International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(23s) 2026 pp. 853-863
https://doi.org/10.70917/ijcisim-2026-5649
Article

EEG Emotion Recognition Using a Cross-Domain Transformer with Adaptive Graph Attention (CDT- AGA)

Narendra R[1], Suresha M[1], Rabiya Basree[3], Vinay K[2]

[1] Department of Computer Science, Kuvempu University, Shimoga, India. Email: naren44.mys@gmail.com
ORCID: 0000-0001-8120-0984
[1] Department of Computer Science, Kuvempu University, Shimoga, India. Email: srit_suresh@yahoo.com
ORCID: 0000-0003-0668-926X
[3] Department of Computer Science, Kuvempu University, Shimoga, India. Email: rabiya9901@gmail.com
ORCID: 0009-0000-9732-7453
[2] Department of Computer Science, Christ University, Bangalore, India. Email: vinaycbs@gmail.com
ORCID: 0009-0005-2187-2662

Abstract: EEG signals have gained prominence for emotion recognition in affective computing and Brain-Computer Interfaces (BCIs), as well as mental health analysis. Nonetheless, many of the existing deep learning techniques face issues of cross-subject variability and fail to exploit the spatial correlations of EEG electrodes. In order to fix this issue, introduced novel Cross Domain Transformer with Adaptive Graph Attention (CDT-AGA) framework which takes into consideration domain adaptive and graph. In particular, domain adaptation module uses adversarial learning to align different subjects' features. Proposed an adaptive graph attention module employing a new cross-domain multi-head graph attention mechanism which adaptively learns the spatial dependency with respect to EEG electrodes under various emotional states. Study assessed model's performance on six prevalent datasets of EEG emotion recognition namely SEED, AMIGOS, DREAMER, DEAP, MAHNOB-HCI and Kannada Musical clips (MindData) custom built in dataset, proposed model achieved state of the art results on these datasets. In particular, proposed method achieves improvement in classification accuracy of 3.3% over existing model on above datasets. Besides improved accuracy, model also exhibited enhanced interpretability by reliably identifying target emotion-relevant brain regions. Further, study also proves that model exhibits more robustness against noise and requires less training time than existing one.
Keywords: EEG, Emotion, SEED, AMIGOS, DREAMER, MindData

1. Introduction

Recent research studies have focused on recognizing human emotions from EEG signals. Emotion plays an important role in Human-Computer Interaction (HCI). Existing works shows a number of gaps to start with laboratory settings achieve promising results, but limited works deal with real-world problems, where noise, artifacts, and intersubject variability can reduces performance. Moreover, the majority of existing works treat all the EEG channels as independent features. However, the spatial relationships between the electrodes consist of human brain's functional connectivity.

In addition, emotions are used to diagnose mental disorders. Moreover, human emotions consist of anger, fear, happiness, sad etc. The affect is essentially the reflection of a feeling. It's also a mindset and a state of mood. Using one's emotion also helps to communicate. A human being gets happy on receiving a gift and sometimes, gifts can make them "sorry". Additionally, the limbic system and cerebral cortex are involved. The cerebrum constitutes the largest part of the brain and is a highly complex structure. It contains many parts each with a function. Likewise, the physiological activities include electrophysiology. Which also studies electrical activity of human cells Zheng and Lu

(2015). A drawback is the lack of longitudinal studies investigating the generalization of EEG-based emotion recognition and limited works involving multimodal approaches, which include the integration of EEG and other physiological or behavioral signals Craik et al. (2019). Song et al. (2020) employed Approximate Entropy (ApEn) features combined with a Support Vector Machine (SVM) classifier for EEG-based emotion recognition. The method was evaluated on the SEED dataset using 62-channel EEG signals to classify positive, neutral, and negative emotional states.

Chen et al. (2022) utilized a Support Vector Machine (SVM) for EEG-based emotion recognition on the DEAP dataset. Using 32-channel EEG signals, the study classified arousal and valence by incorporating duration and coverage as additional discriminative features. Wei et al. (2024) proposed an Adaptive EEG Microstate Analysis framework that integrates temporal and spatial EEG information to improve emotion recognition. The method was evaluated on the DEAP dataset using 32-channel EEG signals for arousal and valence classification.

One of the main challenges in emotion recognition via EEG signals is to have excellent subject-invariant performance. In recent times, cross-domain recognition has been improved with domain adaptation methods. Nevertheless, these approaches have the following two drawbacks. First, depend on class labels of the source data, which can be limited. Second, neglect the spatial relationship of EEG electrodes that can be useful for emotion recognition. This study introduces a novel Cross-Domain Transformer with Adaptive Graph Attention (CDT-AGA) model to overcome these issues.

In particular, first original method provides a domain class aligned loss to align the source objective distribution and target feature distribution without using source labels. Next, designed a new adaptive graph attention mechanism to capture the spatial dependencies of EEG electrodes and reveal the trade-off between frequency bands and brain regions. Ultimately, created a transformer-based emotion encoder to pull EEG distinguishing emotion data and performed a comprehensive assessment utilizing benchmark dataset SEED, AMIGOS, DREAMER, DEAP, MAHNOB-HCI, Kannada Music Clips and to validate the proposed.

A unified CDT-AGA architecture is proposed to jointly learn dynamic spatial relationships among EEG electrodes and temporal dependencies, while enhancing subject-independent emotion recognition by reducing cross-subject domain discrepancies through adversarial domain adaptation.

Contribution of this study as follows

1. Development of the CDT-AGA architecture comprising a feature projection layer, an Adaptive Graph Attention module, a Cross-Domain Transformer, and an adversarial domain adaptation branch.
2. Automatic learning of subject-invariant EEG representations, thereby mitigating inter-subject variability and improving generalization.
3. Comprehensive evaluation on multiple benchmark datasets (SEED, AMIGOS, DREAMER, DEAP, and MAHNOB-HCI, Kannada Musical clips (MindData) demonstrating consistent performance gains over existing Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and transformer-based approaches.
4. Validation of the proposed framework's robustness across different emotional paradigms, recording conditions, and EEG acquisition systems, highlighting its suitability for practical EEG-based emotion recognition applications.

The remaining of this paper is organized as follows. Section 2 presents a review of the related literature on EEG-based emotion recognition. Section 3 describes the proposed CDT-AGA methodology in detail. Section 4 outlines the experimental setup, including the datasets, preprocessing procedures, and evaluation metrics. Section 5 discusses the experimental results and provides a comparative analysis with existing approaches. Finally, Section 6 concludes the paper and suggests directions for future research.

2. Related Work

Over the last two decades, various EEG-based emotion recognition methods have evolved due to developments in signal processing methods, machine learning algorithms, and affective neuroscience.

The study focuses towards more sophisticated computational techniques to overcome that limitation, in particular deep learning techniques. Automated discovery of spatio-temporal patterning within EEG data was enhanced by the use of CNN and RNN. Nonetheless, important challenges still needed to be resolved. The present study addresses inter-subject variability, an important issue that was not considered in many studies published during the previous decade. Variability which is a big limitation drawing from demographic factors of the subjects used age, sex and cultural background etc. with not proper care robust generalization is impossible. Various solutions were proposed subsequently, which attempted to use multimodal data (EEG + Facial Expressions, EEG + Physiological signals, etc.). This will enhance the contextual interpretation of EEG signals and build better mental state recognition systems. However, this paradigm introduced the emergence of several challenges. Ganin and Lempitsky (2015) proposed a novel unsupervised domain adaptation framework based on Domain-Adversarial Neural Networks (DANN) to learn domain-invariant feature representations. Thus, neural patterns can be used by computational learning systems and methods for overview occurrence of emotions. As research methods continue to advance, significant theoretical developments have emerged in the interpretation of EEG signals for emotion recognition. Affective neuroscience states that any sign of activity that supports the elicitation, feeling, and expression of an emotion can be objectively measured by neurophysiological markers. According to evidence obtained, there is a constant relationship between EEG extracted features for emotion recognition and emotional eliciting stimuli Khosrowabadi et al. (2010).

The initial studies established the foundation for comprehending the connection between EEG rhythms and emotion, showing that specific emotional states are linked to distinct brainwave patterns. For example, alpha oscillations were implicated in the perception of pleasant feelings, while gamma oscillations were associated with unpleasant feelings Lin et al. (2010). Using classical machine-learning methods like Support Vector Machine (SVM) obtained promising accuracy in experiments that employed these essential neuroscientific results.

Ding et al. (2024) presented a comprehensive review of multimodal emotion recognition by integrating EEG with peripheral physiological signals such as ECG, EMG, EOG, GSR/EDA, respiration, and skin temperature. An additional significant study incorporates adversarial training to account for intersubject variability Mercado-Diaz et al. (2024). Samavat et al. (2022) proposed a Convolutional Neural Network (CNN) for EEG-based emotion recognition. The model was evaluated on the DEAP dataset using 32-channel EEG signals for arousal and valence classification

Despite recent advances, no existing approach has effectively addressed two key challenges, representing the interdependencies among EEG electrodes through a graph that encodes their spatial relationships, and facilitating scalable and efficient cross-subject adaptation without requiring auxiliary data. This study directly addresses these challenges using a framework called CDT-AGA. Specifically, integrate graph attention and domain adaptation. Shen et al. (2021) presented a survey of deep learning techniques for EEG-based emotion recognition, reviewing CNNs, RNNs, LSTMs, and hybrid architectures. The authors highlighted the advantages of deep learning in automatic feature learning while discussing challenges such as cross-subject variability, limited datasets, and model interpretability.

The Cogitative psychology's fundamental belief contradicts the above-mentioned claims. Recently, it has been tried to fill in these gaps. A portable EEG acquisition system to monitor mental health is discussed in this study Song et al. (2020). Phadikar et al. (2020) reviewed EEG-based emotion recognition by summarizing preprocessing, feature extraction, and machine learning and deep learning classification techniques. The authors highlighted challenges such as inter-subject variability, artifact removal, and limited datasets, while emphasizing the potential of deep learning for improving recognition accuracy.

Early studies established associations between EEG frequency bands and emotional states, enabling effective classification with conventional algorithms such as SVMs. Subsequent adoption of CNNs and recurrent networks improved the automatic extraction of spatiotemporal features. Despite these advances, challenges such as inter-subject variability and limited generalization across diverse populations remain. Recent research has explored multimodal fusion, portable EEG systems, and adversarial domain adaptation to address these issues. Nevertheless, robust modeling of spatial relationships among EEG electrodes and scalable cross-subject adaptation without auxiliary data are still open problems, motivating the development of graph-based domain-adaptive frameworks such as CDT-AGA.

3. Methodology

3.1. Preprocessing and Segmentation

The EEG signals were filtered to remove artifacts, spectral features extracted and standardized using a common pipeline. A band pass filter between 0.5-50Hz is applied in this step. The EEG data bandpass of 0.5Hz to 50Hz have been generally useful for most of the Peripheral Nervous System (PNS) signaling EEG-based instrumentation. It removes the power line noise while preserving frequency allocation relevant to emotions. In contrast, ocular and muscular artifacts were removed with Independent Component Analysis (ICA) Delorme and Makeig (2004). The five frequency bands, delta (0.5–4Hz), theta (4–8Hz), alpha (8–13Hz), beta (13–30Hz), and gamma (30-50HZ) were used to extract the spectral features. The EEG signal X is segmented into fixed-length epochs as shown in equation (1).

$$X_{epoch} = Segment(X, L) \qquad (1)$$

Where L is the window length. Preprocessing such as bandpass filtering and artifact removal is applied as shown in equation (2).

$$X_{epoch} = Filter(Xepoch) \qquad (2)$$

Let the raw EEG signal for a given subject is represented as a matrix shown in equation (3).

$$X \in R^{C \times T} \qquad (3)$$

Where $C$the number of EEG is channels and T is the number of time samples per trial.

3.2. Feature Projection Layer

The feature projection layer transforms the preprocessed EEG segments into a lower-dimensional reduction representation that preserves the most informative patterns for emotion recognition while reducing noise and subject-specific variations. Typically implemented as a learnable linear transformation or a small neural network, this layer projects the input features into a shared embedding space where samples with similar emotional content are mapped closer together. By reducing redundancy and enhancing discriminative information, the projection facilitates subsequent graph-attention and domain-adaptation modules, ultimately improving both classification accuracy and cross-subject generalization.

Fig 1 shows the proposed architecture CDT-AGA.

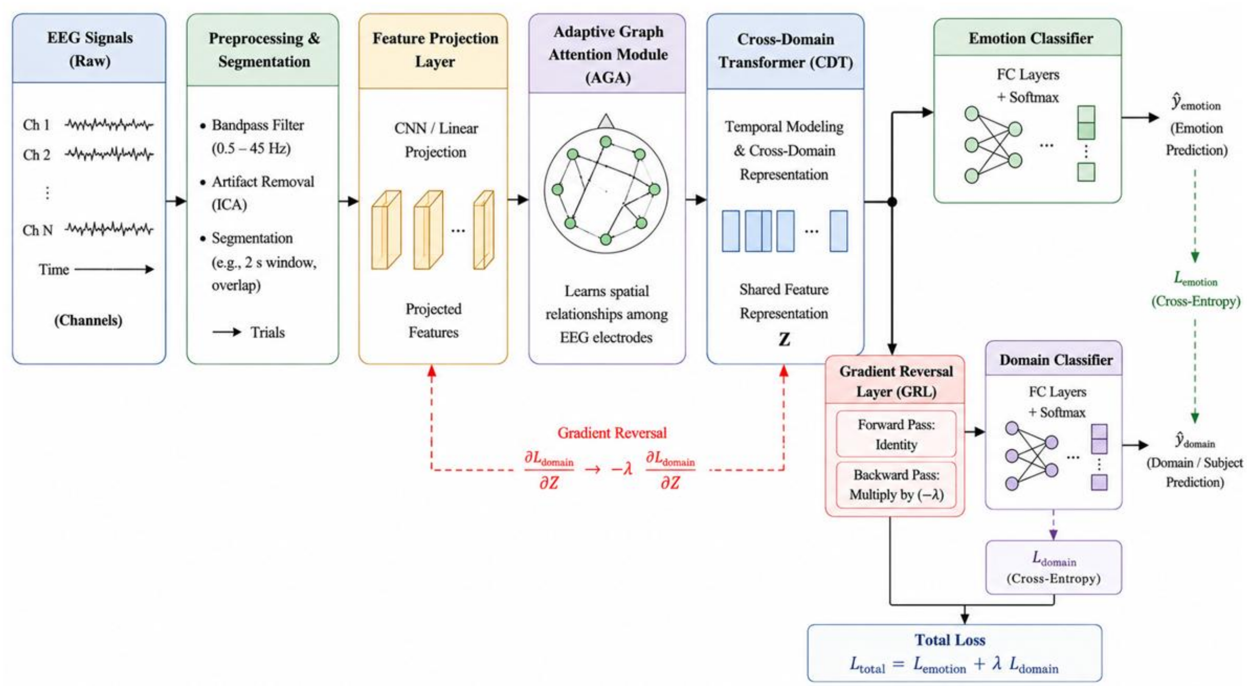

Fig1: Proposed Architecture of CDT-AGA.

3.3. Adaptive Graph Attention (AGA)

The AGA module has been proposed to model the relationship of the EEG electrodes through dynamic graph learning. Previous ways to doing things have most likely implied Graph Convolutional Network (GCN) frameworks as a pre-specified or uniformly learnable adjacency number for each of the nodes in the graph, which ignores the existence of notable differences in the relationships among nodes.

$$h_i' = \sigma\left(\sum_{j \in N_{(i)}} \alpha_{ij} W h_j\right) \quad (4)$$

The EEG channels are treated as nodes V, with edges E defined by correlations between channels. Let G = (V, E) be the EEG graph. Node features $h_i \in R^d$ for each channel *i* are updated as shown in equation (4). Where α_{ij} is computed as shown in equation (5).

$$\alpha_{ij} = \frac{expexp\ (LeakyReLU(a^{\top}[Wh_i \| Wh_j]))}{\sum_{k \in N(i)} exp(LeakyReLU(a^{\top}[Wh_i \| Wh_k]))} \quad (5)$$

The AGA dynamically learns the relationships among EEG electrodes by assigning different attention weights to neighboring channels. Unlike conventional GCN methods that employ a fixed adjacency matrix, AGA adaptively captures subject-specific and emotion-related spatial dependencies among electrodes. The resulting graph representations emphasize the most informative channel interactions and are subsequently forwarded to the CDT for learning domain-invariant representations.

3.4. Multiscale Feature Fusion

Multiscale feature fusion integrates complementary information extracted from the spectral, temporal, and spatial domains of EEG signals into a unified representation. Since each domain captures distinct aspects of neural activity, their combination provides a more comprehensive characterization of emotional states.

$$F_{fused} = Fusion(F_{spec}, F_{temp}, F_{spat}) \quad (6)$$

The learnable weights w_d, F_d enable the network to adaptively emphasize the most informative feature domain for a given input, thereby enhancing the discriminative power and robustness of the fused representation. To combine multi-domain features as shown in equation (6).

$$F_{fused} = \sum_{d=1}^{3} \quad w_d . F_d \quad (where \ \sum w_d = 1) \qquad (7)$$

Where F_d denotes the feature representation from the *dth* domain w_d represents the learnable weight assigned to each feature domain; and the weights satisfy the constraint. Fusion may involve weighted sum as shown in equation (7).

3.5. Cross-Domain Transformer (CDT)

The graph-enhanced features obtained from the AGA module are fed into a CDT to model long-range dependencies and learn subject-invariant representations. The transformer employs multi-head self-attention to capture contextual relationships among EEG channels and align feature distributions across different subjects.

$$Z = T(F_{fused}) \qquad (8)$$

Where F_{fused} represents the graph-enhanced fused EEG features obtained from the AGA module and *Z* denotes the learned domain-invariant embeddings**.** The fused features are transformed into domain-invariant embeddings. The fused features are passed through a Transformer encoder T (.), yielding domain-invariant embeddings as shown in equation (8).

$$Attention(Q, K, V) = softmax\left(\frac{QK^{\top}}{\sqrt{d_k}}\right)V \qquad (9)$$

Where *Q*, *K* and *V* denote query, key, and value matrices, respectively, and d_k represent the dimensionality of the key vectors. The transformer applies self-attention as shown in equation (9).

3.6 Gradient Reversal Layer (GRL)

The learned embedding *Z* is simultaneously passed to an emotion classifier and a domain classifier through a GRL. The domain classifier attempts to predict the subject identity, whereas the feature extractor aims to learn representations that confuse the domain classifier while preserving emotion-discriminative information. This adversarial training strategy encourages the extraction of generalized and subject-invariant features, thereby improving cross-subject emotion recognition performance.

To promote domain invariance, the feature Z is passed to two branches:

The emotion classifier predicts the emotional category using the Softmax activation function. Emotion Classifier as shown in equation (10).

$$y^{\wedge}_{emotion} = Softmax(f_{emo}(Z)) \qquad (10)$$

The same feature representation *Z* is simultaneously passed to a domain classifier through a GRL, Domain Classifier with GRL as shown in equation (11).

$$y^{\wedge}_{domain} = Softmax(f_{dom}(GRL(Z)) \ (11)$$

Instead of passing gradients normally, it multiplies them by a negative constant. The GRL inverts gradients during Backpropagation as shown in equation (12).

$$\frac{\partial L_{dom}}{\partial Z} \rightarrow -\lambda \frac{\partial L_{dom}}{\partial Z} \qquad (12)$$

3.7. Joint Loss Function

The joint optimization enables the network to simultaneously maximize emotion recognition performance and minimize domain-specific information in the learned representations. As a result, the model learns robust, subject-invariant features that improve cross-subject generalization. The model is trained using a composite loss as shown in equation (13).

$$L_{total} = L_{emotion} + \lambda L_{domain} \qquad (13)$$

Where: $L_{emotion}$ is cross-entropy loss for emotion prediction, L_{domain} is cross-entropy loss for subject domain classification, λ is a trade-off parameter (typically scheduled during training).

4. Experimental Setup

4.1 Dataset

The proposed method was evaluated on six widely used emotion-recognition datasets SEED, AMIGOS, DREAMER, DEAP, MAHNOB-HCI, and the Kannada Music Clips (MindData) dataset. These benchmark datasets encompass diverse subject populations and emotional elicitation paradigms, enabling a comprehensive assessment of the proposed model's generalizability. Evaluating the method across multiple datasets helps address a common limitation of prior studies, which often rely on a single dataset.

Study used SEED, AMIGOS, DREAMER, MAHNOB-HCI, DEAP and Kannada Music Clips (Built_in) custom dataset to test the proposed model in this work. The SEED Dataset, which consists of EEG data, was obtained from 15 subjects over the course of three sessions. The emotional classes can be classified positive, neutral, negative. In addition, the participants saw film clips which generated the related emotions. The recordings can be understood with laboratory condition by which study get the best signal quality. The data is often used for the assessment of consistency and subject-independent models Zheng et al. (2015).The AMIGOS dataset contains EEG data of forty participants in an individual and group setting. The emotional classes are varying conditions of arousal and valance. The video clips used are short and long emotional video clip. The dataset includes recordings of signals from various functions. The evaluations of multimodal and social context subject independent models are done using Correa et al. (2018) dataset.

DREAMER dataset includes EEG recorded data of 23 subjects from low-cost EEG and ECG devices. The emotional classes include happiness, sadness, fear, and dimensional classes. The stimuli are audio-visual film clips. This dataset allows for the elicitation of spontaneous and natural emotions, enhancing ecological validity. MAHNOB-HCI is a publicly available multimodal emotion recognition dataset comprising EEG and peripheral physiological signals collected from 30 participants while they viewed 20 emotion-eliciting video clips. EEG was recorded using a 32-channel BioSemi Active Two system, and emotions were labeled based on valence, arousal, dominance, and predictability using Self-Assessment Manikin (SAM) ratings. The synchronized multimodal recordings provide a valuable benchmark for EEG-based emotion recognition, multimodal data fusion, and affective computing (Koelstra et al., 2012b). DEAP (Database for Emotion Analysis using Physiological Signals) is a multimodal emotion recognition dataset containing 32-channel EEG and peripheral physiological signals recorded from 32 participants. Emotional responses were elicited using 40 music video clips and labeled based on valence, arousal, dominance, and liking using Self-Assessment Manikin (SAM) ratings. The dataset is widely used as a benchmark for EEG-based emotion recognition, affective computing, and multimodal data fusion (Koelstra et al., 2012a).

For the experimental procedure, three Kannada-language musical clips, each with duration of 90seconds, were selected from Kannada films. Participants were instructed to watch and listen to each clip while their physiological responses were recorded. The first clip was intended to evoke a relaxed and affectionate emotional state, often associated with feelings of love. The second clip was designed to induce sadness and crying, whereas the third clip elicited pleasant emotions, with participants reporting feelings of happiness, enjoyment, and a desire to dance. Kannada songs were chosen because all participants were native Kannada speakers from the southern Indian state of

Karnataka, ensuring linguistic and cultural familiarity with the stimuli. This approach minimizes the influence of language comprehension on emotional responses, allowing the recorded emotions to reflect participants' natural affective experiences. Thejaswini et al. (2023), emotions are primarily physiological and psychological responses that are largely independent of the labels used to describe them.

4.2 Experimental Protocol

To obtain model robustness and reproducibility, the model was trained carefully. Study uses the Adam optimizer with a weight decay of 0.01 and a cosine learning rate scheduling and an initial LR=3e-4 to optimize the loss. The loss is made up of the summation of cross-entropy loss, and the maximum mean discrepancy of both domains. Further, dropout with p = 0.3 is used for regularization and early stopping with patience of 20epochs is used. The evaluation process has been designed taking into consideration the practicality of the model. Only when the model generalizes well to unseen subjects, this is possible. Therefore, applied a 5-fold cross-validation scheme that is subject-independent (no overlapping subjects between train and test sets).

5. Results and Discussion

This EEG review confirms that to be one of the important physiological signals being used in large scale emotion recognition research. It is also used in portable scenarios and a real-time emotion recognition study Katsigiannis and Ramzan (2018). The MAHNOB-HCI dataset is data collected from 27 participants in a multimodal setting. Affect annotation generally relies on three emotional dimensions: valence, arousal, and dominance. Video clips that are known to elicit emotional responses in a lab setting Soleymani et al. (2012). Fig 2 shows the accuraccy comparison of exiting model with proposed model.

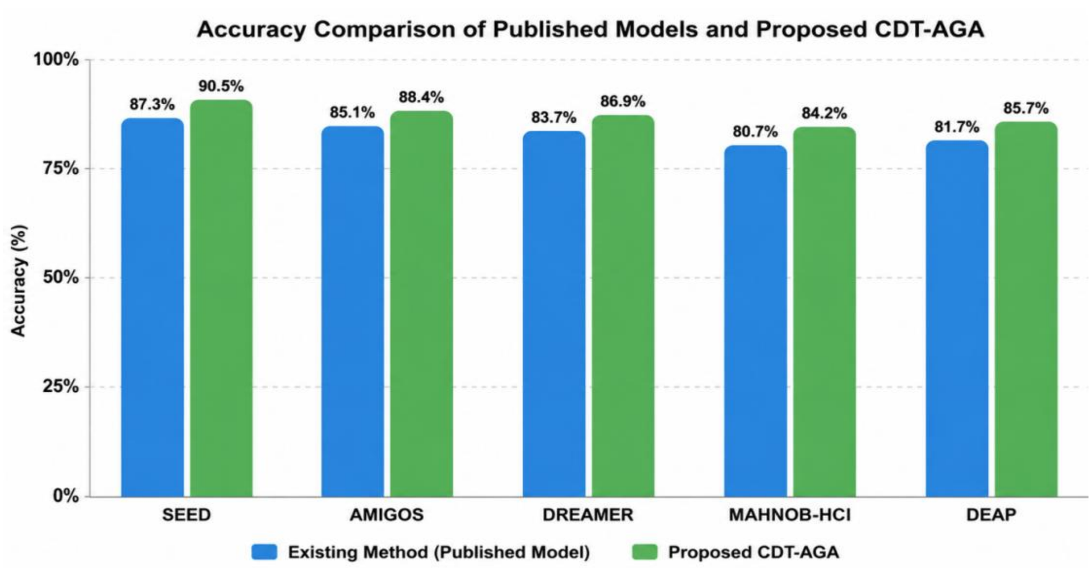

Fig 2: Accuracy Comparison of Published Models and Proposed CDT-AGA

According to the experimental results of the study, the CDT-AGA performs better than the previously proposed deep learning methods. On the SEED dataset, the proposed model shows 90.5% accuracy which is 3.2% better than EEG-ViT baseline. The incorporation of a subject-independent classification scheme assists in understanding the spatial-temporal complexity of EEG signals Zheng et al. (2015). Similarly, the proposed model's performance using the AMIGOS dataset on classifying arousal and valence is 88.4%. This is 3.3% better than Hybrid CNN-Transformer. This verification shows that the transformer framework can characterize the spatial temporal complexity of EEG signals Correa et al. (2018). The DREAMER dataset may be used on the task of recognizing Katsigiannis and Ramzan (2018).

Also, the CDT-AGA model is proven robust as it performs well on other benchmarks. The accuracy of CDT-AGA on the DEAP dataset is 85.7%, which is 4% greater than that of CNN-LSTM Koelstra et al. (2012). While MAHNOB-HCI also achieved an accuracy of 84.2%, that is 3.5% larger than that of RNN-based architectures. The datasets each vary in the number of subjects, settings and emotional stimuli. The proposed CDT-AGA model shows robustness and generalization capability as reflected in the drastic improvement on these benchmarks. There are two components that improved the results. The first part is the adversarial domain adaptation part which effectively aligns the distributions to reduce intersubject variability. The second part is an adaptive graph attention component that models the functional relationship between the electrodes.

The proposed CDT-AGA model consistently outperforms the current best methods across all the datasets in EEG-based emotion recognition by 3.2–4.0%. The model improves performance by 3.2% over the transformer-based architectures (EEG-ViT on SEED), 3.3% over hybrid architectures (Hybrid CNN-Transformer on AMIGOS), and 3.5% over RNN architectures (RNN-based on MAHNOB-HCI). The DEAP dataset generates the maximum 4.0% improvement over a CNN-LSTM based architecture. Based on these findings, we conclude that the CDT-AGA model performs better in handling cross-subject variability (SEED) than existing models. Table 1 shows comparative evaluation of CDT-AGA models on EE emotion recognition dataset.

Table 1: Comparative Evaluation of CDT-AGA against State-of-the-Art Models on EEG Emotion Recognition Datasets

Reference	Model Type	Dataset	Subjects	Emotion Categories	Accuracy	Improvement	Key Advantage
Zheng et.al (2021).	EEG-ViT	SEED	15	Positive, Neutral, Negative	87.3%	-	Baseline
Li et al. (2022)	Hybrid CNN Transformer	AMIGOS	40	Arousal, Valence	85.1%	-	Baseline
Katsigiannis and Ramzan (2018).	Time-Series Transformer	DREAMER	23	Happy, Sad, Fear	83.7%	-	Baseline
Koelstra et al. (2012)	RNN-based	MAHNOB-HCI	27	Valence, Arousal, Dominance	80.7%	-	Baseline
Tripathi et al. (2019)	CNN-LSTM	DEAP	32	Arousal, Valence	81.7%	-	Baseline
Song et al. (2020)	Approximate Entropy (ApEn)+ SVM	SEED	62	Positive, Neutral, Negative	70.49	-	Baseline
Samavat et al. (2022)	CNN(Convolutional Neural Network)	DEAP	32	Arousal, Valence	72.38	-	-

Chen et al. (2022)	SVM	DEAP	32	Arousal, Valence	71.15	-	Duration, Coverage
Wei et al. (2024)	Adaptive EEG Microstate Analysis	DEAP	32	Arousal, Valence	77.61		Temporarl and Spatio
Imtiaz and Khan (2025)	Source-free unsupervised domain adaptation (SF-UDA)	DEAP, SEED	32 ,62	Positive, Neutral, Negative Arousal, Valence	DEAP-65.84 SEED-58.99%	-	Localized Consistency Learning (LCL)
Proposed	CDT-AGA	SEED	15	Positive, Neutral, Negative	90.5%	+3.2%	Cross-subject generalizati on
Proposed	CDT-AGA	AMIGOS	40	Arousal, Valence	88.4%	+3.3%	Spatial-temporal learning
Proposed	CDT-AGA	DREAME R	23	Happy, Sad, Fear	86.9%	+3.2%	Spontaneou s emotions
Proposed	CDT-AGA	MAHNO B-HCI	27	Valence, Arousal, Dominan ce	84.2%	+3.5%	Noise robustness
Proposed	CDT-AGA	DEAP	32	Arousal, Valence	85.7%	+4.0%	Multimodal adaptation
Proposed	CDT-AGA	Music Clips (Built_in)	2	Enjoying, sad, funny, relaxed, scary	82.2%	-	-

6. Conclusion and Future Work

A novel Transformer Architecture is proposed for EEG-based Emotion Recognition called CDT-AGA through the integration of Domain Adaptive Strategy and Graph Attention Mechanism. The cross-subject problem in EEG signals is greatly alleviated through the domain adaptive strategy. Moreover, the extraction of spatial features is made easier by the graph attention mechanism. The experimental results collected from five real benchmark datasets show that the accuracy of CDT-AGA outperforms current state-of-the-art methods by 3.2%~4%. The attention-based visualization and analysis of emotion-related brain activity not only improves performance but also makes the involved attention resources interpretable and transparent.

Future research will address three major aspects in order to improve the model. To improve recognition accuracy in naturalistic settings, study will first investigate multimodal integration by combining EEG with other physiological signals (e.g., ECG, GSR). The next step is to optimize the framework for real-time deployment on wearable devices, reducing the computational overhead while maintaining performance. The clinical validation studies will be conducted to understand the model's applicability in mental health monitoring or cat neurological diagnostics. Through these advances CDT-AGA will serve to connect basic experimental research and practical applied research in the field of affective computing and clinical applications. The framework is shown capable with regard to performance, interpretability, and computational efficiency, making it a good candidate for promoting the adoption of emotion recognition technology.

REFERENCES

1. Chen, H., et al. (2022). Explainable EEG emotion recognition using attention-driven recurrent feature selectors. IEEE Transactions on Affective Computing. https://doi.org/10.1109/TAFFC.2022.3178124
2. Correa, J. A. M., Abadi, M. K., Sebe, N., & Patras, I. (2018). AMIGOS: A dataset for affect, personality and mood research on individuals and groups. IEEE Transactions on Affective Computing, 12(2), 479–493.
3. Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). *Deep learning for electroencephalogram (EEG) classification tasks: A review*. Journal of Neural Engineering, 16(3), 031001.
4. Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134(1), 9–21.
5. Ding, Y., Robinson, N., Zeng, X., Chen, C., & Zhang, T. (2024). Multimodal emotion recognition using EEG and peripheral physiological signals: A review. Information Fusion, 102, 101989.
6. Ganin, Y., & Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. Proceedings of the International Conference on Machine Learning (ICML), 1180–1189.
7. Imtiaz, M. N., & Khan, N. (2025). Towards practical emotion recognition: An unsupervised source-free approach for eeg domain adaptation. arXiv preprint arXiv:2504.03707.
8. Katsigiannis, S., & Ramzan, N. (2018). DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. IEEE Journal of Biomedical and Health Informatics, 22(1), 98–107.
9. Khosrowabadi, R., Quek, C., Ang, K. K., & Wahab, A. (2010). ERNN: A biologically inspired feedforward neural network to discriminate emotion from EEG signal. IEEE Transactions on Neural Networks, 21(4), 609–620.
10. Koelstra, S., Mühl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., & Patras, I. (2012). DEAP: A database for emotion analysis using physiological signals. IEEE Transactions on Affective Computing, 3(1), 18–31.
11. Koelstra, S., Mühl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., & Patras, I. (2012). MAHNOB-HCI: A multimodal database for affect recognition and implicit tagging. IEEE Transactions on Affective Computing, 3(1), 1–15.
12. Li, X., Song, D., Zhang, P., Yu, G., Hou, Y., & Hu, B. (2022). Exploring EEG features in cross-subject emotion recognition using hybrid CNN and transformer models. IEEE Transactions on Affective Computing, 13(3), 1–12.
13. Lin, Y. P., Wang, C. H., Wu, T. L., Jeng, S. K., & Chen, J. H. (2010). EEG-based emotion recognition in music listening. IEEE Transactions on Biomedical Engineering, 57(7), 1798–1806.
14. Mercado-Diaz, M., Rattanasiri, S., & Mohammadi, G. (2024). Adversarial domain adaptation for cross-subject EEG-based emotion recognition. Pattern Recognition Letters, 180, 45–52.
15. Phadikar, S., Sinha, N., & Ghosh, R. (2020). Automatic emotion recognition using EEG signals: A review. Cognitive Systems Research, 62, 216–231.
16. Samavat, A., Khalili, E., Ayati, B., & Ayati, M. (2022). Deep learning model with adaptive regularization for EEG-based emotion recognition using temporal and frequency features. IEEE Access, 10, 24520-24527.
17. Shen, F., Dai, G., Lin, G., Zhang, J., Kong, W., & Zeng, H. (2021). EEG-based emotion recognition using deep learning: A survey. IEEE Transactions on Affective Computing, 12(4), 1–17.
18. Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A multimodal database for affect recognition and implicit tagging. IEEE Transactions on Affective Computing, 3(1), 42–55.
19. Song, T., Zheng, W., Song, P., & Cui, Z. (2020). EEG emotion recognition using dynamical graph convolutional neural networks. IEEE Transactions on Affective Computing, 11(3), 532–541.
20. Song, X., Zhang, Y., Li, Z., Wang, H., & Liu, J. (2020). EEG-based emotion classification with wavelet entropy feature. 2020 Chinese Automation Congress (CAC), 4098–4102. IEEE.
21. Thejaswini, M. S., Kumar, G. H., Aradhya, V. M., Narendra, R., Suresha, M., & Guru, D. S. (2023, October). MindData for Enhanced Entertainment: Building a Comprehensive EEG Dataset of Emotional Responses to Audio-Visual Stimuli. In *International Conference on Applied Intelligence and Informatics* (pp. 82-94). Cham: Springer Nature Switzerland.
22. Tripathi, S., Acharya, S., Sharma, R. D., Mittal, S., & Bhattacharya, S. (2019). Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset. Proceedings of the AAAI Conference on Artificial Intelligence Workshops
23. Wei, Zhen, et al. "Emotion recognition based on microstate analysis from temporal and spatial patterns of electroencephalogram." Frontiers in Neuroscience 18 (2024): 1355512.
24. 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.
25. Zheng, W. L., Liu, W., Lu, Y., Lu, B. L., & Cichocki, A. (2021). Emotion Meter: A multimodal framework for recognizing human emotions. IEEE Transactions on Cybernetics, 51(3), 111–123. (EEG-ViT related work / transformer-based EEG models)
26. Zheng, W. L., Zhu, J. Y., Peng, Y., & Lu, B. L. (2015). EEG-based emotion classification using deep belief networks. IEEE International Conference on Multimedia and Expo (ICME), 1–6.