Real-time EEG-based Neuroadaptive System for Emotion Regulation using Deep Learning in Mixed Reality Environments

Aditya Pandhare
Computer Science, NYUAD
ap7146@nyu.edu

Avinash Gyawali
Computer Science, NYUAD
ag8093@nyu.edu

Dachi Tarughishvili
Computer Science, NYUAD
dt2307@nyu.edu

Mbebo Nonna
Computer Science, NYUAD
mn3001@nyu.edu

Sashank Neupane
Computer Science, NYUAD
sn3006@nyu.edu

Advised by: Mohammed Eid, Hanan Salam

ABSTRACT

Emotion regulation is a vital cognitive process linked to mental well-being, yet traditional interventions often lack real-time responsiveness. Advances in neurofeedback and immersive technologies, such as EEG-based emotion detection and Extended Reality (XR), enable more dynamic support. However, many existing systems rely on static feedback or are confined to Virtual Reality, with limited exploration in Mixed Reality (MR). This research presents a real-time neuroadaptive system that integrates EEG-based emotion classification with MR interventions to support emotional self-regulation. We collect raw EEG data using a wireless 7-electrode X.on headset, streamed in real time into our classification pipeline. The data is processed by the ATCNet model, trained on 12-second EEG segments at 250 Hz with 50% overlap. Using the SEED-V dataset, we restructured five original emotion classes into a binary valence-based task, positive vs. negative, achieving 80% precision. This prediction drives feedback in the MR environment, where users interact with an emotionally intelligent avatar through explicit (e.g., avatar interactions/behaviour) and tactile (e.g., haptic feedback via the bHaptics TactGlove) modulation. Together, these components create a responsive, immersive framework for affective computing and emotion regulation.

This report is submitted to NYUAD's capstone repository in fulfillment of NYUAD's Computer Science major graduation requirements.

Capstone Project 2, Spring 2025, Abu Dhabi, UAE

KEYWORDS

neurofeedback, emotion regulation, neuroadaptive system, emotion detection, EEG, mixed reality, haptics, machine learning, deep learning, transformers

Reference Format:
Aditya Pandhare, Avinash Gyawali, Dachi Tarughishvili, Mbebo Nonna, and Sashank Neupane. 2025. Real-time EEG-based Neuroadaptive System for Emotion Regulation using Deep Learning in Mixed Reality Environments. In *NYUAD Capstone Project 2 Reports, Spring 2025, Abu Dhabi, UAE.* 10 pages.

1 INTRODUCTION

Emotion regulation is a fundamental cognitive process that influences individual well-being, interpersonal relationships, and societal dynamics. Ineffective emotion regulation has been implicated in various psychological disorders, including mood and anxiety disorders, emphasizing the critical need for innovative and accessible intervention strategies [15]. Therefore, it is vital for people to be aware of their emotional state and manage the volatility of extreme emotional responses.

Traditional approaches to emotion regulation, such as cognitive-behavioral therapy (CBT), mindfulness practices, and pharmacological treatments, have demonstrated effectiveness across diverse populations [16]. However, these methods often require substantial time, consistency, and access to trained professionals, posing challenges to their widespread adoption and scalability. As a result, there is increasing interest in technology-enhanced solutions that can complement and enhance these traditional interventions [4].

Among emerging technologies, EEG-based neurofeedback has attracted attention for its ability to provide real-time feedback on brain activity, enabling individuals to moderate their responses associated with emotion regulation [17]. Despite its promise, conventional neurofeedback systems are typically limited by their reliance on static and non-interactive visual or auditory stimuli. These limitations reduce user engagement and fail to replicate the dynamic complexity of real-world emotional experiences [2].

Extended Reality (XR) is another growing cluster of technologies which encapsulates Virtual Reality (VR), Mixed Reality (MR), and Augmented Reality (AR). XR offers an extremely interactive environment, seamlessly blending virtual and real-world elements together. Within the context of emotion regulation, there are a myriad of studies which test emotion regulation in VR, however very few in the other realms such as MR.

Unlike VR which transports the user into an entirely different world, MR integrates virtual digital elements/objects into real-world environments, offering a transformative platform for advancing emotion regulation interventions. MR's adaptability and capacity to create immersive, personalized environments have demonstrated success in applications such as stress reduction and exposure therapy [12]. Therefore, by leveraging MR's unique capabilities, it becomes possible to simulate contextually rich scenarios tailored to individual needs, thereby enhancing user engagement and efficacy in emotion regulation training. Furthermore, the user is present within their own real-world environment, allowing it to remain noninvasive in daily routine life.

Physiological signals provide a window into emotional states, offering critical insights into their underlying neural mechanisms. While techniques such as heart rate variability and galvanic skin response have been widely employed, EEG stands out for its high temporal resolution and ability to capture intricate patterns of brain activity [18]. These characteristics make EEG particularly suitable for emotion regulation research, as it supports real-time monitoring and adaptive feedback.

The development of a neuroadaptive system that integrates EEG-based feedback with sophisticated MR interfaces presents several challenges. Such a system must interpret complex EEG signals accurately while delivering contextually appropriate, personalized feedback in an immersive environment. Furthermore, it must adapt to individual emotional patterns to ensure its efficacy and relevance [3].

This research aims to bridge these challenges by developing an EEG-based neuroadaptive system that leverages deep learning and MR to detect emotional signals and facilitate emotion regulation. By integrating real-time EEG neurofeedback with interactive, user-specific environments, the proposed system aspires to create a more effective and engaging approach to mental health interventions. This work represents a step toward the broader goal of advancing interdisciplinary solutions at the intersection of neuroscience, machine learning, and immersive technologies to improve affective health outcomes.

2 RELATED WORK

2.1 Emotion Regulation

The integration of EEG-based neurofeedback with mixed reality for emotion regulation is a burgeoning field that bridges neuroscience, psychology, and immersive technologies. This section reviews pivotal studies that shape our research approach and methodology.

Neurofeedback and Emotion Regulation: Extensive research has shown that EEG-based neurofeedback can be effective for emotion regulation. Li et al. (2023) demonstrated that neurofeedback training could enhance emotion regulation by employing decoded EEG feedback for cognitive reappraisal tasks [8]. Similarly, Huang et al. (2023) provided evidence that neurofeedback, through a brain-computer interface, could significantly improve the ability to regulate emotions by providing real-time EEG feedback, enhancing individual strategies for emotional control [6].

Extended Reality in Psychological Interventions: The use of XR, including virtual reality (VR), has been explored for psychological interventions. Liang et al. (2023) discussed the effectiveness of an EEG-based VR system that adapts the virtual scenes dynamically based on the user's emotional state, thereby aiding in emotion regulation [9].

2.2 Mixed Reality

Neuroadaptive systems are a group of adaptive systems that employ EEG signals to generate personalized experiences. It is based on the cybernetic theory, which involves steps such as physiological data acquisition and processing, transformation in system response, and shaping the expected psychophysiological response. This system helps users achieve optimal performance, immersion, and engagement. It has potential applications in the healthcare industry to deliver customized therapies suited to patient's psychophysiological conditions. It also aids in technology-based decision-making to assist cognitive, information processing, motivation, and metacognition abilities [3]. ARCADIA is a Mixed Reality platform designed to enhance emotional regulation and self-compassion through gamified therapeutic activities using a virtual agent [14]. This project addresses the Therapeutic Application of Mixed Reality Gamification for user motivation and Personalization through Biofeedback. CAEVR (Context-Aware Empathy in Virtual Reality) explores the integration of biosignals-driven emotion recognition into virtual reality experience and explores the use of empathic virtual agent in modulating positive emotions.[5]

2.3 Deep Learning

Deep Learning has become a cornerstone in emotion recognition research, offering advanced methods to analyze complex physiological signals such as EEG. By leveraging its ability to automatically extract and learn hierarchical features, deep learning has outperformed traditional machine learning approaches in tasks that require the identification of nuanced patterns.

Convolutional Neural Networks (CNNs) are widely used for EEG signal analysis due to their strength in capturing spatial features. CNN-based models have demonstrated significant success in emotion classification tasks by extracting spatial dependencies between EEG channels. For example, models like EEGNet have effectively utilized lightweight convolutional architectures to analyze spatial interactions in brain signals, providing interpretable features relevant to emotion recognition [7].

In recent years, transformers have emerged as a promising alternative for processing sequential data. Initially developed for natural language processing, transformers use self-attention mechanisms to model long-range dependencies, making them particularly suitable for tasks requiring the integration of both spatial and temporal information. Transformers have been adapted for EEG-based emotion recognition by combining spatial information across electrodes with temporal dynamics, yielding state-of-the-art performance in various classification tasks [10].

ATCNet, a physics-informed transformer-based architecture originally designed for motor imagery classification, has demonstrated success in capturing spatial-temporal dependencies in EEG signals. Although its primary application was not in emotion recognition, the core idea behind ATCNet-using attention mechanisms to prioritize relevant features-aligns well with the challenges of classifying emotional states. This ability to dynamically focus on important spatial and temporal features made it a compelling choice for our research. By adapting ATCNet for emotion recognition, we aim to leverage its robust feature extraction capabilities to classify five distinct emotional states-happiness, sadness, neutral, anger, and disgust-using the SEED-V dataset [1].

The adoption of deep learning models like CNNs, transformers, and ATCNet has opened new frontiers in emotion recognition by addressing key challenges such as the complexity of EEG data and the variability in emotional patterns. This research builds upon these advancements by employing a modified version of ATCNet to classify emotional states within a Mixed Reality (MR) environment, providing neuroadaptive feedback for emotion regulation and bridging the gap between advanced computational methods and real-world psychological health applications.

3 PIPELINE ARCHITECTURE

The proposed framework that will be utilized in this study is a closed-loop feedback system which consists of 4 main modules: EEG data collection, Data Processing, Deep Learning Model, and an MR Environment as shown in Figure 1. In this section, we discuss each of these modules in more detail and how they contribute to the overall functionality of the system.

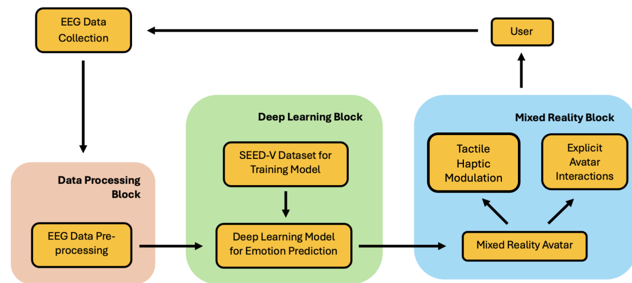

Figure 1: Pipeline for real-time EEG-based emotion classification and MR adaptive environment.

The interface establishes a seamless connection between EEG data acquisition, the emotion classification backend, and the Unity-based Mixed Reality (MR) environment. It operates on a modular TCP/IP client-server architecture, enabling real-time communication across system components. A web-based dashboard (Figure 2) was developed to monitor data streams, classification outputs, and system status in real time.

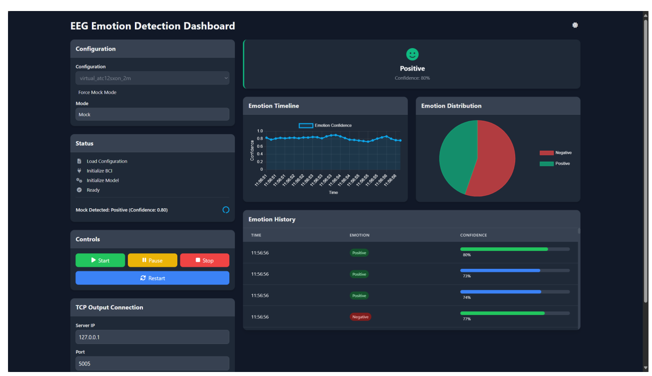

Figure 2: Web interface for monitoring and controlling the emotion detection system.

The data flow is managed through a modular architecture:

(1) The backend handles EEG acquisition and preprocessing either in live mode (LSL) using a wireless X.on headset or a virtual emulator (testing).
(2) The trained model then classifies emotion outputs which are streamed through the TCP server (port 5005), to the MR client.

(3) The Unity MR application runs a TCP client that receives predictions in real time, triggering corresponding visual and haptic feedback.

This modular, bidirectional setup allows the MR environment to adapt dynamically to the user's emotional state, with an end-to-end latency of approximately 6s.

4 DEEP LEARNING METHODOLOGY

4.1 Datasets

The EEG data is gathered from the SEED-V dataset [11]. This dataset contains raw EEG from 16 participants where each participant watched 45 different short movies in total (15 movies/session for 3 sessions). Each movie was aimed to induce a particular emotion from a class of five different emotions: happiness, sadness, neutral, fear, disgust.

4.2 Data Engineering and Pre-processing

The EEG data is preprocessed to ensure optimal signal quality for emotion classification. A notch filter at 50 Hz is applied to remove power line interference, which typically occurs at this frequency in many regions. A bandpass filter with a range of 0.5 to 70 Hz is used to retain relevant brainwave frequencies, removing slow drifts below 0.5 Hz and high-frequency noise above 70 Hz, both of which are unlikely to contain meaningful emotional information. The data is then resampled to 250 Hz to ensure consistent temporal resolution for further analysis as well as to match the streaming rate of the 7 electrode X.on device. Lastly, participants in the SEED-V dataset rated each video clip on a scale from 0 to 5, indicating how well the intended emotion was induced. To ensure higher label reliability, we excluded samples with ratings of 2 or below. This filtering step retained 89.72% of the original dataset, preserving a substantial amount of usable data while improving label quality. To assess the impact of this pre-processing decision, we evaluated one of our benchmark models (ATCNet) on both filtered and unfiltered datasets as shown in Figure 3. This assessment was done on the 5 output classes from the SEED-V dataset. The filtered configuration consistently achieved higher average validation accuracy, demonstrating the effectiveness of this quality-based filtering approach.

In addition, we restructured the emotion classes into two valence-based categories: *positive* and *negative*. This decision was made to simplify real-time classification and reduce the risk of misclassification in a deployed system. We focused specifically on the valence dimension-ranging from unpleasant to pleasant-by grouping all negative emotions (disgust, fear, sadness) into one class, and combining neutral and positive emotions into the other. To ensure class balance, an equal number of samples were drawn from each group during training.

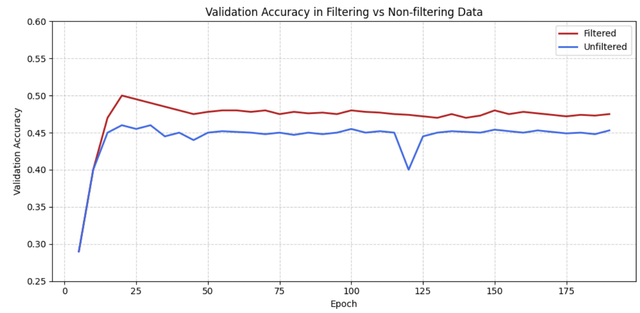

Figure 3: Average Accuracy of ATCNet on SEED-V (5 emotion classes) before and after filtering data.

For model input, overlapping data windows were created to help the network capture temporal dynamics across consecutive EEG segments. Finally, the preprocessed data was formatted into `.h5` datasets for efficient training and storage, enabling high-throughput data handling throughout the pipeline.

4.3 Model Selection and Design

We evaluated multiple deep learning architectures for EEG-based emotion classification, including GRUs, LSTMs, CNN variants, and transformer-based models. Among them, ATCNet was selected for its hybrid CNN-transformer architecture, which effectively captures both spatial and temporal dependencies in EEG signals. Although initially designed for motor imagery classification, ATCNet proved highly adaptable to affective computing tasks.

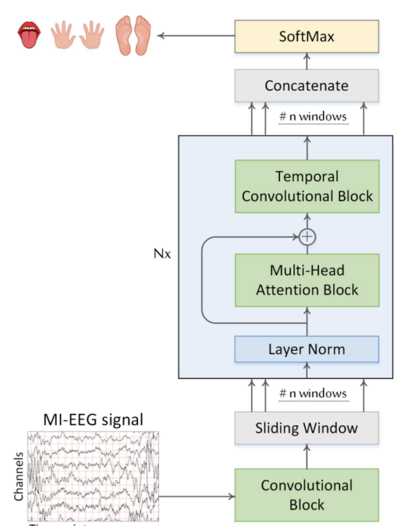

Figure 4: ATCNet architecture consisting of convolutional, attention, and temporal convolutional blocks

Convolutional (CV) Block: The convolutional block serves as the initial stage of spatiotemporal encoding. It applies three successive convolutional operations: a temporal convolution that captures local time-domain signal patterns; a depthwise spatial convolution that learns spatial dependencies across EEG channels; and a final temporal convolution for further temporal abstraction. The use of standard 2D convolutions here enables richer feature representations and pooling layers interspersed between these convolutions reduce the temporal resolution. The aim is to effectively transform raw EEG signals into a high-dimensional temporal sequence of feature vectors capturing patterns across time and space. These representations serve as the input for the subsequent attention module.

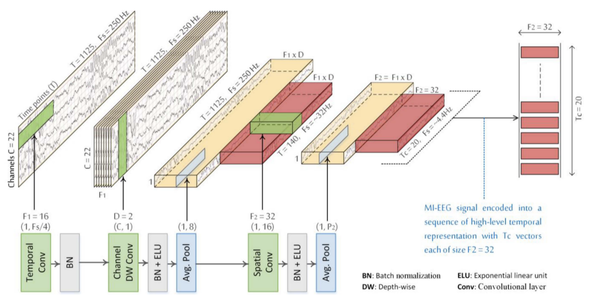

Figure 5: Convolutional block for spatial-temporal feature encoding

Attention (AT) Block: The attention block uses multi-head self-attention (MSA) to dynamically weigh segments of the temporal sequence based on contextual relevance. Each head computes scaled dot-product attention independently, enabling the model to capture diverse temporal dependencies in parallel. The outputs are aggregated into a refined feature representation. Residual connections and layer normalization support stable training and faster convergence. This block helps the model focus on emotionally salient EEG regions without relying on fixed receptive fields.

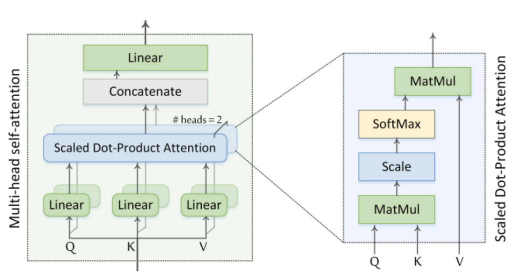

Figure 6: Multi-head self-attention block

Temporal Convolutional (TC) Block: The temporal convolutional block leverages dilated causal convolutions within stacked residual blocks to model long-range temporal dependencies. Causal convolutions enforce temporal order, ensuring that predictions at a given time step do not leak future information-crucial for real-time applications. Dilation exponentially increases the receptive field, enabling the network to capture patterns over extended durations without increasing depth. Residual connections allow stable gradient flow and maintain representational fidelity across layers. With a receptive field calibrated to handle fixed-length input windows, the TC block is particularly well-suited for decoding slow-evolving affective states in EEG data.

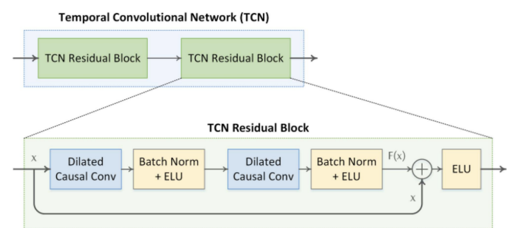

Figure 7: Temporal convolution block with residuals

The concatenated outputs from all windows are passed through a fully connected layer followed by a SoftMax classifier for final emotion prediction.

To validate model performance, we trained each architecture on the SEED-V dataset for 5-class emotion classification. As shown in Figure 8, ATCNet achieved the highest average validation accuracy (46.7%), outperforming baseline models including EEGNet and CNN-LSTM hybrids. This confirms its strong ability to extract discriminative spatial-temporal features from EEG signals.

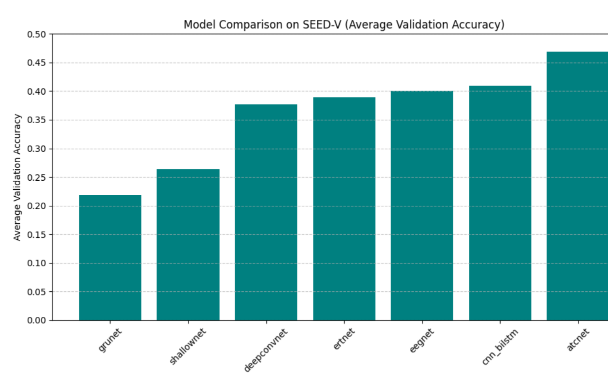

Figure 8: Model comparison based on average validation accuracy on SEED-V

4.4 Hyperparameter Optimization

To further optimize ATCNet's performance, we conducted extensive tuning of key parameters including learning rate, batch size, chunk duration, and overlap ratio. Each parameter was evaluated independently using grid search while holding other settings constant. The results are visualized in the following figures.

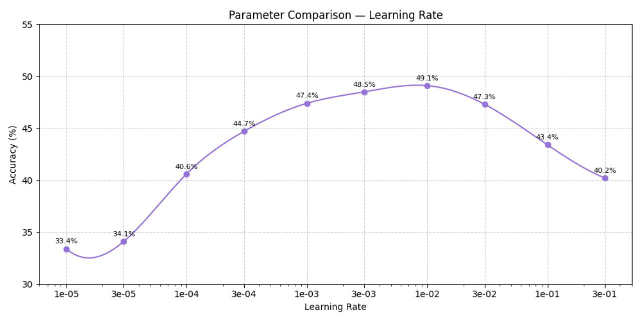

Figure 9: Validation accuracy across different learning rates (log scale)

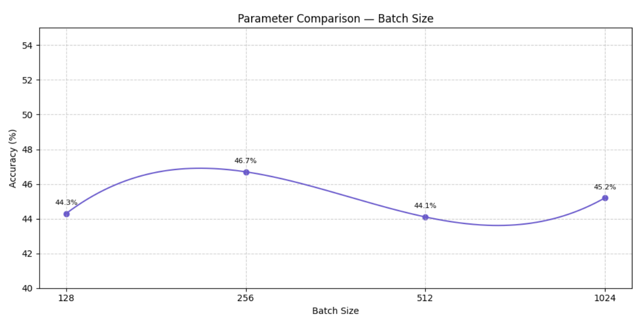

Figure 10: Validation accuracy across different batch sizes

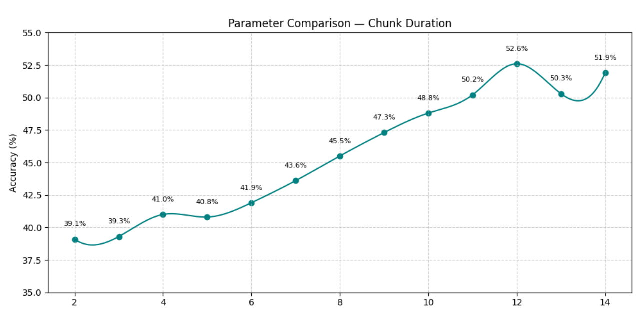

Figure 11: Validation accuracy across varying EEG chunk durations

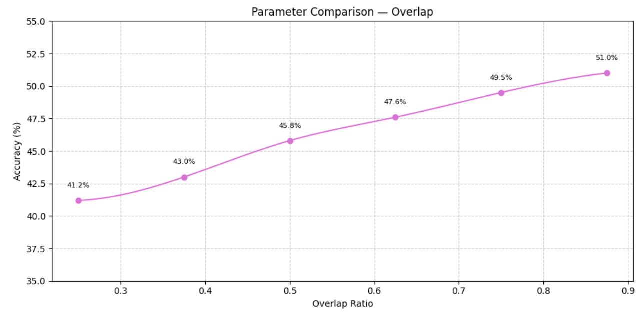

Figure 12: Validation accuracy across different overlap ratios

Based on these hyperparameter optimizations, here are the optimized values used for our training:

Hyperparameter	Selected Value
Chunk Duration	12 seconds
Overlap Ratio	50%
Batch Size	256
Learning Rate	1×10^{-2}

Table 1: Optimized hyperparameter configuration

We selected 12-second chunks as they yielded the highest accuracy and align with the temporal nature of emotional states, which typically unfold over longer time windows. While higher overlap ratios showed slight improvements in accuracy, they significantly increased dataset size and training time. Thus, a 50% overlap was chosen as the most resource-efficient option with minimal accuracy trade-off. The learning rate and batch size were selected based on peak performance during validation, offering a stable balance between convergence speed and generalization.

4.5 Training Setup

All models were implemented using the PyTorch framework and trained on an NVIDIA RTX 4090 GPU running CUDA 12.6. We adopted a structured training pipeline that included logging, early stopping, checkpointing, and dynamic learning rate scheduling.

The dataset was split using a leave-one-subject-out (LOSO) strategy, where the model was trained on data from 15 out of the 16 participants in the SEED-V dataset and tested on the remaining participant. This subject-independent setup ensures generalization by evaluating the model on entirely unseen individuals, making it especially suitable for real-time pipelines where the end users are likely to be individuals the model has never seen before.

We used a batch size of 256 and trained models for up to 200 epochs, evaluating validation every 5 epochs. Early stopping was applied with a patience of 15 epochs to prevent overfitting.

Learning Rate Scheduler: A cosine annealing scheduler was used to progressively decrease the learning rate, allowing for fast convergence during initial epochs and finer adjustments later in training. This cyclic approach helps the model escape local minima and encourages stable convergence (see Figure 13).

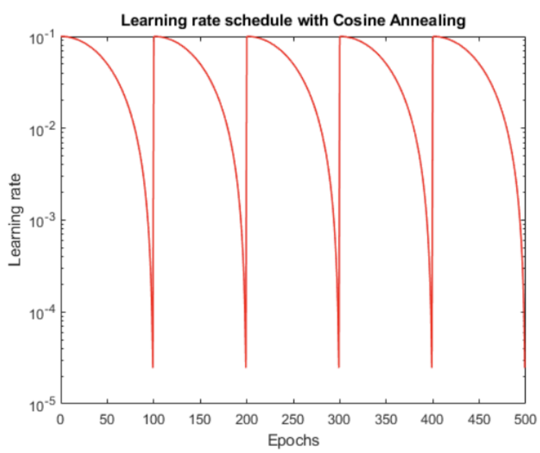

Figure 13: Cosine annealing learning rate schedule

Loss Function: Focal Loss We used the focal loss function to address the class imbalance between positive and negative emotion samples. Compared to cross-entropy loss, focal loss down-weights easy examples and focuses more on hard, misclassified samples, making it well-suited for our task. The focal loss is defined as:

$$\mathcal{L}_{\text{focal}} = -\alpha_t (1 - p_t)^\gamma \log(p_t)$$

where p_t is the model's estimated probability for the true class, α_t is a weighting factor for class t, and γ is the focusing parameter. We used $\gamma = 2$, and class-specific weights of $[1.2, 0.6]$ for the negative and positive classes respectively. This helped the model better distinguish between easily confusable positive chunks and harder negative ones. These weights were manually optimized through iterative experimentation across several training runs to achieve the best class-wise performance balance.

Optimizer: The model was trained using the Adam optimizer with an initial learning rate of 1×10^{-2}. Loss curves and validation accuracy were logged and plotted throughout training. Best model checkpoints were saved automatically based on validation loss and accuracy.

5 MIXED REALITY METHODOLOGY

After the user's physiological signals are classified, the results are sent via a TCP server to the Oculus Quest 3 headset, triggering a personalized gameplay loop based on the user's emotional state. The central interaction involves petting a virtual animal to regulate its emotional parameters—valence and arousal. The simulation features an emotionally intelligent pet avatar with occlusion-aware navigation, enabled through a custom occlusion shader that allows the pet to initially hide behind real-world objects before emerging into the user's field of view. The experience is further enhanced by a physics-driven animation system that blends the user's physical force with the pet's programmed responses, creating fluid and lifelike interactions. A gesture recognition system allows users to raise their palms to beckon the pet, adding an intuitive and embodied control mechanic. The pet's color and facial expression dynamically change based on the quality and intensity of the petting interaction, offering continuous affective feedback. Tactile sensations are delivered via the bHaptics TactGlove, deepening immersion through synchronized physical feedback.

Figure 14: Experimental Setup with EEG headset, Quest 3 and bHaptics TactGloves

This mixed reality simulation draws on principles from socially intelligent agents and therapeutic robotics to create an empathic MR agent capable of adaptive emotional regulation. The design is inspired by social robots and aligns with the three key dimensions of realistic mixed reality experiences, as defined by Skarbez: immersion, coherence, and

(a) Angry avatar

(b) Eating avatar

(c) Happy avatar

(d) Sad avatar

Figure 15: A 2x2 grid of images

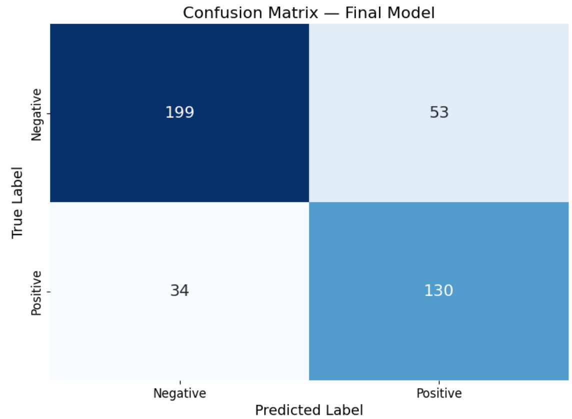

Figure 16: Confusion Matrix of the Binary Classification Model

Class	Precision	Recall	F1 Score
Negative	0.854	0.790	0.821
Positive	0.710	0.793	0.749
Macro Avg	0.782	0.791	0.785
Weighted Avg	0.797	0.791	0.792

Table 2: Precision, Recall, and F1 Scores across emotion classes based on the final model's predictions

extent of world knowledge [13]. Immersion refers to the user's sense of spatial presence; coherence ensures consistency across sensory modalities, enhancing the plausibility of the experience; and extent of world knowledge integrates real-world spatial awareness into the MR environment. The pet agent delivers empathic responses through multimodal cues—verbal, nonverbal, and tactile—creating a believable and emotionally engaging interaction loop. Additional cozy-game interactions, such as collecting coins using a pointing laser, provide low-stress, high-reward activities that further promote emotional well-being and user engagement.

6 RESULTS

Our final model, trained using the optimized pipeline and evaluated under a leave-one-subject-out (LOSO) strategy, demonstrated strong generalization performance across emotion classes. As shown in Figure 16, the confusion matrix reveals a relatively balanced distribution of correct predictions, with the model performing particularly well on negative samples (Precision: 0.854, Recall: 0.790). Positive samples were slightly more challenging, though still predicted with reasonable accuracy (F1 Score: 0.749). The macro average F1 score of 0.785 and weighted F1 score of 0.792 indicate overall consistent performance across both classes.

These results validate the effectiveness of our approach-combining class-weighted focal loss, preprocessing filters, and architecture-level optimizations-to improve classification in a real-time, subject-independent EEG-based emotion recognition setting.

7 DISCUSSION

Our implementation of the ATCNet model for EEG-based emotion classification achieved a weighted F1-score of 0.792 for binary valence classification. This section discusses key findings and potential areas for improvement in our methodology.

7.1 Methodological Strengths

The systematic parameter optimization approach proved highly effective, particularly our exploration of different window sizes and overlap percentages. While our testing indicated that window sizes of 12-14s and high overlap percentages (80-90%) showed promising performance, we ultimately selected a 12s window size and 50% overlap for our final model implementation to balance computational efficiency with classification accuracy. This trade-off was necessary to ensure near real-time processing required for seamless integration with mixed reality environments.

Data filtering significantly improved classification performance, demonstrating the critical importance of preprocessing in EEG analysis. As shown in Figure 3, filtered data consistently outperformed unfiltered data by approximately

2.5% across our experiments, helping to reduce noise and enhance signal quality.

The confusion matrix indicates that the model performs well on the binary classification task, demonstrating strong predictive balance between negative and positive emotional states. Without loss weighting, the model consistently over-predicted positive emotions—likely due to their higher signal clarity and prevalence in the dataset. By introducing class-specific weights [1.2, 0.6] within the focal loss function, we were able to correct this bias, encouraging the model to focus more on the harder-to-classify negative samples. This adjustment significantly improved class-level balance, as reflected in the nearly symmetric precision and recall scores across both classes.

7.2 Limitations and Future Improvements

Despite the promising results, several limitations warrant consideration in future work:

- **Dataset Constraints:** The SEED-V dataset, while valuable for our initial implementation, has inherent limitations in ecological validity and generalizability. It was collected in controlled laboratory conditions with 16 participants using a high-density EEG system (62 electrodes), whereas our deployment target is the X.on 7-electrode EEG headset in less controlled environments. Future iterations should incorporate more refined datasets such as DEAP, which offers multimodal data and greater diversity in emotion elicitation paradigms, or collect custom data using the actual deployment hardware.
- **Emotion Classification Complexity:** While our current binary valence classification (positive/negative) provides a functional foundation, expanding to a more nuanced multi-class emotion recognition system would significantly enhance the system's capabilities. Future work should aim to classify a broader spectrum of emotional states (e.g., the five discrete emotions from SEED-V: happiness, sadness, fear, disgust, and neutral) and adapt the model architecture to handle this increased classification complexity while maintaining real-time performance.
- **Custom Model Architecture:** Although ATCNet performed well in our experiments, developing a custom architecture specifically optimized for our use case could yield improved performance. A tailored model could be designed to be more lightweight for on-device processing on the MR headset, and specifically optimized for the limited spatial resolution of 7-channel EEG data, potentially improving both computational efficiency and classification accuracy.
- **Empirical Validation:** The current system requires comprehensive validation with test subjects to objectively assess its effectiveness in emotion regulation. Future work should include randomized controlled trials measuring pre-post intervention changes using established emotion regulation assessment tools (e.g., Difficulties in Emotion Regulation Scale), alongside physiological measures (HRV, GSR) and subjective experience reports. Such studies would provide crucial insights into real-world efficacy beyond classification accuracy.
- **Enhanced MR Interactions:** The current mixed reality interaction paradigms could be expanded to create more engaging and effective emotion regulation experiences. This could include more sophisticated environmental adaptations based on emotional intensity, personalized feedback mechanisms that evolve over multiple sessions, and integration with other biofeedback modalities such as breathing guidance to create a more holistic approach to emotion regulation.

8 CONCLUSION

This project demonstrates the feasibility of integrating EEG-based emotion recognition with mixed reality environments for emotion regulation training. By achieving an F1-score of 0.792 for binary valence classification while maintaining real-time processing capabilities (approximately 6s latency), we have established a foundation for neuroadaptive systems that respond dynamically to users' emotional states. The successful integration of deep learning, EEG signal processing, and mixed reality technologies represents a promising approach for applications in emotional intelligence training and mental wellbeing interventions. Future work focusing on model optimization for specific hardware constraints, real-world validation studies, enhanced interaction paradigms, and expansion to multi-class emotion recognition will be crucial for advancing this technology toward practical applications.

REFERENCES

[1] Hamdi Altaheri, Ghulam Muhammad, and Mansour Alsulaiman. 2023. Physics-Informed Attention Temporal Convolutional Network for EEG-Based Motor Imagery Classification. *IEEE Transactions on Industrial Informatics* 19, 2 (2023), 2249–2258. https://doi.org/10.1109/TII.2022.3197419

[2] Pasquale Arpaia, Damien Coyle, Giovanni D'Errico, Egidio De Benedetto, Lucio De Paolis, Naomi Du Bois, Sabrina Grassini, Giovanna Mastrati, Nicola Moccaldi, and Ersilia Vallefuoco. 2022. Virtual Reality Enhances EEG-Based Neurofeedback for Emotional Self-regulation. 420–431. https://doi.org/10.1007/978-3-031-15553-6_29

[3] Francesco Chiossi and Sven Mayer. 2023. How Can Mixed Reality Benefit From Physiologically-Adaptive Systems? Challenges and Opportunities for Human Factors Applications. arXiv:cs.HC/2303.17978

[4] Desirée Colombo, Javier Fernández-Álvarez, Azucena García Palacios, Pietro Cipresso, Cristina Botella, and Giuseppe Riva. 2019. New

Technologies for the Understanding, Assessment, and Intervention of Emotion Regulation. *Frontiers in Psychology* 10 (2019). https://doi.org/10.3389/fpsyg.2019.01261

[5] Kunal Gupta, Yuewei Zhang, Tamil Selvan Gunasekaran, Nanditha Krishna, Yun Suen Pai, and Mark Billinghurst. 2024. CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 30, 5 (March 2024), 2671–2681. https://doi.org/10.1109/TVCG.2024.3372130

[6] Weichen Huang, Wei Wu, Molly V. Lucas, Haiyun Huang, Zhenfu Wen, and Yuanqing Li. 2023. Neurofeedback Training With an Electroencephalogram-Based Brain-Computer Interface Enhances Emotion Regulation. *IEEE Transactions on Affective Computing* (2023). https://typeset.io/papers/neurofeedback-training-with-an-electroencephalogram-based-1nk4ojhg

[7] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. 2018. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of neural engineering* 15, 5 (2018), 056013.

[8] Linling Li, Xu Gui, Gang Huang, Lei Zhang, Xue Han, Zhen Liang, and Zhiguo Zhang. 2023. Decoded EEG Neurofeedback-Guided Cognitive Reappraisal Training for Emotion Regulation. *bioRxiv* (2023). https://typeset.io/papers/decoded-eeg-neurofeedback-guided-cognitive-reappraisal-36xohi6p

[9] Hui Liang, Shiqing Liu, Yi Wang, Junjun Pan, and Jialin Fu. 2023. EEG-Based VR Scene Adaptive Generation System for Regulating Emotion. In *International Conference on Virtual Reality*. https://typeset.io/papers/eeg-based-vr-scene-adaptive-generation-system-for-regulating-3bdkwqq8

[10] Ruixiang Liu, Yihu Chao, Xuerui Ma, Xianzheng Sha, Limin Sun, Shuo Li, and Shijie Chang. 2024. ERTNet: an interpretable transformer-based framework for EEG emotion recognition. *Frontiers in Neuroscience* 18 (2024), 1320645.

[11] Wei Liu, Jie-Lin Qiu, Wei-Long Zheng, and Bao-Liang Lu. 2021. Comparing Recognition Performance and Robustness of Multimodal Deep Learning Models for Multimodal Emotion Recognition. *IEEE Transactions on Cognitive and Developmental Systems* (2021).

[12] Eleni Mitsea, Athanasios Drigas, and Charalabos Skianis. 2023. Digitally Assisted Mindfulness in Training Self-Regulation Skills for Sustainable Mental Health: A Systematic Review. *Behavioral Sciences* 13, 12, 1008. http://proxy.library.nyu.edu/login?qurl=https%3A%2F%2Fwww.proquest.com%2Fscholarly-journals%2Fdigitally-assisted-mindfulness-training-self%2Fdocview%2F2904634844%2Fse-2%3Faccountid%3D12768 Copyright - © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License; Last updated - 2024-03-20.

[13] Richard Skarbez, Missie Smith, and Mary C Whitton. 2021. Revisiting Milgram and Kishino's reality-virtuality continuum. *Frontiers in Virtual Reality* 2 (2021), 647997.

[14] Jose Luis Soler-Dominguez, Samuel Navas-Medrano, and Patricia Pons. 2024. ARCADIA: A Gamified Mixed Reality System for Emotional Regulation and Self-Compassion. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 544, 17 pages. https://doi.org/10.1145/3613904.3642123

[15] Franca Tonnaer, Linda van Zutphen, Adrian Raine, and Maaike Cima. 2023. Chapter 7 - Amygdala connectivity and aggression. In *Brain and Crime*, Hanna Swaab and Gerben Meynen (Eds.). Handbook of Clinical Neurology, Vol. 197. Elsevier, 87–106. https://doi.org/10.1016/B978-0-12-821375-9.00002-5

[16] Thomas Llewelyn Webb, E. Miles, and P. Sheeran. 2012. Dealing with feeling: a meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychological bulletin* 138 4 (2012), 775–808. https://doi.org/10.1037/a0027600

[17] Minchang Yu, Yicai Bai, and Yingjie Li. 2023. Emo-regulator: An emotion-regulation training system fusing virtual reality and EEG-based neurofeedback. In *45th Annual International Conference of the IEEE Engineering in Medicine Biology Society, EMBC 2023, Sydney, Australia, July 24-27, 2023*. IEEE, 1–4. https://doi.org/10.1109/EMBC40787.2023.10340975

[18] Jianhua Zhang, Zhong Yin, Peng Chen, and Stefano Nichele. 2020. Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion* 59 (2020), 103–126. https://doi.org/10.1016/j.inffus.2020.01.011