EmoLoop: A Bi-Directional System for Emotion-Driven Interaction Between Remote Audiences and Live Performers

Nicusor Amarie
Transilvania University of Brașov
Brașov, Romania
nicusor.amarie@gmail.com

Gianluca Fadda
Università degli Studi di Cagliari
Cagliari, Italy
gianluca.fadda@unica.it

Anderson Augusto Simiscuka
Dublin City University
Dublin, Ireland
andersonaugusto.simiscuka@dcu.ie

Vlad Popescu
Transilvania University of Brașov
Brașov, Romania
vlad.popescu@unitbv.ro

Maurizio Murroni
Università degli Studi di Cagliari
Cagliari, Italy
maurizio.murroni@unica.it

***Abstract*—This paper presents *EmoLoop*, a bi-directional interactive system designed to create real-time emotional feedback loop between performers and a distributed audience during music or theatre events. The proposed system collects emotional data from the audience using different feedback methods, such as EEG signals from wearable headbands, facial expression analysis through VR headsets, and interaction with smartphones. These emotional states, expressed as valence/arousal values or emotion categories, are used to control various parameters of the performance, like spatial audio, lighting, and scenographic elements, both on stage and in virtual environments. EmoLoop also includes a real-time feedback system that gives performers useful cues based on the audience's emotions. Communication between components is managed with protocols like OSC, MQTT, WebSockets, MIDI, and ArtNet to ensure low-latency transmission. The architecture is modular and flexible, and the paper explains its components, integration, and possible use cases. EmoLoop offers a new approach to participatory performance, where emotions play a central role in shaping the artistic experience, in line with current research in the Internet of Sounds (IoS) and Internet of Musical Things (IoMusT).**

***Index Terms*—Internet of Sounds, emotion recognition, audience feedback, immersive audio, VR performance, real-time audio interaction, affective computing, hybrid performance, Open Sound Control, text-to-emotion mapping, sentiment-based synthesis, spatial audio, networked music performance, artistic installations, human-computer interaction in music**

I. Introduction

The integration of emotion recognition into interactive media and performance arts is gaining relevance as new sensing technologies and machine learning techniques allow deeper audience engagement and adaptive storytelling. Recent work in affective computing has emphasized the potential of emotional feedback loops to influence musical improvisation, virtual environments, and hybrid performances in both artistic and therapeutic settings [1], [2].

EmoLoop introduces a novel architecture that enables real-time, bi-directional interaction between performers and distributed audiences by leveraging *multimodal emotion recognition* and *adaptive media control*. The proposed system intends to integrate a variety of sensing modalities—including facial expression detection, EEG data acquisition, and smartphone-based interactions, to interpret the affective state of audience members and modulate the performance accordingly.

In immersive settings like live concerts or theatrical installations, understanding how emotion manifests in neural, facial, and behavioral patterns is essential to personalize the user experience and enhance presence [3], [4]. Platforms such as Meta Quest and Muse 2 offer real-time access to these biosignals through SDKs and APIs, providing a viable infrastructure for emotion-adaptive systems [5]–[7].

The EmoLoop framework includes: (i) a user emotion detection module that fuses multimodal signals using deep learning technologies, (ii) a mapping engine translating emotional states into real scenographic modulation parameters, and (iii) a real-time feedback channel to performers via audio-visual cues. This modular pipeline enables new forms of audience-performer co-presence, which are especially relevant in hybrid or remote performance contexts as explored in recent studies on networked music and interactive dramaturgy [1], [8]–[10].

This work contributes to a broader research objective: investigating how multimodal emotion sensing can be integrated into live performative ecosystems, shaping expressivity and interaction across physical and virtual spaces. In this first stage, our focus is on designing and implementing the architectural foundation of such a system, defining sensing modalities, communication protocols, and mapping strategies for supporting future evaluative studies. EmoLoop aims to contribute both as a conceptual framework and an operational prototype to address these questions on the *Internet of Sounds (IoS)* and *Internet of Musical Things (IoMusT)* [11] topics.

In addition to affective computing and networked performance research, EmoLoop also aligns with developments in

Technology-Mediated Audience Participation (TMAP) [12], a field that examines how digital, gestural, or sensor-based tools enable audiences to actively influence live performance. Within this spectrum, EmoLoop adopts an affect-driven participation model, where the audience's emotional states modulate the performance rather than serving as explicit inputs.

This paper is structured as follows: Section 2 reviews related work in emotion-driven performance systems, multimodal affective sensing, and networked interaction. Section 3 presents the overall architecture of the EmoLoop system. Section 4 details hardware/software integration and preliminary technical evaluation. Section 5 outlines use-case scenarios. Section 6 discusses the current limitations and future work.

II. Related Work

The EmoLoop system builds upon interdisciplinary research across affective computing, multimodal emotion recognition, immersive performance, and emotion-driven generative systems. This section outlines the specific relevant contributions.

A. Affective Computing and Emotion-Aware Performance

Recent years have seen a growing interest in integrating emotion recognition technologies into performative contexts. Damiano et al. [1] explored strategies for designing affective interaction in live performances, highlighting the narrative and structural implications of emotional modulation. Similarly, Hopkins et al. [10] investigated trust-driven co-creation between audience and performers in interactive musical improvisations, suggesting that affective data can influence performative agency and dramaturgy.

In immersive virtual environments, Lee et al. [3] demonstrated how chat-driven sentiment feedback can enhance shared presence during live VR concerts. Their findings support the use of collective emotion as a modulating signal for scene rendering and audio mixing, which aligns with EmoLoop's real-time feedback architecture.

B. Multimodal Emotion Recognition

Multimodal affect recognition, which combines EEG, facial expression, and interactional data, is central to EmoLoop. Kalansooriya et al. [13] demonstrated EEG-based classification in affective gaming contexts, while Hao et al. [14] proposed a multimodal fusion approach integrating neural and audio features for real-time music emotion recognition.

Recent frameworks such as the Magic XRoom [4] and work by Su et al. [15] stress the importance of combining biosignals across multiple anatomical domains. Meanwhile, Wang et al. [16] examined emotion-triggered scenographic changes in theatrical VR performances, providing a compelling case for fusing emotion and dramaturgy in real-time.

C. Emotion-Driven Generative Systems

Emotion-adaptive systems have been implemented in music recommendation [17], [18], music video emotion tagging [19], and interactive art [20]. Zhang et al. [21] explored deep-learning-based music generation conditioned on emotion, using convolutional and LSTM architectures.

D. Technologies for Real-Time Feedback

Studies such as Colafiglio et al. [7], Kulkarni and Kaur [22] and Geetha et al. [23] provide overviews of real-time emotion recognition pipelines that integrate these technologies. Protocols like Open Sound Control (OSC) [24], ArtNet [25] and MIDI [26] are commonly used in multimedia interaction systems. WebSockets and MQTT are also explored as low-latency control layers in musical contexts [27] and are considered in EmoLoop's planned implementation.

III. System Architecture

EmoLoop advances the IoMusT by integrating audience-side wearable devices as active musical things within a distributed performance network. Through real-time communication protocols like OSC, MQTT, and RTP-MIDI, multimodal biosignals captured from remote users are translated into controllable parameters for sound, lighting, and scenography. This emotional mapping layer extends typical IoMusT interaction models, where control is performer- or instrument-driven, by positioning the audience as an affective source within the same networked ecosystem.

The EmoLoop system is designed as a modular, real-time architecture enabling bi-directional interaction between a live audience and performers by emotion recognition and immersive environment modulation. Figure 1 provides an overview of the architecture. This section details the envisioned system's primary components, organized into five functional modules.

A. Input Devices

1) Meta Quest V3 VR Headset: The Meta Quest V3 provides high-fidelity immersive rendering and captures biosignals via embedded sensors. Eye tracking, IMU, and facial expression data (via infrared camera arrays and Meta Quest SDK) allow real-time acquisition of Facial Expression Activations (FEAs), based on the Facial Action Coding System (FACS) [6]. Recent studies [28] validate the use of these signals for in-VR affect recognition, even under partial occlusion caused by the headset itself.

2) Muse 2 Brain-Sensing Headband: Muse 2 captures EEG signals across frontal and temporal lobes, providing neurophysiological insights into emotional states such as arousal and valence [13]. Integration is supported via the Muse SDK and other open source frameworks, such as BrainFlow and OpenBCI, offering compatibility with existing signal processing pipelines [5], [7], [29].

B. User Emotion Detection Module

This module fuses data from the VR headset and EEG headband. It includes:

- **Preprocessing:** Signal synchronization, noise reduction, and artifact removal across EEG, IMU, and FEAs.
- **Multimodal Emotion Recognition:** Based on deep learning architectures such as the multi-scale attention LSTMs with squeeze-and-excitation (SE) blocks proposed in [5]. Facial expression data is expected to be processed using methods akin to the EfficientNet-B0-based convolutional neural networks exploited in [28].

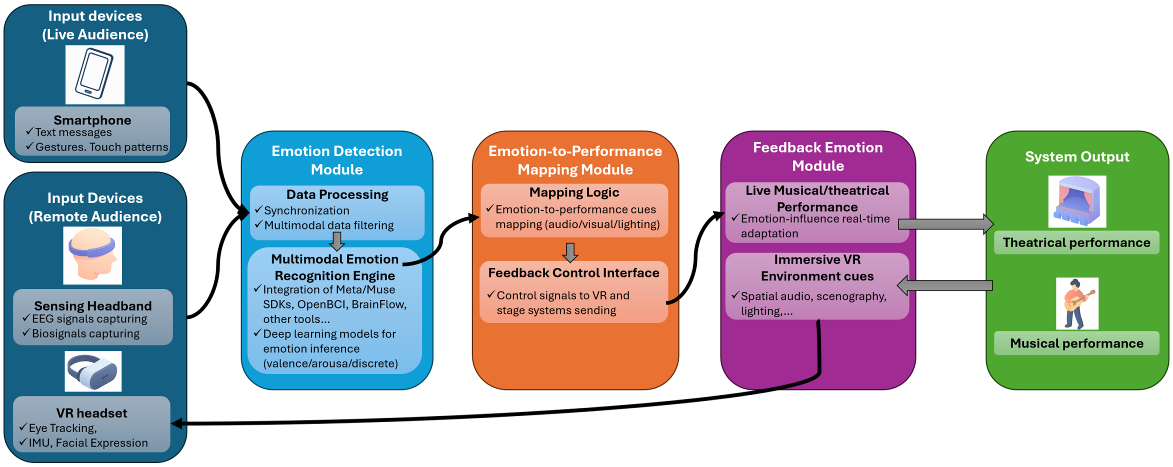

Fig. 1. EmoLoop Architecture

C. Emotion-to-Performance Mapping Module

In this stage, emotion vectors (e.g., valence/arousal, discrete emotion classes) are translated into performance control parameters. Their corresponding control signals are then transmitted on the stage using real-time communication protocols such as ArtNet, MIDI, OSC, and optionally WebSockets or MQTT chosen based on latency requirements and endpoint compatibility [8]. The mapping logic is synced in real-time also in the VR environment (e.g. Unity), using custom plugins and it can include for example:

- Spatial audio modulation (e.g., reverb, panning);
- Dynamic lighting and scenography (e.g., color tone based on collective emotion);
- Scene transitions triggered by the dominant audience feedback.

D. Emotion Feedback Module

Performers receive real-time cues (visual or auditory) reflecting aggregated emotional states of the audience. These could be presented for example via on-stage visual indicators, augmented HUD overlays or audio modulation feedback through in-ear monitors.

The bi-directional link enables co-creative and affect-driven performance modulation, aligning with concepts of participatory dramaturgy [1], [16].

E. System Output

The final output is twofold: (i) immersive VR Experience delivered via VR Headset, adapted to user emotion, (ii) dynamic music/theatre shaped by continuous audience feedback.

This architecture closes the loop between audience perception, emotional response, and performative expression—an innovation grounded in current affective computing frameworks and immersive media interaction.

IV. Implementation and Preliminary Setup

The EmoLoop system is currently under active development, with its modular components designed to support real-time emotional feedback for immersive and interactive performance environments.

A. Hardware Integration and Multimodal Pipeline

The first stage of implementation focuses on establishing a reliable multimodal sensing infrastructure. The Meta Quest V3 headset serves as the primary VR interface and provides streams from built-in eye tracking, inertial measurement unit (IMU), and facial expression sensors. Parallelly, the Muse 2 brain-sensing headband supplies continuous EEG signals representative of user cognitive and affective states.

These data streams are acquired via vendor-specific SDKs, such as Meta Quest SDK for Unity and BrainFlow for Muse, and will be aligned and synchronized in real-time to enable robust cross-modal feature extraction.

B. Software Architecture and Tools

A modular software architecture is being implemented using the following technologies:

- **Unity XR Framework:** for VR scene design, avatar integration, and spatial audio control.
- **Emotion Recognition Engine:** built with PyTorch for EEG and image-based inference pipelines, with plans to integrate multi-scale attention LSTM and EfficientNet-B0-based models [5], [28].
- **Communication Protocols:** Open Sound Control (OSC), MQTT, and MIDI will be evaluated for low-latency message delivery between emotion recognition modules and the VR/performance control subsystems.

C. Preliminary Technical Evaluation

Although a full experimental validation is not yet available, feasibility is corroborated through (i) reference to standard dataset performance, and (ii) simulated latency benchmarks.

For EEG-based recognition, studies on benchmark datasets such as DEAP and AMIGOS report a binary accuracy of around 70% for valence and around 80% for arousal using LSTM-based models. Indeed, Li et al. [30] report 72.06% for valence and 74.12% for arousal on DEAP with a convolutional–recurrent (LSTM) network, while a recent survey [31] summarizes attention-based LSTM-RNN results on AMIGOS of 79.4% and 83.3% for binary classification of valence and arousal, respectively.

Similarly, facial expression recognition in VR settings has been shown to reach accuracies in the 70–82% range even under occlusion using EfficientNet-B0 trained on EmoHeVRDB [6], [28]. These results establish realistic baselines for EmoLoop's emotion detection module.

Latency and jitter analyses from networked performance research and industrial communication studies confirm that, for on-stage feedback where sub-tens-of-milliseconds control is required, OSC over UDP and RTP-MIDI are the most suitable solutions. In [32] and [33] an OSC round-trip performance is reported to be consistently below 10 ms in musical networking contexts, while in [34] is emphasized that RTP-MIDI is designed for real-time musical interaction. By contrast, in hybrid or remote scenarios, WebSockets and MQTT become more viable: empirical evaluations such as [35] show that WebSocket typically remains within 40–90 ms depending on cloud routing, whereas MQTT ranges from 90–130 ms under comparable conditions. These latency bands remain acceptable for non-time-critical feedback and asynchronous control tasks.

Furthermore, previous implementations of multimodal interactive frameworks report end-to-end latencies which aligns with EmoLoop's design goals [11].

V. Artistic Use Case: Emotion-Driven Performance Scenario

To validate the EmoLoop system within a real-world creative context, we envision its deployment in a hybrid musical-theatrical performance scenario, where audience emotion becomes an integral part of the artistic language.

The use case centers around a live immersive concert set in a virtual scenography, where the EmoLoop architecture enables a dynamic feedback loop between audience and performers. Multiple audience members wear the Meta Quest V3 VR headset and the Muse 2 EEG band, exploring different scenes in a shared virtual environment.

Each scene corresponds to a narrative segment with its own musical and visual aesthetics. Emotional states derived from the audience (e.g., valence, arousal, discrete labels) dynamically influence:

- Spatial audio mix (reverb depth, source positioning);
- Ambient lighting and scenographic elements;
- Textural evolution and expressive variations in the music played live.

A. Interaction Mechanism

The emotional vectors inferred from EEG and facial signals are mapped in real-time to audiovisual parameters through the Emotion-to-Performance Mapping Module using OSC/MQTT messages to interface with the Unity environment, where the system updates scene visuals, audio, and stage lighting.

Performers receive simplified visual cues via stage displays or auditory feedback, helping them to modulate dynamics, tempo, or timbral expressions based on the collective emotional feedback of the audience.

B. Role of the Director and Audience Participation

An optional human director can mediate dramaturgy by choosing which user's view is projected onto the holographic screen. Alternatively, the system may automate viewpoint selection using aggregated emotional data or entropy-based metrics. Audience members act as co-creators, since their affective responses shape the evolution of the performance. Currently, emotional inputs are captured individually but fused into a single collective feedback stream for both performers and the immersive environment. Personalized feedback is planned but not yet implemented.

C. Creative Objectives

This scenario explores how emotion-driven interaction can enhance audience immersion, enable performer adaptability and improvisation, and reshape dramaturgical structures by considering emotion as a dynamic variable. By making affective states perceptible and manipulable in real time, EmoLoop supports expressive relationships among musicians, audiences, scenographers, and computational systems. Audience immersion will be evaluated through self-report questionnaires, biosignal and behavioral indicators such as EEG patterns and gaze, and responsiveness to performative modulation. Dramaturgical effects will be assessed through logged scene transitions, performer feedback, and analysis of how emotional triggers influence musical or narrative development. These measures will help determine how the emotional loop increases presence and co-creation across virtual and physical settings.

VI. Conclusion and Future Work

This paper introduces EmoLoop, an interactive framework that enables a bi-directional emotional feedback loop between performers and audiences in immersive musical and theatrical contexts. By combining multimodal emotion recognition from EEG signals, facial expressions, and behavioral cues with real-time performance control, the system supports new forms of affect-based co-creation in hybrid environments.

The architecture relies on wearable sensors such as the Meta Quest V3 and Muse 2, together with emotion inference models and low-latency protocols, to dynamically influence sound, scenography, and spatialization. The work is positioned within Affective Computing, Networked Performance, and Internet of Sounds, offering a feasible model for participatory art.

Future work will focus on finalizing system integration of emotion sensing, VR interaction, and performance control, conducting user studies to assess emotional accuracy and expressiveness, testing scalability in hybrid online–offline setups and co-developing new dramaturgies with artists and designers. The long-term ambition is to connect human affect and computational performance systems, enabling scenarios where emotions do not only inspire but actively shape performances in real time.

Acknowledgment

The research activities described in this paper were carried out within the "HEAT – Hybrid Extended reAliTy" Project GA 101135637 funded by the EU Horizon Europe Framework Programme (HORIZON).

References

[1] R. Damiano, V. Lombardo, G. Monticone, and c. A. Pizzo, "Studying and designing emotions in live interactions with the audience," *Multimedia Tools and Applications*, vol. 80, pp. 6711–6736, 2021.

[2] Y. Zhang, M. Li, and S. Pan, "Deep Learning-based Emotion Recognition Algorithms in Music Performance," *Scalable Computing: Practice and Experience*, vol. 25, no. 6, pp. 4712–4719, Oct. 2024.

[3] S. Lee, Y. Cho, and J. Lee, "Concert Interaction Translation: Augmenting VR Live Concert Experience using Chat-Driven Artificial Collective Reactions," in *Proc. of the Ext. Abstracts of the CHI Conf. on Human Factors in Computing Systems*, ser. CHI EA '25. New York, NY, USA: Association for Computing Machinery, Apr. 2025, pp. 1–7.

[4] S. M. H. Mousavi, M. Besenzoni, D. Andreoletti, A. Peternier, and S. Giordano, "The Magic XRoom: A Flexible VR Platform for Controlled Emotion Elicitation and Recognition," in *Proc. of the 25th Int. Conf. on Mobile Human-Computer Interaction*. Athens Greece: ACM, Sep. 2023, pp. 1–5.

[5] P. L. Indrasiri, B. Kashyap, C. Kolambahewage, B. Nakisa, K. Ijaz, and P. N. Pathirana. VR Based Emotion Recognition Using Deep Multimodal Fusion With Biosignals Across Multiple Anatomical Domains. Preprint.

[6] T. Ortmann, Q. Wang, and L. Putzar, "Unimodal and Multimodal Static Facial Expression Recognition for Virtual Reality Users with EmoHeVRDB," in *2025 IEEE Int. Conf. on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, Jan. 2025, pp. 252–256.

[7] T. Colafiglio, D. Lofù, P. Sorino, A. Lombardi, F. Narducci, F. Festa, and T. Di Noia, "EmoSynth Real Time Emotion-Driven Sound Texture Synthesis via Brain-Computer Interface," in *Adjunct Proc. of the 32nd ACM Conf. on User Modeling, Adaptation and Personalization*, ser. UMAP Adjunct '24, pp. 616–621.

[8] L. Turchet and M. De Cet, "A web-based distributed system for integrating mobile music in choral performance," *Personal and Ubiquitous Computing*, vol. 27, no. 5, pp. 1829–1842, Oct. 2023.

[9] M. Dima and H. Maples, "Affectual Dramaturgy for Augmented Reality Immersive Heritage Performance," *Body, Space & Technology*, vol. 20, no. 1, Mar. 2021.

[10] T. Hopkins, E. Doherty, N. Ofer, S. C. C. Weng, P. Gyrory, C. Tobin, L. Hirshfield, and E. Y.-L. Do, "Stringesthesia: Dynamically Shifting Musical Agency Between Audience and Performer Based on Trust in an Interactive and Improvised Performance," *ACM Int. Conf. Proc. Series*, vol. 1, pp. 9–16, Sep. 2023.

[11] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of Musical Things: Vision and Challenges," *IEEE Access*, vol. 6, pp. 61 994–62 017, 2018.

[12] O. Hödl and L. Turchet, "The TMAP Framework: Describing the Design Space of Technology-Mediated Audience Participation in Live Music," *IEEE Access*, vol. 13, pp. 38 593–38 608, 2025.

[13] P. Kalansooriya, G. A. D. Ganepola, and T. S. Thalagala, "Affective gaming in real-time emotion detection and Smart Computing music emotion recognition: Implementation approach with electroencephalogram," in *2020 Int. Research Conf. on Smart Computing and Systems Engineering (SCSE)*, Sep. 2020, pp. 111–116.

[14] X. Hao, H. Li, and Y. Wen, "Real-time music emotion recognition based on multimodal fusion," *Alexandria Engineering Journal*, vol. 116, pp. 586–600, Mar. 2025.

[15] Y. Su, Y. Liu, Y. Xiao, J. Ma, and D. Li, "A review of artificial intelligence methods enabled music-evoked EEG emotion recognition and their applications," *Frontiers in Neuroscience*, vol. 18, Sep. 2024.

[16] R. Wang, "Exploring an emotion-based interactive system in live theatre," PhD thesis, De Montfort University, 2024.

[17] S. K. Et. Al., "Facial Emotion Based Music Recommendation System using computer vision and machine learning techiniques," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 2, pp. 912–917, Apr. 2021.

[18] Mr. B. Sreedhar, S. Shahida, P. Tanmai, M. Nithin, and G. S. Neelima, "Music Recommendation System Based on Facial Emotion," *Journal of Humanities,Music and Dance*, no. 33, pp. 11–22, Apr. 2023.

[19] Y. R. Pandeya, B. Bhattarai, and J. Lee, "Deep-Learning-Based Multimodal Emotion Classification for Music Videos," *Sensors*, vol. 21, no. 14, p. 4927, Jul. 2021.

[20] Y. Liang, X. Zhang, C. Wu, and M. Yang, "Emotional responses in ai-enhanced interactive installation art," *Sustainability*, vol. 15, no. 22, p. 15830, 2023.

[21] Y. Zhang, M. Li, and S. Pan, "Deep Learning-based Emotion Recognition Algorithms in Music Performance," *Scalable Computing: Practice and Experience*, vol. 25, no. 6, pp. 4712–4719–4712–4719, Oct. 2024.

[22] N. Kulkarni and S. Kaur, "Recent trends and challenges in human computer interaction using automatic emotion recognition: A review," vol. 1, no. 1, p. 1.

[23] G. A.V., M. T., P. D., and U. E., "Multimodal emotion recognition with deep learning: Advancements, challenges, and future directions," *Information Fusion*, vol. 105, p. 102218, 2024.

[24] S. Scarani, A. Muñoz, J. Serquera, J. Sastre, and R. B. Dannenberg, "Software for Interactive and Collaborative Creation in the Classroom and Beyond: An Overview of the Soundcool Software," *Computer Music Journal*, vol. 43, no. 4, pp. 12–24, Dec. 2019.

[25] X. Wu, "Integration and Application of Multimedia Communication Systems in Stage "Sound, Light, and Video" Linkage Control," in *2024 5th Int. Conf. on Electronic Communication and Artificial Intelligence (ICECAI)*, pp. 130–134.

[26] A. Carôt, A. B. Renaud, and P. Rebelo, "Networked Music Performance: State of the Art: AES 30th Int. Conf." in *Proc. of the AES 30th Int. Conf.*, pp. 1–7.

[27] H. Fakhrurroja, C. Machbub, A. S. Prihatmanto, and A. Purwarianti, "Multimodal Interaction System for Home Appliances Control," vol. 14, no. 15, pp. 44–67.

[28] T. Ortmann, Q. Wang, and L. Putzar, "EmojiHeroVR: A Study on Facial Expression Recognition Under Partial Occlusion from Head-Mounted Displays," in *2024 12th Int. Conf. on Affective Computing and Intelligent Interaction (ACII)*, Sep. 2024, pp. 80–88.

[29] Y. N. Cardona-Álvarez, A. M. Álvarez-Meza, D. A. Cárdenas-Peña, G. A. Castaño-Duque, and G. Castellanos-Dominguez, "A Novel OpenBCI Framework for EEG-Based Neurophysiological Experiments," *Sensors (Basel, Switzerland)*, vol. 23, no. 7, p. 3763, Apr. 2023.

[30] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, and B. Hu, "Emotion recognition from multi-channel eeg data through convolutional recurrent neural network," in *2016 IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM)*, 2016, pp. 352–359.

[31] M. N. Dar, M. U. Akram, S. G. Khawaja, and A. N. Pujari, "Cnn and lstm-based emotion charting using physiological signals," *Sensors*, vol. 20, no. 16, 2020.

[32] M. Wright, "Open Sound Control: An enabling technology for musical networking," *Organised Sound*, vol. 10, no. 3, pp. 193–200, Dec. 2005.

[33] A. Schmeder and A. Freed, "Implementation and applications of open sound control timestamps," in *ICMC 2008*, ICMA. Belfast, Ireland: ICMA, 25/08/2008 2008, pp. 655–658.

[34] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An Overview on Networked Music Performance Technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.

[35] M. Hlayel, H. Mahdin, and H. A. Mohd Adam, "Latency Analysis of WebSocket and Industrial Protocols in Real-Time Digital Twin Integration," *Int. Journal of Engineering Trends and Technology*, vol. 73, no. 1, pp. 120–135, Jan. 2025.