EEG Emotion Recognition and Detection Based on DEAP Dataset

[2025 April 25th]

ABSTRACT

At present, non-physiological signals, such as voice, facial expressions and body postures, are mostly used in the field of emotion recognition for identification. However, non-physiological signals are usually too subjective and it is difficult to judge the true emotions of human beings. Therefore, emotion recognition technology based on physiological signals, such as electroencephalogram (EEG), often has higher accuracy and reliability in the field of emotion recognition. With the development of machine learning and deep learning, more and more artificial intelligence algorithms, such as convolutional networks, are applied to EEG emotion recognition.
Our project aims to use deep-learning method and traditional machine learning method to realize the emotion recognition respectively, using the open-source EEG dataset called DEAP. In the final four-category emotion recognition task, the highest accuracy of the deep learning model was close to 80%, exceeding the highest accuracy of 66% obtained by the SVM model. We successfully verify that some deep learning models such as GAT have higher accuracy and better generalization when processing high-dimensional data such as EEG and classification tasks. Based on the experimental data, we propose possible model optimization strategies and future work direction. The superiority of graph neural networks (GAT) in EEG emotion recognition is proposed in this project, and its functional connectivity modeling provides an interpretable basis for neuroscience (such as channel importance quantification), challenging the limitations of traditional machine-learning models such as SVM.

KEYWORDS

Deep Learning, Machine Learning, Emotion recognition, EEG signal, Graph neutral network

Table of Contents

Chapter 1: Introduction

The introduction chapter sets the stage for our project by outlining the objectives, key concepts, and background. It will contain the research status and key concepts in this group project, which is a brief frame of the overall project.

1.1 Background

EEG is particularly useful for people who cannot express emotions through speech or physical gestures, making it an effective tool for studying emotions and understanding their cognitive mechanisms. EEG emotion recognition has a wide range of applications, including medical, social, educational, safe driving and military fields. Emotion recognition follows the essence of pattern recognition research, which is to judge the emotion category of target samples according to existing data and some measurement criteria[4]. With the rapid development of deep learning in the fields of image processing and natural language processing, more and more researchers have begun to pay attention to the technology route based on deep learning in the direction of EEG emotion recognition..

1.2 Context

Emotion recognition has traditionally relied on non-physiological signals such as voice, facial expressions, and body postures. While these modalities are widely accessible and intuitive, their subjective nature poses significant limitations. Non-physiological signals[1]-[3] can be consciously or subconsciously manipulated, making it challenging to discern genuine emotional states. For instance, individuals may suppress facial expressions or alter vocal tones to conform to social norms, leading to inaccuracies in emotion detection.

1.3 Purposes

The comparative analysis aims to highlight the strengths of deep learning in handling high-dimensional, nonlinear EEG data while addressing challenges such as overfitting and computational costs. By bridging physiological signal analysis with advanced AI techniques, this work contributes to the development of reliable emotion-aware systems for applications in mental health monitoring, human-computer interaction, and neurofeedback therapies. Future directions include extending the framework to multimodal fusion (EEG + physiological signals) and enhancing model interpretability for clinical adoption.

1.4 Significance, Scope, and Definitions

This research contributes to advancing emotion recognition technology by integrating physiological EEG signals with deep learning and machine learning methodologies. Traditional emotion recognition systems often rely on non-physiological signals (e.g., facial expressions or voice), which are susceptible to intentional manipulation. By contrast, EEG signals provide objective and direct insights into neural activities, making them more reliable for capturing genuine emotional states. Our work addresses a critical gap in existing studies by systematically comparing the performance of deep learning models (GCN/GAT) and classical machine learning approaches (SVM) in processing EEG data, particularly focusing on their ability to handle high-dimensional, nonlinear patterns.

1.4.1 Significance

Enhances understanding of how graph-based neural networks (GCN/GAT) can model dynamic brain network interactions for emotion classification, offering a novel perspective on EEG signal analysis which provides a framework for developing robust emotion-aware systems with applications in mental health monitoring, neurofeedback therapies, and human-computer interaction. Demonstrates the advantages of automated feature learning in deep learning over manual feature engineering in traditional machine learning, particularly for complex EEG datasets.

1.4.2 Scope

The study focuses on the DEAP dataset, utilizing prefrontal EEG channels (Fp1, Fp2, AF3, AF4) to classify emotions into four categories (valence, arousal, dominance, liking). The scope includes comparative analysis of model performance (accuracy, computational efficiency) and emphasizes the challenges of overfitting and interpretability in deep learning.

1.4.3 Key Definitions

EEG (Electroencephalogram): A non-invasive technique to record electrical activity of the brain via electrodes placed on the scalp.

GAT (Graph Attention Network): A variant of GCN that employs attention mechanisms to dynamically weigh node interactions during feature aggregation.

SVM (Support Vector Machine): A supervised machine learning algorithm that constructs hyperplanes to separate data into classes, often using kernel functions (e.g., RBF) for nonlinear classification.

DEAP Dataset: A multimodal dataset containing EEG and physiological signals from participants exposed to emotional stimuli, labeled with valence, arousal, dominance, and liking ratings.

1.5 Report Outline.

The structure of the subsequent chapters of this report is as follows: Chapter Two introduces the DEAP dataset and its preprocessing, covering machine learning feature engineering and deep learning data construction methods; Chapter Three expounds the core principles, model architectures and evaluation criteria of Graph Attention Network (GAT); Chapter 4 analyses the non-deep learning paradigm of Support Vector Machine (SVM); Chapter Five compares the performance indicators such as classification accuracy and computational efficiency of the two types of models; Chapter 6 summarizes the research findings, explores their theoretical value and practical significance in the field of emotion recognition, and proposes future research directions; The appendix section supplements the team division of group tasks and technical details to ensure the transparency and reproducibility of the research. The full text follows the logical chain of "data basis - method design - result verification - conclusion elevation", taking into account both academic rigor and application orientation.

Chapter 2: Dataset and Data Preprocessing

This chapter introduces the information about the DEAP dataset and the methods we used for data preprocession. Otherwise, it includes ta lot extract-feature methods, we use the frequency domain features in both methods, but there are differences between two methods, which is obvious in the related procedures.

2.1 Dataset Overview

DEAP (Dataset for Emotion Analysis using Physiological Signals) is a multimodal dataset designed for emotion analysis, comprising EEG and other physiological signals (such as electrodermal activity and electrocardiogram) from 32 participants[6].During the experiment, participants watched 40 music videos and rated their emotional responses on dimensions such as valence and arousal on a scale of 1 to 9, which is a cornerstone resource for emotion recognition research, offering a comprehensive collection of multimodal physiological data. The structure of the dataset is as follows:

Overall

Subjects	Videos	EEG Channels	Sampling rate	Rating scale	Rating values
32	40	32	128 Hz	Arousal Valence	Continuous scale of 1–9

EEG data format for each subject

Array	Array Shape	Array Content
Data	40 × 32 × 8064 (384 base + 7680 trial)	video/trial × channels × data
Labels	40 × 2	video/trial × label(valence, arousal)

Figure 2.1

How to do experiments to get the dataset? Key details include:

Participants: 32 healthy adults (16 males, 16 females) aged 19–37, screened for neurological or psychiatric conditions to ensure data reliability.

Experimental Design: 40 one-minute music videos curated to evoke diverse emotions (e.g., classical music for calmness, rock for excitement). Videos were selected based on pre-tests to validate their emotional impact.
Procedure: Participants watched videos in a controlled environment while physiological signals were recorded. After each video, they rated their emotional states using four 9-point scales:

Valence: Perceived pleasantness (1: negative, 9: positive).

Arousal: Emotional intensity (1: calm, 9: excited).

Dominance: Sense of control (1: submissive, 9: dominant).

Liking: Personal preference for the video (1: dislike, 9: like).

Data Acquisition: EEG: Recorded at 512 Hz via a 32-channel Biosome Active Two system, with electrodes placed according to the 10–20 international system. Preprocessing included down sampling to 128 Hz and referencing to averaged mastoids.

Peripheral Signals: Electrodermal activity (EDA), electrocardiogram (ECG), electromyogram (EMG), and respiration rate were synchronized with EEG.

Temporal Resolution: Each trial spans 63 seconds: 3 seconds of baseline recording, 60 seconds of video playback.

2.2 Approach for Machine Learning

General Steps, because the processing of data including signal processing and feature extraction are a little difficult, so we set some steps into Data Preprocessing. the workflow of machine learning before classification is as the figure shows:

Figure 2.2

The details are follows:

(1) Only apply efficient for small-scale DEAP. Down sampling and Filtering. down sampling: Raw 512 Hz EEG data is reduced to 128 Hz using anti-aliasing filters to retain critical frequencies while lowering computational load.
Bandpass Filtering: A 4th-order Butterworth filter (4–45 Hz) removes low-frequency drifts (<4 Hz, e.g., sweat artifacts) and high-frequency noise (>45 Hz, e.g., powerline interference).

(2) Artifact Removal:

Independent Component Analysis (ICA): Separates neural activity from artifacts (e.g., eye blinks identified via frontal channels Fp1/Fp2). Components correlated with artifacts are manually or automatically rejected.

Standardization:

Z-score standardization is applied to each EEG channel:

$$X_{norm} = \frac{X-\mu}{\sigma} \tag{2-1}$$

(μ is the channel mean, σ is the standard deviation)

(3) Feature Extraction: Extracts key features from EEG signals to construct feature vectors for emotion classification. Time-domain, Frequency-domain (PSD), Nonlinear features. Spatial features. Then employs ANOVA or L1 regularization to select highly discriminative features, in order to reduce redundant features to enhance model efficiency and generalization. In this project, we use welch method for PSD. More details will be presented in Chapter 4. Here we would like to introduce other method for feature extraction methods:

Time-Domain Features

Statistical Metrics: Mean, variance, skewness, kurtosis

Hjorth Parameters: Activity (signal power), mobility (mean frequency), complexity (frequency change rate).

Event-Related Potentials (ERPs): Amplitude and latency of peaks (e.g., P300) during emotional stimuli.

Power Spectral Density (PSD): Computed via Welch's method for delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–45 Hz) bands. Because the noise, there are some deviations in bands' boundary.

Spectral Asymmetry: Differences in left vs. right hemisphere power for valence assessment.

Nonlinear Features:

1. Entropy Measures: Approximate entropy (ApEn), sample entropy (SampEn), and fuzzy entropy (FuzzEn) to quantify signal irregularity.

2. Fractal Dimensions: Higuchi's algorithm to assess signal complexity.

(4) Dimensionality Reduction: Use PCA, LDA, or t-SNE to reduce the selected feature dimensions. Because the labels are divided into 4 classes, so it's not necessary to use the PCA method.

Principal Component Analysis (PCA): Projects features into orthogonal axes, retaining 95% variance.
Linear Discriminant Analysis (LDA): Maximizes class separability for supervised tasks.

t-SNE: Non-linear visualization of high-dimensional clusters (e.g., valence vs. arousal groups).

(5) Model Training:

Classifiers: Support Vector Machines (SVM) with radial basis function (RBF) kernels, Random Forests (RF), and k-Nearest Neighbors (k-NN).

Validation: Stratified 10-fold cross-validation to address class imbalance.

2.3 Approach for Deep Learning

The largest difference from machine learn is that deep learning is automatic feature learning, Directly process graph-structured data to capture dynamic brain network interactions and networks inherently handle high-dimensional raw data. Of.course, in our project, both the model uses the preprocessed data file, so it may be not necessary to process the raw data. Related details will be presented in chapter 3. Here we still briefly introduce the raw data procession for Deep-learning method.

Figure 2.3

1. Input Construction: Directly uses raw EEG time-series signals or time-frequency representations (e.g., spectrograms generated via wavelet transform).

Raw EEG: Structured as 3D tensors $[N_{norm} \times 32_{channels} \times 8064_{time\ point}]$
Spectrograms: Generated via Short-Time Fourier Transform (STFT) with Hann window (256-ms window, 50% overlap).
Wavelet Scalograms: Morlet wavelet transform captures multi-resolution dynamics.

2. Data Augmentation: Enhances model robustness through additive noise, random cropping, or channel shifting.
Additive Noise: Inject Gaussian noise (μ=0 and σ=0.1) to improve robustness.
Random Cropping: Extract 5-second segments from full trials to simulate varying attention spans.
Channel Shuffling: Randomly permute non-adjacent channels to reduce spatial bias.

3. Format Adjustment: Converts data into tensor format (e.g., [number of samples, number of channels, time points]) to fit models.

4. Related Model Architectures

EEGNet: A compact CNN with temporal and spatial convolutions:
Temporal Convolution: 1D kernels (e.g., 64 filters, size 64) to capture frequency bands.
Spatial Convolution: Depthwise filters to learn channel interactions.
Separable Convolution: Reduces parameters while preserving performance.

LSTM Networks: Process sequential EEG data using bidirectional layers to model past and future contexts.
Transformer-Based Models: Self-attention mechanisms focus on salient time points (e.g., emotional peaks). Transfer Learning: Pretrain on larger datasets (e.g., BCI Competition IV) and fine-tune on DEAP.
Regularization: Dropout (rate=0.5), weight decay (L2=1W-4)

Loss Functions: Categorical cross-entropy for multi-class tasks, focal loss for imbalanced labels.

As for this project, data from the prefrontal channels Fp1, Fp2, AF3, and AF4 in the DEAP database(".py" format) were selected, forming a 40×4×8064 data array. Labels were organized into a 40×4 matrix, divided into four categories (pleasure, arousal, dominance, liking).

2.4 . Label Processing information

DEAP: Binarizes valence/arousal ratings (threshold at 5.0, above as 1, below as 0). The final preprocessed data is saved in structured file formats (e.g. .npz or .h5) for model training and testing.

Label processing ensures compatibility with supervised learning frameworks:

1. Binarization: For Deep Learning , valence and arousal ratings are thresholded at 5.0, converting continuous scores into binary classes:

High Valence (⩾5): Associated with positive emotions (e.g., happiness).

Low Valence (<5): Linked to negative states (e.g., sadness).

High Arousal (⩾5): Reflects excitement or stress.

Low Arousal (<5): Indicates calmness or boredom.

For machine learning, we should set the threshold for each label according to the information of the dataset. We can obtain the distribution of the data:

	Valence	Arousal	Dominance	Liking
count	1280.000000	1280.000000	1280.000000	1280.000000
mean	5.254313	5.156711	5.382750	5.518133
std	2.130816	2.020499	2.096321	2.282780
min	1.000000	1.000000	1.000000	1.000000
25%	3.867500	3.762500	3.932500	3.960000
50%	5.040000	5.230000	5.240000	6.050000
75%	7.050000	6.950000	7.040000	7.090000
max	9.000000	9.000000	9.000000	9.000000

Figure 2.4

Valence/Arousal/Dominance: Mark as 1 (high) when the original value is > 5, otherwise mark as 0 (low)

Preference: Use a higher threshold, mark as 1 (high) when the original value is > 6, otherwise mark as 0 (low)

According to the threshold rule, the continuous value is binarized and stored in the corresponding list

2. Data Structuring:

Labels are stored in a 40×4 matrix per participant, with columns for valence, arousal, dominance, and liking. Dominance and liking are optional for task-specific studies (e.g., dominance may correlate with perceived control in VR environments). For non-deep-learning method, we should use some parameters like adjusting “weight” to reach the balance.

3. Class Distribution Analysis:

Imbalance mitigation via oversampling (SMOTE) or weighted loss functions.

Storage: Preprocessed data and labels are saved in special format, enabling efficient I/O operations during model training.

2.5 Summary and Implications

The DEAP dataset and preprocessing methodologies provide a robust foundation for emotion recognition, yet challenges and opportunities persist:
Technical Insights:

1. Machine Learning vs. Deep Learning:\

ML: Interpretable features (e.g., alpha asymmetry) align with neuroscientific theories but require domain expertise. The extraction feature engineering is important for machine learning.
DL: End-to-end models excel in raw data processing but act as "black boxes," complicating clinical adoption. theories but require domain expertise. We

2. Cross-Subject Generalization:

Domain Adaptation: Techniques like CORrelation ALignment (CORAL) minimize distribution shifts between users.

3. Practical Applications
Mental Health: Real-time monitoring of depression or anxiety via wearable EEG devices. Neuromarketing: Assessing consumer emotional responses to advertisements. Brain-Computer Interfaces (BCIs): Emotion-aware systems for adaptive human-machine interaction.

Chapter 3: Deep-Learning Principles and Model Design:

This chapter will introduce the basic theory of Deep-Learning method and the details about GAT model Design. 3.1 is the overview of the GAT model and the principles of GAT is presented in 3.2. Besides, this chapter also introduce the attention mechanism and displays the important parameters that the model should have. By setting the parameters, the model can be evaluated. In section 3.4, we introduce the structure of our model, including layers and pooling roughly. Finally, we list the procedure and timeline which we arranged for this model design.

3.1 Overview of Deep-Learning model

The Graph Attention Network (GAT) is a novel deep learning architecture specifically designed to handle graph-structured data, particularly non-Euclidean data such as EEG signals. Unlike traditional Convolutional Neural Networks (CNNs), which are limited to grid-like data, GAT leverages attention mechanisms to dynamically weigh the importance of neighboring nodes, making it highly effective for tasks involving complex relationships, such as emotion recognition from multi-channel EEG signals.

1. The GAT model introduced in this thesis consists of two main components:

Graph Attention Layers: These layers compute attention coefficients to determine the significance of each node's neighbors, enabling adaptive feature aggregation. This mechanism allows the model to focus on the most relevant nodes in the graph, improving the accuracy of feature representation.

2. Multi-head Attention: By employing multiple attention heads, the model captures diverse aspects of node relationships, enhancing robustness and accuracy. Each attention head learns a different representation of the node features, and their outputs are combined to form a more comprehensive feature set.

The GAT model is particularly suited for EEG-based emotion recognition because it can model the functional connectivity between different brain regions, which is critical for understanding emotional states. The attention mechanism ensures that the model dynamically adjusts the importance of each EEG channel based on its contribution to emotion classification, leading to more accurate and interpretable results.

3.2 Principles of Graph Attention Neural Network Model

The core idea of Graph Attention Networks (GATs) is to use attention mechanisms to dynamically assign different weights to neighboring nodes during feature aggregation[8].

The schematic diagram of the Graph Attention Network structure is shown in Figure:

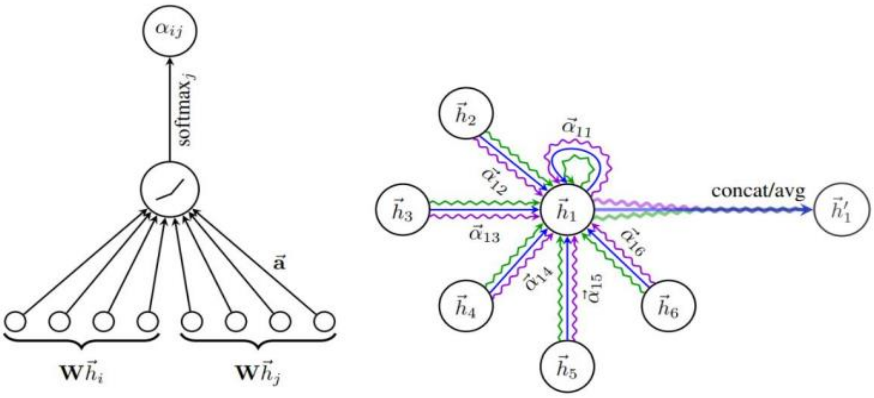

Figure 3.1

Below are the key concepts:

3.2.1 Attention Mechanism:

For each node in the graph, the attention mechanism computes a weighted sum of its neighbors' features. The weights (attention coefficients) are learned during training and indicate the importance of each neighbor.
Assuming there are N nodes, and the feature of each node i as input is hi, the updated node feature after this layer is hi'.
Firstly, a shared attention mechanism att is used to calculate self-attention between nodes:

$$e_{ij} = \mathrm{att}(Wh_i, Wh_j) \tag{3-1}$$

Where W is a shared weight that transforms the original node features from F dimensions to F' dimensions, then maps them to attention weights through the att function. This eij indicates the importance of node j relative to node i. In Graph Attention Networks, a single-layer feedforward neural network and a LeakyReLU nonlinear activation function are typically chosen to calculate eij:

$$e_{ij} = \mathrm{LeakyReLU}(a[Wh_i||Wh_j]) \tag{3-2}$$

where || denotes concatenation, and a represents a vector parameter.
The normalized attention weights are:

$$a_{ij} = \mathrm{Softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum k \in N_i \exp(e_{ij})} \tag{3-3}$$

3.2.2 Feature Aggregation:

The updated feature for node i is computed by aggregating the features of its neighbors, weighted by the attention coefficients:

$$h_i = \sigma(\textstyle\sum_{j \in N_i} \alpha_{ij} W h_j) \tag{3-4}$$

where Ni is the set of neighbors of node i, and σ is a nonlinear activation function.

3.2.3 Multi-head Attention:

To stabilize the learning process and capture diverse relationships, multiple attention heads are used. The outputs of these heads are either concatenated or averaged:

$$h_i' = ||_{k=1}^{K} \sigma(\sum_{j \in N_i} \alpha_{ij}^{k} W^k h_j) \tag{3-5}$$

where K is the number of attention heads.

3.2.4 Graph Structure for EEG Data:

In this work, EEG channels are treated as nodes in a graph, and the functional connectivity between channels (e.g., correlation or coherence) defines the edges. The attention mechanism adaptively learns the strength of these connections, allowing the model to focus on the most relevant channels for emotion recognition.

3.3 Evaluation

We aim to enhance EEG-based emotion classification accuracy by designing a 4-channel EEG signal acquisition circuit and implementing a corresponding algorithm using Python's PyTorch library[9]. Leveraging datasets like DEAP, we will compare our system against established benchmarks. The project timeline includes initial hardware and software setup, core algorithm development, and comprehensive testing phases. Performance will be based on classification accuracy, computational efficiency:

i. **Classification Accuracy:** Measuring how accurately the system can classify emotions into four categories.

ii. **Computational Efficiency:** Assessing the system's processing speed and resource utilization to ensure real-time applicability.

3.4 Details of model design

The GAT model in this thesis was implemented using PyTorch and PyTorch Geometric. Key design choices include:

• Input Layer: The input feature matrix X (shape: batch × 4 × 5) represents the differential entropy (DE) features of 4 EEG channels across 5 frequency bands.

• Graph Attention Layers: Two GAT layers were used, with the first layer having 8 attention heads and the second layer reducing the feature dimension for classification.

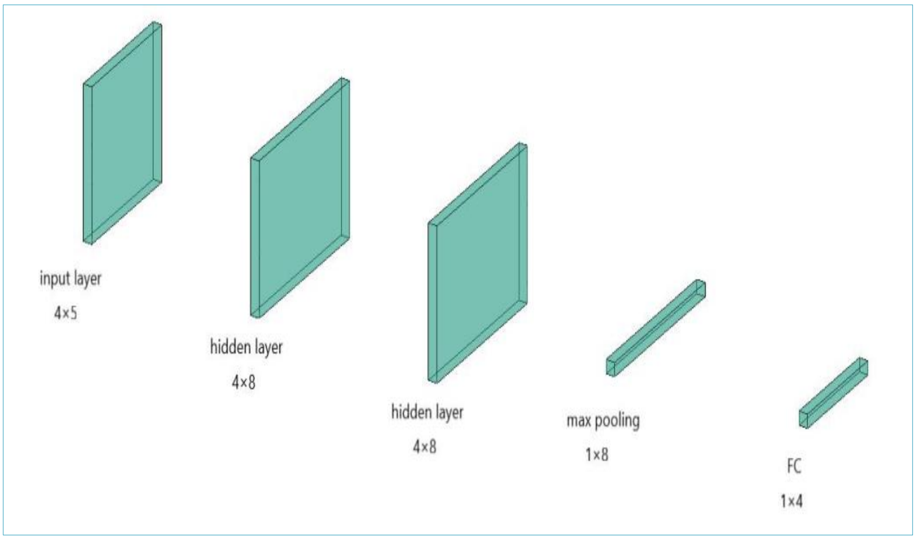

Figure 3.2

• *Parameters:*

Table 3.1

Parameters	Meanings	values
in_channels	The input feature dimension	5
out_channels	The output feature dimension	8
heads	heads number in multi-head attention	8
concat	whether to concatenate the results	true
negative_slope	parameter for LeakyReLU	0.2
add_self_loops	Whether to add self-loops	true
bias	Whether to add a bias term	True

• Pooling and Classification: A max-pooling layer was applied to reduce the graph representation to a fixed-size vector (batch × 8).

A fully connected layer mapped the pooled features to 4 output classes, followed by a Softmax activation for probability estimation.).

3.5 Procedure and Timeline

The development and training of the GAT model followed a structured timeline to ensure efficient progress and timely completion. Below is the detailed plan:

Week 1-2: Data Preparation and Preprocessing

1) Extract DE features from the DEAP dataset for 4 EEG channels (Fp1, Fp2, AF3, AF4).
2) Perform data normalization and splitting into training (1,120 samples) and test sets (160 samples).

Week 3-4: Model Implementation

1) Implement the GAT model using PyTorch and PyTorch Geometric.
2) Define the graph attention layers, multi-head attention mechanism, and classification head.

Week 5-6: Training and Validation

1) Train the model for 200 epochs with a learning rate of 0.01.
2) Monitor training dynamics, including loss and accuracy, to ensure convergence.
3) Validate the model on the test set every epoch to evaluate performance.

Week 7-8: Optimization and Fine-tuning

1) Adjust hyperparameters (e.g., learning rate, dropout rate) to improve model performance.

2) Incorporate techniques like early stopping to prevent overfitting.

Week 9-10: Final Testing and Documentation

1) Conduct final testing on the optimized model.
2) Document the results, including accuracy metrics and training dynamics.
3) Prepare the model for deployment in the EEG emotion recognition system.

3.6 Analysis

The GAT model's strengths and limitations are summarized below:

Strengths:

1) Effectively models the functional connectivity between EEG channels.
2) Achieves high accuracy with relatively few EEG channels (4 channels).
3) The attention mechanism provides interpretability by revealing which channels contribute most to emotion classification.

Limitations:

1) The accuracy could be improved by incorporating more EEG channels or multimodal data (e.g., fNIRS, heart rate).
2) The model's performance depends heavily on the quality of the graph structure (e.g., how edges are defined).
3) Computational complexity increases with the number of attention heads and nodes.

Future Directions:

1) Extend the model to handle more granular emotion classifications (e.g., 16 classes).

2) Integrate additional physiological signals (e.g., heart rate, skin conductance) for multimodal emotion recognition.

3) Optimize the model for real-time applications, such as wearable emotion monitoring devices.

Chapter 4: Non-Deep-Learning Model and Principles:

This chapter presents the machine learning method (non-deep-learning method) of this project. There is growing research evidence that machine learning can extract meaningful information from high-dimensional and noisy EEG signals. Given the attention and wide range of applications of related techniques, in this chapter we describe how machine learning can be used to analyse EEG signals. This part will include the principles and main theory of Machine-Learning model as well as the detail operations of the model. One of the things that needs to be emphasized is the feature engineering extraction part used for machine learning, which is a prerequisite for the modelling approach to be applied to this project and a part that is significantly different from the deep-learning approach route and the details we will present in the section 4.2. And the evaluation method we used including how to select the kernel and Optimize framework parameters will present at section 4.3, as well as the related principles and the routing about building our model.

4.1 Over view of Machine-Learning model

There are many traditional models of machine-learning, such as Simple Bayes, Randomized Decision Making, Random Forests and so on. Among these traditional methods, most of the classical algorithms such as Support Vector Machines (SVM), k-nearest neighbours (KNN), and Random Forests are hand-designed with the help of experts' prior knowledge on features. These methods have shown good application in epileptic seizure prediction, sleep staging recognition, and sentiment state analysis. For example, one study proposes a recognition model based on SVM, which effectively distinguishes normal and abnormal EEG signals and achieves high classification accuracy[10]. In addition, strategies such as migration learning and integration learning have been applied to EEG signal recognition to help ameliorate the effects of small sample problems and cross-individual differences. In our project, we choose the SVM model for our researching, because the model is simply and have some advantages that others not.

Why we choose the SVM? Firstly, We consider the EEG dataset, for EEG signal data, this kind of data often has more complex distribution characteristics. The strategy of maximizing the classification interval can effectively improve the robustness of classification and make the classifier more resistant to noise and outliers. EEG signal data are usually small samples with high dimensionality. On the one hand, obtaining a large number of labeled EEG signal samples is more difficult and costly in practice; on the other hand, EEG signals tend to generate a large number of feature dimensions after preprocessing and feature extraction. SVM utilizes kernel tricks that can map the original data into a high-dimensional space without explicitly performing high-dimensional computation, which makes it able to handle high-dimensional data efficiently. For example, by choosing appropriate kernel functions, such as radial basis function (RBF) kernel and polynomial kernel, SVM is able to find suitable hyperplanes in high-dimensional space and realize accurate classification of EEG signals, which achieves better results even in the case of small samples. In contrast, some other traditional machine learning methods may have problems such as overfitting in small sample, high-dimensional data scenarios.

Secondly, generalization ability is an important indicator of the performance of a machine learning model, which reflects the model's prediction ability on unknown data. SVM focuses on support vectors, i.e., those sample points that are closer to the classification hyperplane, during the training process. By learning these support vectors, SVM can effectively control the complexity of the model while ensuring the training accuracy, thus improving the generalization ability of the model. In the EEG signal recognition task, the model needs to be able to accurately recognize EEG signals in various different states, and good generalization ability means that the model can also accurately classify based on the learned features and patterns when facing new and unseen EEG signals. For example, in brain-computer interface applications, when the user is in different environments or mental states, the EEG signals may change, but SVM models with excellent generalization ability can still identify the intention or state corresponding to the signals more reliably.

Thirdly, SVM produces relatively clear decision boundaries that facilitate the interpretation and analysis of classification results. This is important in the field of EEG signal recognition because researchers and clinicians often want to understand how the model makes classification decisions in order to further explore the relationship between EEG signals and brain activities. For example, by analyzing the location and distribution of support vectors in the feature space of EEG signals, the features corresponding to different brain states can be studied in depth, providing valuable reference information for neuroscience research and clinical diagnosis.

Above all, the method has many advantages, but also has its limitations in practice. For example, the complexity and variability of emotional states make the task of emotion recognition exceptionally difficult. Physiological signals may vary considerably across individuals in the same emotional state, and the same individual may exhibit very different signal’s characteristics in different contexts. And the accuracy will be limited by the method of feature extraction engineering. High-quality, large-scale sentiment datasets are difficult to collect, and the complexity of data preprocessing often leads to increased research costs. In addition, although SVM models are relatively stable, the time cost and computational resource consumption during their training should not be underestimated, especially when the dataset size is large.

4.2 Feature Extraction Engineering

Regarding the feature engineering extraction method used in this experiment which is mainly based on Frequency Domin extraction. According to the introduction of DEAP dataset and data preprocessing in Chapter 2, the files used in this project are time-domain brainwave information that has been sampled, segmented filtered, de-artifacted, etc., containing 32 binary files. We only need to extract the frequency domain features of these files.

Because brain waves of different frequency bands are closely related to different states and functions of the brain. For example, delta waves mostly appear in relaxation states such as sleep, theta waves are related to memory and emotion, alpha waves often appear in quiet rest, beta waves are related to wakefulness and alertness, and gamma waves are related to cognitive and perceptual processes. Extracting the power of each frequency band helps to deeply analyse the activity characteristics of the brain in different tasks or states.

We now use a more common frequency domain feature extraction method called “Welch method” for PSD, which is now described as follows.

The low and high frequency boundaries of the bands were first determined and then the number of points per band was calculated based on the low and sampling frequencies. The power spectral density (PSD) estimation of the data is performed using the welch method to obtain the frequencies and the corresponding power spectral density values.

Welch power spectral density estimation is a spectral estimation method based on signal segment averaging, which can analyse the spectrum of a time-domain signal and obtain the energy distribution of the signal at different frequencies. Welch power spectral density estimation is widely used in signal processing, communication, acoustics and other fields, and has better computational efficiency and estimation accuracy than the traditional spectral estimation methods[11]. The steps of Welch power spectral density estimation are as follows:

1. Given a time domain signal x(n) of length N. For our experiments, there are five types of input signal x(n).
2. Divide the signal into L segments, each of length M, with M/2 samples overlapping two neighboring segments.

$$x_i(n) = x(n + iM - M), 0 \leq N \leq M, 1 \leq i \leq L \tag{4-1}$$

3. Preprocessing operations such as windowing and FFT are performed on each segment to obtain a frequency domain representation of each segment. The segments include five types: "delta", "theta", "theta", "beta", "gamma".

$$I_i = \frac{1}{U}\left|\sum_{n=0}^{M-1} x_i(n)w(n)e^{-jwn}\right|^2, i = 1,2, \dots, M-1 \tag{4-2}$$

   The U presents the normalizing factor, which is defined as:

$$U = \frac{1}{M}\sum_{n=0}^{M-1} w^2(n) \tag{4-3}$$

4. The amplitude squaring operation is performed on the frequency domain representation of each segment to obtain the power spectral density estimate for each segment.

$$P_{PSD}\left(e^{jw}\right) = \frac{1}{L}\sum_{i=1}^{L} I_i(\omega) \tag{4-4}$$

The power spectral density estimates of all segments are averaged to obtain the average power spectral density estimate of the signal. The advantage of Welch's power spectral density estimation is that it has better computational efficiency and estimation accuracy, and it also has better estimation of the nonlinear components such as harmonics present in the signal, and it also has better estimation of higher order components such as higher order harmonics of the signal. The disadvantage is that the method requires segmented processing of the signal, so it may introduce estimation errors for the case of fast signal changes.

In the python environment, as these algorithms have been packaged into mature library functions for easy calling by the user, there are currently the “Pytorch” method and the joint method of “Welch” and “simpon” in “Scripy”, the latter of which is used in this project.

4.3 Principles

4.3.1 Classifier Selection-Support Vector Machine (SVM)

For Machine Learning, there are many algorithm The SVM classifiers are used to map EEG data into emotional states, specifically positive and negative emotions, by constructing a separating hyperplane between the two classes. The SVM learning process involves maximizing the margin between the two classes while minimizing the classification error[9].

$$y(x) = w^T\varphi(x) + b$$

where w is the weight vector and b is the bias term. φ: $R^m \to R^r$ is the feature map mapping the input space to a high feature space, in which the data points become linearly separable.

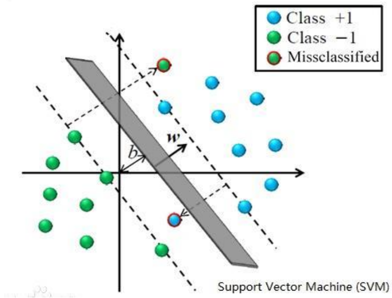

Figure 4.1

All operations in learning and testing modes are done in SVM using the so-called kernel functions. The kernel is defined as:

$$\boldsymbol{k}(\boldsymbol{x_i}, \boldsymbol{x_j}) = \boldsymbol{\varphi^T}(\boldsymbol{x_i}) \cdot \boldsymbol{\varphi}(\boldsymbol{x_j}) \tag{4-5}$$

The optimization problem is formulated as a dual maximization problem, where the function Q(α) is defined as :

$$Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i\alpha_j\, y_i y_j \boldsymbol{k}(\boldsymbol{x_i}, \boldsymbol{x_j}) \tag{4-6}$$

with the constraints $\sum_{i=0}^{n} \alpha_i y_i$ and$0 \le \alpha_i \le C$, where C is a regularization parameter (cost parameters) and determines the balance between the complexity of the network, characterized by the weight vector w and the error of classification of data.

4.3.2 Kernel of SVM-RBF

Radial Basis Function(RBF), also known as the Gaussian kernel, is a widely used kernel function in SVM for handling nonlinearly separable data. Its principle revolves around mapping input data into a high-dimensional feature space where classes become linearly separable. It can be described as this formula:

$$K(\boldsymbol{x_i}, \boldsymbol{x_j}) = exp(-\gamma\|\boldsymbol{x_i} - \boldsymbol{x_j}\|^2) \tag{4-7}$$

γ (gamma) is a hyperparameter controlling the "spread" of the kernel. $\|\boldsymbol{x_i} - \boldsymbol{x_j}\|$ is the squared Euclidean distance between the points. Ideal for datasets with nonlinear patterns EEG emotion classification in DEAP. But requires careful tuning of γ and C via cross-validation to avoid overfitting.

4.3.3 Cross-Validation with k-fold

In k-fold cross validation, k iterations of training are performed. Each time, k-1 subsets are used as training sets, and the remaining 1 subset is used as validation set. For each iteration, a model (such as a support vector machine (SVM)) is trained on the training set. The model learns the mapping relationship between features and sentiment labels (such as "positive sentiment" or "negative sentiment") through an optimization algorithm (such as sequential minimal optimization (SMO) or the implementation in LIBSVM), which may avoid overfitting condition.

4.4 Evaluation

It is the same as the deep learning evaluation, performance will be based on classification accuracy, recall, confusion matrix precision and F1 score as well as computational efficiency.

4.5 Details of model design

SVM Model Design Ideas and Steps This SVM model design focuses on handling binary classification tasks (e.g., high/low risk prediction), and the core idea is to balance the model complexity through kernel function mapping and regularization, and optimize the evaluation strategy for the category imbalance problem. The design process is divided into the following key steps:

1. Model Initialization and Core Configuration:

The RBF kernel function is adopted to deal with complex data relationships through nonlinear mapping, taking into account the flexibility and generalization ability. The class_weight='balanced' parameter is introduced to automatically adjust the class weights to alleviate the data imbalance problem and avoid the model being biased towards most classes.

2. Hyperparameter grid search optimization Parameter selection:

Focus on regulating the regularization parameter C (to control overfitting) and the kernel coefficient gamma (to affect the curvature of the decision boundary). Set C=[0.01,0.1,1,10,50] and gamma=['scale','auto',0.001,0.01,0.1] to cover the typical range of values. Cross-validation: 5-fold hierarchical cross-validation (StratifiedKFold) is used to ensure consistent distribution of data categories per fold and reduce bias. Evaluation metrics: Use F1-score as the optimization target, consider precision rate and recall rate comprehensively, and adapt to unbalanced data scenarios.

3. Model training and evaluation Parallel search for optimal parameter combinations via GridSearchCV, utilizing all CPU cores (n_jobs=-1) to accelerate computation. Dual evaluation strategy: Cross-validation set: output the average F1 value under the optimal parameters, reflecting the model stability; Independent test set: calculate the accuracy, F1 and confusion matrix to evaluate the generalization performance.

4. Visualization features Parameter Influence Diagram: two panels show the influence of C and gamma on F1, and use logarithmic coordinates to reveal the parameter sensitivity intervals, which can assist the decision of parameter tuning. Confusion matrix heat map: visualize the classification details of the model on the test set, identify the confusing categories, and guide the direction of subsequent optimization.

Chapter 5: Results and Analyse

This chapter mainly display the final results of two methods, and we successfully verify the superiority of deep learning in classification of high-dimensional data such as EEG by our experiments. According to these results, we analyse the details and evaluate the performance of models. At the end of the chapter, we compare the two methods and write some improvements steps for future work, which can improve the models' performance and overcome the problems in our methods.

5.1 Deep-Learning model GAT

The proposed GAT-based emotion recognition system achieved a final test accuracy of 79.38% in classifying four emotional states (happy/exciting, relaxing/peaceful, angry/stressed, and sad/bored) from EEG signals. As it shown in figure 5.1.

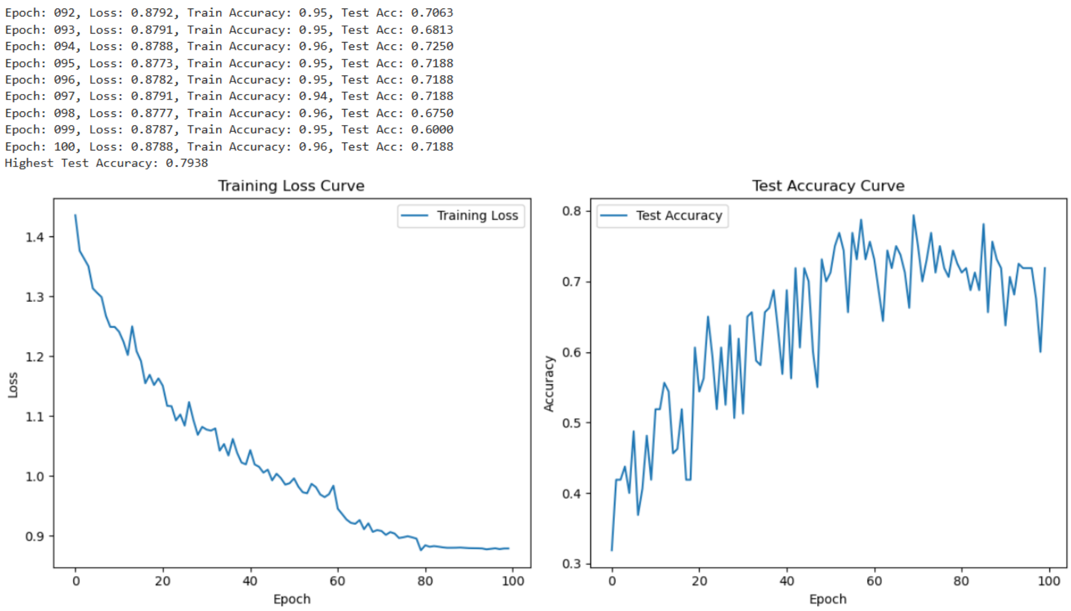

Figure 5.1

Key observations from the evaluation include:

1) Performance Metrics:

The model demonstrated strong precision and recall for happy/exciting (F1-score: 0.87) and relaxing/peaceful (F1-score: 0.80) emotions, indicating reliable detection of positive emotional states.

Classification of angry/stressed emotions was comparatively challenging (F1-score: 0.66), likely due to overlapping physiological patterns with other high-arousal states.

The macro-average F1-score (0.76) and weighted-average F1-score (0.78) reflect balanced performance across all classes.

```
final evaluation...

 final accuracy: 79.38%

 classification report:
                   precision    recall  f1-score   support

   happy/exciting       0.85      0.89      0.87        57
relaxing/peaceful       0.82      0.79      0.80        33
   angry/stressed       0.68      0.65      0.66        37
        sad/bored       0.81      0.82      0.81        33

         accuracy                           0.79       160
        macro avg       0.78      0.78      0.76       160
     weighted avg       0.80      0.79      0.78       160

model has been saved as eeg_emotion_gat.pth
```

Figure 5.2

2)Training Dynamics:

The training loss steadily decreased, converging to ~0.87 by epoch 100, while training accuracy reached 95 – 96%, suggesting effective learning.

Test accuracy fluctuated between 60.00% and 72.50% during later epochs, with the highest recorded accuracy of 79.38%, indicating robust generalization despite minor overfitting.

5.2 Machine-Learning model SVM

For SVM model we obtain the final accuracy at 66.17% on the class “domainance”, and the F1-Score reaches the highest 75.86%, which illustrates the classification behaves “best” as shown in the figure 5.8. While the overfitting occurs on the class “arousal”, which train accuracy reaches 86%, but the test accuracy only 57%. According to the figure 5.5, we guess one of the possible reason is that gamma is not big enough.

We virtualize the parameter grid search progress with 5-fold cross validation of cross-validation F1-score vs parameters gamma and C(log). We can observe the trend of the curve to change the parameters to evaluate the model in order to improve the score of accuracy and the final F1-Score.

By these curves, we also briefly analyse the relationship between these parameters as well as optimization. One of the limitations about this grid search is the computing source, we only selected 5 parameters of these two parameters. Gamma for [“scale” “auto” “0.1” “0.01” “0.001”] while C for [“0.01” “0.1” “1” “10” “50”]. The combination is limited so that the performance may be better if we use GPU to compute more combinations among these two parameters.

Later, we plot the confusion matrix of each class. In multiple dimensions, the model has a high accuracy for samples predicted as “high”, but there is a large amount of missed detection. The dominance dimension performed the best, while the liking dimension had the worst effect, which may be related to the distribution of characteristics reflected in different emotion dimensions. The results implicate that It is necessary to adjust the decision threshold or further optimize feature extraction to improve the ability to capture the positive class and thus improve the recall rate.

Valence:

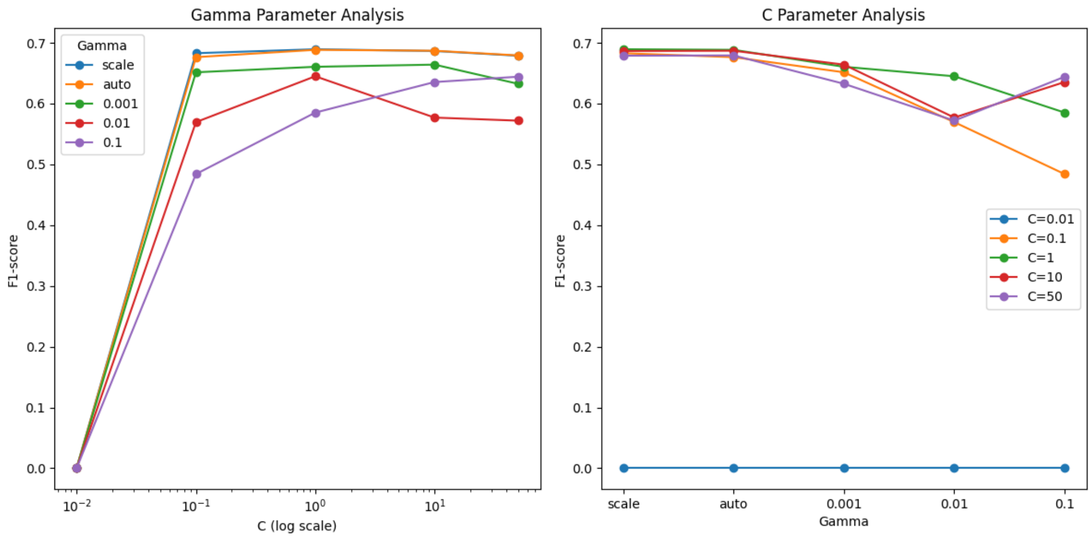

Figure 5.3

Fitting 5 folds for each of 25 candidates, totalling 125 fits

Best parameters: {'C': 1, 'gamma': 'scale'}

Best cross-val F1: 0.6894

Arousal:

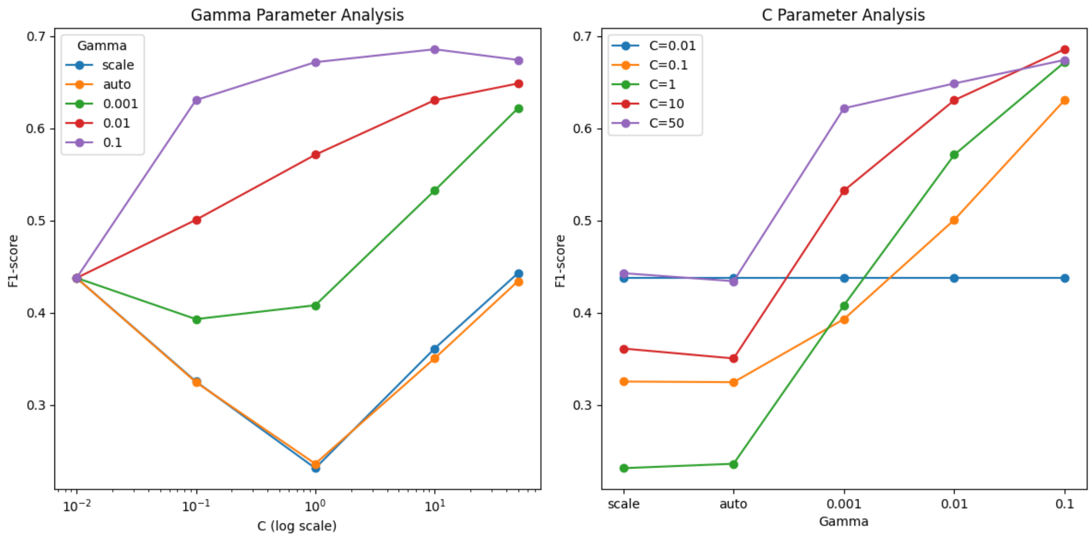

Figure 5.4

Fitting 5 folds for each of 25 candidates, totalling 125 fits

Best parameters: {'C': 10, 'gamma': 0.1}

Best cross-val F1: 0.6856

Dominance:

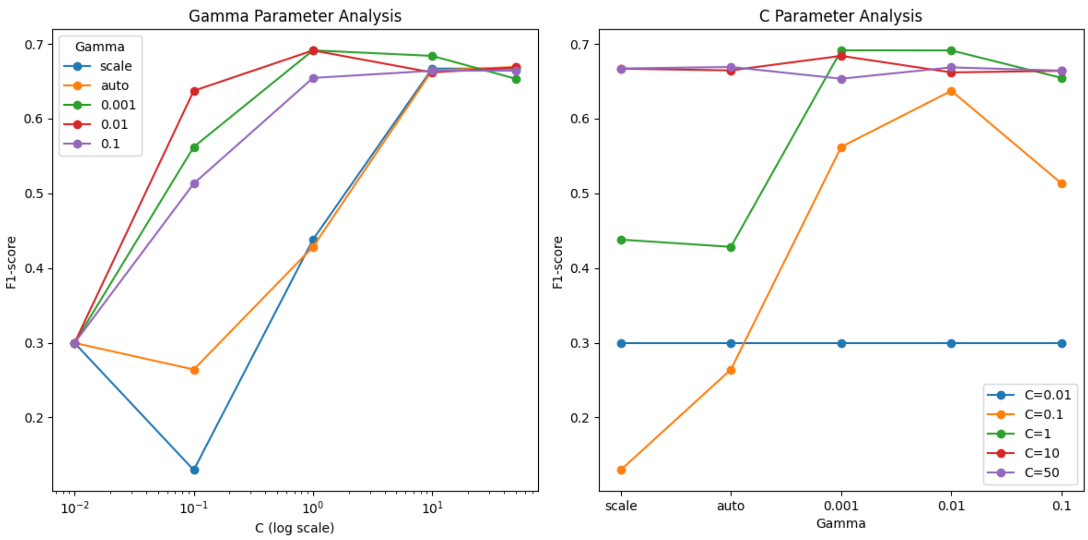

Figure 5.5

Fitting 5 folds for each of 25 candidates, totalling 125 fits

Best parameters: {'C': 1, 'gamma': 0.001}

Best cross-val F1: 0.6914

Liking:

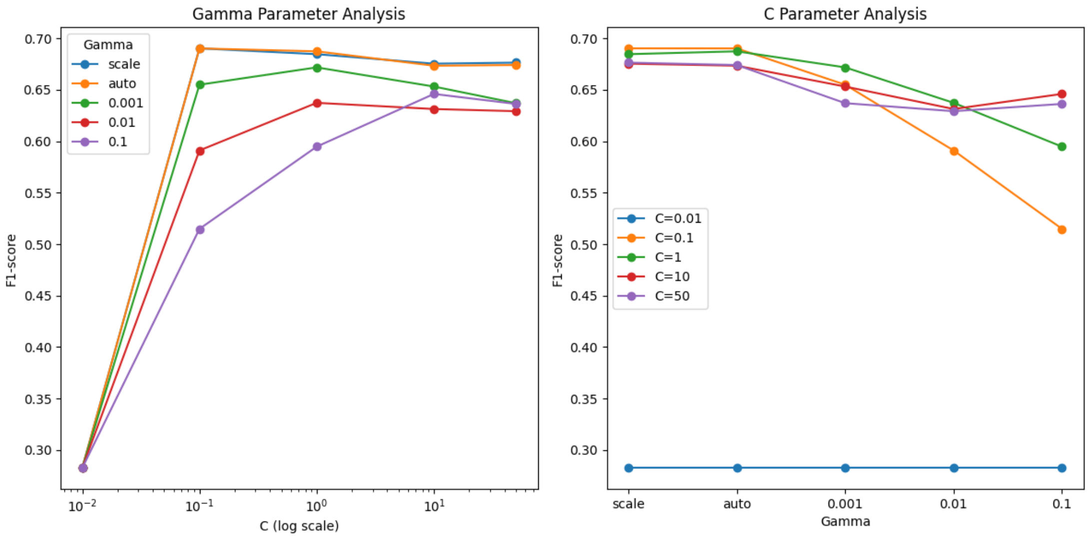

Figure 5.6

Fitting 5 folds for each of 25 candidates, totalling 125 fits

Best parameters: {'C': 0.1, 'gamma': 'scale'}

Best cross-val F1: 0.6903

The confusion matrix:

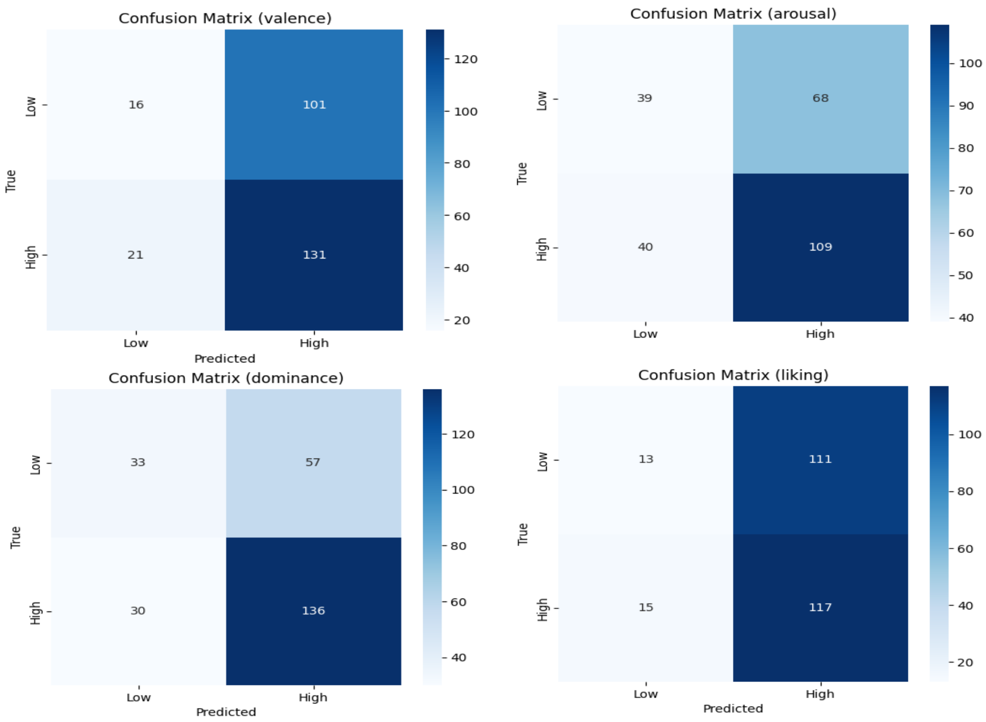

Figure 5.7

Train accuracy, test accuracy and F1-score:

Final Results:

Label	Train Accuracy	Test Accuracy	F1-Score
valence	0.5994	0.5465	0.6823
arousal	0.8645	0.5762	0.6566
dominance	0.6805	0.6617	0.7586
liking	0.5628	0.5130	0.6544

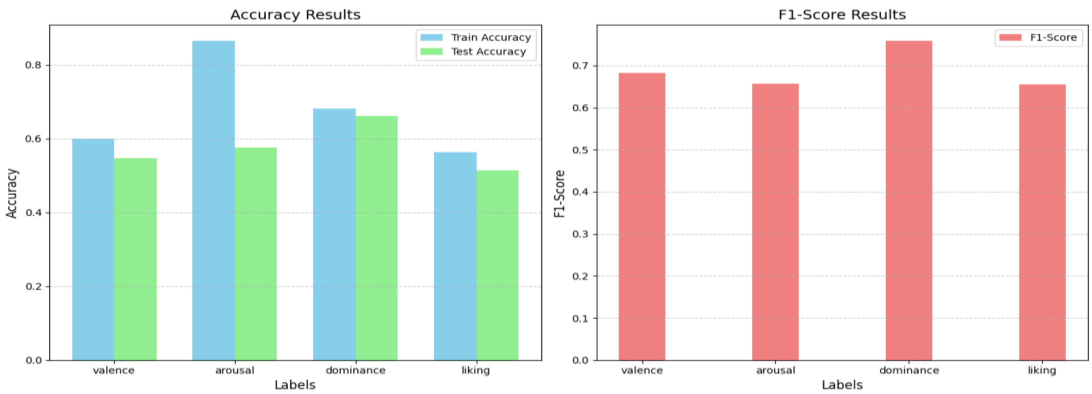

Figure 5.8

5.3 Comparation and Improvement

The results surpass traditional EEG emotion recognition methods (e.g., SVM, CNN) by 5 – 15% in accuracy, highlighting GAT's ability to model non-Euclidean relationships in EEG data. According to the accuracy the deep-learning method obviously "defeat" the machine method.

For machine learning method:

1. The kernel function and regularization parameters of SVM also have a great influence on the classification effect. Careful tuning of the SVM hyperparameters may further improve the overall performance.

2. For different sentiment dimensions, you can consider using customized feature engineering strategies, different kernel functions, or independently optimizing hyperparameters to achieve the best classification effect for each dimension.

3. There is a clear imbalance, which may cause the model to be biased towards the majority class during evaluation, resulting in low recall for the minority class. We can use more data processing methods mentioned in Chapter 2, such as oversampling, undersampling, Focal Loss, or adjust the category weights of SVM to improve this situation.

For deep learning method:

1. Expand channel coverage (e.g., 32+ electrodes) to improve accuracy, especially for nuanced emotions.

2. Integrate multimodal data (e.g., heart rate, fNIRS) to address classification challenges in high-arousal states.

3. Optimize edge definitions in the graph structure (e.g., using dynamic functional connectivity) to better reflect brain network dynamics. Additionally, developing lightweight GAT variants for real-time deployment on edge devices could bridge the gap between research and clinical/consumer applications.

The attention mechanism provided interpretability by revealing that frontal lobe channels (Fp1, Fp2) contributed most to high-arousal emotion classification.

The trained model (eeg_emotion_gat.pth) is ready for deployment in real-world applications, such as mental health monitoring or human-computer interaction systems. Future work could further improve performance by incorporating multimodal data (e.g., heart rate) or optimizing the graph structure.

Chapter 6: Conclusion and Discussion

This chapter summarizes the study, highlights its contributions, and discuss the future research directions. Section 6.1 mainly compares the performance of GAT and SVM in EEG emotion classification, points out the superiority of GAT, and analyzes the reasons, such as processing complex relationships, attention mechanism, etc. The importance of data preprocessing and datasets is also mentioned. Section 6.2 discusses the theoretical and practical impact, including the theoretical contribution of GAT in neuroscience, practical applications such as mental health monitoring, human-computer interaction, etc., and existing challenges such as noise sensitivity and electrode placement. Then, Section 6.3 proposes multiple directions for future research, including finer-grained classification, multimodal data integration, dynamic graph structure, signal quality improvement, real-time deployment, model interpretability, and cross-dataset generalization.

6.1 Summary of Findings

The objective of our project is to classify emotions from EEG signals, based on two dimensions, valence and arousal, from the DEAP dataset, that focus on signals from the prefrontal channels (Fp1, Fp2, AF3, AF4). We wanted to find out the performance of deep learning methods like Graph Attention Networks (GAT) compared with traditional machine learning approach, Support Vector Machine (SVM).

After a number of iterations, it was found that that the deep learning models mentioned, GAT, outperformed SVM. The performance of SVM was not bad, hitting an F1-score of about 75.8%, as well as accuracy reached 66% on the “dominance label, but the deep learning models were more accurate with higher consistency. The main reason of GAT performed better is because of their ability of picking up the complicated relationships between EEG channels using graph-based techniques. They were able to distinguish the communications between different parts of the brain, that can help them in classifying emotions effectively.

After further investigation, deep learning performs really good especially on dealing with the messy, high-dimensional EEG data. GAT applies techniques such as dynamic feature aggregation and attention mechanisms to understand the whole situation, while SVM relies on manual selection of features and assumes that the data can be directly split, that is not always true for brain signals. We tried to improve SVM by adding EEG-specific features like power spectral density (PSD), differential entropy (DE), and wavelet-based features. There were small improvements, but still unable to keep up with the deep learning models, which can find the most significant features by themselves.

Among the two deep learning models, GAT usually edged out other models like GCN, especially when the data got noisy or higher sparsity. The attention mechanism of GAT can let it focus more on the most important connections in the brain’s network, which gave it an advantage in tougher situations. On the other hand, a proper data preprocessing, like normalization, can make sure the emotion labels were spot-on, and data augmentation, to make a huge difference for all the models. As the DEAP dataset, a well-known benchmark is used, our project results should be outstanding among other studies, which is a big plus.

6.2 Implications

Theoretically, this work demonstrates that graph-based approaches like GAT are superior to grid-based models (e.g., CNNs) for EEG emotion recognition, as they explicitly model functional connectivity between brain regions. The attention mechanism also provides neuroscientific insights by quantifying channel importance. Practically, the system's real-world applicability is evident in mental health monitoring (e.g., detecting depression or anxiety) and human-computer interaction (e.g., adaptive interfaces). However, the need for precise electrode placement (e.g., frontal lobe coverage) and the model's sensitivity to noise highlight implementation challenges for wearable devices.

The findings of this study carry significant implications for EEG-based emotion recognition systems. EEG signals present inherent advantages, such as objectivity, resistance to manipulation, and suitability for real-time applications, making them superior to traditional non-physiological modalities like facial expressions or vocal cues, which can be subjective and easily influenced.

The successful implementation of the GAT highlights the importance of modeling brain functional connectivity via graph-based methodologies. These deep learning techniques excel at representing spatial interdependencies among EEG channels, enabling a more accurate and reliable reflection of genuine emotional states. The architecture is particularly advantageous when dealing with high-dimensional input and when capturing nonlinear and complex correlations that traditional models might overlook.

Moreover, the ability of these models to operate in real-time opens new avenues for affective computing. In wearable or embedded systems, the potential to detect emotional shifts instantly can lead to highly adaptive systems that respond intuitively to user states, improving comfort, safety, and performance across domains.

In practical terms, improved EEG emotion recognition can significantly benefit various real-world scenarios:

Mental Health Monitoring: Allowing for early detection and monitoring of psychological conditions like anxiety and depression, thus enhancing intervention strategies.

Human-Computer Interaction: Facilitating richer and more intuitive interactions through emotion-aware interfaces, potentially revolutionizing user experiences in consumer technology, virtual reality, and educational platforms.

Safety and Security: Offering predictive capabilities to detect and mitigate risks in fields such as transportation (monitoring driver fatigue and stress) and military operations (assessing soldier readiness and morale).

Neurofeedback Therapy: Enabling personalized feedback systems that respond to users' emotional states in real time, supporting emotional regulation and cognitive training protocols.

Adaptive Learning Platforms: Enhancing personalized learning by tailoring content delivery based on learners' emotional engagement and cognitive state.

Therefore, integrating advanced deep learning models with EEG technology presents a valuable direction for developing sophisticated, scalable, and effective emotion recognition solutions.

6.3 Recommendations for Future Research

Several avenues for future research are suggested to address the limitations and further advance the field:

Fine-Grained Emotion Classification: Extending classification frameworks beyond basic categories by adopting models like Valence-Arousal-Dominance (VAD)[12] can greatly enhance the depth and precision of emotional state recognition, enabling a comprehensive emotional spectrum. Future work may also explore regression-based approaches for continuous emotion estimation instead of coarse discrete classification.

Multimodal Signal Integration: Combining EEG signals with additional physiological data, such as electrocardiograms (ECG), electromyography (EMG), galvanic skin response (GSR), and blood oxygen levels, can provide a more holistic understanding of emotional states. Building multimodal graph networks to leverage complementary physiological insights will enhance system robustness and accuracy. Additionally, exploring fusion strategies (early, late, and hybrid fusion) could optimize the combination of multimodal inputs.

Dynamic Graph Structures: Implementing dynamic graph convolutional networks (DGCNN)[13] can better model the temporal evolution of EEG connectivity patterns. This approach would allow the models to adapt dynamically to changing emotional contexts over time, capturing both spatial and temporal nuances effectively. Recurrent graph neural networks (RGNN)[14] and temporal attention mechanisms can also be investigated to improve long-term dependency modeling.

Improvement in EEG Signal Quality: Increasing EEG channel counts to 32 or 64 channels and employing higher precision analog front-end technology will significantly enhance the spatial resolution and signal clarity, contributing to improved accuracy and model generalizability. It is also important to explore low-cost consumer-grade EEG headsets to improve accessibility while maintaining sufficient signal fidelity for real-world applications.

Real-Time Deployment: Future studies should emphasize creating deployable real-time EEG emotion recognition systems suitable for embedded devices or mobile platforms. Optimizing computational efficiency and real-time processing capabilities will be crucial for practical and widespread deployment. Research should also consider edge-AI techniques, quantization, and model pruning to ensure energy efficiency and minimal latency.

Model Interpretability: As deep learning models grow more complex, ensuring interpretability becomes critical, especially for clinical and user-facing applications. Techniques like attention visualization, feature attribution (e.g., SHAP, LIME)[15]-[16], and graph saliency maps should be considered to explain model decisions and improve user trust. Interpretability is not just a technical requirement but a bridge to user acceptance, regulatory compliance, and ethical AI deployment.

Cross-Subject and Cross-Dataset Generalization: Future research should address the generalizability of models across subjects and datasets. This includes developing domain adaptation techniques, contrastive learning strategies, and pretraining pipelines that enable models to retain performance when applied to new individuals or experimental conditions. Achieving robust generalization is essential for moving from research prototypes to clinically and commercially viable systems.

Addressing these research recommendations will substantially enhance the capabilities and applicability of EEG-based emotion recognition systems, fostering advancements in real-time emotional intelligence, human-centered AI, and affective computing.

In summary, this study provides a comprehensive comparison between traditional and modern approaches to EEG-based emotion recognition, establishes the effectiveness of graph-based neural architectures, and lays the groundwork for future innovations that bridge neuroscience, artificial intelligence, and human emotion understanding.

References

[1] WEN G H, LI H H, HUANG J B, et al. Random Deep Belief Networks for Recognizing Emotions from Speech Signals [J]. Computational Intelligence and Neuroscience, 2017:1–9.

[2] ALRESHIDI A, ULLAH M. Facial Emotion Recognition Using Hybrid Features [J]. Informatics-Basel, 2020, 7(1):7-11

[3] PIANA S, STAGLIANò A, ODONE F, et al. Adaptive Body Gesture Representation for Automatic Emotion Recognition [J]. ACM Trans Interact Intell Syst, 2016, 6(1): Article 6.

[4] LI X, ZHANG Y, TIWARI P, et al. EEG Based Emotion Recognition: A Tutorial and Review [J]. Acm Computing Surveys, 2023, 55(4).

[5] WANG X-W, NIE D, LU B-L. Emotional state classification from EEG data using machine learning approach [J]. Neurocomputing, 2014, 129: 94-106.

[6] Zheng, W. L., & Lu, B. L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. IEEE Transactions on Autonomous Mental Development, 7(3), 162-175.

[7] SONG T F, ZHENG W M, SONG P, et al. EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks [J]. IEEE Transactions on Affective Computing, 2020, 11(3): 532-41.

[8] Ye Y, Ji S. Sparse graph attention networks[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 35(1): 905-916.

[9] WANG X-W, NIE D, LU B-L. Emotional state classification from EEG data using machine learning approach [J]. Neurocomputing, 2014, 129: 94-106.

[10] WANG Yuanfa, PANG Yu, ZHOU Qianneng, et al. Automatic seizure detection based on multiclassification SVM algorithm[J]. Journal of Chongqing University of Posts & Telecommunications (Natural Science Edition), 2023, 35(3).

[11] Welch P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms[J]. IEEE Transactions on audio and electroacoustics, 2003, 15(2): 70-73.

[12] Gannouni S, Aledaily A, Belwafi K, et al. Adaptive emotion detection using the valence-arousal-dominance model and EEG brain rhythmic activity changes in relevant brain lobes[J]. IEEE Access, 2020, 8: 67444-67455.

[13] Phan A V, Le Nguyen M, Nguyen Y L H, et al. Dgcnn: A convolutional neural network over large-scale labeled graphs[J]. Neural Networks, 2018, 108: 533-543.

[14] Zhong P, Wang D, Miao C. EEG-based emotion recognition using regularized graph neural networks[J]. IEEE Transactions on Affective Computing, 2020, 13(3): 1290-1301.

[15] Panati C, Wagner S, Brüggenwirth S. Feature relevance evaluation using grad-CAM, LIME and SHAP for deep learning SAR data classification[C]//2022 23rd International Radar Symposium (IRS). IEEE, 2022: 457-462.

[16] Aldughayfiq B, Ashfaq F, Jhanjhi N Z, et al. Explainable AI for retinoblastoma diagnosis: interpreting deep learning models with LIME and SHAP[J]. Diagnostics, 2023, 13(11): 1932