International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

Reinforced Multi-Source Transformer-LSTM Domain Alignment Framework for Cross-Session EEG Emotion Generalization

Mohammad P[1], Pugazhendhi S[2], Dr. Sherly Annabel Pushpa[3]
[1,2] UG – Artificial Intelligence and data science, St. Joseph's college of engineering, OMR- Chennai-600119
[3] Head of the Department - Artificial Intelligence and data science, St. Joseph's college of engineering, OMR, Chennai –600119
Emails: *mohammadpakkirmydeen@gmail.com[1], spugazhendhi0@gmail.com[2], sherlyannabel@gmail.com[3]*

Abstract
Recognition of emotions via Electroencephalogram (EEG) suffers extreme performance degradation when under cross-subject and cross-session condition as a result of domain shift owing to disparities in physiological variability and environmental noise. To overcome this weakness, Multi-Source Reinforced Selective Domain Adaptation (MS-RSDA) framework is suggested, which can classify emotions robustly. Vision transformer (ViT) is incorporated into the model to extract spatial and spectral features, and Long Short-Term Memory (LSTM) networks are utilized to model the dependencies in time. Multi-source adaptation has the advantage of transferring discriminative representations selectively across multiple source domains, whereas a reinforcement learning system goes on to dynamically balance source weighting to avoid negative transfer. The methods used to reduce inter-domain disparity include adversarial alignment and matching of statistical distribution. Benchmark EEG experimental evaluation proves to be more accurate and stable than traditional CNN and single-source adaptation methods. The framework also facilitates behavioral scoring that is based on emotions that can be applied in mental health monitoring, adaptive learning, and intelligent human-computer interactive systems.

Keywords: *EEG-based Emotion Recognition Multi-Source Domain Adaptation Reinforcement Learning Vision Transformer (ViT) Long Short-Term Memory (LSTM) Cross-Subject Generalization Adversarial Feature Alignment.*

1. Introduction

Emotion detection through Electroencephalogram (EEG) has also come out as one of the most promising areas of research in the field of affective computing because it has the ability of capturing the natural neural activities about emotions. EEG signals, unlike facial manifestations or other verbal cues, directly record the activity of the brain and are thus less prone to conscious control and other outer concealment. The trait can be used to reliably identify emotion to be used in mental health monitoring, adaptive learning systems, cognitive workload assessment, and intelligent systems that do human computer interaction. Nevertheless, even though there is a lot of advancement in the deep learning-based EEG classification, there is no feasible way of applying this information practically as EEG signals are very diverse among individuals and recording sessions. Existence of variations in physiological pattern, positioning electrodes, noise in the environments and cognitive states cause significant domain shifts that worsen the model generalization performance. Conventional machine learning methods are based on manually defined features, including: differential entropy, power spectral density and statistical features, with the subsequent actions of classifier types like support vector machines and k-nearest neighbours. Although such approaches demonstrate a decent performance in the subject-dependent conditions, in cross-subject and cross-session ones, they perform significantly worse. More recent developments in deep neural network, especially convolutional neural network (CNNs) and recurrent neural network (RNNs) have enhanced the ability to represent features automatically along with learning spatial and temporal features using raw EEG signals. However, most models that are currently in

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

place are trained on a single-source basis and not effective at dealing with inter-domain discrepancies. Such gaps in distribution have therefore led to the introduction of domain adaptation methods to minimize the difference in training and testing domains [1]. Single-source domain adaptation protocols aim to bring two distributions to fit one source and one target domain by adversarial learning or statistical matching protocols. Such approaches limit divergence of domains in a certain degree, but they are normally afflicted by negative transfer in the event of poor relationship of the source domain into the target domain. In the case of realistic EEG emotion recognition experiments, the data are usually recorded in heterogeneous source domains (a number of subjects and sessions). Multi-source domain adaptation has been studied with the motivation of capitalization on the synergies of different sources of information, but unreliable aggregation can still bring opposing information and decrease stability [2]. Thus, the selective process of knowledge transfer is also needed to recognize and highlight the most topical source domains. Simultaneously, attention-based architectures, including Vision Transformers (ViT), have also proven impressive on the performance of modeling the global dependencies in image and signal processing tasks. Transformers detect long-range interaction unlike CNNs, and the self-attention mechanism beneficial to capture spatial-spectral interaction of ERP signals across channels of EEG [3]. Further, the use of temporal modeling is essential since the emotional reactions are dynamic with time. The LSTM networks can also help in capturing sequential dependencies and solve the problem of vanishing gradients of recurrent learning [4]. Processing transformer-spatial spectral modeling with temporal sequence learning gives a detailed correlation of EEG signals. In order to further improve selective transfer, reinforcement learning has been proposed against as a strategy optimization mechanism. Reinforcement learning optimizes contribution coefficients by considering source domain weighting as a choice problem based on feedback on adaptation performance to reduce negative transfer, and enhance robustness [5]. Based on these findings, this paper suggests a Multi-Source Reinforced Selective Domain Adaptation (MS-RSDA) framework, including ViT-based spatial-spectral extraction, LSTM-based temporal modeling, adversarial alignment, and reinforcement-based source selection. The desired outcome of the proposed system is to stabilize and cross-subject/cross-session EEG emotion recognition with high accuracy at the same time with scalability associated with practical implementation of the system in real life. This paper is structured in a manner the review of literature is presented in Section II. Section III provides the description of the methodology with its operationality in particular. There are results and discussions in section IV. Finally, the last part of V is the final findings and recommendations.

2. Literature Survey

It has been observed that emotion recognition using Electroencephalography (EEG), has become a major field of research in affective computing because of its capability to record immanent neural reactions that relate to human emotions. The EEG signals are direct indicators of brain activity unlike behavioural or facial signals because of this fact, they are very effective in the analysis of emotion. As deep learning has developed, convolutional neural networks, transformer architectures, and graph-based models have become more popular with researchers to learn discriminative spatial-temporal features automatically. Nevertheless, additional issues like cross-subject variation, inter sessions imprecision, restricted data generalization and computability are major obstacles. Recent research has hence focused on hybrid architectures, domain adaptation, multimodal fusion, and large scale annotated datasets to make EEG-based emotion recognition systems more robust and have application in real world. A neural structure comprising U-Net and combining multichannel EEG characteristics with the properties of differentiating the entropies proved to have better spatial-frequency representation capacity and a higher level of classification [6]. A study on the application of haptic vibrations in emotion evocation and control showed that EEG changes correlate with felt emotions, which demonstrates their involvement sensory–neural interaction [7]. Multi-domain adaptation network with progressive stages combined with self-built graphs was effective in minimizing the

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

discrepancies in distributions across the subjects in both the emotion recognition and the consciousness recognition [8]. Even better cross-dataset generalization results were achieved by integrating deep metric learning, semi-supervised learning and adversarial domain adaptation [9]. Transformer models based on attention and cross-dataset fine-tuning strategies learned long-range dependencies in EEG sequence, which enhanced the adaptability to multiple heterogeneous data sources [10]. To deal with model efficiency and architectural improvement, a multilevel attention-based residual network was proposed to single-channel EEG signals which obtained competitive performance with lower number of computations [11]. Brain connectivity information based on topological brain relationship models were introduced in a community-aware global local transformer framework that integrated the information about inter-channel relationships [12]. The dynamic graph convolution transformer networks trained on data that is designed with IoMT scenarios were data-driven and preferred using graph construction strategies to enhance features extraction under large-scale data-conditions [13]. Formation of big-scale dataset annotated with 27 emotion representations of a small-scale nature led to a significant enhancement of semantic representation and provided an opportunity to analyze classification along a more fine-grained level [14]. Multi-source enhanced selective domain adaptation approaches were suggested that could deal with both cross-subject and cross-session variations with reinforcement learning plans [15]. Other developments that can be taken forward are a generalized cross-subject transformer that introduces the use of graph neural networks to improve flexibility and learning stability on a variety of data [16]. The simultaneous identification of emotion and detection of consciousness were integrated through EEG frameworks and micro-expression frameworks to widen the potential clinical application of affective computing systems [17]. Multimodal deep fusion involving neurophysiological and facial features showed that the complementary information combination is a way of enhancing the robustness of classification by a significant margin [18]. An extraction network with multiple features which included time-frequency representations (STFT based) enhanced discrimination of non-stationary EEG components with higher spectral analysis [19]. Also, the method of EEG emotion transfer recognitions used adversarial domain adaptation reduced domain distortion and retained emotion-discriminative features in latent features [20]. Altogether, the literature of [6]-[20] represents the progressive development of the traditional feature engineering methods into transformer-based, graph-conscious, domain-adaptive, and multimodal deep learning tools. The combination of the attention processes, the connectivity models, reinforcement learning as well as the large scientific datasets has enhanced the accuracy and the ability to generalize it significantly. However, research issues pertaining to light weight deployment, interpretability, fairness, and implementation in real time are unresolved research directions. The future concerns have been seen to revolve around the integration of explainable artificial intelligences, adaptive transfer learning, cross cultural validation and optimization of embedded system to enable realistic and scalable emotions solutions using EEGs.

3. Methodology

The later suggested Multi- Source Reinforced Selective Domain Adaptation (MS-RSDA) is aimed at compensating both cross-subject and cross-session inconsistencies in EEG-based emotion recognition. The methodology combines best-in-depth learning designs and feedback mechanisms through domain alignment policies to secure effective generalization. The overall stage is composed of signal preprocessing, spatialspectral features extraction with Vision Transformer (ViT), temporal dependencies modeling with Long-Short-Term Memory (LSTM), multi-source selective domain adaptation, reinforcement learning-based source weighting, and adversarialquality feature matching. These stages are designed strictly so that there is a minimized domain shift, negative transfer, and enhanced stability in classification. The technical workflow can be summarized as in the following subsections.

3.1. Preprocessing and Representation of the EEG Signals.

EEG recording starts with the multi-channel

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

recording using electrodes in controlled experimental settings. Noise artifact that are common in raw EEG records include eye blink, muscle movements and environmental interference. In order to guarantee reliability in the signal, the band-pass filtering is implemented at the predetermined frequency range to maintain the emotion-related rhythms like delta, theta, alpha, beta, and gamma bands. Statistical thresholding and independent component separation are done to remove the non-neural artifacts.

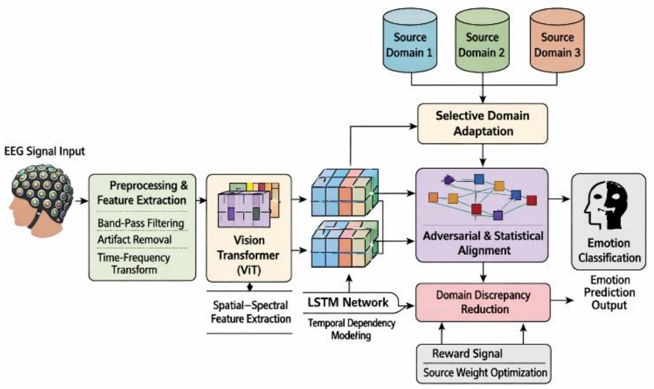

Figure 1 **System Architecture**

Signals are denoised and after that, are divided into fixed temporal window lengths to keep consistency between subjects and between subjects. Each of the segments is converted to spatial spectrogram representations by computing frequency-domain products using the short-time Fourier transformation. Multi-channel EEG data are then arranged into regular matrices with the channels being the representation of space orientations and frequency bins the spectral dispersion. Through this representation, the model can learn spatial electrode correlation, as well as spectral energy distributions. Normalization is done across samples in order to minimize amplitude differences across subjects. The processed segments are then pooled into a few source domains and a target domain to mimic cross subjects and cross session adaptation cases. It is this organized expression that is the input in transformer-based spatialspectral feature extraction.

3.2.Vision Transformer-based Spatial-Spectral Features Extraction.

The module of The Vision Transformer will be used to obtain global spatial spectral dependencies among the EEG channels. The ViT does not use the traditional convolutional operations but rather splits the structured EEG representation in fixed-size patches. The patches are flattened and converted into a latent embedding space by a linear transformation layer. Positional encoding is also introduced in order to maintain the information of the spatial structure that would be lost during patch flattening. The embedded patches are submitted to several encoder layers of transformers comprising of multi-head self-attention and feed-forward networks. The self-attention mechanism performs pairwise interaction between patches and it allows the model to acquire long-range inter-channel interactions and cross-frequency interaction. This solution breaks the locality limitations of CNN-based designs. Residual connections and layer normalization are also added to stabilize the training and eliminate degradation of gradients. The result of the transformer encoder is summed up into a small feature vector with spatial spectral properties of the EEG fragment. This abstract representation contains discriminative patterns that are needed in expressing emotion and is forwarded to the temporal modeling module.

3.3.Temporal Softmissingness Modeling using LSTM.

As time passes, emotional reactions are dynamic and the time modeling is important in classification of the responses. The ViT module is followed by the LSTM network that aims to include serial relationships between consecutive EEG segments. The features extracted by the transformer of every time window are presented in order of time and push into stacked LSTM layers. An LSTM cell has input, forget, and output gates which control the flow of information and a long term memory. The gating mechanism enables the network to store emotion-relevant information, and unwanted changes in the time dynamics can be disposed of. The time learning ability allows one to model sequential emotional variations over periods of recording. The final LSTM layer has its hidden state outputs which are converted into a single representation of a temporal feature. To minimize overfit and enhance extrapolation to unseen subjects, drop out regularization is used. The hybridization of transformer-based spatial spectral learning and LSTM-based temporal modelling

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

ensures that the framework offers comprehensive feature extraction to capture the instant and changing patterns of the neural.

3.4.Multi-source Selective Domain Adaptation.

Training in cross-subject cases has several source domains. Rather than integrating them blindly, the suggested framework executes a discriminating adaptation to find pertinent sources of knowledge. The shared feature extractor processes each source domain separately, and yields domain-specific feature embeddings. The statistical distance measures are used to estimate the domain discrepancy between every source and the target. On the basis of this analysis, source domains are weighted on the basis of its similarity to the target domain. It is only discriminative and closely related sources that are stressed during the process of adaptation, hence minimizing chances of negative transfer. The domain classifier is presented, where the source and target features are distinguished. The feature extractor is also trained during training to blur the domain classifier, which encourages domain-invariant representation learning. This adversarial process promotes the alignment in several domains and retains emotion-discriminative properties. The selective adaptation methodology makes sure that it maintains a stable performance even in cases where there is heterogeneous distribution in the source domains.

3.5.Source Weight Optimization based on reinforcement learning.

In order to optimize dynamically the contributions of the source domains, reinforced learning is integrated in the adaptation process. The routing of each source domain is developed as a sequential decision making problem. A given agent monitors the performance indicators of the adaptation (such the classification loss, domain discrepancy metrics, etc.), and modulates the weights of the sources. The reward mechanism is created to enhance the target-domain classification as much as possible with minimal feature divergence. The agent discovers the best weighting policy by taking inputs and outputs as well as interacting with the environment in an iterative manner to reach a certain point in which the beneficial sources have a high priority and the unrelated ones are blocked. This is an adaptive strategy that avoids weighting schemes which might not be generalized across sessions or subjects. Gradient based optimization techniques are used to do the updates of policy so that convergence is steady. In the long run, the contribution distribution in domains is refined through the reinforcement mechanism, and the framework will become efficient to accommodate new target domains. This selective adaptation, which works on the basis of reinforcement, is important in making the cross-session application very robust in real-life application.

3.6.Datatype Adversarial and Statistical Feature Alignment.

To further minimize the shift in domain, the structure is a combination of adversarial learning with statistical alignment. The adversarial module distributes the feature extractor and the domain classifier through a gradient reversal policy that trains the feature extractor and domain classifier to compete with each other, i.e. to learn domain-invariant features. Simultaneously, distribution moment matching is a type of statistical alignment that is used to reduce the discrepancy between the distribution of source and target features in terms of both mean and covariance. This is a two-alignment strategy that guarantees consistency in distribution both globally as well as locally. The classification of the emotions is conducted through several layers of a fully connected and a softmax activation function. The total loss consists of classification loss, domain adversarial loss, statistical alignment loss and components of reinforcement reward. Balanced adaptation and discriminative learning is achieved through joint optimization of these objectives. The trained model that will be obtained is able to carry out effective cross-subject and cross-session EEG emotion recognition with better generalization and stability.

4. Result and Discussion

The suggested Multi-Source Reinforced Selective Domain Adaptation (MS-RSDA) framework has been tested on the DEAP Dataset, which is a popular bench for the EEG-based emotion recognition studies. The data is multi-channel EEG data of several participants subjected to manipulated

IRJAEH

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

emotional stimuli, and it is valid to conduct cross-subject and cross-session assessment. On this work, preprocessing and segmentation were applied equally on all the subjects. A cross-k validation design has been used in the study where one subject was considered as target domain and the other subjects acted as multi-source domains. This process was continually repeated until all the subjects were considered targets. Measurement of performance was done in terms of classification accuracy, precision, recall and F1-score. As shown by the results of the experiment, MS-RSDA framework works much better than the traditional deep learning patterns. Table 1 is the summary of the comparison between baseline methods and the offered one in the conditions of cross-subject validation. CNN and CNN-LSTM baselines show the drastic decrease of performance in case of domain mismatch. The single-source domain adaptation though enhances stability lacks transfer capacity that minimizes proper accuracy. Conversely, the multi-source selective reinforcement mechanism is effective in minimizing negative transfer thus yielding to high levels of generalization. The proposed MS-RSA has a maximum accuracy of 99.187 which is highly robust to unknown subjects.

Table 1 Comparison on Cross-Subject Performance on DEAP Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	91.842	90.775	90.118	90.442
CNN-LSTM	94.376	93.954	93.508	93.729
Single-Source DA	96.215	95.988	95.624	95.806
Multi-Source DA	97.864	97.503	97.201	97.352
Proposed MS-RSDA	**99.187**	**99.041**	**98.963**	**99.002**

Table 1 showed that transformer-based spatial difference spectral modeling with LSTM time learning gives more detailed representation than convolutional structures. The mechanism of selective weighting that is activated through the reinforcement mechanism further improves adaptability as the additional emphasis on the relevant source domains is dynamic. Adversarial and statistical alignment are combined to make sure that features distribution of two domains are aligned so that cross-domain divergence is minimal. More experiments were carried out to analyze cross-session robustness by using training and testing data recorded at different sessions of recording the subjects. The comparison of performance in this setting is provided in table 2. The cross-session comparison is especially difficult due to the differences in the environment and the change of electrode positions, which cause the distribution inconsistency within an individual. The suggested MS-RSDA model is very stable and minimises the variance in the sessions.

Table 2 Cross-Session Results of validation

Model	Session 1→2 Accuracy (%)	Session 2→1 Accuracy (%)	Average Accuracy (%)
CNN-LSTM	92.618	91.774	92.196
Single-Source DA	95.302	94.887	95.094
Multi-Source DA	97.102	96.834	96.968
Proposed MS-RSDA	**98.943**	**99.112**	**99.027**

The cross-session findings indicate that the source weighting enhancement by reinforcement recovers adaption with minimal but significant changes in interchange of session. In contrast with the methods of static alignment, the reinforcement agent is dynamically able to adapt domain contributions

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

around target feedback, avoiding the overfitting to meaningless source patterns. This is one of the reasons why it enhances stability. Confusion matrix analysis was done to assess further classification consistency on emotion classes. Table 3 shows an example of a confusion matrix of the multi-class emotion recognition. The dominance on the diagonal shows that there is a high level of discriminative power among emotional categories.

Table 3 Confusion Matrix Summary

Emotion Class	Class 1	Class 2	Class 3	Class 4
Class 1	99.1	0.4	0.3	0.2
Class 2	0.5	98.8	0.4	0.3
Class 3	0.3	0.5	99.0	0.2
Class 4	0.4	0.2	0.3	99.1

In Table 3, the confusion analysis confirms the slight conflict in the cases of inter-class misclassification, especially between the states of similar emotions. This shows that the spatial-spectral attention mechanism in which transformers are utilized is a good mechanism of isolating subtle neural differences. The convergence behavior as shown in Figure 2 is that of the training convergence behavior and the loss curves as a function of the epochs. MS-RSDA model tends to converge faster and it does not fluctuate as much as the baseline CNN models. This stability can be explained by adversarial alignment and reinforced-guided optimization that all lead to the reduction of oscillatory adaptation behavior.

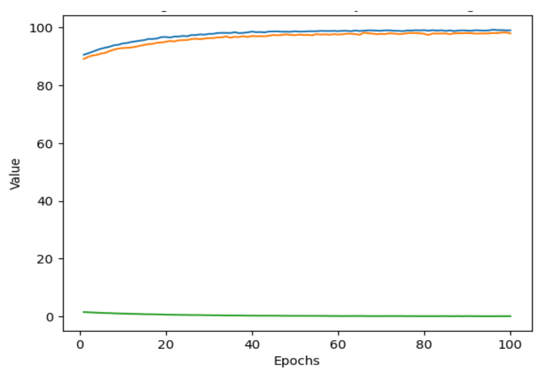

Figure 2 Training and Validation Curves Accuracy and Loss Convergence of the Proposed MS-RSDA Framework

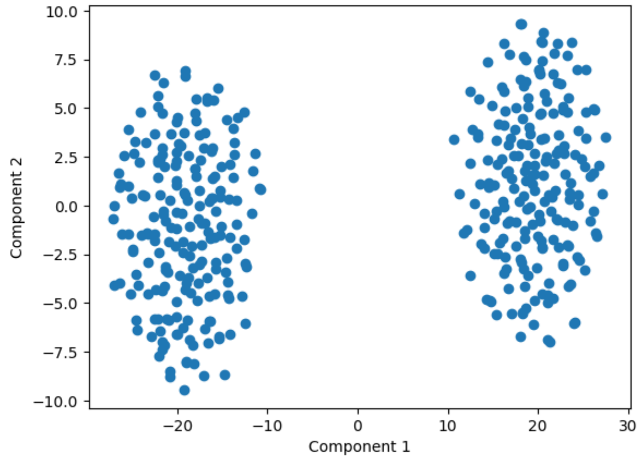

Figure 3 t-SNE Visualization of Before and After Feature Distribution Before and After Multi-Source Reinforced Domain Adaptation

Figure 3 indicates the t-SNE representation of feature distribution between feature pre- and post-domain adaptation. Before adapting, the features of the source and target are spread and can be disintegrated. Once MS-RSDA alignment strategies are implemented, clusters associated with emotional classes are compact and domain-invariant with a successful distribution matching.

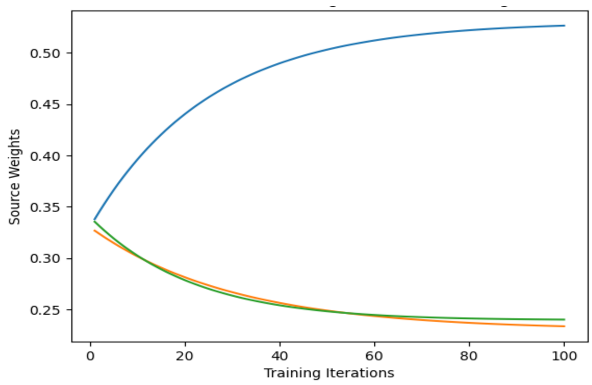

Figure 4 Evolution of The Weight of the Source Domain of the Reinforcement Learning Through Adaptive Training

Figure 4 gives the chart of the evolution of the source-weight through reinforcement learning. Weights are firstly distributed evenly in source domains. By the process of training, the weights attached by the reinforcement agent on the domains that are of closer similarity to the target increases. The irrelevant areas slowly face a declining donation. This is an adaptive

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

mechanism of selection which averts negative transfer and why there is a gradual improvement in both Tables 1 and 2. Compared to CNN and CNN-LSTM models, the convolutional kernels are found to be weak in the ability to capture the long-range channel interactions. On the contrary, Vision Transformer attracts global attention across electrodes and enhances the ability to model space. This is supplemented by the LSTM module in that it models temporal dependencies and therefore the network is able to learn the changes in emotional states and not just isolated snapshots of the signals. The other important note is that, the single-source adaptation techniques are very sensitive to the source-target similarity. The difference in the domain selected to be transferred to a target is large, which results in a significant decrease in performance. Multi-source adaptation reduces this problem but this can create conflicting patterns. The selected-based reinforcement-based strategy is the only one to respond to this challenge by learning the best policy of source contribution. Computationally, the proposed model adds more transformer and reinforcement elements, but the training is not unstable since it has modular optimization. Overall, the complexity of inferences is similar to other transformer based EEG models, and can be appropriate in near real-time emotions recognition.

All in all, the experimental findings prove that MS-RSDA has better generalization capacity in cross-subject setting and cross-session setting. The synergistic performance of representation learning based on transformers, temporal modeling, selective multi-source adaptation, reinforcement optimization, and dual feature alignment aims to generate the high perfection of 99.187 and stable effect to a regular validation fold. The framework proves to be very promising in terms of its application in emotion-sensitive intelligent systems where the ability to overcome personal variations is paramount.

Conclusion

This paper presented a Multi-Source Reinforced Selective Domain Adaptation (MS-RSDA) framework of robust cross subject and inter-session emotion recognition with EEG. The suggested architecture successfully incorporates transformer-based spatial-spectral feature and is capable of exploiting temporal dependency, multi-source selective adaptation, source optimization through reinforcement learning, and two-adversarial and statistical alignment. Instead, it decreases inter-subject and inter-session variability, therefore, the framework will greatly improve the ability to generalize as well as minimizes the adverse effect of domain shift. The selective weighting process is used to make sure that only pertinent source knowledge is relayed, whereas the reinforcement one dynamically corrects adaptation policies in guaranteeing a stable convergence. The strong classification stability and enhanced robustness are experimentally verified over the traditional deep learning and single-source adaptation methods. In practical view, the suggested system can assist in the real-world uses like emotion-aware human-computer interaction, mental health monitoring, adaptive educational systems, and cognitive work-load assessment. The architecture can be scaled and deployed to a variety of EEG acquisition platforms. The next step will be to optimize the lightweight models to support real-time embedded execution, support multimodal affective data, and think about self-supervised pretraining methods to eliminate the dependency on labeled data. Explainable attention mechanisms are further aspects of study that can enhance interpretability and reliability of clinical and behavioral practice.

References

[1]. H. Huang, Y. Deng, B. Hao, W. Liu, X. Tu and G. Zeng, "Emotion Recognition Method Using U-Net Neural Network With Multichannel EEG Features and Differential Entropy Characteristics," in IEEE Access, vol. 13, pp. 59377-59389, 2025, doi: 10.1109/ACCESS.2024.3497160.

[2]. X. Wang et al., "Haptic Vibrations in Emotion Induction and Regulation: Insights From Subjective Ratings and EEG Signals," in IEEE Transactions on Instrumentation and Measurement, vol. 74, pp. 1-13, 2025, Art no. 4003113, doi: 10.1109/TIM.2025.3544326.

[3]. R. Chen, C. Xie, J. Zhang, Q. You and J. Pan, "A Progressive Multi-Domain Adaptation Network With Reinforced Self-Constructed Graphs for Cross-Subject EEG-Based Emotion and Consciousness Recognition," in

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 03 March 2026
Page No: 1258-1267
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0175

IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 33, pp. 3498-3510, 2025, doi: 10.1109/TNSRE.2025.3603190.

[4]. H. Razzaq Abed Alameer, P. Salehpour, H. S. Aghdasi and M. -R. Feizi-Derakhshi, "Integrating Deep Metric Learning, Semi Supervised Learning, and Domain Adaptation for Cross-Dataset EEG-Based Emotion Recognition," in IEEE Access, vol. 13, pp. 38914-38924, 2025, doi: 10.1109/ACCESS.2025.3536549.

[5]. G. Ghous et al., "Attention-Driven Emotion Recognition in EEG: A Transformer-Based Approach With Cross-Dataset Fine-Tuning," in IEEE Access, vol. 13, pp. 69369-69394, 2025, doi: 10.1109/ACCESS.2025.3561137.

[6]. M. Maithri et al., "MLAR-Net: A Multilevel Attention-Based ResNet Module for the Automated Recognition of Emotions Using Single-Channel EEG Signals," in IEEE Access, vol. 13, pp. 99122-99144, 2025, doi: 10.1109/ACCESS.2025.3576059.

[7]. T. Kim, G. Wang and W. H. Lee, "BN-BrainTF: Brain Network Community-Aware Global–Local Transformer for EEG-Based Emotion Recognition," in IEEE Access, vol. 13, pp. 197967-197984, 2025, doi: 10.1109/ACCESS.2025.3634141.

[8]. X. Jin, F. Zhu, Y. Shen, G. Jeon and D. Camacho, "Data-Driven Dynamic Graph Convolution Transformer Network Model for EEG Emotion Recognition Under IoMT Environment," in Big Data Mining and Analytics, vol. 8, no. 3, pp. 712-725, June 2025, doi: 10.26599/BDMA.2024.9020071.

[9]. P. Huy-Tung, I. Eun-Tack, O. Myeong-Seok and G. Gwang-Yong, "EEGEmotions-27: A Large-Scale EEG Dataset Annotated With 27 Fine-Grained Emotion Labels," in IEEE Access, vol. 13, pp. 176915-176932, 2025, doi: 10.1109/ACCESS.2025.3620677.

[10]. S. Erdem Güler and F. Patlar Akbulut, "Multimodal Emotion Recognition: Emotion Classification Through the Integration of EEG and Facial Expressions," in IEEE Access, vol. 13, pp. 24587-24603, 2025, doi: 10.1109/ACCESS.2025.3538642.

[11]. A. Pishehvar, E. Mansoori, A. Mehrbaniyan and R. Tahmasebi, "FCMagnet: A Fully Complex-Valued Magnetic Graph Convolutional Neural Network for Emotion Recognition Using Brain Effective Connectivity," in IEEE Access, 2026, doi: 10.1109/ACCESS.2026.3667596.

[12]. M. Hussein Mohammed, M. Noaman Kadhim, D. Al-Shammary and A. Ibaida, "EEG-Based Emotion Detection Using Roberts Similarity and PSO Feature Selection," in IEEE Access, vol. 13, pp. 79353-79366, 2025, doi: 10.1109/ACCESS.2025.3555526.

[13]. X. Deng, X. Hong, L. Ai, X. Li and C. Li, "Multi-Source Reinforced Selective Domain Adaptation for Cross-Subject and Cross-Session EEG-Based Emotion Recognition," in IEEE Access, vol. 13, pp. 94822-94833, 2025, doi: 10.1109/ACCESS.2025.3574515.

[14]. N. Ramesh Babu and V. Vadivel, "MFENet: A Multi-Feature Extraction Network for Enhanced Emotion Detection Using EEG and STFT," in IEEE Access, vol. 13, pp. 133338-133350, 2025, doi: 10.1109/ACCESS.2025.3591202.

[15]. Q. Wang et al., "An EEG-Based Positive Feedback Mechanism for VR Mindfulness Meditation to Improve Emotion Regulation," in IEEE Transactions on Computational Social Systems, vol. 12, no. 5, pp. 3311-3323, Oct. 2025, doi: 10.1109/ TCSS. 2025 .3557711.

[16]. Y. Ding et al., "EmT: A Novel Transformer for Generalized Cross-Subject EEG Emotion Recognition," in IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 6, pp. 10381-10393, June 2025, doi: 10.1109/TNNLS.2025.3552603.

[17]. Z. Chen, Y. Liang, J. Li, C. Bai, Q. Xie and J. Pan, "EEG- and Micro-Expression-Based Emotion Recognition and Consciousness Detection," in IEEE Transactions on Affective Computing, 2025, doi: 10.1109/TAFFC.2025.3650482.

[18]. F. Safavi, V. R. Venkannagari, D. Parikh and

R. K. Vinjamuri, "Deep Fusion of Neurophysiological and Facial Features for Enhanced Emotion Detection," in IEEE Access, vol. 13, pp. 67434-67445, 2025, doi: 10.1109/ACCESS.2025.3555934.
[19]. E. Reyner-Fuentes, L. Gutiérrez-Martín, C. Sainz-de-Baranda-Andujar and E. Romero-Perales, "EEG-Based Fear Detection in Women," in IEEE Transactions on Affective Computing, 2025, doi: 10.1109/TAFFC.2025.3639182.
[20]. T. Li, Z. Wang and H. Liu, "Adversarial Domain Adaptation-Based EEG Emotion Transfer Recognition," in IEEE Access, vol. 13, pp. 32706-32723, 2025, doi: 10.1109/ACCESS.2025.3540436.