Attention-Driven Adaptive Deep Canonical Correlation Analysis for Multimodal Sentiment Analysis with EEG and Eye Movement Data

Yunhong Liao, Zhengyi Lu, Geoffrey Louie, Jia Li
Oakland University, Rochester, MI 48309, USA
{yunhongliao, zhengyilu, louie, li4}@oakland.edu

Mark Brudnak
Ground Vehicle Systems Center
mark.j.brudnak.civ@army.mil

2025 IEEE World AI IoT Congress (AIIoT) | 979-8-3315-2508-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/AIIoT65859.2025.11105339

***Abstract*—Multimodal sentiment analysis (MSA) has gained prominence for its ability to more comprehensively capture human emotions by leveraging different data sources. In this paper, we propose Adaptive Deep Canonical Correlation Analysis (Adaptive DCCA), a novel framework that integrates electroencephalography (EEG) and eye movement signals for improved sentiment classification. Unlike traditional DCCA, our approach incorporates an attention-based adaptive weighting mechanism, allowing each modality to dynamically contribute according to its relative importance. We evaluate the proposed model on two standard datasets, SEED-IV and SEED-V, under both subject-dependent and subject-independent conditions. Experimental results demonstrate that Adaptive DCCA substantially outperforms existing methods, achieving high accuracy and robustness across varying emotional categories. Notably, the adaptive weighting proves especially beneficial in exploiting the complementary information embedded in EEG and eye movement data, although eye movement signals remain more susceptible to cross-subject variability. The comprehensive confusion matrix analysis further corroborates the effectiveness of our approach in distinguishing fine-grained emotional states. Overall, this work highlights the untapped potential of combining physiological signals for emotion recognition and offers a flexible, high-performance framework that can be extended to other affective computing applications.**

***Index Terms*—Multimodal sentiment analysis; EEG; Eye Movement; Deep Canonical Correlation Analysis; Attention-based fusion; Physiological data**

I. INTRODUCTION

The study of human emotions and sentiments forms a foundational aspect of affective computing and human-computer interaction. While traditional sentiment analysis has predominantly focused on unimodal data, such as text or speech, the emergence of multimodal sentiment analysis (MSA) has highlighted the importance of integrating diverse data sources, including audio, video, and physiological signals [1]. This evolution reflects the inherent complexity of human emotions, which are expressed through intricate combinations of verbal, acoustic, and behavioral cues. By exploiting the complementary strengths of these modalities, multimodal analysis provides a more nuanced and holistic understanding of sentiments compared to unimodal approaches.

However, the integration and effective representation of heterogeneous modalities present significant challenges within the domain of MSA. Text-based features often dominate due to the advanced state of natural language processing models, whereas audio and visual features frequently suffer from noise and variability. Moreover, physiological signals, such as electroencephalography (EEG) and eye movements, have garnered increasing attention for their ability to capture subconscious emotional states [2]. These signals offer a robust alternative as they are less susceptible to intentional modulation, making them particularly valuable in diverse and dynamic application scenarios. Despite their promise, the integration of such physiological data into MSA frameworks remains underexplored.

A deeper look into the field of affective computing reveals several real-world applications that necessitate robust emotional state detection. For instance, in healthcare and mental wellness monitoring, continuous emotion tracking can provide insight into the onset of stress or depression, allowing timely interventions. In education, adaptive learning systems can benefit from real-time feedback on students' engagement levels, enabling personalized teaching approaches. Similarly, in marketing or entertainment analytics, understanding audience sentiment through physiological cues helps

979-8-3315-2508-8/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: OAKLAND UNIVERSITY. Downloaded on September 21,2025 at 16:47:12 UTC from IEEE Xplore. Restrictions apply.

creators optimize content delivery. These expanding areas underscore the need for MSA techniques that can effectively unify diverse signal types, including EEG and eye movements, in a manner that remains stable across different individuals and contexts.

To address these challenges, DCCA has emerged as a transformative approach for learning correlated representations of multimodal data [3]. Unlike traditional canonical correlation analysis (CCA), DCCA leverages deep neural networks to extract nonlinear transformations for each modality, thereby maximizing their correlations in a shared latent space. While effective, traditional DCCA treats all modalities equally without accounting for their varying contributions, potentially overlooking critical inter-modal dynamics and the unique characteristics of specific modalities.

Building upon these advancements, this paper introduces Adaptive DCCA, a novel extension of the DCCA framework designed specifically for multimodal sentiment analysis using physiological data [4]. Unlike traditional DCCA, our model uses an attention-based adaptive weighting mechanism to fuse features more effectively. This mechanism dynamically adjusts the contributions of each modality, allowing the framework to prioritize the most informative features from EEG and eye movement signals. By doing so, Adaptive DCCA effectively captures modality-specific and cross-modal relationships, ensuring a more robust and accurate representation of the data. The framework follows a three-stage architecture: feature extraction, feature fusion, and sentiment classification. During feature extraction, EEG signals are processed using Long Short-Term Memory (LSTM) [5] networks, while eye movement data is analyzed using Fully Connected Networks (FCNs) [6]. These features are then integrated using the attention-enhanced CCA constraints, where the adaptive mechanism plays a pivotal role in optimizing the fusion process. Finally, the fused features are passed through a neural network for sentiment classification into emotional categories such as Happy, Sad, Neutral, Disgust, and Fear.

The remainder of this paper is structured as follows. Section 2 provides an in-depth review of related work in multimodal sentiment analysis and the integration of physiological data. Section 3 elaborates on the proposed Adaptive DCCA framework, detailing its architectural components and optimization strategies. Section 4 presents experimental results on benchmark datasets, illustrating the effectiveness and robustness of the proposed approach. Finally, Section 5 concludes the paper and discusses potential avenues for future research.

II. RELATED WORK

The field of MSA has gained significant attention in recent years due to its potential to provide comprehensive insights into human emotions by leveraging multiple modalities such as text, audio, video, and physiological signals. This section reviews key advancements in MSA, focusing on traditional approaches, the integration of physiological signals, and recent developments in feature fusion techniques, including DCCA.

A. Traditional Approaches to Multimodal Sentiment Analysis

Early efforts in MSA primarily relied on unimodal data, such as text or audio, with each modality analyzed in isolation. However, the inherent limitations of unimodal analysis, including its inability to capture the richness of emotional expressions, spurred interest in multimodal approaches. Studies such as those by Zadeh et al. introduced tensor-based fusion methods, which combined modalities at the feature level to improve sentiment classification accuracy [7]. Other works explored attention mechanisms and graph-based fusion models, enabling more sophisticated integration of text, audio, and visual features.

Despite these advancements, text-based features often dominated due to the availability of large pre-trained language models like BERT. In contrast, audio and video features frequently suffered from noise and variability, highlighting the need for robust fusion methods that can effectively balance the contributions of different modalities [8].

B. Physiological Signals in Multimodal Sentiment Analysis

Physiological signals, such as electroencephalography (EEG) and eye movements, have gained traction in MSA due to their ability to capture subconscious emotional states. Unlike behavioral cues, physiological signals are less prone to deliberate manipulation, making them particularly useful for applications requiring robust sentiment detection. For instance, studies on the SEED-V and DREAMER datasets demonstrated the effectiveness of EEG in emotion recognition tasks, achieving high accuracy rates through feature-level and decision-level fusion methods [9].

Eye movements, which provide complementary information about cognitive and emotional states, have also been integrated into multimodal frameworks.

Authorized licensed use limited to: OAKLAND UNIVERSITY. Downloaded on September 21,2025 at 16:47:12 UTC from IEEE Xplore. Restrictions apply.

Zheng et al. proposed a multimodal framework that combines EEG and eye movement features, revealing their complementary characteristics in emotion recognition tasks [10]. However, the integration of physiological data into multimodal systems remains underexplored, particularly in the context of balancing their contributions with more traditional modalities like text and audio.

C. Deep Canonical Correlation Analysis for Multimodal Fusion

DCCA has emerged as a powerful tool for multimodal fusion by learning nonlinear transformations of multiple modalities to maximize their correlations in a shared latent space. Unlike traditional CCA, DCCA leverages deep neural networks to handle the complex relationships inherent in multimodal data [11]. Recent works have demonstrated the superiority of DCCA over traditional fusion methods, achieving state-of-the-art performance on datasets such as CMU-MOSI and CMU-MOSEI.

Furthermore, certain variants of DCCA have begun to incorporate constraints tailored to each modality, aiming to preserve unique characteristics while maximizing shared features. For example, domain adaptation layers and modality-specific dropout have been introduced to accommodate noise and missing data in signals like EEG or audio. These techniques allow for more flexible architectures that retain meaningful modality-specific nuances without compromising the fused representation. Nevertheless, many of these methods still treat each modality's importance as fixed throughout training, an assumption that may not hold when dealing with variable-quality signals or dynamically changing experimental conditions.

D. Contributions of Adaptive DCCA

Building on these advancements, Adaptive DCCA extends the traditional DCCA framework by introducing an attention-driven adaptive weighting mechanism. This mechanism dynamically adjusts the contributions of EEG and eye movement signals during the feature fusion stage, ensuring a robust and context-aware representation of the data. By addressing the limitations of existing DCCA models, Adaptive DCCA represents a significant step forward in the integration of physiological signals for multimodal sentiment analysis. The proposed framework is evaluated on benchmark datasets, demonstrating its ability to achieve superior robustness and accuracy compared to state-of-the-art methods.

III. PROPOSED FRAMEWORK

This section introduces the Adaptive DCCA framework for multimodal sentiment analysis. This framework integrates EEG and eye movement signals to leverage their complementary characteristics for more accurate sentiment classification. The methodology is structured into three main parts: introduction to the method, framework description, and algorithmic process.

A. Method Overview

Adaptive DCCA aims to jointly extract and fuse features from two distinct modalities—EEG signals and eye movement signals—to enhance sentiment analysis. The framework utilizes a Long Short-Term Memory (LSTM) network for EEG feature extraction, as it is well-suited for processing sequential data, and a Fully Convolutional Network (FCN) for eye movement signals, which excels at capturing spatial and temporal features. The features extracted, denoted as f_{EEG} for EEG and f_{EM} for eye movements, are subsequently aligned using a Deep Canonical Correlation Analysis (DCCA) network. This alignment ensures that the features from both modalities are correlated and capture complementary aspects of the input data.

To refine this fusion, an adaptive attention mechanism is applied to dynamically adjust the contribution of each modality based on their relative importance in the given context. The attention weights guide how the features are fused, enabling the model to adaptively weigh each modality for optimal performance. Finally, the fused representation is fed into a fully connected network for sentiment classification, enabling the prediction of emotional states such as happiness, sadness, neutrality, disgust, and fear.

B. Framework

The Adaptive DCCA framework, illustrated in Fig. 1, consists of three interconnected stages: feature extraction, feature fusion, and sentiment analysis. EEG signals (X_{EEG}) are processed through an LSTM network to capture sequential patterns, resulting in the extracted features f_{EEG}. Similarly, eye movement signals (X_{EM}) are processed using a FCN to capture spatial and temporal correlations, resulting in the extracted features f_{EM}.

Authorized licensed use limited to: OAKLAND UNIVERSITY. Downloaded on September 21,2025 at 16:47:12 UTC from IEEE Xplore. Restrictions apply.

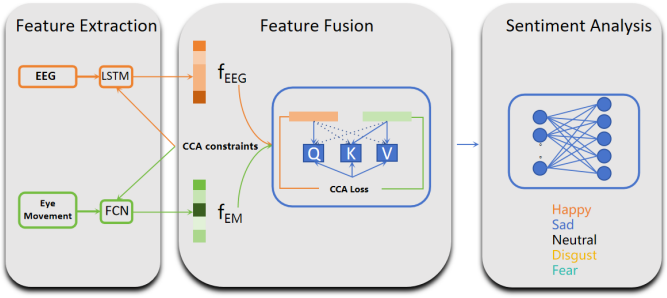

Fig. 1: Overview of the proposed Adaptive DCCA Framework. EEG and Eye movement features are extracted using LSTM and FCN respectively. The adaptive weight mechanism dynamically adjusts each modality's contribution during fusion, enhancing sentiment classification performance by prioritizing informative features.

In our experiments, we also incorporate a simple but effective data preprocessing step. For EEG signals, we apply a band-pass filter in the range of 1–50 Hz to remove high-frequency noise and baseline drift. Eye movement data undergoes a normalization procedure that scales raw gaze coordinates and pupil size metrics into a consistent range. This ensures that outliers or abrupt changes in fixation do not disproportionately affect the downstream fusion process. By harmonizing the input representations of these two modalities, the proposed framework can better learn consistent embeddings during the DCCA alignment stage.

The extracted features f_{EEG} and f_{EM} are aligned in a shared latent space using a Deep Canonical Correlation Analysis (DCCA) network. The alignment is optimized by minimizing the canonical correlation loss:

$$\mathcal{L}_{\text{CCA}} = -\operatorname{corr}\Big(O_{\text{EEG}}, O_{\text{EM}}\Big) \tag{1}$$

To refine the fusion process, an adaptive attention mechanism dynamically adjusts the contributions of each modality. The final fused feature representation is:

$$O_{\text{fusion}} = \alpha \cdot \text{Attention}_{\text{EEG}} \cdot O_{\text{EEG}} + \beta \cdot \text{Attention}_{\text{EM}} \cdot O_{\text{EM}} \tag{2}$$

where α and β are learnable parameters balancing the contributions of EEG and eye movement features.

Finally, the fused features O_{fusion} are passed through a fully connected network C for sentiment classification. The predicted sentiment label $\hat{Y}$ is computed as:

$$\hat{Y} = C\big(O_{\text{fusion}}\big) \tag{3}$$

Rather than relying on simple weighted addition, our attention mechanism uses a soft alignment strategy: each latent representation O_{EEG} and O_{EM} is first projected into an attention space using a small multi-layer perceptron, which outputs modality-specific attention scores. These scores are then normalized via a softmax function before multiplying each modality's representation. This approach allows the network to focus on salient features within each modality and to suppress less informative aspects, especially when facing noisy or incomplete inputs.

IV. EXPERIMENTS

In order to validate the effectiveness and robustness of our proposed Adaptive DCCA framework, we conducted extensive experiments on two widely-used multimodal sentiment analysis datasets, SEED-IV [10] and SEED-V [12]. These datasets provide EEG and eye movement data collected under controlled emotional stimuli, making them ideal benchmarks to comprehensively assess the capability of our model in both subject-dependent and subject-independent scenarios.

A. Dataset Description

The SEED-IV dataset contains four emotional categories—Happy, Sad, Neutral, and Fear—with physiological signals including EEG and eye movement data collected during the experiment. SEED-V extends this framework to five emotional states: Happy, Sad, Neutral, Disgust, and Fear. Both datasets include sessions recorded over multiple days to account for variability in emotional responses.

B. Experimental Setup

For both datasets, we utilized EEG features extracted using LSTM networks and eye movement features processed by FCNs. These modality-specific features were fused using the attention-based adaptive mechanism in Adaptive DCCA.

C. Results on SEED-IV

The performance of Adaptive DCCA on SEED-IV is summarized in Table I. For the subject-dependent scenario, Adaptive DCCA achieved the highest mean accuracy of 87.5% with a standard deviation of 10.2%, outperforming state-of-the-art methods such as Bimodal Deep AutoEncoder (BDAE) (85.1%) and Pyramidal Graph Convolutional Network (PGCN) (82.2%). In the subject-independent setup, Adaptive DCCA also demonstrated competitive performance, achieving 70.7% mean accuracy compared to 73.7% for PGCN, highlighting the robustness of our approach.

Authorized licensed use limited to: OAKLAND UNIVERSITY. Downloaded on September 21,2025 at 16:47:12 UTC from IEEE Xplore. Restrictions apply.

TABLE I: SEED-IV Subject-Dependent

Method	Mean Acc. (%)	Std. Dev. (%)
Concatenation [13]	77.6	16.4
MAX [13]	60.0	17.1
FuzzyIntegral [13]	73.6	16.7
BDAE [14]	85.1	11.8
PGCN [15]	82.2	14.9
Adaptive DCCA	**87.5**	**10.2**

TABLE II: SEED-IV Subject-Independent

Method	Mean Acc. (%)	Std. Dev. (%)
DGCNN [16]	52.8	9.2
BiDANN-S [17]	65.6	10.4
BiHDM [18]	69.0	8.7
PGCN [15]	**73.7**	**7.2**
Adaptive DCCA	70.7	8.9

D. Results on SEED-V

Table III presents the results of experiments on SEED-V. Adaptive DCCA achieved a mean accuracy of 88.4% with a standard deviation of 7.3% in the subject-dependent scenario, outperforming traditional DCCA (85.4%) and other baseline methods. In the subject-independent setup, Adaptive DCCA demonstrated its robustness with a mean accuracy of 55.23%, showing competitive performance against PGCN (61.78%).

TABLE III: SEED-V Subject-Dependent

Method	Mean Acc. (%)	Std. Dev. (%)
DCCA [9]	85.4	**7.1**
Adaptive DCCA	**88.4**	7.3

TABLE IV: SEED-V Subject-Independent

Method	Mean Acc. (%)	Std. Dev. (%)
PGCN [15]	**61.78**	**8.59**
Adaptive DCCA	55.23	10.87

The experimental results on SEED-V reveal an interesting observation: while Adaptive DCCA demonstrates superior performance in the subject-dependent scenario, it struggles in the subject-independent setting, achieving a mean accuracy of 55.23%, lower than PGCN (61.78%). This discrepancy suggests that eye movement signals contribute effectively in subject-dependent experiments but exhibit limited generalizability across subjects. Unlike EEG, which captures intrinsic physiological responses to emotions, eye movement patterns are highly influenced by individual habits and visual attention biases, making them less consistent across different participants. The reduced effectiveness of eye movement features in cross-subject settings highlights the challenge of modeling visual attention variability. Future improvements could involve domain adaptation techniques or personalized feature normalization to enhance the robustness of eye movement signals in subject-independent sentiment analysis.

E. Confusion Matrix Analysis on SEED-V

To further evaluate the performance of Adaptive DCCA, we analyzed the confusion matrices generated during the experiments on SEED-V under the subject-dependent setup. The confusion matrices for EEG-only, eye movement-only, BDAE, and Adaptive DCCA reveal important insights. EEG features performed well in identifying Fear and Disgust, but showed confusion in distinguishing Happy and Neutral. Similarly, eye movement features were effective for Fear but struggled with Sad and Disgust, resulting in lower precision for these categories. BDAE, which combines both modalities, demonstrated improvements in classifying Neutral and Disgust but was unable to fully leverage the complementary nature of EEG and eye movement features.

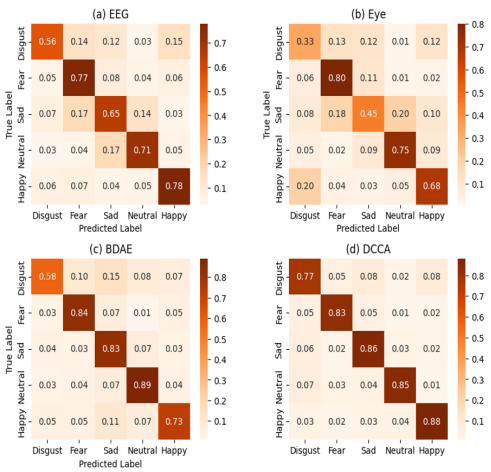

Fig. 2: Confusion matrices comparison for multimodal emotion recognition on SEED-V dataset (subject-dependent setting). Adaptive DCCA effectively reduces confusion by adaptively weighting EEG and eye movement signals.

Authorized licensed use limited to: OAKLAND UNIVERSITY. Downloaded on September 21,2025 at 16:47:12 UTC from IEEE Xplore. Restrictions apply.

Adaptive DCCA, however, outperformed these methods by significantly reducing misclassification rates across all emotion categories. The dynamic attention mechanism in Adaptive DCCA allowed the model to prioritize the most informative features from each modality, thereby enhancing overall precision and recall. Notably, Happy and Neutral—categories that were challenging for other methods—achieved much higher classification accuracy, highlighting the effectiveness of the adaptive fusion strategy. These results demonstrate the robustness and adaptability of Adaptive DCCA in integrating physiological modalities for emotion recognition.

V. CONCLUSION

In this study, we introduced Adaptive DCCA, a novel framework for multimodal sentiment analysis that integrates EEG and eye movement signals through an attention-based adaptive weighting mechanism. The proposed framework ef- fectively addresses the limitations of traditional DCCA by dynamically weighting the contributions of each modality, resulting in more robust and accurate feature fusion. Extensive experiments conducted on SEED-IV and SEED-V demon- strated the superiority of Adaptive DCCA under both subject- dependent and subject-independent conditions. Particularly, Adaptive DCCA achieved state-of-the-art performance by sig- nificantly enhancing classification accuracy and effectively reducing misclassification rates compared to baseline methods.

Looking ahead, several directions can further strengthen the capabilities of Adaptive DCCA. First, incorporating additional physiological signals—such as galvanic skin response (GSR) or photoplethysmography (PPG)—could potentially yield a more comprehensive picture of emotional states. Second, future work may explore meta-learning or transfer learning approaches, enabling models trained on one cohort of subjects to better generalize to others, especially in the context of variable eye movement patterns. Finally, a more detailed investigation into modality-specific attention patterns could reveal how attention weights shift in real time, offering deeper insight into individual differences and paving the way for personalized affective computing solutions.

The results emphasize the practical value of incorporating physiological signals, such as EEG and eye movement data, into sentiment analysis. Notably, our method opens avenues for numerous real-world affective computing applications, including improved human-computer interaction, adaptive user interfaces, emotional state monitoring in healthcare, and im- mersive virtual reality experiences. The demonstrated effec- tiveness in handling physiological data highlights potential extensions of Adaptive DCCA to other affective computing tasks requiring multimodal fusion.

ACKNOWLEDGMENT

The authors would like to acknowledge the technical and financial support provided by the Air Force Office of Scientific Research under grant FA9550-21-1-0224, and the Automotive Research Center (ARC) through Cooperative Agreement W56HZV-24-2-0001 with the U.S. Army DEVCOM Ground Vehicle Systems Center (GVSC) in Warren, MI. DISTRIBUTION STATEMENT A. Approved for public release; distribution is unlimited. OPSEC9558.

REFERENCES

[1] R. Das and T. D. Singh, "Multimodal sentiment analysis: A survey of methods, trends, and challenges," *ACM Comput. Surv.*, vol. 55, no. 13s, 2023.

[2] M. Zhu, Q. Wu, Z. Bai, Y. Song, and Q. Gao, "EEG-eye movement based subject dependence, cross-subject, and cross-session emotion recognition with multidimensional homogeneous encoding space alignment," *Expert Systems with Applications*, vol. 251, p. 124001, 2024.

[3] G. Andrew, R. Arora, J. Bilmes, and K. Livescu, "Deep canonical correlation analysis," in *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28*, ser. ICML'13, 2013, p. III–1247–III–1255.

[4] Z. Sun, P. K. Sarma, W. A. Sethares, and E. P. Bucy, "Multi-modal sentiment analysis using deep canonical correlation analysis," *CoRR*, vol. abs/1907.08696, 2019. [Online]. Available: http://arxiv.org/abs/1907.08696

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[6] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017.

[7] Y.-H. H. Tsai, P. P. Liang, A. Zadeh, L.-P. Morency, and R. Salakhutdinov, "Learning factorized multimodal representations," 2019. [Online]. Available: https://arxiv.org/abs/1806.06176

[8] Z. Sun, P. Sarma, W. Sethares, and Y. Liang, "Learning relationships between text, audio, and video via deep canonical correlation for multimodal language analysis," 2019. [Online]. Available: https://arxiv.org/abs/1911.05544

[9] W. Liu, J.-L. Qiu, W.-L. Zheng, and B.-L. Lu, "Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 715–729, 2022.

[10] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 1110–1122, 2019.

Authorized licensed use limited to: OAKLAND UNIVERSITY. Downloaded on September 21,2025 at 16:47:12 UTC from IEEE Xplore. Restrictions apply.

[11] K. Zhang, Y. Li, J. Wang, Z. Wang, and X. Li, "Feature fusion for multimodal emotion recognition based on deep canonical correlation analysis," *IEEE Signal Processing Letters*, vol. 28, pp. 1898–1902, 2021.
[12] T.-H. Li, W. Liu, W.-L. Zheng, and B.-L. Lu, "Classification of five emotions from eeg and eye movement signals: Discrimination ability and stability over time," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2019, pp. 607–610.
[13] Y. Lu, W.-L. Zheng, B. Li, and B.-L. Lu, "Combining eye movements and eeg to enhance emotion recognition." in *IJCAI*, vol. 15. Buenos Aires, 2015, pp. 1170–1176.
[14] W. Liu, W.-L. Zheng, and B.-L. Lu, "Emotion recognition using multimodal deep learning," in *Neural Information Processing: 23rd International Conference, ICONIP 2016, Kyoto, Japan, October 16–21, 2016, Proceedings, Part II 23*. Springer, 2016, pp. 521–529.
[15] M. Jin, C. Du, H. He, T. Cai, and J. Li, "Pgcn: Pyramidal graph convolutional network for eeg emotion recognition," *IEEE Transactions on Multimedia*, 2024.
[16] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532–541, 2018.
[17] Y. Li, W. Zheng, Y. Zong, Z. Cui, T. Zhang, and X. Zhou, "A bi-hemisphere domain adversarial neural network model for eeg emotion recognition," *IEEE Transactions on Affective Computing*, vol. 12, no. 2, pp. 494–504, 2018.
[18] Y. Li, L. Wang, W. Zheng, Y. Zong, L. Qi, Z. Cui, T. Zhang, and T. Song, "A novel bi-hemispheric discrepancy model for eeg emotion recognition," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 2, pp. 354–367, 2020.

Authorized licensed use limited to: OAKLAND UNIVERSITY. Downloaded on September 21,2025 at 16:47:12 UTC from IEEE Xplore. Restrictions apply.