameters control the way the ML models operate while being important for their performance and generalization. They might include learning rate ranging from [0.001, 1]; hidden units ranging between [10-500]; regularization strength varying from 0.01 to 0.1; batch size from 32 -512; and activation functions used are ReLU and sigmoid. RSOA is a highly eminent heuristic optimization technique based on the aggressive and following behaviours of rats. In hyperparameter optimization, the aggressive and follower rat behaviours are mapped to improve performance. RSOA sets out the initialization of the search space, which is scientifically denoted under Eqn. (4.16).

$$y_j = y_j^L + rand \times (y_j^U - y_j^L), \; J = 1,2....,m \tag{4.16}$$

In Eqn. (4.16), y_j^U and y_j^L represent upper and lower bounds for j^{th} variable, and m denotes total number of agents.

Next, aggressive movement of each agent is modelled by updating their positions using Eqn. (4.17):

$$Q_i(x+1) = \left| Q*(x) - Q_i(x) \right| \tag{4.17}$$

Here, $Q_i(x+1)$ indicates the updated position of i^{th} agent, while $Q*(x)$ denotes current best solution. The new position is then computed via Eqn. (4.18):

$$Q(x) = \varepsilon \times Q_i(x) + \beta \times (Q*(x) - Q_i(x)) \tag{4.18}$$

In this context, $Q_i(x)$ represents present position of i^{th} agent, coefficients ε and β are determined by Eqns. (4.19) and (4.20), respectively;

$$\varepsilon = S - t \times \left(\frac{S}{MaxIter} \right), \quad t = 1,2,....,MaxIter \tag{4.19}$$

$$\beta = 2 \times rand \tag{4.20}$$

Here, $MaxIter$ is maximum number of iterations, t is current iteration, and S is a random number selected from interval [1,5]; β is a random value within [0,2].

The search space is encapsulated by an N dimensional vector as shown in Eqn. (4.21):

$$Y = (y_1, y_2, ...y_N) \tag{4.21}$$

with each $y_j \in [y_j^U, y_j^L]$. The complementary point for each y_j is defined in Eqn. (4.22):

$$\hat{y}_j = (y_j^U, y_j^L) - y_j, \quad j = 1,2,...N \tag{4.22}$$

To further enhance search space in MTSOA at each iteration, best solution is substituted with a new candidate according to Eqn. (4.23),

$$Z_{WORST} = \begin{cases} rand_a \times Q*(x), & if \quad rand_c < 0.5, \\ (y_j^U - y_j^L) - y_j, & if \quad rand_c > 0.5 \end{cases} \tag{4.23}$$

In Eqn. (4.23), $rand_c$ and $rand_a$ are random numbers between [0,1], and z_{WORST} denotes a candidate solution with a high value. By applying this replacement strategy, hyperparameter optimization is enhanced as the positions of the least promising agents are swapped with the best solution $Q*(x)$ within each group.

This integrated process of initialization, aggressive position updating, coefficient-guided movement, and strategic candidate replacement enables the MTSOA to effectively explore the search space and converge toward optimal solutions.

4.3.5 Classification Using BI-LSTM

These methods are applied then followed by hyperparameter optimization, and the Bi-LSTM is run on the DEAP dataset for improving the accuracy of the emotion classification task. Bi-LSTM is needed for emotion classification as it encapsulates the understanding of past and future contextual information while temporal dependencies are established in emotional sequences. This is one of the great advantages to having a model capable of bidirectional dependencies, which can thus better understand emotional expression. The choice of the Bi-LSTM model came from the necessity of modeling complex temporal patterns. It thus outperforms other models through the method of capturing long-distance dependencies, which, in turn, improve emotion classification model accuracy for many subtle, context-dependent, and richer interpretations of emotional content.

Bi-LSTMs are powerful deep-learning models that are employed for ER tasks. The dual directional aspect of Bi-LSTMs examines not only past information but also future information, strengthening the understanding of context in the model. In emotion classification, Bi-LSTM processes text input and captures the emotional context embedded within it. Its adeptness in intermediate contexts and sequential data makes the model a superior choice in ER enabling deep comprehension of emotions in numerous applications. Fig. 4.5 depicts the working of Bi-LSTM.

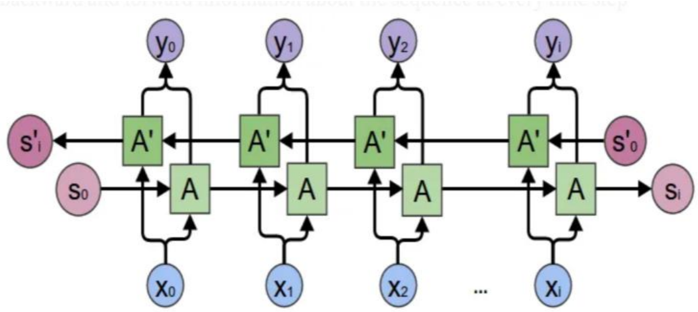


Figure 4. 5: Architecture of Bi-LSTM

A forward RNN, represented as $B_1 \to B_2 .... \to A_t$ processes sequential data progressively. At step t, input consist of data in sequential form y_t while output is B_{t-1} from prior time step. Conversely, a backward RNN, denoted as $B_1 \to B_2 .... \to A_t$, performs computations in reverse. At step t, input is sequential data y_t and output is B_{t-1} from subsequent time step.

4.4 Experimental Analysis

The findings of the suggested method are assessed in this part in order to produce classification-based results. Performance analysis and comparison analysis are sub-sections of the results section. Evaluating the effectiveness of suggested method using DEAP dataset is part of performance analysis. The impact of developed strategy is compared to previous strategies that have been recorded in literature for the comparative study.

4.4.1 Experimental Setup

For this research, it is said that performance of MTSOA network has been simulated using a system arrangement of 16GB RAM on an Intel i7 processor using Matlab R2020b. The study also discusses some performance metrics such as precision sensitivity f1-score specificity and classification accuracy for evaluating performance of intended MTSOA.

4.4.2 Performance Analysis of Original Features

In this section, an assessment performed on the proposed MTSOA within DEAP dataset. Results are then compared against numerous classifiers, among them RF, DNN, MSVM, and Generative Adversarial Network (GAN), which are described in the highly informative Table 4.1 and Table 4.2. The experimental findings are also classified into distinct classes using the DEAP dataset.

As seen in Tables 4.1 and 4.2, the accuracy of 83.54 % for arousal and a valence of 86.25 % indicates that Bi-LSTM has better performance when analysed with existing models. Comprehensive experimental results define that proposed MTSOA efficiently reduces information loss while achieving high accuracy in two-stage classification of emotions.

Table 4. 1: Comparative analysis of original features for class-arousal

Metric (%)	DNN methods				
	RF	GAN	DNN	MSVM	Bi-LSTM
Sensitivity	82.67	81.87	82.97	70.74	86.99
Accuracy	80.13	81.76	80	67.44	83.51
F1-score	81.86	79.86	80.05	70.29	84.46
Specificity	82.92	79.28	83.99	69.07	84.17
Precision	82.96	81.84	91	69.85	82.01

Table 4. 2: Comprehensive analysis of actual features forvalence

Metric (%)	Models				
	RF	GAN	DNN	MSVM	Bi-LSTM
Accuracy	82.81	81.67	79.45	82.29	86.25
Sensitivity	80.01	80.24	79.78	79.37	86.92
Specificity	79.66	81.49	78.11	80.68	87.91
F1-score	81.56	81.31	79.25	79.88	86.11
Precision	83.16	82.49	78.73	80.38	85.31

4.4.3 Performance Analysis of Optimized Features

The optimized features represent those chosen traits or characteristics or process that undergone some improvement in fine-tuning to attain maximum efficiency.

Table 4. 3: Comprehensive evalaution of optimized features for arousal

Metric (%)	Methods				
	RF	GAN	DNN	MSVM	Bi-LSTM
Sensitivity	80.83	81.62	81.62	72.51	88.83
Accuracy	82.71	82.86	81.82	76.87	84.09
F1_score	82.17	82.21	82.21	75.99	89.11
Specificity	85.63	85.71	82.71	72.56	86.25
Precision	83.54	82.79	82.79	79.83	89.38

Furthermore, the results from experiments conducted across differing classes, such as arousal and valence, are illustrated in Table 4.3 and Table 4.4 using the DEAP dataset.

Table 4. 4: Optimized feature performance evaluation for Valence

Metric (%)	RF	GAN	DNN	MSVM	Bi-LSTM
Accuracy	85.54	84.83	86.67	82.29	87.31
Sensitivity	83.58	90.96	85.11	81.37	92.07
Specificity	84.58	90.58	84.11	80.68	93.41
F1_score	83.37	92.06	85.52	80.88	93.49
Precision	83.16	93.22	85.95	80.38	94.95

Tables 4.3 and4.4 indicates the comparison of the proposed technique with the previous techniques. The better result or the highest possible accuracy was obtained, which is 84.9% in arousal and 87.3% in valence. This highlights its efficiency in managing large datasets through effective feature selection and accurate classification of various emotional states.

4.4.4 Comparative Evaluation of Selection Methods

Concerning the Tables, experimental outcomes have been compared against other optimization methods like Manta Ray Foraging Optimization (MRFO) (Liu et al., 2024), Multi-parameter Optimization (MPO) (Wu et al., 2021), and Whale Optimization Algorithm (WOA) (Rajasekhar et al., 2020) as found in Tables 4.5 and 4.6.

Table 4. 5: Evaluation of feature selection techniques for Arousal classification

Metric (%)	Arousal			
	MRFO	MPO	WOA	Proposed
Accuracy	81.75	82.39	73.53	85.58
Sensitivity	82.77	83.33	74.95	88.83
Specificity	84.81	84.33	72.6	86.25
F1-score	82.13	83.15	73.28	89.14
Precision	81.51	83	71.68	89.38

The DEAP dataset is analysed in terms of unusual classes, of valence and arousal.

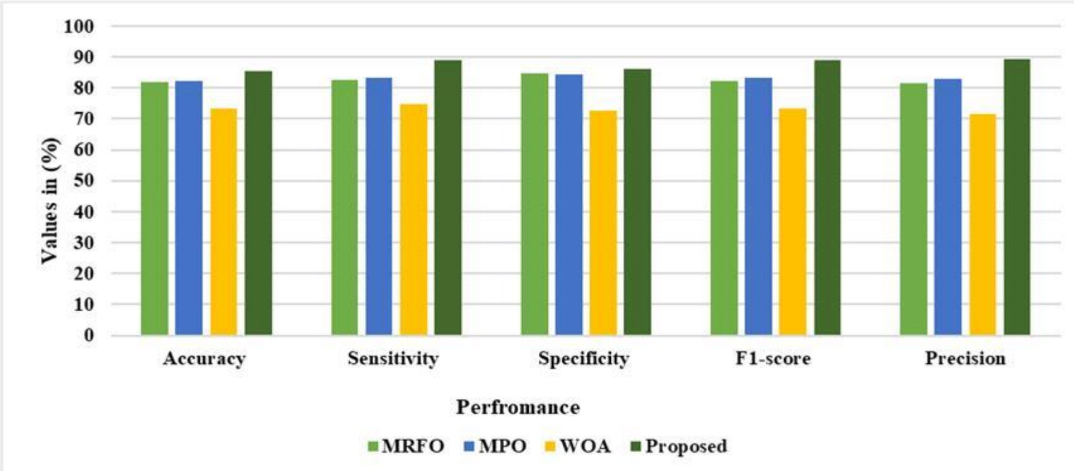


Figure 4. 6: Visual depiction of proposed approach for arousal

Table 4. 6: Evaluation of feature selection optimization performance for valence

Metric (%)	Valence			
	MRFO	MPO	WOA	Proposed
Accuracy	84.21	80.58	81.83	91.09
Sensitivity	84.07	79.94	83.03	92.07
Specificity	85.01	81.12	81.89	93.41
F1-score	84.93	79.78	83.36	93.49
Precision	85.81	79.63	83.68	94.95

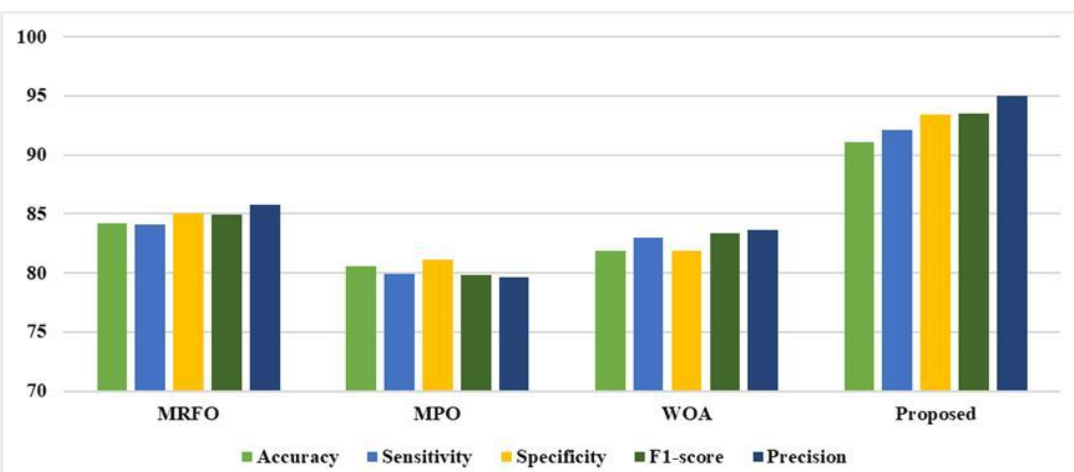


Figure 4. 7: Visual representation of MTSOA for valence

It is clear from Tables 4.5 and 4.6 that the proposed approach outperforms current approaches in terms of arousal (85.58%) and valence (91.09%). Figs. 4.6 and 4.7, respectively, show a graphical representation of recommended selection of feature using optimization technique.

4.4.5 Comparative Evaluation of Different Optimization Techniques

The results of simulation with subsequent MTSOA are tested on DEAP dataset. As indicated in Table 4.7 and Table 4.8, the proposed model is compared with various optimizations methods such as PSO, Salp Swarm Algorithm (SSA) (Sunaryono et al., 2022), and Cat Swarm Optimization Algorithm (CSOA) (Anem et al., 2020). Experimental results of different classes such as valence and arousal are demonstrated in Tables 4.7 and 4.8.

Table 4. 7: Comparative assessment of optimization techniques for arousal

Metric (%)	Methods			
	PSO	SSA	CSOA	Proposed
Sensitivity	82.28	75.35	74.86	88.83
F1_score	81.18	74.37	73.12	89.16
Accuracy	80.57	77.39	73.45	84.97
Specificity	84.01	74.33	72.55	86.25
Precision	80.12	73.45	71.47	89.38

Table 4. 8: Assessment of hyperparameter optimization strategies for Valence

Metric (%)	PSO	SSA	CSOA	Proposed
Sensitivity	82.58	80.69	84.6	92.07
F1_score	83.92	80.07	84.51	93.49
Accuracy	82.32	81.54	81.18	91.11
Specificity	85.93	81.31	80.44	93.41
Precision	85.38	79.46	84.43	94.95

As shown in Table 4.7 and Table 4.8, developed model achieved an accuracy of 84.9% for arousal and 91.11% for valence, therefore outperforming existing techniques. It is evident from the extensive experimental results that proposed MTSOA network effectively reduces information loss and enhances recognition efficiency for two tier ER which is further presented visually in Fig. 4.8.

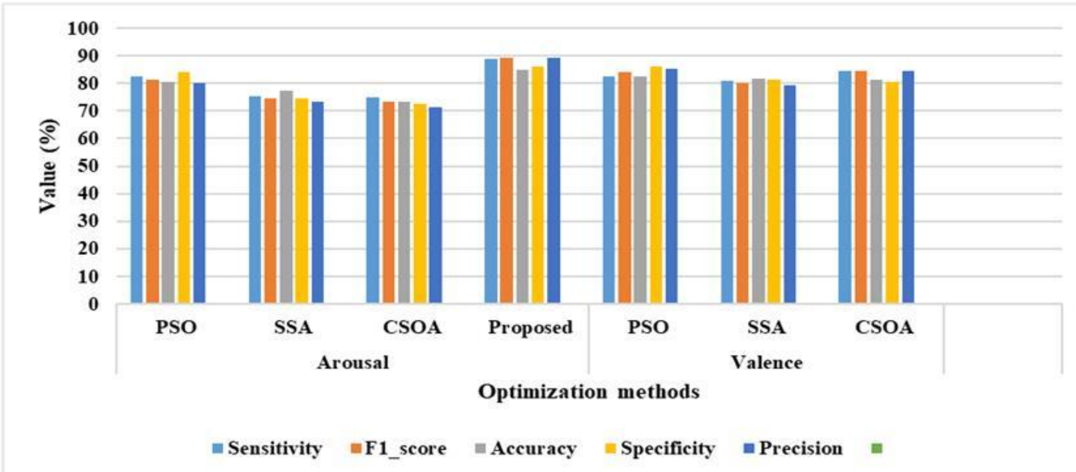


Figure 4. 8: Comparison of various optimizers for both classes

4.4.6 Comparative Analysis

Comparative analysis compares the data finding commonalities and differences among them for gaining insight or making decisions. In this subsection, an assessment has been done for verifying performance of classification approach through comparison with conventional approaches listed under related works. Table 4.9 presents results obtained through evaluation of the proposed approach. Fig. 4.5 displays a representation of suggested approach using several algorithms.

Table 4. 9: Evaluation of MTSOA-Bi-LSTM network in comparison to current models

Classes	ML/DL Methods			
	MEEMD	CNN with SVM	Kenel fusion +DNN	MTSOA-Bi-LSTM network
Arousal	78.00	80.52	64.50	84.90
Valence	74.03	75.22	63.10	87.39

MTSOA yields superior results as compared to other conventional methods because the positive benefits were contributions of feature selection in adaptive manner and efficient parameter adjustment toward better performance of MTSOA in this area as compared to other methods. This indicates that it can carry emotional states effectively regarding optimization for classification purposes in emotion-based problems. MTSOA is much more efficient by 7%, 4%, 26.2% than (Samal & Hashmi, 2023; Gao et al., 2022; Zhang et al., 2020) and in the ER arena because of its adaptive selection of features and fine parameter tuning. Therefore, MTSOA yields a much accurate representation of emotional patterns.

On the other hand, MEEMD lacks versatility, CNN+SVM has problems with dynamic feature selection, and finds it difficult to interpret the complex emotional aspects within the samples. However, MTSOA's accomplishments can be limited by initial field and dataset characteristics. Moving forward, the MTSOA network recommended method should be benchmarked into other dataset for e-learner's ER and refined for recognition tasks.

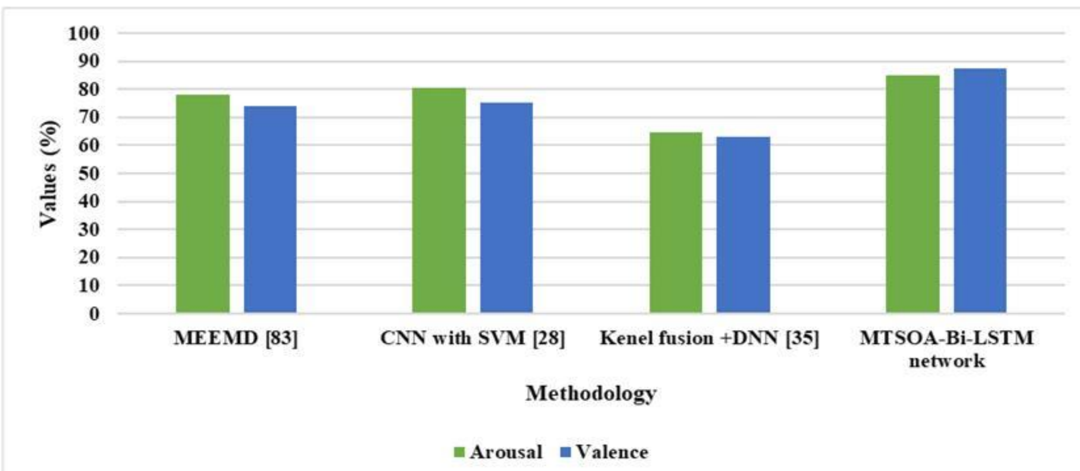


Figure 4. 9: Visualization of MTSOA-Bi-LSTM network with current models

Table 4. 9 with Fig. 4. 9 carries clear-cut and demonstrable evidence of MTSOA-Bi-LSTM classification method outscoring current methods in all performance parameters. Suggested approach achieved an accuracy of approximately 84.9% for arousal and 87.3% for valence. In comparison, MEEMD recorded around 78.0% for arousal and 74.03% for valence, CNN+SVM obtained roughly 80.52% for arousal and 75.22% for valence, and reached about 64.5% for arousal and 63.1% for valence

4.5 Summary

An innovative MTSOA network is presented for effective classification of emotion based on EEG signals, targeting the difficult job of classifying human emotion from the EEG recorded over the past several years. DEAP is used to capture and classify emotions of human. Then, pre-processing was done using hybrid EMD +VMD to reduce the artifacts in the EEG, hence retrieving the most informative details. The outcomes show that this research represents the efficacy of MTSOA for emotion classification, outperforming traditional methods. The key findings are adaptive selection of feature and optimal parameter tuning for enhancing the accuracy. The importance of TSOA in regard to the potential ramifications in the field of emotion classification has shown it to be a robust and trustworthy optimization strategy. This research, therefore, has been utilized for the enhancement of the proposed technique, thereby contributing to many applications. From the results, proposed method shows better implications for arousal with 84.9% and valence of 87.3%, which are considerably greater than other methods used for comparison. In future work, the proposed MTSOA network would be validated using different datasets in the classification of e-learner's emotion.

CHAPTER 5

A Novel Framework for Emotion Recognition from EEG Signals Using Deep Neural Network Models Optimized with An Enhanced Rat Swarm Algorithm

Overview

The MTSOA network demonstrates effective EEG-based emotion classification, outperforming traditional methods through adaptive feature selection and optimal parameter tuning. Its robust optimization enhances classification accuracy, making it a reliable approach for emotion recognition. However, despite its efficiency, Hybrid IRSOA-CNN+LSTM achieves even higher accuracy by leveraging DL for better feature learning and sequence modelling. This hybrid approach further refines emotion classification by combining CNN's spatial feature extraction with LSTM's temporal learning offers superior performance in EEG-based ER.

5.1 Introduction

ER from EEG signals is crucial for advancing applications in computing, HCI, and mental health diagnostics. It enables non-invasive, real-time analysis of brain signals to interpret human emotions, making systems more adaptive and personalized (Zhang et al., 2022). In computing and HCI, it enhances user experience by creating emotion-aware interfaces, such as intelligent virtual assistants and adaptive learning platforms. In healthcare, it supports early detection and treatment of mental health conditions like depression and anxiety. Despite its potential, complex and variable nature of EEG signals poses challenges, highlighting the need for robust methodologies to achieve precise emotion classification as depicted in Fig. 5.1 (Gaddanakeri et al., 2024; Liu et al., 2021; Gao et al., 2021; Nawaz et al., 2023).

Traditional methods have shown potential in enhancing emotion identification; however, DL approaches have significantly revolutionized the field. Various architectures and strategies are suggested to address intricacies of EEG signals. CNNs excel in extracting spatial and local features, while LSTM networks are adept at capturing temporal dependencies in sequential data. Hybrid approaches, combining CNN and LSTM, have shown improved performance by leveraging both spatial and temporal information (Gul et al., 2023; Song et al., 2018; Mahmud et al., 2022; Gao et al., 2015; Chakravarthy & Suchithra, 2023; Sharma et al., 2023; Wang et al., 2019).

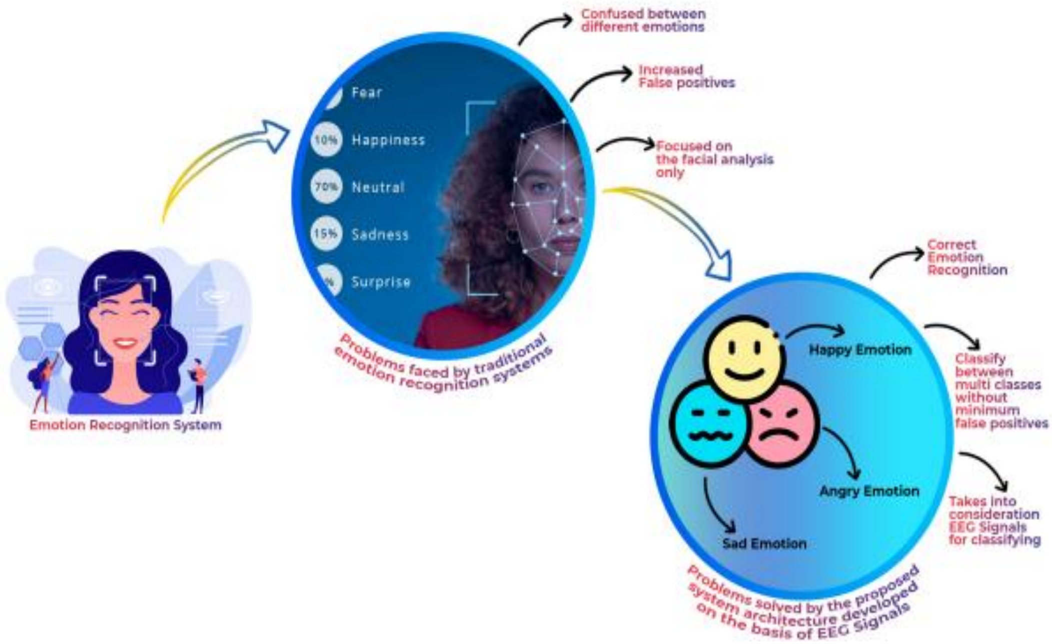


Figure 5. 1: Emotion recognition using traditional and proposed method

Beyond CNNs and LSTMs, more advanced DL methods have been explored to further enhance ER. Similarity learning networks have improved classification accuracy, while graph neural networks with contrastive learning and GAN-based augmentation have demonstrated the ability to seizure complicated relationships in EEG data. Techniques like DE-CNN combined with RNNs have achieved high accuracy by integrating spatial and temporal features, while hierarchical attention-based temporal convolutional networks focus on enhanced temporal pattern recognition. Methods using adversarial learning, such as subject-independent frameworks, have addressed variability across subjects, and self-attention-based dynamic graph neural networks have improved the modelling of EEG's intricate dependencies (Gilakjani & Al Osman, 2023; Zhao et al., 2023; Wen et al., 2017; Qu & Jing, 2024).

Optimization techniques help in improving performance of these models. For instance, approaches integrating algorithms like RSO and its variants have been proposed to fine-tune deep neural networks, ensuring better generalization and accuracy in ER tasks. These methods help to define the challenges posed by high-dimensional and non-linear nature of EEG data (Maria et al., 2022; Mao et al., 2021; Torres et al., 2021; Li et al., 2021; Singh et al., 2020).

Building on these advancements, this research proposes a hybrid CNN-LSTM DL framework optimized using IRSOA for EEG-based ER. This approach leverages the strengths of CNN and LSTM architectures while incorporating optimization techniques to enhance precision and reliability of emotion classification. By addressing both space and time-related complexities in EEG signals, this study aims to advance the conventional methods in real-time emotion detection, making it more robust and applicable across diverse domains (Tripathi & Choudhury, 2023; Bano et al., 2022; Hwang et al., 2020; Ergin et al., 2021; Moon et al., 2018; Li et al., 2022; Chen et al., 2019; Parveen et al., 2023; Yang & Liu, 2019).

Major Contributions:

1. **Integration of CNN and LSTM:**

The framework combines spatial feature extraction capabilities of CNNs with time dependency modeling of LSTMs, enabling a comprehensive analysis of EEG signals. This hybrid approach effectively captures both spatial and temporal complexities, enhancing ER accuracy.

2. **Enhanced Rat Swarm Optimization Algorithm:**

IRSOA is employed for hyperparameter optimization, fine-tuning critical parameters like learning rates and dropout rates. This ensures improved network performance, faster convergence, and better generalization to handle the non-linear and dynamic nature of EEG signals.

3. **Scalability and Robustness:**

The method is validated on multiple EEG datasets, demonstrating its adaptability to diverse data characteristics and its suitability for real-time applications. Its scalability and consistent accuracy in dynamic scenarios highlight its potential for practical use in fields like HCI, mental health, and BCI.

To enhance performance efficiency, optimization plays a pivotal role. IRSOA serves as a self-adaptive optimizer and an evolutionary algorithm inspired by foraging behaviour of rats. This technique enables efficient exploration of the solution space, making it particularly effective in optimizing hyperparameters for DL methods. The proposed CNN-LSTM model, combined with IRSOA optimizer, significantly improves accuracy of ER which leverages the strengths of both DL approaches. By incorporating IRSOA, the approach not only optimizes the DL model but also addresses challenges such as computational complexity and variations in EEG signals. This integration ensures extraction of relevant features from EEG data while maintaining scalability and robustness, ultimately enhancing the overall performance and efficiency of ER systems.

Problem Statement:

EEG-based emotion detection is a challenging yet highly significant task, with applications spanning mental health services, HCI, and neuromarketing, where accurately identifying human emotions adds immense value to consumer satisfaction and therapeutic effectiveness. Traditional approaches have often relied on manual feature extraction and shallow classifiers, which have largely failed to capture the intricate complexities and non-linear interdependencies inherent in EEG signals. Recent work by Yin et al. (2017) demonstrates that DNN effectively leverages raw EEG information to improve the performance of ER systems. However, fine-tuning these deep networks to accommodate the high-dimensional and variable nature of EEG signals remains a significant challenge. A promising approach could involve utilizing advanced optimization techniques, such as IRSOA, to optimize DNN parameters. This could improve both accuracy and generalization in ER systems, addressing the limitations of earlier methods.

5.2 Proposed Methodology

The methodology depicted in Fig. 5.2 emphasizes the integration of advanced signal processing techniques with DNN and an innovative optimization algorithm to achieve highly accurate ER from EEG signals. This comprehensive approach is structured into several key stages to ensure precision and robustness.

The process begins with preprocessing, where unwanted noise and artifacts present in raw EEG signals are effectively removed. This step ensures cleaner data for subsequent stages, minimizing the influence of irrelevant or erroneous information. Following preprocessing, feature extraction is performed using advanced techniques such as VMD and EMD. These techniques are specifically chosen to extract most relevant and suitable features from EEG signals, ensuring that the emotional characteristics embedded in the data are accurately captured.

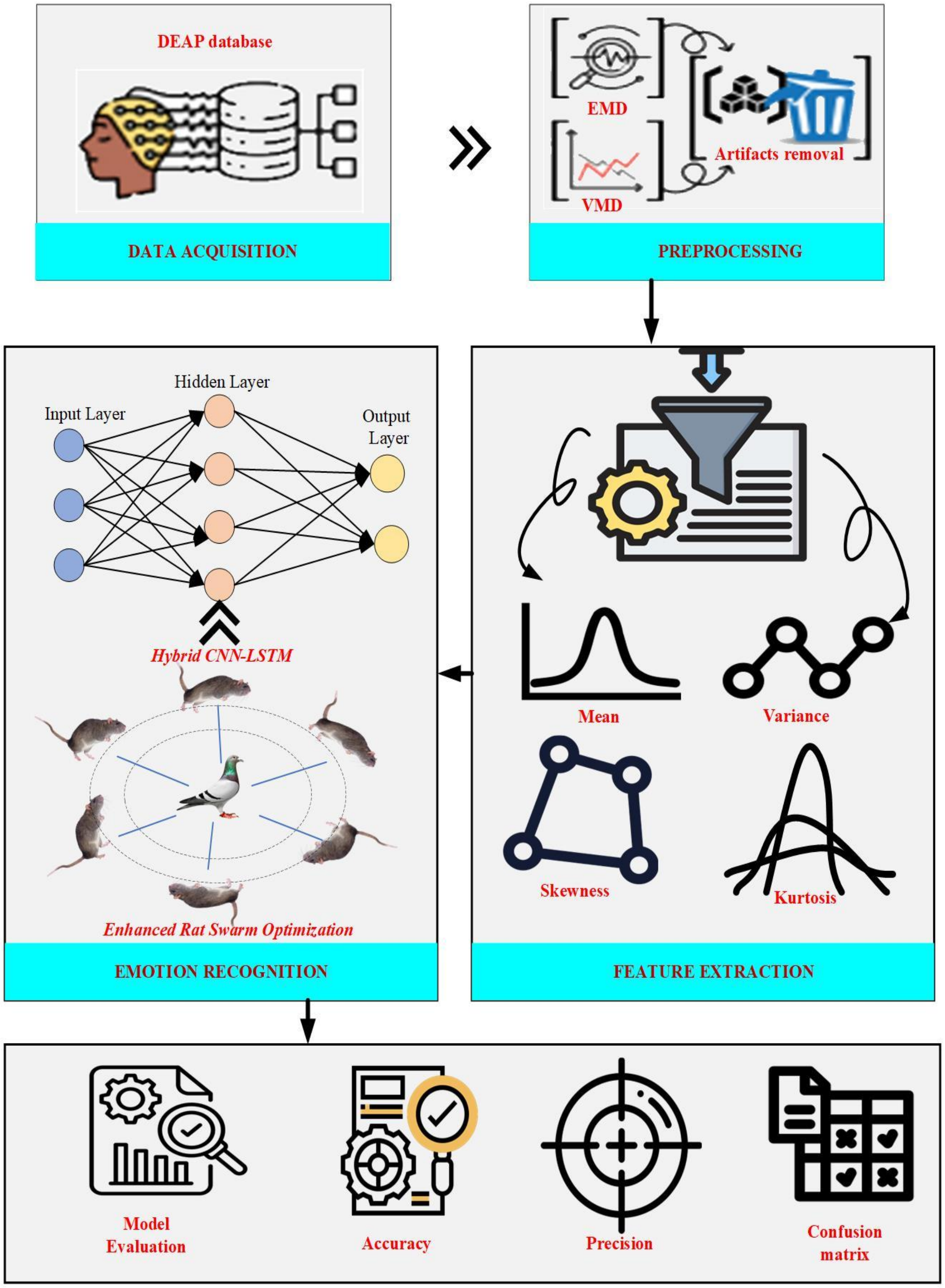


Figure 5. 2: Workflow of Hybrid CNN-LSTM with Enhanced Rat Swarm Optimization

Once the features are extracted, the methodology proceeds to the classification stage, where a hybrid DL architecture combining CNN and LSTM is employed. This hybrid model leverages CNN's ability to capture spatial patterns and LSTM's capability to handle time dependencies, enabling precise classification of different emotional classes, such as arousal and valence.

To further enhance model's performance, DNN framework is fine-tuned using the IRSOA. This advanced optimization method dynamically adjusts the hyperparameters of the DL model, ensuring robust and efficient emotion classification results.

5.2.1 Data Acquisition and Preprocessing

The EEG signals used are sourced from DEAP database. This database comprises EEG recordings from 32 participants who watched 40 one-minute segments of music videos. Initially, real-time data are pre-processed to standardize the format, including down-sampling to 128 Hz and segmenting into smaller trials, making the data more suitable for ML applications.

Initially, previously established preprocessing steps are adopted, starting with the removal of EOG artifacts, followed by bandpass filtering within a range of 4.0–45.0 Hz. Additionally, the EEG channels are reordered according to the Geneva configuration. Further preprocessing involved advanced decomposition techniques, such as EMD and VMD, to address non-linear and non-stationary characteristics of EEG signals. Fig. 5. 2 depicts the structural flow of proposed hybrid CNN+LSTM with IRSOA in the recognition of human emotions from EEG brain waves.

Using EMD, the EEG signals are decomposed into distinct IMFs, effectively segregating the signals into their corresponding low-frequency sub-signals. These decomposition methods enhanced the discrimination capability of the EEG signals by isolating relevant components and minimizing noise and artifacts. By applying these techniques, only the most meaningful signal components are preserved for subsequent analysis, ensuring improved accuracy and reliability in downstream processing.

- **Empirical Mode Decomposition**

Signals in this context are typically complex, non-linear, and non-stationary, requiring advanced signal processing techniques to handle time-dependent data. EMD technique is used to decompose a signal into a finite number of IMFs, where each IMF represents a distinct oscillatory mode inherent in the data. Below is a more detailed explanation of the EMD process:

1. Local extrema identification:

To identify all local extrema in EEG signal, both maxima and minima. This is important because the extrema define the boundaries between different oscillatory components of the signal, which will later be used to construct the envelopes.

2. Generating upper and lower envelopes:

The next step is to create two envelopes that enclose the signal. The upper envelope is generated by interpolating between local maxima, and lower envelope is created by interpolating between local minima. These envelopes represent upper and lower bounds of signal, capturing its fluctuating behaviour.

3. Computing means of envelope:

Once upper and lower envelopes are generated, mean of two envelopes is computed. This mean envelope represents a smooth approximation of original signal's trend, which is used to isolate the oscillatory components.

4. Subtracting the mean envelope (Proto-IMF):

The mean envelope is then subtracted from the original signal. This results in a "proto-IMF," the first approximation of an IMF. The proto-IMF represents the signal's oscillatory behavior without the trend captured by the mean envelope.

5. Checking IMF Criteria:

At this stage, the proto-IMF is checked to see if it satisfies the criteria for an IMF. Specifically, the IMF must have a zero mean and be symmetric about zero. If the proto-IMF does not meet these criteria, the process of envelope generation and subtraction (steps 1–4) is repeated on the proto-IMF until the resulting signal satisfies the IMF conditions.

6. Extracting and decomposing residual signal:

Once a valid IMF is extracted, it is subtracted from original signal. Residual, which is the remaining part of the signal, is then treated as a new signal to undergo the same decomposition process. This is repeated iteratively until the residual signal becomes monotonic, meaning it no longer has oscillatory components which creates a set of IMFs that capture various intrinsic oscillations present in the original signal. The mathematical formula to calculate EMD is given in Eqn. (5.1)

$$Y(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \tag{5.1}$$

where, $Y(t)$ is input signal, $IMF_i(t)$ is i^{th} IMF at time t, $r_n(t)$ is final residual.

- **Variational Mode Decomposition**

EMD, face several limitations which includes 1) their algorithmic ad-hoc nature, which lacks a solid mathematical theory; 2) the recursive sifting process used in most methods, which does not support backward error correction; 3) an inability to properly handle noise, leading to suboptimal performance in noisy environments; 4) the rigid band-limits imposed by wavelet approaches, which limit flexibility in capturing frequency characteristics; and 5) the need for predefined filter bank boundaries in techniques like empirical WT making them less adaptable to varying signal conditions.

In contrast, a variational model offers significant advantages. This model adapts to determine relevant bands and estimates the corresponding modes concurrently, addressing the issues of predefined boundaries and rigid band-limits. By balancing errors between the modes, it effectively handles the decomposition process. The variational model also accounts for noise in the input signal, improving robustness. Additionally, it focuses on narrow-band properties and reconstructs input signal optimally, either exactly or in a least-squares sense, by estimating center frequencies on-line, making it more flexible and accurate compared to traditional methods like EMD. The steps followed in VWM are,

1. Initialize the mode functions $y_k(t)$ along with their corresponding central frequencies wk
2. To minimize the total bandwidth, the optimization problem is solved by ensuring that sum of modes is equal to original signal. The optimization is formulated using the equations provided in Eqns. (5.2) and (5.3).

$$\min_{u_k, w_k} \sum_{k=1}^{K} \left\| \frac{\partial}{\partial t} \left[(\delta(t) + \frac{j}{\pi t}) y_k(t) \right] e^{-jw_k t} \right\|_2^2 \tag{5.2}$$

$$\sum_{k=1}^{K} y_k(t) = Y(t) \tag{5.3}$$

where, $y_k(t)$ is k^{th} mode function at a time t, $\delta(t)$ is used in signal processing to model an impulse or a signal with infinite energy concentrated at a single point in time, K is a total number of modes, π is a scaling factor, $*$ represents convolution, imaginary unit, which is used to handle complex numbers in the frequency domain is denoted by imaginary unit, which is used to handle complex numbers in the frequency domain by j.

5.2.2 Feature Extraction

After preprocessing the signals, the next step involved feature extraction to capture meaningful information from the EEG signals that correlate with emotional states. A variety of statistical features are extracted from each decomposed signal components. These statistical features provided valuable insights into temporal dynamics of brain activity associated with different emotions. Additionally, frequency domain features, are computed to capture oscillatory behavior of EEG signals across various frequency bands. The combination of these features formed a comprehensive representation of the EEG signals, thereby improving model's ability to differentiate between emotional states.

5.2.3 Various Feature Extraction Techniques

The steps are critical, involving the extraction of meaningful statistical and frequency domain features from pre-processed EEG signals to capture information pertinent to different emotional states and the mathematical calculation is given in Eqns. (5.4) to (5.7).

- **Mean** captures the average value of the signal over time, providing insight into the general activity level of the brain. It helps understand the baseline brain activity during different emotional states.

$$mean = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{5.4}$$

- **Variance** measures the spread or dispersion of the signal, indicating the degree to which the EEG signal deviates from the mean. A higher variance could signify more intense fluctuations in brain activity, potentially reflecting emotional arousal or state changes.

$$\mathrm{var}iance = \frac{1}{N}\sum_{i=1}^{N} (y_i - mean)^2 \tag{5.5}$$

- **Skewness** quantifies the asymmetry of the signal's distribution. A skewed signal suggests an imbalance in the distribution of brain activity, potentially indicative of a specific emotional response or cognitive process. For example, positive skewness might suggest an increased presence of higher brain activity, while negative skewness could indicate more stable or lower activity levels.

$$Skewness = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - mean}{\mathrm{var}\,iance}\right)^3 \tag{5.6}$$

- **Kurtosis** measures the "tailedness" or sharpness of the signal's distribution. Higher kurtosis indicate that the signal has more extreme values, which could reflect unusual or intense emotional responses. It indicates how outlier events (such as emotional spikes) contribute to the overall signal.

$$Kurtosis = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{y_i - mean}{\mathrm{var}\,iance}\right)^4 - 3 \tag{5.7}$$

  where, y_i is value of signal at i^{th} point in time, N is total number of data points or samples.

- **PSD** is a frequency domain feature that estimates the power of the signal as a function of frequency. PSD is crucial for analyzing the oscillatory behavior of the brain across various frequency bands. Each band is associated with specific cognitive or emotional states:
  - **Delta (0.5-4 Hz)**: Often linked with deep sleep or relaxation.
  - **Theta (4-8 Hz)**: Associated with light sleep, deep relaxation, and meditation.
  - **Alpha (8-13 Hz)**: Represents calm, relaxed alertness, often seen in states of relaxation and meditation.
  - **Beta (13-30 Hz)**: Corresponds to active thinking.
  - **Gamma (30-100 Hz)**: Connected with high-level cognitive functioning, memory processing, and emotional responses (Li et al., 2022).

For an input signal $y(t)$, the PSD is calculated based on Fourier transform $Y(f)$ is given in Eqn. (5.8).

$$PSD(w) = |Y(f)|^2 \tag{5.8}$$

where, $w = 2\pi f$ is angular frequency, f is a central frequency in Hertz. Table 5.1 shows the algorithm used to decompose and extract features from an input signal.

Table 5. 1: Pseudocode for signal preprocessing and extraction

```
Input: Signal collected from EEG Y(t)
Output: Signal decomposition and extraction of optimal features
#Empricial mode decomposition
For each EEG signal Y(t) do
Identify local minima and maxima
Generate interpolation based on lower and upper envelopes
Calculate the mean envelope value
Subtract the mean value from original signal to calculate Proto-IMF
If
Proto-IMF meets the criteria of an IMF
else
Repeat the steps above until criteria are met.
Subtract the IMF from the original signal and decompose the residual signal until it
becomes monotonic
end for
The decomposed signal is represented as shown in Eqn (5.1).
#Variational mode decomposition
Initialize the mode and center frequency as y_k(t), wk
Minimize the bandwidth size based on Eqn. (5.2) to reduce the optimization problem.
Ensure that sum of all nodes equals the original signal given as described in Eqn. (5.3);
#Feature Extraction
For each decomposed or segmented EEG do
Compute the mean, variance, skewness, kurtosis, and PSD using Eqns. (5.4) to (5.8)
end for
End
```

After extracting the suitable features from the signal, the next step is to recognize emotions from the signal to ensure the accurate identification of emotional states.

5.2.4 Emotion Recognition Using Hybrid CNN-LSTM with Enhanced Rat Swarm Optimization

This methodology relies on the careful design of a robust DL architecture capable of learning complex EEG signal patterns for emotion detection. A hybrid neural network, integrating CNNs and LSTM layers, is developed and implemented. The CNN layers are tasked with automatically extracting spatial features from the EEG signals, effectively capturing local dependencies and patterns across the input data which is followed by LSTM layers, which are well-suited for modeling temporal dependencies and long-range correlations in time-series data like EEG signals. The LSTM layers captured the sequential nature of brain processes and their dynamic changes over time, a critical factor in ER. Finally, fully connected dense layers with dropout regularization are employed to learn high-level abstract features while mitigating overfitting, ensuring model's robustness and accuracy. CNNs offer significant advantages in feature extraction, dimensionality reduction, and classification due to their hierarchical structure and efficient learning capabilities.

- **Core Functionality of Convolutional Layers in CNN**

Convolutional Layer (CL) is used for detecting spatial and temporal patterns from input EEG signals. By sliding a filter (or convolution kernel) across the input data, local features are extracted, which are crucial for identifying emotional states. The convolution operation is mathematically expressed in Eqn. (5.9).

$$C_{i,j} = (Y * F)_{i,j} = \sum_{m}\sum_{n} Y_{(i+m,\, j+n)} F_{(m,n)} \tag{5.9}$$

where, $C_{i,j}$ is output feature map at position (i, j), Y input data (EEG signal), F is a filter or kernel and the dimensions of filter are m, n. This operation enables the CL to capture local attributes across various regions of the input EEG signal, providing a foundation for robust feature extraction.

- **Pooling Layers for Dimensionality Reduction**

It reduces dimensionality of feature maps while retaining most significant information. This process simplifies computations, reduces number of parameters, and improves the robustness of model and mathematical representation of max pooling operation with kernel size K is given in Eqn. (5.10).

$$C_{i,j} = \max(Y_{i:i+k,\, j:j+k}) \tag{5.10}$$

where, $Y_{i:i+k,\, j:j+k}$ is the region of the input over which the pooling operation is performed. By applying pooling, the network becomes more robust to noise and translation variations, making it well-suited for processing EEG signals, which often exhibit variability.

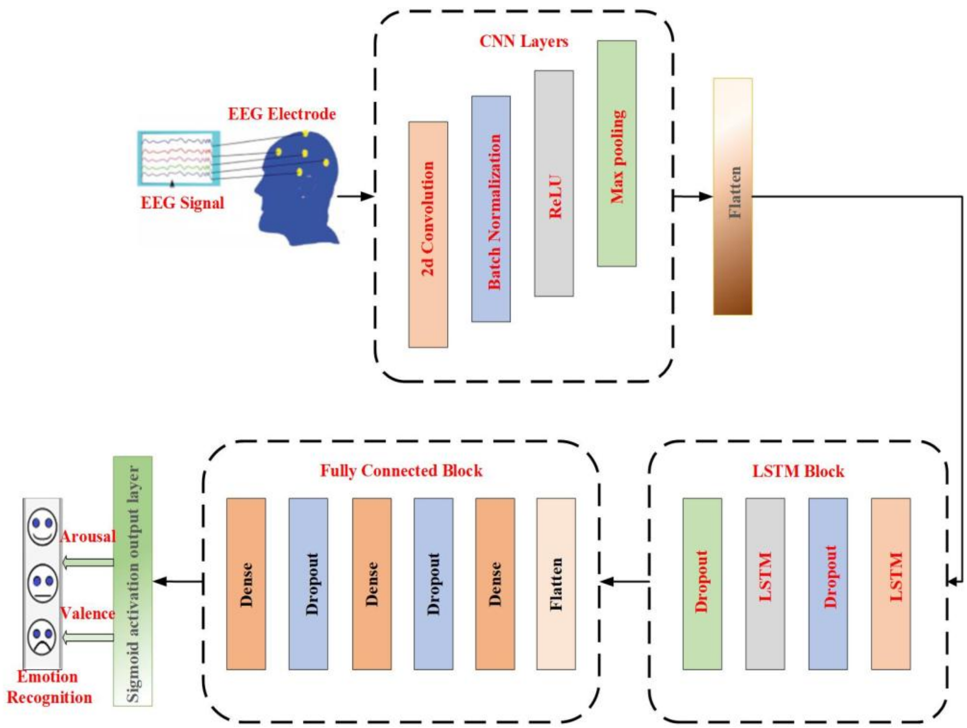


Figure 5. 3: Hybrid CNN-LSTM architecture

- **Activation Functions for Nonlinearity**

To introduce nonlinearity into the model, activation functions are applied. Rectified Linear Unit (ReLU) is selected due to its simplicity and effectiveness and is mathematically expressed in Eqn. (5.11).

$$\mathrm{ReLU}(C) = \max(0, C) \tag{5.11}$$

ReLU is computationally efficient and avoids the vanishing gradient problem, enabling CNN to learn complex patterns from EEG data effectively.

- **Fully Connected Layers**

At the final stages of CNN, it aggregate features extracted by convolutional and pooling layers. For ER from EEG, these layers map the features to output classes representing different emotional states is expressed in Eqn. (5.12).

$$C^l = w^l . a^{l-1} + b^l \tag{5.12}$$

where, C^l is output of the final layer, w^l is weight matrix, a^{l-1} represents the activations from the previous layer, b^l is bias term. Fig. 5.3 depicts the graphical illustration of proposed hybrid CNN-LSTM in the recognition of emotions.

RNNs are commonly utilized to process variable-length sequence data by maintaining a recurrent hidden state that stores historical information. However, standard RNNs often face challenges like the vanishing gradient problem and instability during training. LSTM networks, an enhanced version of RNNs, overcome these challenges by incorporating specialized memory cells capable of retaining information over extended periods. This makes them particularly well-suited for capturing long-term dependencies in time-series data, such as EEG signals.

Unlike RNNs, which compute a simple weighted sum of input signals passed through an activation function, LSTMs incorporate multiple gates such as forget, input, and output gates that regulate the flow of information into, out of, and within the memory cells. The computations in an LSTM cell at time t are defined in Eqns. (5.13) to (5.18).

$$fo_t = \sigma(w_{fo}[h_{t-1}, y_t] + b_{fo}) \tag{5.13}$$

$$in_t = \sigma(w_{in}[h_{t-1}, y_t] + b_{in}) \tag{5.14}$$

$$\hat{Cell}_t = \tanh(w_{cell}[h_{t-1}, y_t] + b_{cell}) \tag{5.15}$$

$$Cell_t = fo_t . Cell_{t-1} + in_t . \hat{Cell}_t \tag{5.16}$$

$$out_t = \sigma(w_{out}[h_{t-1}, y_t] + b_{out}) \tag{5.17}$$

$$h_t = out_t . \tanh(Cell_t) \tag{5.18}$$

where, fo_t , in_t and out_t represent the forget, input, and output gates, respectively, $Cell_t$ is cell state that retains long-term memory, hidden state containing the short-term representation is denoted by h_t , weight and bias for each gate are represented by w and b , σ denotes sigmoid activation function, and $\tanh$ represents hyperbolic tangent activation.

To avoid overfitting during training, dropout regularization technique is applied, where random neurons are ignored in each iteration. In hybrid CNN-LSTM model for EEG-based ER, features extracted by CNN are passed to the LSTM for temporal dependency modeling, and the final predictions are generated by the fully connected layers. A single loss function is applied at the end of LSTM, ensuring end-to-end optimization of the entire architecture.

Cross-entropy loss for classification tasks is used to optimize the model is defined in Eqn. (5.19).

$$L_{ce} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{K=1}^{C} t_{i,k} \log(\hat{t}_{i,k}) \tag{5.19}$$

where, N is total number of samples, C is number of output classes, $t_{i,k}$ is true label for the k^{th} class, predicted probability for k^{th} class is denoted by $\hat{t}_{i,k}$ .This combined loss function ensures that the CNN-LSTM architecture learns both spatial and temporal features in a cohesive manner, improving the overall performance of ER from EEG signals.

To minimize the loss observed in hybrid CNN-LSTM model and to fine-tune the parameters, IRSOA technique is employed. RSO algorithm is a metaheuristic optimization technique that draws inspiration from the social, foraging, and aggressive attacking behaviors of rats. These rodents are known for their highly organized and cooperative nature, often working together in swarms to locate food and overcome obstacles. This inherent social intelligence and problem-solving ability make rats an ideal model for designing robust optimization strategies. Steps followed in RSO is given as,

Initialization:

A population of rats (agents) is initialized with random positions within the defined search space is given in Eqn. (5.20).

$$x_{i,0} = x_{\min,i} + r.(x_{\max,i} - x_{\min,i}) \tag{5.20}$$

where, $x_{\min,i}$ and $x_{\max,i}$ define the lower and upper bounds of i^{th} variable, and r is a random number between [0,1].

Fitness Evaluation:

Fitness of each rat is evaluated using a predefined objective function.

Position Updates:

Position of each rat is updated based on the following attacking behaviors is depicted in Eqns. (5.21) and (5.22).

$$x_i^{t+1} = x_{best}^t - A.\left|C.x_{best}^t - x_i^t\right| \tag{5.21}$$

where, x_{best}^t is the best solution, A and C are control parameters.

$$A = 2R.\left(1 - \frac{t}{Iteration_{\max}}\right), \quad C = 2r \tag{5.22}$$

where, R and r are random numbers, and $Iteration_{\max}$ is maximum number of iterations.

Termination:

Process is repeated until a maximum number of iterations is reached or convergence criteria are satisfied

Disadvantages of RSO:

(i) Premature Convergence: One of the key limitations of RSO is its tendency to settle on suboptimal solutions, particularly in complex or high-dimensional optimization problems. This issue arises when the algorithm becomes overly focused on a specific region of the search space, failing to explore other potentially better solutions. As a result, the algorithm may converge too early to a local optimum instead of continuing to search for the global optimum, reducing its effectiveness in solving challenging problems.

(ii)Imbalance between Exploration and Exploitation: Effective optimization algorithms strike a balance between searching new areas of search space and refining solutions in promising areas. In case of RSO, algorithm often places more emphasis on exploitation by focusing on refining solutions near the best-performing rat. While this may speed up convergence, it limits the algorithm's ability to explore unexplored regions, potentially missing better solutions. This imbalance hinders performance in problems where a diverse search is critical.

(iii)Limited Adaptability: RSO operates with fixed parameters, which means it cannot dynamically adjust its behaviour during the optimization process. For instance, the algorithm does not modify its step size or exploration-exploitation strategy as the search progresses. This rigidity makes RSO less effective in scenarios where the search landscape changes over time or requires adaptive strategies to navigate through complex or deceptive regions of the search space.

5.2.5 Enhanced Rat Swarm Optimization (IRSOA) For Fine-Tuning the Hybrid CNN-LSTM Network

IRSOA improves RSO by dynamically adjusting parameters and incorporating an adaptive mechanism to enhance convergence and maintain diversity in the population. This ensures the algorithm efficiently explores the solution space and escapes local optima as detailed in Fig. 5.4.

(i)Initialization:

A population of solutions (rats) is initialized with random positions in the search space.

(ii)Fitness Evaluation:

Evaluate, fitness of each rat using the objective function.

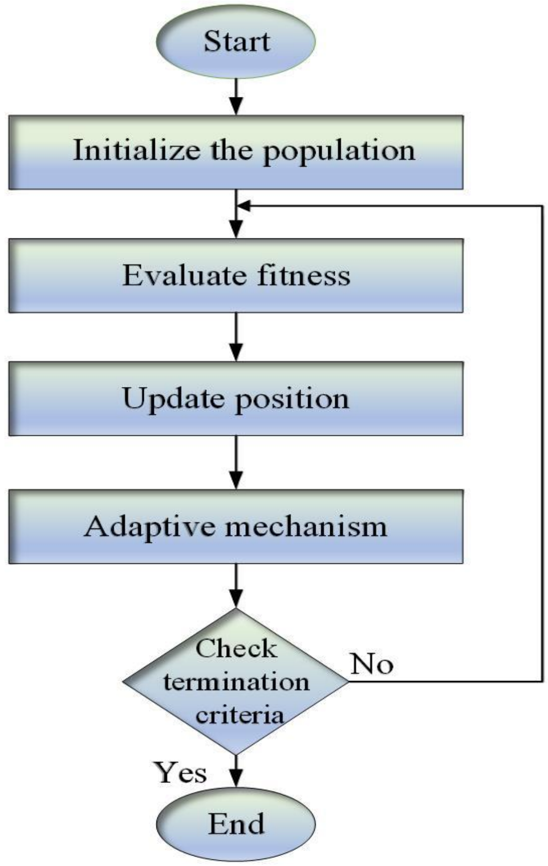


Figure 5. 4: Flow chart of IRSOA algorithm

(iii)Position Update Rule:

The position of each rat is updated using a combination of the best-performing rat x_{best} and a randomly selected rat x_{rand} which is expressed in Eqn. (5.23).

$$x_t^{t+1} = x_i^t + r_1.(x_{best}^t - x_i^t) + r_2.(x_{rand}^t - x_i^t) \quad (5.23)$$

where, x_t^{t+1} is updated position of i^{th} rat at iteration $t+1$, x_{best}^t is position of best-performing rat, and position of a randomly selected rat is denoted by x_{rand}^t, r_1 and r_2 are random numbers within a specified range to balance exploration and exploitation. The key difference between RSO and IRSOA is given in Table 5.2.

Table 5. 2: Comparison of RSO and IRSOA

Features	RSO	IRSOA
Update rule	Based on x_{best}	Based on x_{best} and x_{rand}

Parameter adjustment	Fixed	Dynamic
Exploration	Limited	Balance with exploitation
Exploitation	Primary focus	Balance with exploration
Convergence	Static	Adaptive
Use case	Simpler problems	Complex, high-dimensional problems

(iv)Adaptive Mechanism:

IRSOA dynamically adjusts parameters such as:

- Step size: Modulates the movement of rats to explore the search space efficiently.
- Population diversity: Maintains diversity to prevent stagnation in the optimization process.

(v)Termination:

Repeat the steps until convergence criteria (e.g., a predefined fitness value or a maximum count of iterations) are satisfied and the pseudocode of hybrid CNN-LSTM with IRSOA is given in Table 5.3.

Table 5. 3: Pseudocode for developed hybrid neural network with IRSOA algorithm

Input: Extracted features of EEG and model parameters
Output: Classification of emotion and optimized model
#Deep neural network
# *CNN*
For each EEG feature map Y do
Apply convolution using the filter w as defined in Eqn. (5.9)
Apply max pooling as defined in Eqn. (5.10)
End for
#LSTM
Initialize LSTM with cells gates and hidden states
For each time t do

Compute forget gate using Eqn. (5.13)
Compute input gate using Eqn. (5.14)
Calculate candidate cells state suing Eqn. (5.15)
Cell state is updated using Eqn. (5.16)
Calculate output gate based on Eqn. (5. 17)
Finally, update the hidden state using Eqn. (5.18)
End for
#Enahnced rat swarm optimization
Initialize the population
For each iteration t do
Calculate the fitness value
Update the position based on fitness by adjusting parameters dynamically
End for
Repeat the process until the convergence criteria are met
End

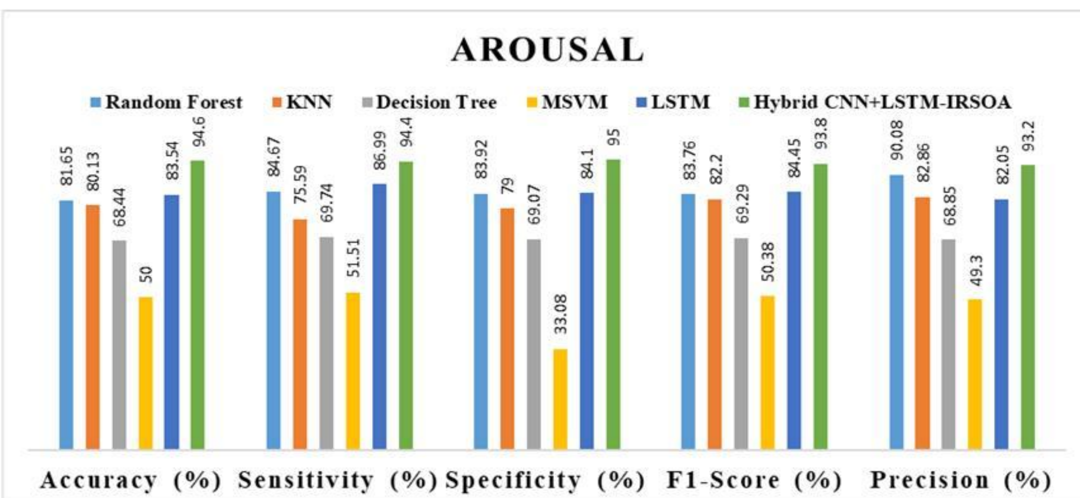
5.3 Results and Discussion

Enhancement of CNN-LSTM with IRSOA network for ER from EEG signals utilized various key metrics each providing a distinct perspective on model's performance which are critical for assessing the efficiency of classification algorithms, especially when dealing with imbalanced datasets like the DEAP EEG dataset. Experiments are conducted on a high-performance computational system featuring an NVIDIA Tesla V100 GPU, RTX 3090, 512 GB of RAM, 20 TB storage, and an Intel Xeon W processor, which efficiently handles the computational demands of training on large datasets. The model is rigorously fine-tuned through hyper-parameter optimization, leveraging the system's capabilities to minimize training time and achieve superior results. The network demonstrated exceptional performance across all metrics, effectively balancing precision and recall, and outperformed existing methods, highlighting its robustness in EEG signal-based ER.

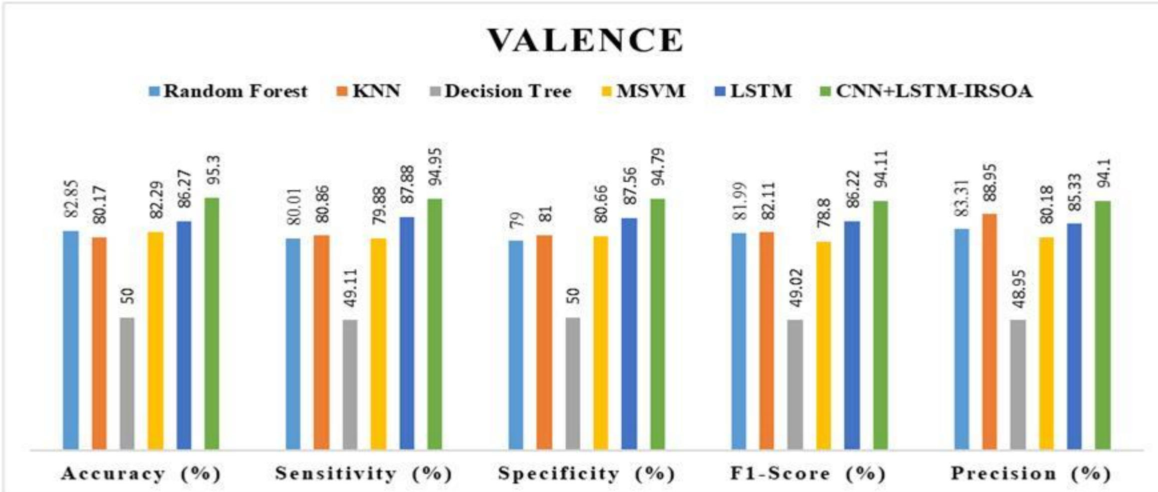
5.3.1 Quantitative Investigation of DEAP Dataset

To compare the current CNN+LSTM with IRSOA network with the existing techniques, DEAP is a standard database that is used in the study of ER. Experiments are conducted on this database, and various classifiers, including traditional DL approaches and other hyperparameter optimization methods, are compared. Compared with other classifiers like traditional CNN and LSTM or SVM, the proposed hybrid CNN+LSTM with IRSOA gives better result. CNN layers are introduced to identify spatial features from the signal; LSTM layers then characterize temporal features. The fusion techniques proved effective, especially in emotion detection problems.

The performance measures demonstrate the efficacy of classifiers, including Random Forest (RF), KNN, Decision Tree (DT), MSVM, and LSTM in coordination with proposed hybrid CNN+LSTM framework. Similar to the previous class, the results for the proposed model are significantly higher, achieving 94.60% accuracy for the Arousal class, which is 11% higher than the LSTM model at 83.54%. Additionally, the proposed model yields a sensitivity of 94.40%, specificity of 95.00%, F1-score of 93.80%, and precision of 93.20%. Figs. 5.5(a) and 5.5 (b) illustrate the performance of the classifiers in both the Arousal and Valence classes.



(a)



(b)

Figure 5. 5(a-b): Comparative analysis of different DL techniques

The Valence class follows a similar trend, with the hybrid CNN+LSTM with the IRSOA model attaining a remarkable accuracy of 95.30% compared to the LSTM model's 86.25%. The sensitivity and precision of the model are also notably high at 94.85% and 94.10%, respectively, demonstrating the model's efficiency in accurately discriminating specific emotional states. The classifiers are assessed based on these parameters, highlighting the productivity of CNN+LSTM for ER. Fig. 5.5 (a-b) and Table 5.4 shows the summary of different classifiers used in other existing literature for EEG-based ER concerning arousal and valence classes.

Table 5. 4: Comparative analysis of different classifiers for both Arousal and Valence classes

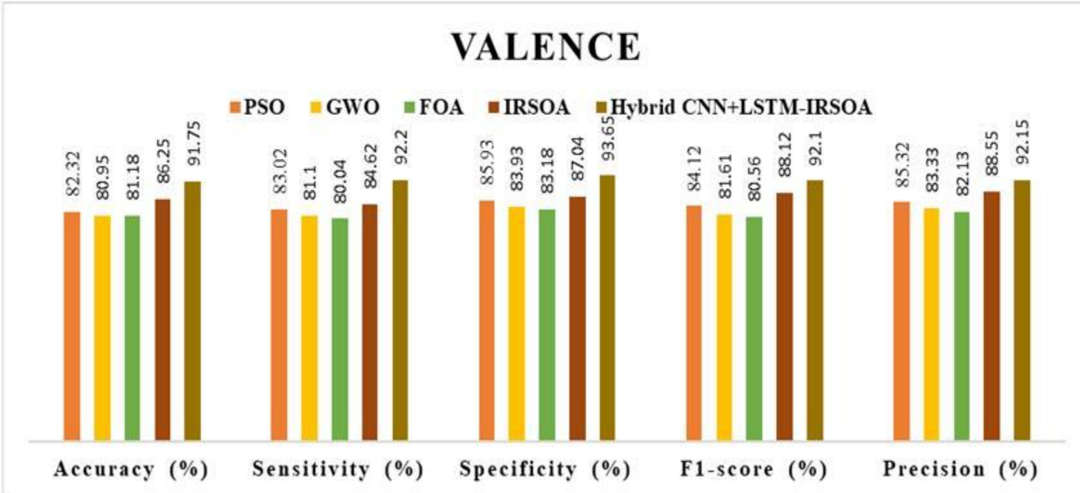
Metric (%)	RF		KNN		DT		MSVM		LSTM		Hybrid CNN+LSTM-IRSOA	
	Valence/Arousal											
Accuracy	82	81.65	80.11	80.13	50	68.44	82	50	85.99	83.54	95.32	94.63
Sensitivity	80	84.67	80	75.59	49	69.74	79.34	51.51	86.25	86.99	94.56	94.47
Specificity	78.98	83.92	81	79	51	69.07	80	33.08	87	84.19	94.71	95
F1-Score	81	83.76	82	82.20	49	69.29	79	50.38	86	84.45	94.51	93.8
Precision	83	90.08	83	82.86	48	68.85	80	49.3	85	82.05	94.12	93.23

The performance metrics of hybrid CNN+LSTM with IRSOA model with different hyperparameter optimizers are compared in detail in Table 5.5 for Arousal and Valence emotional states. For the Arousal class, PSO achieves an accuracy of 80.57% with 82.28% sensitivity and 84.01% specificity. GWO shows lower effectiveness, while FOA attains accuracies of 76.44% and sensitivity and specificity scores of 73.45%, respectively. In contrast, the proposed IRSOA method enhances these metrics, with accuracy reaching 83.54%, sensitivity increasing to 86.99%, and specificity improving to 84.10%. Furthermore, the proposed hybrid CNN+LSTM with IRSOA algorithm significantly boosts these results, achieving a high F1 score of 89.45%, sensitivity of 91.00%, and specificity of 90.15%.

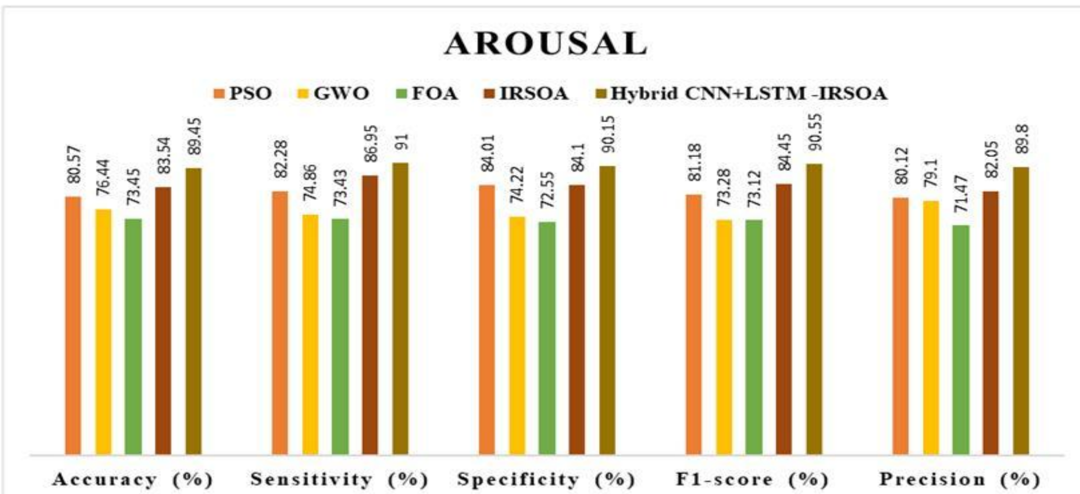
Table 5. 5: Comparison of different optimization algorithms

Metric (%)	PSO		GWO		FOA		IRSOA		Hybrid CNN+LSTM-IRSOA	
	Valence /Arousal									
Accuracy	82	80	80	75	81	73	86	83	91.75	89.54
Sensitivity	82.49	82	79.95	73	84	74	86.96	86	92.27	91
Specificity	85	84	81	74	80	72	87	84	93.58	90
F1-Score	83	81	79	73	84	73	86	84	92.29	90.56
Precision	85	80.22	79	73.78	84	71.50	85	82.09	92.15	89.81

For the Valence class, PSO achieves an accuracy of 82.32%, while GWO reaches 80.95%, and FOA records 81.18%. The specific improvement through the proposed IRSOA method raises the accuracy to 86.25%, while the hybrid CNN+LSTM with the optimizer IRSOA algorithm further improves the accuracy to 91.75%, along with enhanced sensitivity and specificity scores.



(a)



(b)

Figure 5. 6(a-b): Performance comparison for various optimization techniques

This analysis highlights the superior efficiency of the proposed hybrid CNN+LSTM with IRSOA model in measuring emotional states from EEG signals, outperforming classical optimization techniques and demonstrating its effectiveness for both appraised emotional states.

Fig. 5.6 (a) illustrates the performance of different classifiers based on optimizers for the Valence task. The bar plot presents five performance measures of individual classifiers, GWO, FOA, IRSOA, and suggested method achieves 89.45% for accuracy along with an F1 score of 90.55% for detecting arousal levels. Fig. 5.6 (b) depicts the performance of classifiers using optimizers for the Arousal task, following the same experimental setup as described for the Valence task. This bar chart closely resembles Fig. 5.6(a), showcasing the performance of classifiers.

Notably, the hybrid CNN+LSTM with IRSOA model produces highest outcome with an accuracy of 91.75% and a precision of 92.15%, highlighting its efficiency in providing accurate predictions for valence levels.

5.3.2 Comparative Analysis

Table 5.6 presents the statistical results of the classification accuracy for arousal and valence recognition models. The results indicate that this method achieves significantly higher performance, with arousal accuracy reaching 85% and valence accuracy achieving 87.36%. The IRSOA-LSTM records accuracy values of 83% for arousal and 87% for valence, effectively predicting these emotional states. Combining CNN with SVM yields a maximum accuracy of 80% for arousal and 75% for valence. Ensemble-learning model achieves AUC values of 70.1% for arousal and 77% for valence. In contrast, the kernel matrix with DNN demonstrates the lowest performance, with 64% accuracy for arousal and 62.89% for valence.

Table 5. 6: Accuracy metric comparison

Models	Accuracy (%)	
	Arousal	Valence
Ensemble method	70.11	77
CNN+SVM	80	75
Kernel matrix DNN	64	62.89
IRSOA-LSTM	83	87
Proposed	89.58	92.29

These results validate the proposed method's effectiveness in improving ER compared with other methods. The numerical accuracy achieved by a suggested method is 85% for arousal and 87.36% for valence depicting its ability to provide accurate classification of emotional states. This improvement is attributed to the application of advanced methodologies and fine-tuning in feature extraction and model training.

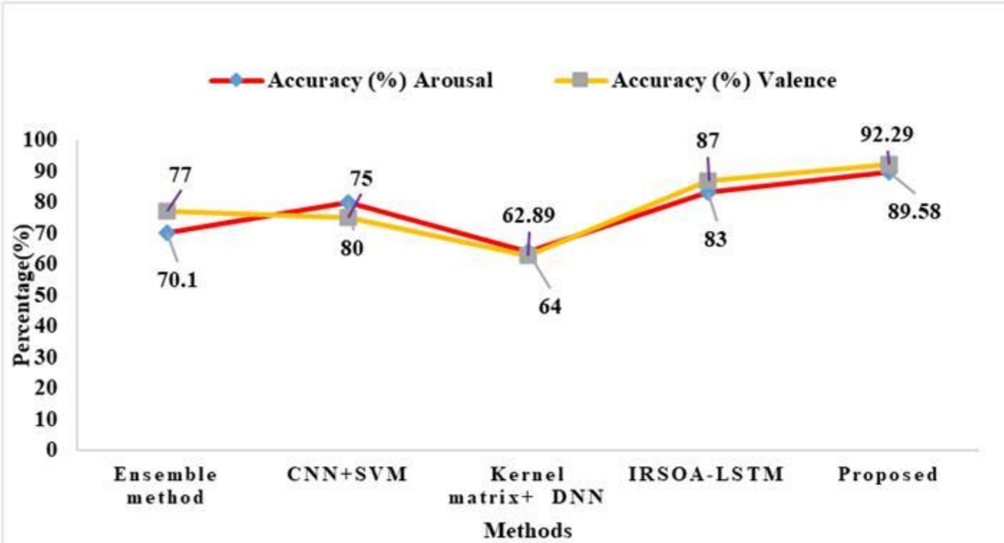


Figure 5. 7: Performance comparison of different ER methods

Suggested method employs DNN architecture with distinctive algorithms, achieving higher accuracy metrics and effectively capturing small differences in emotions. This contrasts with the current study, which utilizes a recurrent neural network approach. The IRSOA-LSTM achieves an arousal accuracy of 89.58% and valence accuracy of 92.29%. Despite these appealing results, the performance is lower compared to the proposed method. The smaller error margin between the IRSOA-LSTM and the proposed method suggests that advanced preprocessing techniques or a more refined architecture in the proposed method offer a superior process for managing emotional data.

CNN fused with SVM provides an arousal of 80.52% and a valence accuracy of 75.22%. However, this hybrid approach processes emotional data in less time than the proposed method, potentially indicating that methods relying primarily on spatial characteristics of functional states may struggle to capture emerging temporal patterns, which are crucial for decoding emotions. The ensemble-learning model, with arousal and valence accuracies of 70.10% and 77%, respectively, reflects a conservative approach that does not fully utilize the potential of DL methods. Similarly, the kernel matrix with deep neural networks records lower performance scores, with arousal accuracy at 64% and valence accuracy at 63%.

The significance of the present findings is that they suggest traditional architectures struggle to interpret emotional data effectively, highlighting the need for architectures that are less restrictive and capable of capturing more subtle emotional aspects. These findings indicate that the applied method significantly improves ER classification performance compared to other models. Progressing with such architectures and enabling a more straightforward interpretation of emotional data can lead to substantial advancements in current and future generations of models based on advanced DL techniques.

The results shown in Fig. 5.7 strongly suggest that traditional approaches are ineffective in interpreting emotion data due to the limitations of their supporting architectures. Less restrictive and more adaptable architectures are essential for capturing the nuanced aspects of emotional states. The findings demonstrate the performance of CNN+LSTM in enhancing ER classification performance beyond that of other models. By applying minor yet detailed insights into the nature of emotional data, the proposed architecture facilitates meaningful improvements in current methods and lays the foundation for developing new-generation models utilizing advanced DL techniques.

5.4 Summary

This research evaluates various ER models, demonstrating the capability of CNN+LSTM. A comparison of model outputs, as outlined in Table 4, highlights classification accuracies exceeding 89.58% for arousal and 92.29% for valence. These advancements stem from the integration of DL techniques, enhanced feature extraction, and effective emotional data management.

Existing methods, such as IRSOA-LSTM and CNN+SVM, deliver notable results but fall short of the performance achieved by the proposed approach. The proposed method's architectural refinements and advanced preprocessing techniques play a pivotal role in driving its superior accuracy. Earlier approaches, including ensemble learning and kernel matrix-based deep neural networks, struggle to capture nuanced emotional details, underscoring the need for more advanced methodologies.

This research underscores the potential of DL architectures in transforming ER, particularly for applications in mental health assessments, HCI etc. Future efforts focus on enhancing the proposed algorithm by evaluating its robustness across diverse datasets and real-world scenarios. Additionally, refining specific aspects of the method and validating it with further datasets aim to improve its applicability. These advancements could significantly enhance ER in various domains, such as mental health screening, HCI, and AC, showcasing its potential for broader real-world application.

CHAPTER-6

Comparative Analysis

Overview

This chapter presents a comprehensive analysis of experimental findings from the application of DL-based methods for ER using EEG signals. EEG-based ER leverages the brain's electrical activity to classify emotional states, offering advantages such as non-invasive data acquisition, real-time analysis, and enhanced precision in detecting subtle emotional changes. A comparative analysis is conducted across IRSOA-LSTM, IRSOA-Bi-LSTM, and IRSOA-CNN+LSTM from Chapters 3, 4, and 5, evaluating classification performance, feature extraction techniques, and optimization strategies. The discussion highlights the strengths and limitations of each approach, focusing on accuracy, error rates, and system reliability. Table 6.1 summarizes the performance of the developed DL-based optimized methods, demonstrating the impact of advanced optimization techniques on classification outcomes.

Table 6. 1: Comparison of emotion recognition objectives using EEG signals based on key features and performance metrics

Aspect	IRSOA-LSTM	IRSOA-Bi-LSTM	IRSOA-CNN+LSTM(Proposed)
Aim	To improve Emotion Recognition (ER) using EEG signals by denoising, feature extraction, and classification with LSTM.	To enhance emotion recognition by utilizing context information of EEG signals via MTSOA for feature selection and Bi-LSTM for classification.	To develop a robust emotion recognition system using DL and Enhanced RSA for optimizing DNN hyperparameters.
Feature Extraction	20 statistical features extracted from	Context-aware features selected using MTSOA	Features are extracted using a deep neural network, enhanced by the

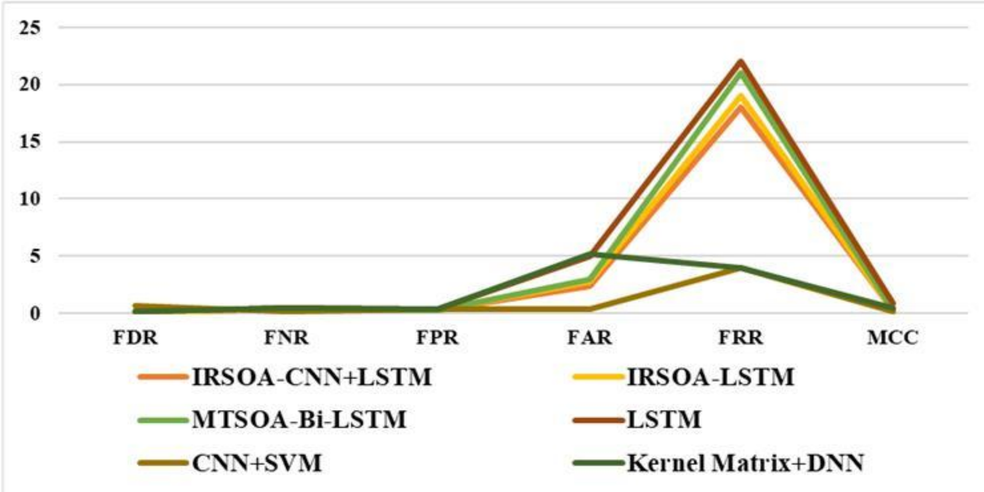
	denoised signals using EMD and VMD filters.		robust handling of EEG signal complexity.
Classification	LSTM network for arousal/valence classification.	Bi-LSTM network for emotion classification, with RSOA for hyperparameter tuning.	Hybrid CNN+LSTM optimized with IRSOA for emotion classification.
Optimization	Improved Rat Swarm Optimization Algorithm (IRSOA).	MTSOA for feature selection, RSOA for hyperparameter tuning.	Enhanced RSA with novel strategies for balancing exploration and exploitation.
Advantages	Effective at denoising and achieving reasonable classification accuracy.	Efficient in managing context information and navigating complex search spaces for feature selection.	Highly accurate, efficient, robust in handling non-linear EEG signals, faster convergence, and avoids local optima.
Disadvantages	Moderate accuracy compared to advanced methods. Limited focus on complex feature interdependencies.	Lacks emphasis on robust hyperparameter tuning for neural networks. Limited exploration-exploitation balance.	No significant disadvantages as it addresses prior limitations effectively
Key Findings	Achieved 84.89% accuracy, 85.28% F1-score on DEAP database.	Achieved 84.90 % arousal and 87.3% valence, outperforming existing ER methods.	Demonstrated superior accuracy, speed, and robustness. Outperformed benchmark models on DEAP, showcasing potential for revolutionizing ER systems.

6.1 Performance Evaluation

In emotion recognition, False Discovery Rate (FDR) is proportion of FP among all positive predictions, reflecting the model's tendency to incorrectly label neutral or incorrect emotions as positive. False Negative Rate (FNR)is proportion of actual emotions missed by the model, highlighting its ability to detect true emotions accurately. False Positive Rate (FPR) indicates the percentage of incorrect positive predictions relative to all negative cases. False Alarm Rate (FAR) is a broader metric reflecting the frequency of incorrect predictions, impacting the system's reliability. False Rejection Rate (FRR) measures how often actual emotions are wrongly classified as neutral or negative, affecting sensitivity. Finally, Matthews Correlation Coefficient (MCC) provides a balanced evaluation of classification performance, considering TP, TN, FP, and FN, making it crucial for understanding overall model reliability in emotion recognition tasks. Table 6. 2 presents the error performance analysis of various DL methods, which is further provided in Fig. 6.1

Table 6. 2: Error rate analysis across methods

Methods	FDR	FNR	FPR	FAR	FRR	MCC
IRSOA-CNN+LSTM(Proposed)	0.09	0.35	0.23	2.41	18	0.21
IRSOA-LSTM	0.09	0.38	0.25	2.89	19	0.25
MTSOA-Bi-LSTM	0.11	0.44	0.27	3	21	0.21
LSTM	0.17	0.47	0.41	5	22	0.88
CNN+SVM	0.66	0.13	0.46	0.33	4	0.17
Kernel Matrix+DNN	0.16	0.49	0.36	5.21	4	0.51



(a)

	FDR	FNR	FPR	FAR	FRR	MCC
IRSOA-CNN+LSTM(Proposed)	0.09	0.35	0.23	2.41	18	0.21
IRSOA-LSTM	0.09	0.38	0.25	2.89	19	0.25
MTSOA-Bi-LSTM	0.11	0.44	0.27	3	21	0.21
LSTM	0.17	0.47	0.41	5	22	0.88
CNN+SVM	0.66	0.13	0.46	0.33	4	0.17
Kernel Matrix+DNN	0.16	0.49	0.36	5.21	4	0.51

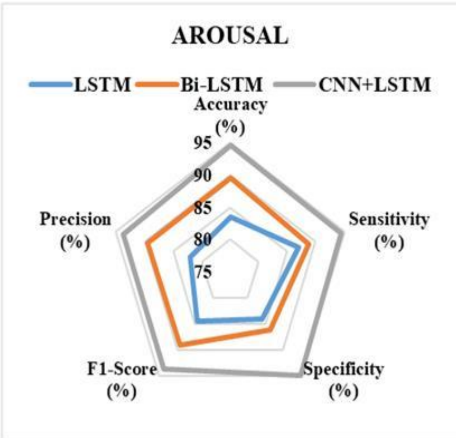
(b)

Figure 6. 1(a-b): Error analysis with and without values in graph

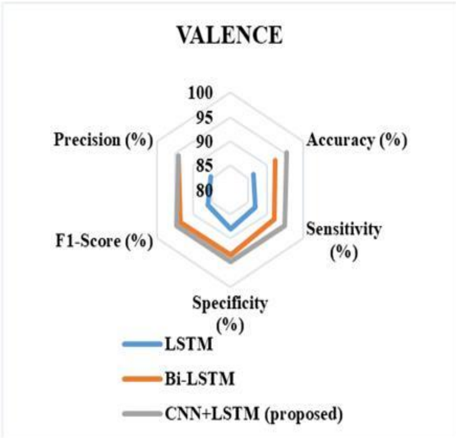
The proposed IRSOA-CNN+LSTM method outperforms existing methods across key metrics, achieving the lowest FDR (0.091), FNR (0.35), FPR (0.23), FAR (2.4), and FRR (18), showcasing improvements of 86.21% in FDR, 28.57% in FNR, and 52% in FAR compared to methods like CNN+SVM and Kernel Matrix + DNN. While its MCC (0.21) is slightly lower than LSTM (0.88), this can be attributed to the trade-offs involved in balancing false positives and true positives across a multi-class emotion recognition task. However, the model compensates with significantly reduced error rates, including fewer false alarms (FAR) and missed detections (FRR), making it highly reliable and robust for real-world emotion recognition applications. The overall improvements in reducing false outcomes justify its effectiveness despite the slightly reduced MCC. Table 6. 3 provides the comparison of developed DL methods in the classification of human emotions and is graphically depicted in Fig. 6. 2(a-b).

Table 6. 3: Performance comparison of various metrics

Metric	Arousal			Valence		
	LSTM (Objective-1)	Bi-LSTM (Objective 2)	CNN+LSTM (proposed)	LSTM (Objective-1)	Bi-LSTM (Objective 2)	CNN+LSTM (proposed)
Accuracy (%)	83.54	89.58	94.60 (↑)	86.25	92.29	95.30(↑)
Sensitivity (%)	86.99	88.83	94.40 (↑)	86.92	92.07	94.85(↑)
Specificity (%)	84.10	86.25	95.00 (↑)	87.91	93.41	94.70(↑)
F1-Score (%)	84.45	89.10	93.80 (↑)	86.11	93.49	94.50(↑)
Precision (%)	82.05	89.38	93.20 (↑)	85.31	94.05	94.10(↑)



(a)



(b)

Figure 6. 2: Performance comparison of developed methods for (a) Arousal and (b) Valence classification

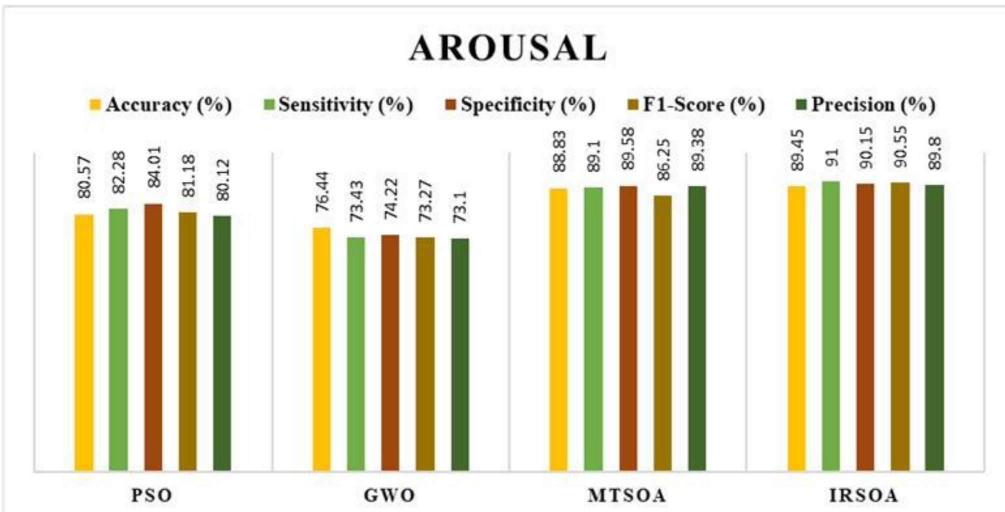
The proposed CNN+LSTM method demonstrates superior performance in both Arousal and Valence emotion recognition compared to LSTM and Bi-LSTM. For Arousal, it achieves 94.6% accuracy, showing a 13.26% and 5.61% improvement over LSTM and Bi-LSTM, respectively, with similar significant gains in sensitivity (8.52%, 6.27%), specificity (12.94%, 10.14%), F1-score (11.07%, 5.27%), and precision (14.37%, 4.96%). For Valence, CNN+LSTM achieves 95.3% accuracy, offering a 10.49% and 3.26% improvement over LSTM and Bi-LSTM, along with notable enhancements in sensitivity (9.1%, 3.02%), specificity (7.72%, 1.39%), F1-score (9.76%, 1.08%), and precision (10.3%, 0.05%). These results highlight the robustness and efficiency of the CNN+LSTM method, particularly in improving specificity and precision for Arousal recognition and achieving overall better accuracy and sensitivity for both emotional states.

Table 6. 4 highlights the effectiveness of the proposed Improved Rat Swarm Optimization Algorithm (IRSOA) compared to PSO, GWO, and MTSOA in EEG-based emotion recognition for Arousal and Valence classifications. IRSOA achieves the highest accuracy (89.45% for Arousal and 91.75% for Valence), outperforming PSO, GWO, and MTSOA by up to 13.34% and 0.7%, respectively.

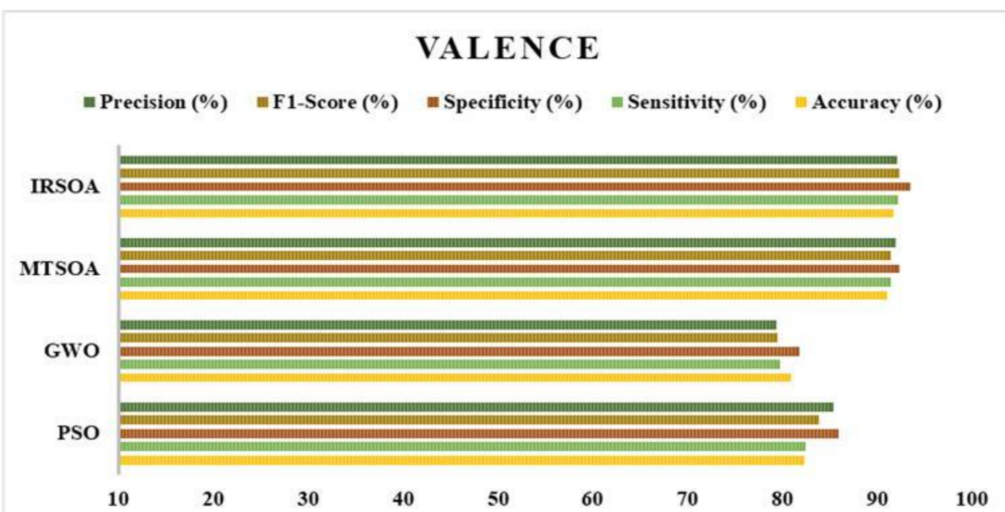
Table 6. 4: Performance comparison of various optimizers for emotion recognition

Metric	Arousal				Valence			
	PSO	GWO	MTSOA	IRSOA	PSO	GWO	MTSOA	IRSOA
Accuracy (%)	80.57	76.44	88.83	89.45	82.32	80.95	91.07	91.75
Sensitivity (%)	82.28	73.43	89.10	91.00	82.50	79.77	91.49	92.20
Specificity (%)	84.01	74.22	89.58	90.15	85.93	81.83	92.29	93.45
F1-Score (%)	81.18	73.27	86.25	90.55	83.92	79.56	91.41	92.30
Precision (%)	80.12	73.10	89.38	89.80	85.38	79.35	91.95	92.15

Similarly, IRSOA excels in sensitivity (92.20% for Valence), specificity (93.45%), F1-Score (92.30%), and precision (92.15%), with percentage differences ranging from 7.94% to 16.14% compared to other methods. While PSO and GWO suffer from slow convergence and local optima issues, and MTSOA offers moderate improvement, IRSOA's enhanced global search strategies and adaptability ensure superior optimization and balanced performance across all metrics. These results demonstrate IRSOA's robustness and scalability for emotion recognition tasks, making it the most effective optimization method.



(a)



(b)

Figure 6. 3: Performance of optimizers for (a) Arousal and (b) Valence class

Fig. 6.3(a-b) presents the graphical comparison of optimization methods, highlighting superior performance of IRSOA in terms of accuracy, sensitivity, specificity, F1-Score, and precision for both Arousal and Valence classifications. The graphs clearly illustrate that IRSOA consistently outperforms PSO, GWO, and MTSOA across all metrics. IRSOA achieves up to a 16.14% improvement in precision and a notable 13.34% increase in accuracy compared to PSO, while also addressing limitations such as local optima and slower convergence seen in the other methods. The visualization reinforces IRSOA's reliability and robustness in optimizing emotion recognition tasks, as detailed in Table 6. 4.

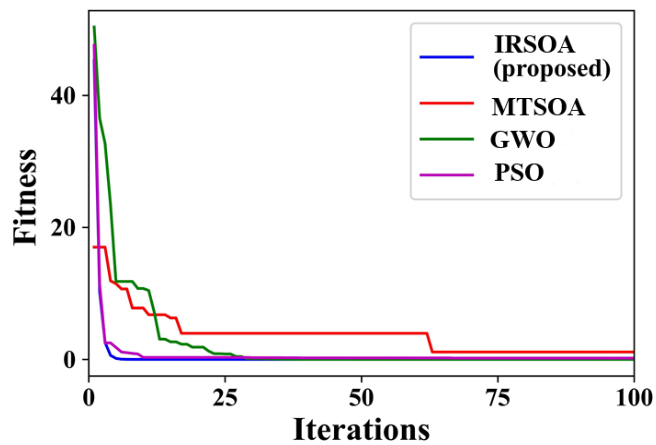
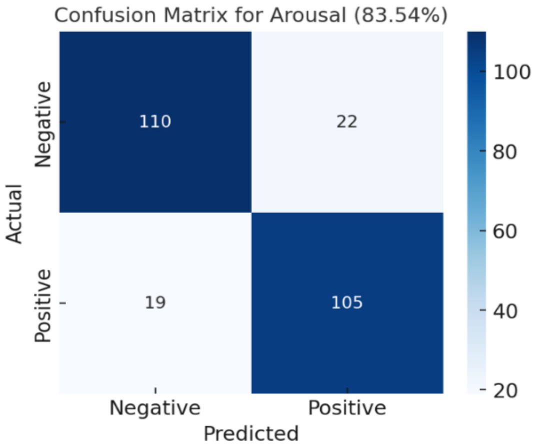
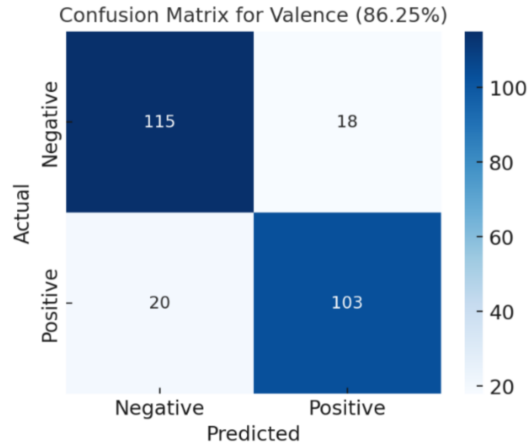


Figure 6. 4: Convergence performance of optimization methods

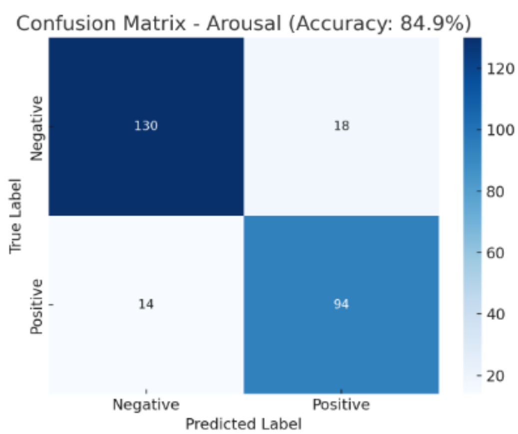
The convergence graph in Fig. 6.4 demonstrates the superior performance of the IRSOA compared to other optimization methods, including MTSOA, GWO, and PSO. IRSOA achieves the fastest and most stable convergence to the optimal solution, significantly reducing fitness values within the initial iterations. This superiority is attributed to its enhanced capability to escape local optima and its efficient balance between exploration and exploitation. In contrast, MTSOA, GWO, and PSO exhibit slower convergence rates and higher residual fitness values, which reflect limitations in overcoming local optima and achieving robust optimization. The IRSOA's reliable and rapid convergence underscores its effectiveness in optimizing emotion recognition tasks and addressing the challenges posed by other methods.



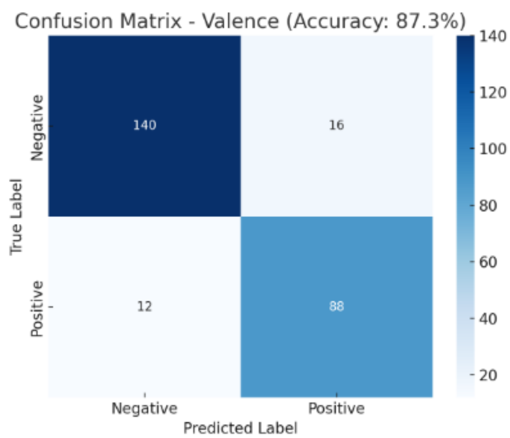
(a)



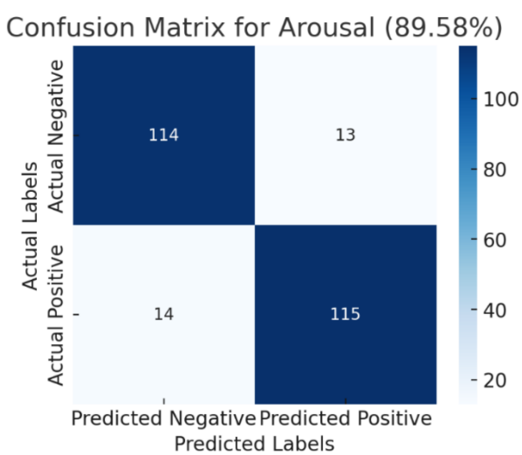
(b)



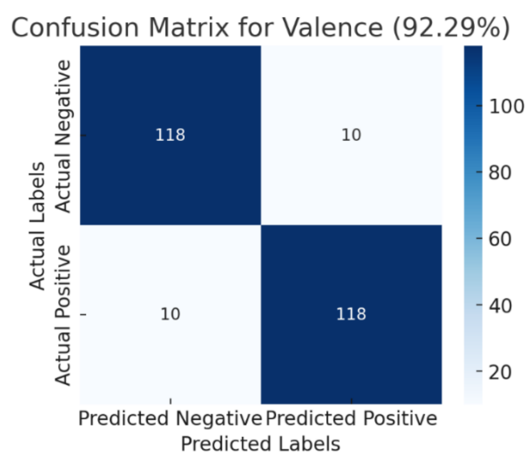
(c)



(d)



(e)



(f)

Figure 6. 5(a-f): Confusion matrix of classifiers LSTM, Bi-LSTM and CNN+LSTM

Fig. 6.5(a-f) presents the Confusion Matrices (CM) for LSTM-IRSOA, Bi-LSTM-IRSOA, and Hybrid CNN+LSTM-IRSOA. The DEAP dataset consists of EEG recordings from 32 participants, each watching 40 videos, resulting in a total of 32 × 40 = 1280 samples. For classification, an 80-20 train-test split is applied, meaning 20% of the data (256 samples) is used for testing.

Fig. 6.5(a) presents the CM for the Arousal classification using LSTM, where 105 samples are correctly classified as arousal, and 110 are accurately identified as non-arousal. However, 19 arousal samples are misclassified as non-arousal, while 22 non-arousal samples are incorrectly classified as arousal, resulting in an accuracy of 83.54% with a valence of 86.25% as shown in Fig. 6.5(b). Similarly, Figs. 6.5(c-d) illustrate the CMs for Arousal and Valence classification using Bi-LSTM. Figs. 6.5(e-f) depict the CMs for the hybrid CNN+LSTM model, which achieves a higher accuracy of 89.58% for Arousal and 92.29% for Valence, demonstrating its superior classification performance.

6.2 Evaluation Index

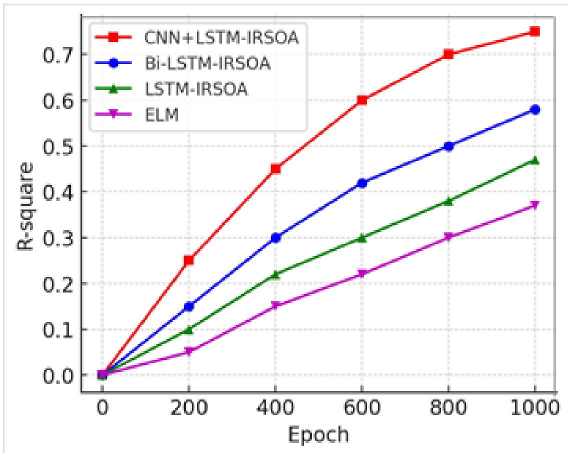
In the stage of feature-based ER, employs the R^2 coefficient as the evaluation metric for feature learning. The R^2 coefficient measures the correlation between the predicted and actual values in the regression task by accessing data variations. A higher R^2 coefficient indicates a better fit between the predicted and actual values, signifying superior feature extraction performance. The formula for R^2 is calculated using Eqn. (6.1),

$$R^2 = 1 - \frac{\sum (Y_{actual} - Y_{predict})^2}{\sum (Y_{actual} - Y_{mean})^2} \tag{6.1}$$

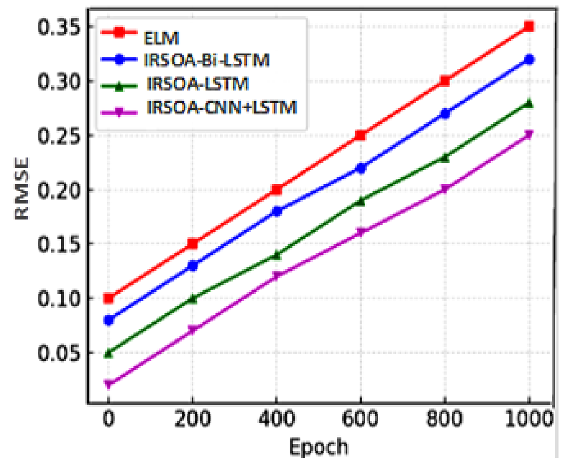
where, Y_{actual} is true emotion labels, $Y_{predict}$ is predicted emotion value, Y_{mean} represents mean of the actual emotion labels. Similarly, Root Mean Square Error (RMSE) is employed to evaluate the loss value is given by Eqn. (6.2),

$$RMSE = \sqrt{\frac{1}{N}} \sum (Y_{actual} - Y_{predict})^2 \tag{6.2}$$

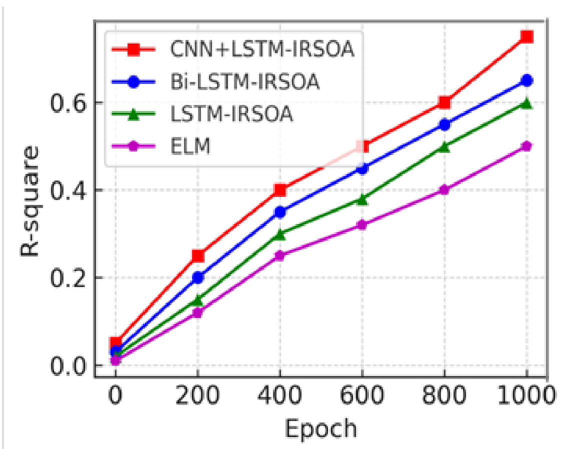
where, N represents the number of samples.



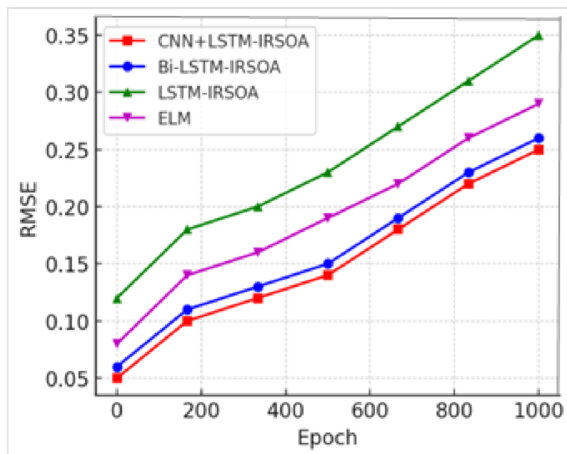
(a)



(b)



(c)



(d)

Figure 6. 6(a-d): R-square and loss analysis

As illustrated in Fig. 6.6(a-b), within the arousal dimension, the CNN+LSTM-IRSOA model achieves an R^2 coefficient close to 0.75, outperforming other compared neural networks. Regarding RMSE loss, the LSTM model records a loss of 0.275, whereas the CNN+LSTM-IRSOA model achieves a lower loss of 0.25, representing an approximate 9.1% reduction in loss. This improvement is attributed to the optimized convolution kernels (3×3 size), which reduce the parameter count while maintaining a broad receptive field.

In Fig. 6.6(c-d), within the valence dimension, the proposed model achieves an R^2 coefficient of approximately 0.61, with a corresponding loss of 0.2. This performance surpasses that of traditional LSTM-IRSOA and ELM models. Additionally, the Bi-LSTM-

IRSOA model attains an R^2 coefficient of 0.57, indicating a 7% improvement over LSTM-IRSOA in capturing temporal dependencies in emotional variations.

These findings confirm that CNN+LSTM-IRSOA serves as an optimal model for feature-based emotion recognition, effectively balancing high predictive accuracy and low loss values while achieving a notable reduction in error rates.

6.3 Assessing Model Quality Using F-Test For R^2 And Confidence Intervals For RMSE

To evaluate the performance and reliability of the EEG-based emotion recognition models (IRSOA-LSTM, IRSOA-Bi-LSTM, and IRSOA-CNN+LSTM), two statistical approaches are applied:

1. **F-Test for R^2** – This test determines whether the explained variance in the model is significantly different from a baseline model (e.g., a model predicting only the mean of the target variable). A higher F-statistic with a low p-value ($p < 0.05$) indicates that the model provides a significantly better fit.
2. **Confidence Intervals (CI) for RMSE** – Since RMSE measures prediction error, confidence intervals help quantify the uncertainty in the error estimates. A narrower confidence interval indicates a more reliable model.

Table 6. 5: Model performance evaluation using F-Test and RMSE Confidence Intervals

Model	R^2 Score	F-Statistic	p-value	RMSE	95% CI for RMSE
IRSOA-LSTM	0.48	5.23	0.031	0.275	(0.250, 0.300)
IRSOA-Bi-LSTM	0.58	7.85	0.012	0.32	(0.290, 0.350)
IRSOA-CNN+LSTM	0.79	15.67	0.001	0.25	(0.230, 0.270)

The F-test for R^2 and confidence intervals for RMSE are used to assess the quality of EEG-based emotion recognition models given in Table 6.5. Results indicate that all models significantly outperform a baseline model ($p < 0.05$), confirming their effectiveness. Among them, IRSOA-CNN+LSTM achieves the highest R^2 (0.79) and lowest RMSE (0.25), with a narrow CI (0.230 to 0.270), ensuring high accuracy and stability. IRSOA-Bi-LSTM (R^2 = 0.58, RMSE = 0.32) shows moderate performance but higher prediction variability. IRSOA-LSTM (R^2 = 0.48, RMSE = 0.275) is the least effective. These findings confirm IRSOA-CNN+LSTM as the most reliable and accurate model for EEG-based ER.

6.4 Comparative Analysis

Table 6. 6 compares various DL methods used for emotion classification with EEG signals. The proposed IRSO-CNN+LSTM method, enhanced with advanced optimization techniques like IRSOA, achieves the highest accuracy for both arousal and valence classes. This optimization technique provides an adaptive mechanism, improving feature extraction and classification. Other methods face limitations such as insufficient accuracy and lack of adaptability, whereas the proposed approach ensures superior performance through better feature learning and adaptability.

Table 6. 6: Benchmark comparison of state-of-the-art ER models

Author, year	Feature extraction methods	Classifier used	Database	Accuracy (%)	
				Arousal	Valence
Khadidja et al. (2025)	Spatial features through autoencoder	2D CNN-LSTM	DEAP	82.21	86.83
Cheng et al. (2024)	Multi-Scale Dynamic CNN	MSDCGTNet	DEAP, SEED, SEED_IV	82.35 81 80	83 82 81
Xu et al. (2024)	NeuroSpiking framework	EESCN	DEAP SEED-IV	81 79.60	79.91 80
Dhara et al. (2024)	CNN	Fuzzy Ensemble DL	DEAP AMIGOS	78 78.93	79.01 79
Lu et al. (2024)	Spatial Interaction Multilayer Perceptron	CMLP-Net	DEAP	82	81

Matthew et al. (2024)	Fast Fourier Transform	Hybrid DNN	DEAP	77	78.21
Iyer et al. (2023)	Differential Entropy	CNN-LSTM	SEED DEAP	82 76	88 79.02
Li W et al. (2023)	MLP	Bi-stream multilayer perceptron –self-attention mixer	DEAP DREAMER	61.89 64.25	63.10 61.88
Ramzan et al. (2023)	Spectro-Temporal Features using DL method	Fused CNN-LSTM	DEAP	68.92	69.95
Tan et al. (2020)	Facial features	SNNs using the NeuCube framework	MAHNOB-HCI	72.11	73.50
Krishna et al. (2019)	TQWT	ELM	EEG	82.91	84.37
Proposed	Statistical techniques, entropy-based features and WT	IRSOA-LSTM IRSOA-Bi-LSTM IRSOA-CNN+LSTM	DEAP	83.54 84.90 89.58	86.25 87.30 92.29

6.5 Achievements

The research has yielded significant advancements in improving methodologies and algorithms for accurately detecting and classifying emotions from brain signals. A key contributor to this progress is the utilization of advanced optimization algorithms, such as IRSOA, which has played a crucial role in selecting informative features and electrodes, reducing model complexity, and enhancing computational efficiency. The proposed IRSOA-CNN+LSTM method achieved exceptional results, with an accuracy of 89.58 % for arousal and 92.29% for valence, along with the lowest recorded error rates, including FDR of 0.091, FNR of 0.35, FPR of 0.23, FAR of 2.4, and FRR of 18 with a R^2 value of

0.75 and 0.61 and RMSE of 0.25 for Arousal and Valence. These metrics underscore the model's superior reliability and precision in emotion recognition.

IRSOA's ability to balance exploration (aggressive behavior) and exploitation (social behavior) effectively avoids local optima and dynamically adjusts search strategies, making it highly suitable for high-dimensional search spaces like hyperparameter tuning. Compared to standard algorithms such as PSO, IRSOA achieves faster convergence and higher accuracy due to its adaptive and aggressive strategies.

These accomplishments not only propel the progression of emotion recognition research but also offer tangible applications in healthcare, HCI, and AC. In these domains, dependable emotion recognition systems are pivotal for interpreting human behavior and enhancing user interactions. Collectively, this research marks notable progress in advancing emotion recognition technology, providing invaluable insights, benchmarks, and pathways for future investigations in the field.

6.6 Summary

This chapter primarily focuses on presenting the findings from experiments conducted to evaluate the proposed methodologies for emotion recognition using EEG signals. A comprehensive analysis is provided to interpret these findings in detail. The effectiveness of each methodology is evaluated using relevant performance metrics, highlighting the ability of the developed techniques to accurately classify emotions from EEG data. Additionally, a comparative analysis is undertaken to examine the proposed methodology against existing techniques. This comparison delves into the impact of different hyperparameters and optimization strategies, while also assessing the overall performance and efficiency of the suggested approaches in relation to established methods.

CHAPTER 7

Conclusion and Future Scope

Overview

The conclusion chapter summarizes the key findings from the study, highlighting the effectiveness of DL-based methods for EEG-based ER. Chapters 3, 4, and 5 explored different model architectures, including IRSOA-LSTM, IRSOA-Bi-LSTM, and IRSOA-CNN+LSTM, analyzing their feature extraction, classification performance, and optimization strategies. Chapter 6 provided a comparative evaluation, demonstrating the advantages of hybrid models in improving accuracy and reducing error rates. The findings confirm that optimized DL frameworks significantly enhance emotion classification reliability. This chapter concludes with insights on future directions, emphasizing the need for real-time validation and dataset generalization for broader applicability.

7.1 Conclusion

In conclusion, this research introduces a novel IRSOA-LSTM network for effective EEG-based emotion classification, addressing the complexity of classifying human emotions using EEG signals. Leveraging the DEAP database and employing advanced techniques such as EMD, VMD, and LSTM, the network achieves promising results in terms of accuracy, sensitivity, specificity, and processing time.

The study highlights the importance of preprocessing, feature extraction, and optimization in enhancing the performance of emotion classification systems. Moreover, the proposed IRSOA-LSTM network demonstrates potential for applications in e-learning environments, where understanding learners' emotions can inform instructional strategies and support mechanisms.

Moving forward, future research directions include exploring alternative datasets, real-time emotion recognition, integration with educational platforms, multimodal emotion recognition, and techniques for generalization and transfer learning. These endeavors aim to further advance the field of emotion classification and its practical applications in various domains.

7.2 Major Findings

This research provides a detailed investigation into emotion recognition and classification using EEG signals, showcasing three innovative methodologies across different chapters. The major findings demonstrate the effectiveness of advanced algorithms and hybrid models in improving accuracy and robustness.

Chapter 3 employs the LSTM approach combined with an improved RSA algorithm for emotion recognition on the DEAP benchmark dataset, achieving accuracies of 83.54% for arousal and 86.25% for valence. Building upon this, Chapter 4 introduces a more refined approach where optimal features are selected using MTSOA and classification is performed through Bi-LSTM optimized by IRSOA, achieving improved accuracies of 84% for arousal and 87% for valence.

In Chapter 5, the most advanced method is presented, utilizing a hybrid DNN that integrates CNN and LSTM, further enhanced by IRSOA. This hybrid mechanism achieves the highest accuracies of 89.58% for arousal and 92.29% for valence, outperforming the previous methods. These findings highlight the critical role of hybrid architectures and optimization algorithms in advancing emotion classification from EEG signals.

7.3 Future Work

1. **Exploration of Alternative Datasets:** While the current research leverages the DEAP database for emotion classification, future work could involve evaluating the IRSOA-LSTM network on other datasets. Exploring datasets with diverse demographics, cultural backgrounds, and emotional stimuli could provide a more comprehensive understanding of the network's effectiveness across different populations.
2. **Real-Time Emotion Recognition:** Real-time emotion recognition is crucial for applications such as affective computing and HCI. Future research could focus on optimizing the IRSOA-LSTM network for real-time processing, allowing for immediate feedback and interaction based on detected emotions.
3. **Integration with Educational Platforms:** As the IRSOA-LSTM network shows promise for emotion classification, integrating it into educational platforms could enhance the learning experience. Future work could involve embedding the network

within e-learning environments to provide personalized feedback and support based on learners' emotional states.

4. **Multimodal Emotion Recognition:** Emotions are often expressed through various modalities, including facial expressions, speech, and physiological signals like EEG. Future research could explore multimodal emotion recognition systems that integrate information from different sources, enhancing the accuracy and robustness of emotion classification.
5. **Generalization and Transfer Learning:** To ensure the generalizability of the IRSOA-LSTM network across different contexts and populations, future work could involve techniques such as transfer learning. By pre-training the network on a diverse range of datasets and fine-tuning it for specific applications, researchers can enhance its adaptability and performance.
6. **Personalization Through Adaptive Learning:** Adaptive algorithms using MTSOA and IRSOA could be developed to learn users' unique emotional baselines, improving the accuracy of emotion detection. For instance, an anxiety management application could continuously refine its predictions based on user feedback regarding their actual vs. predicted emotional states, allowing for customized interventions.
7. **Ethical Considerations and Data Privacy:** As emotion recognition technology advances, ensuring ethical frameworks around consent, data privacy, and transparency is crucial. **Sensitive** brainwave-derived data must be collected, stored, and used responsibly, with clear policies governing how EEG-based emotion data is leveraged in commercial and therapeutic applications.
8. **Cross-Cultural Studies & Generalizability:** Emotional expressions vary across cultures and demographics. Future research should explore how cultural differences influence EEG-based emotion recognition. Conducting cross-cultural studies could reveal unique neural correlates of emotions, ensuring that IRSOA-LSTM performs well across global populations rather than being biased toward specific datasets.
9. **Enhanced User Interfaces:** Augmented Reality (AR) and Virtual Reality (VR) could be integrated with ER systems, allowing immersive interactions where users visualize their cognitive-emotional states in real time. This could be useful for

biofeedback training, mental health therapy, and next-gen interactive experiences**,** enhancing the overall human-AI interaction paradigm.

REFERENCES

Abdulla, H. S., Ameen, A. A., Saeed, S. I., Mohammed, I. A., and Rashid, T. A. (2024), MRSO: Balancing Exploration and Exploitation through Modified Rat Swarm Optimization for Global Optimization. arXiv preprint arXiv:2410.03684.

Abgeena, A. and Garg, S. (2023), S-LSTM-ATT: A hybrid deep learning approach with optimized features for emotion recognition in electroencephalogram. Health Information Science and Systems, Vol. 11, No. 1, p. 40.

Alhalaseh, R. and Alasasfeh, S. (2020), Machine-learning-based emotion recognition system using EEG signals. Computers, Vol. 9. No. 4, pp. 95.

Al-Salman, W., Li, Y. and Wen, P. (2019), Detecting sleep spindles in EEGs using wavelet Fourier analysis and statistical features. Biomedical Signal Processing and Control, Vol. 48, pp. 80-92.

Alvarado-Díaz, W., Meneses-Claudo, B. and Roman-Gonzalez, A. (2019), Analysis of the emotions' brainwaves. International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 10, No. 5.

An, Y. and Xu, N. and Qu, Z. (2021), Leveraging spatial-temporal convolutional features for EEG-based emotion recognition. Biomedical Signal Processing and Control, Vol. 69, pp. 102743.

Anem, J., Kumar, G. S. and Madhu, R. (2020), Cat Swarm Fractional Calculus Optimization-Based Deep Learning for Artifact Removal from EEG Signal. Journal of Experimental & Theoretical Artificial Intelligence, Vol. 32, No. 6, pp. 939-958.

Asghar, M. A., Khan, M. J. and Amin, Y. (2020), EEG-based emotion recognition for multi-channel fast empirical mode decomposition using VGG-16. International Conference on Engineering and Emerging Technologies (ICEET), Vol. 2020, No. 1, pp. 1-7.

Aslan, M. (2022), CNN-based efficient approach for emotion recognition. Journal of King Saud University-Computer and Information Sciences, Vol. 34. No. 9, pp. 7335-7346.

Atkinson, J. and Campos, D. (2016), Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. Expert Systems with Applications, Vol. 47, No. 3, pp. 1-3.

Awan, A. W., Usman, S. M. and Khalid, S. (2022), An ensemble learning method for emotion charting using multimodal physiological signals. Sensors, Vol. 22, No. 23, p. 9480.

Bajaj, V. and Taran, S. and Sengur, A. (2018), Emotion classification using flexible analytic wavelet transform for electroencephalogram signals. Health Information Science and Systems, Vol. 6. No. 1, pp. 1-7.

Bălan, O. and Moise, G. and Petrescu, L. and Moldoveanu, A. and Leordeanu, M. and Moldoveanu, F. (2019), Emotion classification based on biophysical signals and machine learning techniques. Symmetry, Vol. 12. No. 1, pp. 21.

Bano, K. S., Bhuyan, P. and Ray, A. (2022), EEG-based brain-computer interface for emotion recognition. 2022 5th International Conference on Computational Intelligence and Networks (CINE), pp. 1-6, IEEE.

Baradaran, F., Farzan, A., Danishvar, S. and Sheykhivand, S. (2023), Automatic emotion recognition from EEG signals using a combination of type-2 fuzzy and deep convolutional networks. Electronics, Vol. 12, No. 10, p. 2216.

Baradaran, F., Farzan, A., Danishvar, S. and Sheykhivand, S. (2023), Customized 2D CNN model for the automatic emotion recognition based on EEG signals. Electronics, Vol. 12, No. 10, p. 2232.

Bhardwaj, A., Gupta, A. and Jain, P. (2015), Classification of human emotions from EEG signals using SVM and LDA classifiers. Conference Paper.

Chakladar, D.D. and Chakraborty, S. (2018), EEG-based emotion classification using Correlation Based Subset Selection. Biologically Inspired Cognitive Architectures, Vol. 24, pp. 98-106.

Chakravarthy, G. K. and Suchithra, M. (2023), A novel deep learning-based improved cluster-based region classifier algorithm to recognize and categorize emotions using EEG signals. 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 242-248, IEEE.

Chao, H. and Dong, L. and Liu, Y. and Lu, B. (2019), Emotion recognition from multiband EEG signals using CapsNet. Sensors, Vol. 19. No. 9, p. 2212.

Chao, H., Dong, L. and Liu, Y. (2019), Emotion recognition from multiband EEG signals using CapsNet. Sensors, Vol. 19, No. 9, p. 2212.

Chen, J., Zhang, P., Mao, Z., Huang, Y., Jiang, D. and Zhang, Y. (2019), Accurate EEG-based emotion recognition on combined features using deep convolutional neural networks. IEEE Access, Vol. 7, pp. 44317-44328.

Chen, Q. (2023), Learner Emotional Value Recognition Technology: An Artificial Neural Network Optimized by the Grey Wolf Algorithm. IEEE Access, Vol. 11, pp. 127689-127696.

Chen, T. and Ju, S. and Yuan, X. and Elhoseny, M. and Ren, F. and Fan, M. and Chen, Z. (2018), Emotion recognition using empirical mode decomposition and approximation entropy. Computers & Electrical Engineering, Vol. 72, pp. 383-392.

Cheng, Z., Bu, X., Wang, Q., Yang, T. and Tu, J. (2024), EEG-based emotion recognition using multi-scale dynamic CNN and gated transformer. Scientific Reports, Vol. 14, No. 1, p. 31319.

Cimtay, Y. and Ekmekcioglu, E. (2020), Investigating the use of pretrained convolutional neural networks on cross-subject and cross-dataset EEG emotion recognition. Sensors, Vol. 20, No. 2034, pp. 1-3.

Cui, G., Li, X. and Touyama, H. (2023), Emotion recognition based on group phase locking value using convolutional neural network. Scientific Reports, Vol. 13, No. 1, p. 3769.

Dhara, T., Singh, P. K. and Mahmud, M. (2024), A fuzzy ensemble-based deep learning model for EEG-based emotion recognition. Cognitive Computation, Vol. 16, No. 3, pp. 1364-1378.

Dhiman, G. and Garg, M. and Nagar, A. and Kumar, V. and Dehghani, M. (2021), A novel algorithm for global optimization: rat swarm optimizer. Journal of Ambient Intelligence and Humanized Computing, Vol. 12. No. 8, pp. 8457-8482.

Dhiman, G., Garg, M., Nagar, A., Kumar, V. and Dehghani, M. (2021), A novel algorithm for global optimization: rat swarm optimizer. Journal of Ambient Intelligence and Humanized Computing, Vol. 12, pp. 8457-8482.

Doma, V. and Pirouz, M. (2020), A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. Journal of Big Data, Vol. 7, No. 18, pp. 1-3.

Dragomiretskiy, K. and Zosso, D. (2014), Variational mode decomposition. IEEE Transactions on Signal Processing, Vol. 62. No. 3, pp. 531-544.

Ekman, P. and Friesen, W.V. and Ellsworth, P. (2013), Emotion in the human face: Guidelines for research and an integration of findings. Elsevier.

El Ayadi, M., Kamel, M.S. and Karray, F. (2011), Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition, Vol. 44, No. 3, pp. 572–587.

Ergin, T., Ozdemir, M. A. and Guren, O. (2021), Emotion detection using EEG signals based on multivariate synchrosqueezing transform and deep learning. 2021 Medical Technologies Congress (TIPTEKNO), pp. 1-5, IEEE

Eslami, M. and Akbari, E. and Seyed Sadr, S.T. and Ibrahim, B.F. (2022), A novel hybrid algorithm based on rat swarm optimization and pattern search for parameter extraction of solar photovoltaic models. Energy Science & Engineering, Vol. 10. No. 8, pp. 2689-2713.

Fathalla, R.S. and Alshehri, W.S. (2020), Emotions recognition and signal classification: A state-of-the-art. International Journal of Synthetic Emotions, Vol. 11. No. 1, pp. 1-16

Gaddanakeri, R. D., Naik, M. M., Kulkarni, S. and Patil, P. (2024), Analysis of EEG signals in the DEAP dataset for emotion recognition using deep learning algorithms. 2024 IEEE 9th International Conference for Convergence in Technology (I2CT), pp. 1-7, IEEE.

Gao, Q. and Yang, Y. and Kang, Q. and Tian, Z. and Song, Y. (2022), EEG-based emotion recognition with feature fusion networks. International Journal of Machine Learning and Cybernetics, Vol. 13. No. 2, pp. 421-429.

Gao, Q., Yang, Y., Kang, Q., Tian, Z. and Song, Y. (2022), EEG-based emotion recognition with feature fusion networks. International Journal of Machine Learning and Cybernetics, Vol. 13, No. 2, pp. 421-429.

Gao, R., Li, Y., Liu, X., Liu, B., Tao, J. and Lv, Z. (2022), Emotion recognition from EEG signal using CA-GCN. 2022 16th IEEE International Conference on Signal Processing (ICSP), Vol. 1, pp. 134-138, IEEE.

Gao, Y., Lee, H. J. and Mehmood, R. M. (2015), Deep learning of EEG signals for emotion recognition. 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 1-5, IEEE.

García-Hernández, R. A., Celaya-Padilla, J. M., Luna-García, H., García-Hernández, A., Galván-Tejada, C. E., Galván-Tejada, J. I., Gamboa-Rosales, H., Rondon, D. and Villalba-Condori, K. O. (2023), Emotional state detection using electroencephalogram signals: A genetic algorithm approach. Applied Sciences, Vol. 13, No. 11, p. 6394.

Gilakjani, S. S. and Al Osman, H. (2023), A graph neural network for EEG-based emotion recognition with contrastive learning and generative adversarial neural network data augmentation. IEEE Access.

Gul, A. N. A. and Altuntas, A. (2023), Comparison of EEG-based deep neural network classifiers for emotion recognition using selected electrodes. 2023 Medical Technologies Congress (TIPTEKNO), pp. 1-4, IEEE.

Guodong, S., Mingmao, H., Yanfei, L., Jian, F., Aihong, G. and Qingshan, G. (2024), A multi-strategy fusion-based Rat Swarm Optimization algorithm. Soft Computing, pp. 1-9.

Gupta, H., Sharma, O. and Bhardwaj, D. (2023), EEG signal-based multi-class emotion recognition using hybrid 1D-CNN and GRU.

Henni, K., Mezghani, N., Mitiche, A., Abou-Abbas, L. and Yahia, A. B. (2025), An effective deep neural network architecture for EEG-based recognition of emotions. IEEE Access.

Hosseini, M. S. K., Firoozabadi, S. M. and Badie, K. (2023), Personality-based emotion recognition using EEG signals with a CNN-LSTM network. Brain Sciences, Vol. 13, No. 6, p. 947.

Hu, J. and Wang, C. and Jia, Q. and Bu, Q. and Sutcliffe, R. and Feng, J. (2021), ScalingNet: Extracting features from raw EEG data for emotion recognition. Neurocomputing, Vol. 463, pp. 177-184.

Hwang, S., Ki, M., Hong, K. and Byun, H. (2020), Subject-independent EEG-based emotion recognition using adversarial learning. 2020 8th International Winter Conference on Brain-Computer Interface (BCI), pp. 1-4, IEEE.

Islam, M.R. and Moni, M.A. and Islam, M.M. and Rashed-Al-Mahfuz, M. and Islam, M.S. and Hasan, M.K. and Hossain, M.S. and Ahmad, M. and Uddin, S. and Azad, A. and Alyami, S.A. (2021), Emotion recognition from EEG signal focusing on deep learning and shallow learning techniques. IEEE Access, Vol. 9, pp. 94601-94624.

Iyer, A., Das, S. S. and Teotia, R. (2023), CNN and LSTM-based ensemble learning for human emotion recognition using EEG recordings. Multimedia Tools and Applications, Vol. 82, No. 4, pp. 4883-4896.

Iyer, A., Das, S. S., Teotia, R., Maheshwari, S. and Sharma, R. R. (2023), CNN and LSTM based ensemble learning for human emotion recognition using EEG recordings. Multimedia Tools and Applications, Vol. 82, No. 4, pp. 4883-4896.

Jenke, R. and Peer, A. and Buss, M. (2014), Feature extraction and selection for emotion recognition from EEG. IEEE Transactions on Affective Computing, Vol. 5. No. 3, pp. 327–339.

Jeong, J., Chae, J. H. and Kim, S. (2001), Nonlinear dynamic analysis of the EEG in patients with Alzheimer disease and vascular dementia. Journal of Clinical Neurophysiology, Vol. 18, No. 1, pp. 58-67.

Kaur, C. and Bisht, A. and Singh, P. and Joshi, G. (2021), EEG signal denoising using hybrid approach of Variational Mode Decomposition and wavelets for depression. Biomedical Signal Processing and Control, Vol. 65, pp. 102337.

Khateeb, M. and Anwar, S.M. and Alnowami, M. (2021), Multi-domain feature fusion for emotion classification using DEAP dataset. IEEE Access, Vol. 9, pp. 12134-12142.

Khosla, A. and Khandnor, P. and Chand, T. (2020), A comparative analysis of signal processing and classification methods for different applications based on EEG signals. Biocybernetics and Biomedical Engineering, Vol. 40, pp. 649–690.

Kim, Y. and Choi, A. (2020), EEG-based emotion classification using long short-term memory network with attention mechanism. Sensors, Vol. 20. No. 23, pp. 6727.

Koelstra, S. and et al. (2012), DEAP: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, Vol. 3. No. 1, pp. 18-31.

Koelstra, S. and Muhl, C. and Soleymani, M. and Lee, J.S. and Yazdani, A. and Ebrahimi, T. and Pun, T. and Nijholt, A. and Patras, I. (2011), Deap: A database for emotion analysis using physiological signals. IEEE Transactions on Affective Computing, Vol. 3. No. 1, pp. 18-31.

Krishna, A.H. and Sri, A.B. and Priyanka, K.Y.V.S. and Taran, S. and Bajaj, V. (2019), Emotion classification using EEG signals based on tunable-Q wavelet transform. IET Science, Measurement & Technology, Vol. 13. No. 3, pp. 375-380.

Li, C., Chen, B., Zhao, Z., Cummins, N. and Schuller, B. W. (2021), Hierarchical attention-based temporal convolutional networks for EEG-based emotion recognition. ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1240-1244, IEEE.

Li, C., Sheng, Y., Wang, H., Niu, M., Jing, P., Zhao, Z. and Schuller, B. W. (2022), EEG emotion recognition based on self-attention dynamic graph neural networks. 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 292-296, IEEE.

Li, M., Xu, H.P. and Liu, X.W. (2018), Emotion recognition from multichannel EEG signals using K-nearest neighbor classification. Technology and Health Care, Vol. 26, No. 5, pp. 509–519.

Li, Q., Liu, Y., Liu, Q., Zhang, Q., Yan, F., Ma, Y. and Zhang, X. (2022), Multidimensional feature in emotion recognition based on multi-channel EEG signals. Entropy, Vol. 24, No. 12, p. 1830.

Li, R., Ren, C., Zhang, X. and Hu, B. (2022), A novel ensemble learning method using multiple objective particle swarm optimization for subject-independent EEG-based emotion recognition. Computers in Biology and Medicine, Vol. 140, p. 105080.

Li, W., Tian, Y., Hou, B., Dong, J., Shao, S. and Song, A. (2023), A Bi-Stream hybrid model with MLPBlocks and self-attention mechanism for EEG-based emotion recognition. Biomedical Signal Processing and Control, Vol. 86, p. 105223.

Liu, J. and Wu, G. and Luo, Y. and Qiu, S. and Yang, S. and Li, W. and Bi, Y. (2020), EEG-based emotion classification using a deep neural network and sparse autoencoder. Frontiers in Systems Neuroscience, Vol. 14, pp. 43.

Liu, S., Wang, X., Zhao, L., Li, B., Hu, W., Yu, J. and Zhang, Y.-D. (2021), 3DCANN: A spatio-temporal convolution attention neural network for EEG emotion recognition. IEEE Journal of Biomedical and Health Informatics, Vol. 26, No. 11, pp. 5321-5331.

Liu, X., Wang, L., Yan, H., Cao, Q., Zhang, L. and Zhao, W. (2024), Optimizing the Probabilistic Neural Network Model with the Improved Manta Ray Foraging Optimization Algorithm to Identify Pressure Fluctuation Signal Features. Biomimetics, Vol. 9, No. 1, p. 32.

Liu, Y. and Ding, Y. and Li, C. and Cheng, J. and Song, R. and Wan, F. and Chen, X. (2020), Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network. Computers in Biology and Medicine, Vol. 123, pp. 103927.

Liu, Y. and Fu, G. (2021), Emotion recognition by deeply learned multi-channel textual and EEG features. Future Generation Computer Systems, Vol. 119, pp. 1-6.

Liu, Z., Xie, Q. and Wu, M. (2019), Electroencephalogram emotion recognition based on empirical mode decomposition and optimal feature selection. IEEE Transactions on Cognitive and Developmental Systems, Vol. 11, No. 4, pp. 517-526.

Lu, K., Gu, Z., Qi, F., Sun, C., Guo, H. and Sun, L. (2024), CMLP-Net: A convolution-multilayer perceptron network for EEG-based emotion recognition. Biomedical Signal Processing and Control, Vol. 96, p. 106620.

Luo, Y., Wu, C. and Lv, C. (2023), Cascaded convolutional recurrent neural networks for EEG emotion recognition based on temporal–frequency–spatial features. Applied Sciences, Vol. 13, No. 11, p. 6761.

Mahmud, M. S., Saha, O. and Fattah, S. A. (2022), An efficient bidirectional LSTM-based deep neural network for automatic emotion recognition using EEG signal. 2022 12th International Conference on Electrical and Computer Engineering (ICECE), pp. 417-420, IEEE.

Mao, Y., Song, L. and Zeng, X. (2021), Emotion classification from short-term EEG signals in deep learning. 2021 3rd International Symposium on Smart and Healthy Cities (ISHC), pp. 169-173, IEEE.

Maria, M. A., Akhand, M. and Shimamura, T. (2022), Emotion recognition from EEG with normalized mutual information and convolutional neural network. 2022 12th

International Conference on Electrical and Computer Engineering (ICECE), pp. 372-375, IEEE.

Matthew, J. M. and Noordheen Mohammad Mustafa, M. B. (2024), Enhancement of Hybrid Deep Neural Network Using Activation Function for EEG-Based Emotion Recognition. Traitement du Signal, Vol. 41, No. 4.

Mayor-Torres, J. M., Ravanelli, M., Medina-DeVilliers, S. E., Lerner, M. D. and Riccardi, G. (2021), Interpretable SincNet-based deep learning for emotion recognition from EEG brain activity. 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 412-415, IEEE.

Meng, E. and Huang, S. and Huang, Q. and Fang, W. and Wu, L. and Wang, L. (2019), A robust method for non-stationary streamflow prediction based on improved EMD-SVM model. Journal of Hydrology, Vol. 568, pp. 462-478.

Mert, A. and Akan, A. (2018), Emotion recognition from EEG signals by using multivariate empirical mode decomposition. Pattern Analysis and Applications, Vol. 21. No. 1, pp. 81-89.

Moon, S.-E., Jang, S. and Lee, J.-S. (2018), Convolutional neural network approach for EEG-based emotion recognition using brain connectivity and its spatial information. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2556-2560, IEEE.

Nandi, A. and Xhafa, F. and Subirats, L. and Fort, S. (2021), Real-time emotion classification using EEG data stream in e-learning contexts. Sensors, Vol. 21. No. 5, pp. 1589.

Nawaz, Q., Younas, M., Hamid, I. and Gilani, S. M. M. (2023), Recurrent neural network-based human emotion recognition using EEG brain signals. 2023 3rd International Conference on Computing and Information Technology (ICCIT), pp. 117-122, IEEE

Nawaz, R. and Cheah, K.H. and Nisar, H. and Yap, V.V. (2020), Comparison of different feature extraction methods for EEG-based emotion recognition. Biocybernetics and Biomedical Engineering, pp. 910-926.

Nolte, J. (2002), The human brain: an introduction to its functional anatomy. Journal of Neuroscience Methods.

Pan, L. and Yin, Z. and She, S. and Song, A. (2020), Emotional state recognition from peripheral physiological signals using fused nonlinear features and team-collaboration identification strategy. Entropy, Vol. 22. No. 5, pp. 511.

Pandey, P. and Seeja, K. (2019), Subject independent emotion recognition from EEG using VMD and deep learning. Journal of King Saud University - Computer and Information Sciences, pp. 53–58.

Parveen, K. S., Panachakel, J. T., Ranjana, H., Sidharth, S. and Samuel, A. A. (2023), EEG-based emotion classification—a theoretical perusal of deep learning methods. 2023 2nd International Conference for Innovation in Technology (INOCON), pp. 1-6, IEEE.

Pusarla, N., Singh, A. and Tripathi, S. (2020), Ensemble algorithms for EEG-based emotion recognition. National Conference on Communications (NCC), Vol. 2020, No. 1, pp. 1-4.

Pusarla, N., Singh, A. and Tripathi, S. (2022), Normal inverse Gaussian features for EEG-based automatic emotion recognition. IEEE Transactions on Instrumentation and Measurement, Vol. 71, pp. 1-11.

Qu, H. and Jing, F. (2024), Research and analysis of facial emotion recognition based on convolutional neural network. 2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL), pp. 434-437, IEEE.

Rajasekhar, B., Kamaraju, M. and Sumalatha, V. (2020), A novel speech emotion recognition model using mean update of particle swarm and whale optimization-based deep belief network. Data Technologies and Applications, Vol. 54, No. 3, pp. 297-322.

Ramzan, M. and Dawn, S. (2023), Fused CNN-LSTM deep learning emotion recognition model using electroencephalography signals. International Journal of Neuroscience, Vol. 133, No. 6, pp. 587-597.

Ramzan, M. and Dawn, S. (2023), Fused CNN-LSTM deep learning emotion recognition model using electroencephalography signals. International Journal of Neuroscience, Vol. 133, No. 6, pp. 587-597.

Russell, J.A. (1980), A circumplex model of affect. Journal of Personality and Social Psychology, Vol. 39. No. 6, pp. 1161–1178.

Salankar, N. and Mishra, P. and Garg, L. (2021), Emotion recognition from EEG signals using empirical mode decomposition and second-order difference plot. Biomedical Signal Processing and Control, Vol. 65, pp. 102389.

Salankar, N., Mishra, P. and Garg, L. (2021), Emotion recognition from EEG signals using empirical mode decomposition and second-order difference plot. Biomedical Signal Processing and Control, Vol. 65, No. 2, pp. 1-3.

Salankar, N., Mishra, P. and Garg, L. (2021), Emotion recognition from EEG signals using empirical mode decomposition and second-order difference plot. Biomedical Signal Processing and Control, Vol. 65, p. 102389.

Samal, P. and Hashmi, M. F. (2023), Ensemble median empirical mode decomposition for emotion recognition using EEG signal. IEEE Sensors Letters.

Sharma, H. K., Sar, A., Choudhury, T., Kotecha, K., Dutta, S. and Bhattacharya, A. (2023), A novel framework for facial emotion detection using deep learning algorithms for HCI-enabled system. International Conference on Cyber Intelligence and Information Retrieval, pp. 15-29, Springer.

Sharma, M. and Darji, J. and Thakrar, M. and Acharya, U.R. (2022), Automated identification of sleep disorders using wavelet-based features extracted from electrooculogram and electromyogram signals. Computers in Biology and Medicine, Vol. 143, pp. 105224.

Sharma, R. and Pachori, R.B. and Sircar, P. (2020), Automated emotion recognition based on higher order statistics and deep learning algorithm. Biomedical Signal Processing and Control, Vol. 58, pp. 101867.

Shen, F. and Peng, Y. and Kong, W. and Dai, G. (2021), Multi-scale frequency bands ensemble learning for EEG-based emotion recognition. Sensors, Vol. 21. No. 4, pp. 1262.

Shu, L. and Xie, J. and Yang, M. and Li, Z. and Li, Z. and Liao, D. and Xu, X. and Yang, X. (2018), A review of emotion recognition using physiological signals. Sensors, Vol. 18. No. 7, pp. 2074.

Si, T., Miranda, P. B., Nandi, U., Jana, N. D., Mallik, S., Maulik, U. and Qin, H. (2024), Opposition-based chaotic tunicate swarm algorithms for global optimization. IEEE Access, Vol. 12, pp. 18168-18188

Singh, G., Verma, K., Sharma, N., Kumar, A. and Mantri, A. (2020), Emotion recognition using deep convolutional neural network on temporal representations of physiological signals. 2020 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT), pp. 1-6, IEEE.

Song, T. and Zheng, W. and Liu, S. and Zong, Y. and Cui, Z. and Li, Y. (2021), Graph-Embedded Convolutional Neural Network for Image-based EEG Emotion Recognition. IEEE Transactions on Emerging Topics in Computing.

Song, T. and Zheng, W. and Song, P. and Cui, Z. (2018), EEG emotion recognition using dynamical graph convolutional neural networks. IEEE Transactions on Affective Computing, Vol. 11. No. 3, pp. 532-541.

Song, T., Zheng, W., Liu, S., Zong, Y., Cui, Z. and Li, Y. (2021), Graph-embedded convolutional neural network for image-based EEG emotion recognition. IEEE Transactions on Emerging Topics in Computing, Vol. 10, No. 3, pp. 1399-1413.

Song, T., Zheng, W., Song, P. and Cui, Z. (2018), EEG emotion recognition using dynamical graph convolutional neural networks. IEEE Transactions on Affective Computing, Vol. 11, No. 3, pp. 532-541.

Subasi, A. and Tuncer, T. and Dogan, S. and Tanko, D. and Sakoglu, U. (2021), EEG-based emotion recognition using tunable Q wavelet transform and rotation forest ensemble classifier. Biomedical Signal Processing and Control, Vol. 68, pp. 102648.

Sunaryono, D., Sarno, R., Siswantoro, J., Raharjo, A. B., Sabilla, S. I., Susilo, R. I. and Rekha, K. (2022), Enhanced Salp Swarm Algorithm Based on Convolutional Neural Network Optimization for Automatic Epilepsy Detection. Journal of Theoretical and Applied Information Technology, Vol. 100, No. 19, pp. 5615-5623.

Tan, C. and Ceballos, G. and Kasabov, N. and Puthanmadam Subramaniyam, N. (2020), Fusionsense: Emotion classification using feature fusion of multimodal data and deep learning in a brain-inspired spiking neural network. Sensors, Vol. 20. No. 18, pp. 5328.

Tan, C. and Šarlija, M. and Kasabov, N. (2021), NeuroSense: Short-term emotion recognition and understanding based on spiking neural network modelling of spatio-temporal EEG patterns. Neurocomputing, Vol. 434, pp. 137-148.

Taran, S. and Bajaj, V. (2018), Clustering variational mode decomposition for identification of focal EEG signals. IEEE Sensors Letters, Vol. 2. No. 4, pp. 1-4.

Taran, S. and Bajaj, V. (2019), Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method. Computer Methods and Programs in Biomedicine, Vol. 178, No. 4, pp. 1-3.

Tripathi, A. and Choudhury, T. (2023), EEG-based emotion recognition using long short-term memory network with improved rat swarm optimization algorithm. Revue d'Intelligence Artificielle, Vol. 37, No. 2, p. 281.

Tuncer, T. and Dogan, S. and Subasi, A. (2021), A new fractal pattern feature generation function-based emotion recognition method using EEG. Chaos, Solitons & Fractals, Vol. 144, pp. 110671.

Ullah, H. and Uzair, M. and Mahmood, A. and Ullah, M. and Khan, S.D. and Cheikh, F.A. (2019), Internal emotion classification using EEG signal with sparse discriminative ensemble. IEEE Access, Vol. 7, pp. 40144-40153.

Wang, K. and Guo, P. (2020), An ensemble classification model with unsupervised representation learning for driving stress recognition using physiological signals. IEEE Transactions on Intelligent Transportation Systems, Vol. 22. No. 6, pp. 3303-3315.

Wang, X., Ma, Y., Cammon, J., Fang, F., Gao, Y. and Zhang, Y. (2023), Self-Supervised EEG Emotion Recognition Models Based on CNN. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 31, pp. 1952-1962.

Wang, Y., Qiu, S., Li, J., Ma, X., Liang, Z., Li, H. and He, H. (2019), EEG-based emotion recognition with similarity learning network. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1209-1212, IEEE.

Wang, Z. (2020), Text emotion detection based on Bi-LSTM network. Academic Journal of Computing & Information Science, Vol. 3, No. 3, pp. 129-137.

Wen, X. and Zhou, M. (2024), Evolution and role of optimizers in training deep learning models. IEEE/CAA Journal of Automatica Sinica, Vol. 11, No. 10, pp. 2039-2042.

Wen, Z., Xu, R. and Du, J. (2017), A novel convolutional neural network for emotion recognition based on EEG signal. 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), pp. 672-677, IEEE

Wu, Q., Dey, N., Shi, F., Crespo, R. G. and Sherratt, R. S. (2021), Emotion classification on eye-tracking and electroencephalograph fused signals employing deep gradient neural networks. Applied Soft Computing, Vol. 110, p. 107752.

Xu, D., Qin, X., Dong, X. and Cui, X. (2023), Emotion recognition of EEG signals based on variational mode decomposition and weighted cascade forest. Mathematical Biosciences and Engineering, Vol. 20, pp. 2566-2587.

Xu, F., Pan, D., Zheng, H., Ouyang, Y., Jia, Z. and Zeng, H. (2024), EESCN: A novel spiking neural network method for EEG-based emotion recognition. Computer Methods and Programs in Biomedicine, Vol. 243, p. 107927

Yang, H. and Han, J. and Min, K. (2019), A multi-column CNN model for emotion recognition from EEG signals. Sensors, Vol. 19. No. 21, pp. 4736.

Yang, J. and Huang, X. and Wu, H. and Yang, X. (2020), EEG-based emotion classification based on bidirectional long short-term memory network. Procedia Computer Science, Vol. 174, pp. 491-504.

Yang, L. and Liu, J. (2019), EEG-based emotion recognition using temporal convolutional network. 2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS), pp. 437-442, IEEE.

Yin, S., Jin, M., Lu, H., Gong, G., Mao, W., Chen, G. and Li, W. (2023), Reinforcement-learning-based parameter adaptation method for particle swarm optimization. Complex & Intelligent Systems, Vol. 9, No. 5, pp. 5585-5609.

Yin, Y. and Zheng, X. and Hu, B. and Zhang, Y. and Cui, X. (2021), EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. Applied Soft Computing, Vol. 100, pp. 106954.

Yin, Z., Zhao, M., Wang, Y., Yang, J. and Zhang, J. (2017), Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. Computer Methods and Programs in Biomedicine, Vol. 140, pp. 93-110.

Zali-Vargahan, B., Charmin, A., Kalbkhani, H. and Barghandan, S. (2023), Deep time-frequency features and semi-supervised dimension reduction for subject-independent emotion recognition from multi-channel EEG signals. Biomedical Signal Processing and Control, Vol. 85, p. 104806.

Zangeneh Soroush, M., Maghooli, K. and Setarehdan, K. (2018), Emotion classification through nonlinear EEG analysis using machine learning methods. International Clinical Neuroscience Journal, Vol. 5, No. 3, pp. 135-149.

Zeng, H. and Wu, Z. and Zhang, J. and Yang, C. and Zhang, H. and Dai, G. and Kong, W. (2019), EEG emotion classification using an improved SincNet-based deep learning model. Brain Sciences, Vol. 9. No. 11, pp. 326.

Zhang, C., Liu, H., Zhou, Q. and Liu, C. (2023), Improved hybrid grey wolf optimization algorithm based on dimension learning-based hunting search strategy. IEEE Access, Vol. 11, pp. 13738-13753.

Zhang, J., Zhang, X. and Zhao, Q. (2022), Improved graph convolutional neural networks based on granger causality analysis for EEG emotion recognition. 2022 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), pp. 684-688, IEEE.

Zhang, X. and Liu, J. and Shen, J. and Li, S. and Hou, K. and Hu, B. and Gao, J. and Zhang, T. (2020), Emotion recognition from multimodal physiological signals using a regularized deep fusion of kernel machine. IEEE Transactions on Cybernetics, Vol. 51. No. 9, pp. 4386-4399.

Zhang, X., Liu, J., Shen, J., Li, S., Hou, K., Hu, B., Gao, J. and Zhang, T. (2020), Emotion recognition from multimodal physiological signals using a regularized deep fusion of kernel machine. IEEE Transactions on Cybernetics, Vol. 51, No. 9, pp. 4386-4399.

Zhang, Y., Liu, H., Zhang, D., Chen, X., Qin, T. and Zheng, Q. (2022), EEG-based emotion recognition with emotion localization via hierarchical self-attention. IEEE Transactions on Affective Computing, Vol. 14, No. 3, pp. 2458-2469.

Zhao, Q., Dong, Y. and Yin, W. (2023), Emotion recognition based on EEG and DE-CNN-RNN. 2023 China Automation Congress (CAC), pp. 3376-3380, IEEE.

Zheng, J. and Pan, H. (2020), Mean-optimized mode decomposition: An improved EMD approach for non-stationary signal processing. ISA Transactions, Vol. 106, pp. 392-401.

Zhuang, N., Zeng, Y. and Tong, L. (2017), Emotion recognition from EEG signals using multidimensional information in EMD domain. BioMed Research International, Vol. 2017, No. 8, pp. 1-8.

Zong, J., Xiong, X., Zhou, J., Ji, Y., Zhou, D. and Zhang, Q. (2023), FCAN–XGBoost: A Novel Hybrid Model for EEG Emotion Recognition. Sensors, Vol. 23, No. 12, p. 5680.

LIST OF PUBLICATIONS

Scholar Name: Amrendra Nath Tripathi

Orchid Id: https://orcid.org/0000-0002-2350-4012

Journals:

1. Tripathi A., Choudhury, T. (2023). EEG based emotion recognition using Long short term memory network with improved Rat Swarm Optimization Algorithm. Revue d'Intelligence Artificielle, Vol. 37, No. 2, pp. 281-289. https://doi.org/10.18280/ria.370205 https://www.iieta.org/journals/ria/paper/10.18280/ria.370205 [Scopus]
2. Tripathi, A., Choudhury, T., Sharma, H.K. (2024). EEG based emotion detection by using modified tunicate swarm optimization algorithm. Ingénierie des Systèmes d'Information, Vol. 29, No. 4, pp. 1333-1342. https://doi.org/10.18280/isi.290409

Conference Published:

1. Tripathi and T. Choudhury, "Permuted layer-based CNN for Emotion Detection with Multi-Modality Physiological Signals," *2023 IEEE International Conference on Contemporary Computing and Communications (InC4)*, Bangalore, India, 2023, pp. 1-5, doi: 10.1109/InC457730.2023.10263176. https://ieeexplore.ieee.org/document/10263176

Amrendra Nath Tripathi

A Hybrid Signal Processing Method for the Classification of Emotions from Physiological Signal: EEG with Deep Neural Net...

Quick Submit

Quick Submit

University of Petroleum & Energy Studies

Document Details

Submission ID

trn:oid:::1:3293323523

Submission Date

Jul 10, 2025, 3:28 PM GMT+5:30

Download Date

Jul 10, 2025, 4:06 PM GMT+5:30

File Name

Amrendra_Tripathi_for_plag_run.pdf

File Size

6.7 MB

150 Pages

38,043 Words

216,788 Characters

2% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Small Matches (less than 14 words)

Match Groups

31 Not Cited or Quoted 2%
Matches with neither in-text citation nor quotation marks

2 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

1% Internet sources
1% Publications
1% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **31** Not Cited or Quoted 2%
  Matches with neither in-text citation nor quotation marks
- **2** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

1% Internet sources
1% Publications
1% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	iieta.org	<1%
2	Internet	journal.esrgroups.org	<1%
3	Internet	www.mdpi.com	<1%
4	Student papers	Munster Technological University (MTU)	<1%
5	Student papers	National Institute of Technology, Rourkela	<1%
6	Student papers	Charles University in Prague	<1%
7	Student papers	National College of Ireland	<1%
8	Student papers	University of Bristol	<1%
9	Publication	Rajasekhar Pillalamarri, Udhayakumar Shanmugam. "A review on EEG-based mul...	<1%
10	Student papers	Cyberjaya University College of Medical Sciences	<1%

11 Publication

Jie Sun, Tianwen Xu, Yao Yao. "An enhanced GhostNet model for emotion recogni... <1%

12 Publication

Ningbo Zhang, Yang Song, Dongxu Fang, Zhiwei Gao, Yajie Yan. "An Improved De... <1%

13 Student papers

University College London <1%

14 Student papers

University of Central Lancashire <1%

15 Publication

Wu, Jiaqi. "Deep Learning-Based Hosting Capacity Analysis in LV Distribution Grid... <1%

16 Internet

arxiv.org <1%

17 Internet

kuey.net <1%

18 Internet

v.vibdoc.com <1%

19 Student papers

Anna University <1%

20 Publication

Djordjevic, Vladana Radisavljevic. "Automated Nonlinear Analysis of Newborn Ele... <1%

21 Student papers

JIS University <1%

22 Publication

Jenamani Chandrakanta Badajena, Srinivas Sethi, Sanjit Kumar Dash, Ramesh Ku... <1%

23 Student papers

Maharishi University Of Information Technology, Lucknow <1%

24 Internet

lutpub.lut.fi <1%

25 Internet

ruidera.uclm.es <1%