SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Electroencephalogram (Eeg)-Conformer For Emotion Recognition And Theory-Guided Music Prompt Generation: A Prototype Brain-Computer Music Interface

By Yanlin Zhou

AUTHOR BIOGRAPHY

Yanlin Zhou is a student at Concord Academy in Massachusetts. Her interests include machine learning, neuroscience, and music technology. She hopes to continue conducting research in affective computing and brain-computer interfaces.

ABSTRACT

There is a robust body of research demonstrating that emotional states profoundly affect both physical and mental health. Acknowledging this, researchers have sought to examine methods for detecting an individual's emotional state and thus promoting their emotional and psychological well-being. In recent years, scholars working in this area have turned to computer science and machine learning to serve these objectives. In this same vein, the current study examines electroencephalogram (EEG) signals as a means for emotion recognition and for facilitating brain-computer interfaces. To assess the potential of EEG signals for such purposes, this article reviews the principles underlying methods for measuring EEG signals, including frequency bands and electrode distributions, to examine how EEG data can be used for emotion recognition. To inquire into the patterns inherent in brainwave recordings and determine whether a person was experiencing positive, negative, or neutral emotions, this study used a specialized artificial intelligence (AI) model called an EEG-Conformer. Once an emotional state has been recognized, the system pipeline translates the classification result into a text prompt, which is then passed to an AI music-generation model that produces music aligned with the detected state. Testing on a widely used EEG emotion dataset (the SEED dataset) showed that the model was over 95% accurate in identifying emotional states, and visual analyses further confirmed that the model was able to ascertain meaningful patterns in the brain activity associated with different emotions. In this way, findings from this study point to how AI and neuroscience can be integrated in such a way that it supports human health and well-being.

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Keywords: *EEG signal, machine learning, emotion recognition, EEG-Conformer, music generation, affective computing*

INTRODUCTION

Emotion is at the foundation of all human psychological activity, influencing anything from cognition to decision-making to overall physiological state. The latter of these has received special attention within research into the neuro- and health sciences, as prolonged negative emotional states have been shown to lead to anxiety, depression, and even severe cardiovascular disorders (Koelstra et al., 2012; Steptoe & Kivimäki, 2012). In light of this, it has become increasingly important to better recognize emotional states so as to inform timely and targeted intervention.

Historically speaking, emotion recognition has relied largely on self-report or third-party observation, two methods that are known to be susceptible to the follies of subjective interpretation, marked by a slow timeline to progress, and difficult to continuously monitor. Yet, electroencephalogram (EEG) signals (i.e., non-invasive recordings of electrical activity within the brain) have emerged as an effective method for subtly capturing subtle changes in neurophysiological features during changes to emotional states (Chen et al., 2021). This, coupled with its ability to offer high temporal resolution and brain activity reflections, has also prompted the technology to garner significant attention within the realm of affective computing (Picard, 1997; Zheng & Lu, 2015).

Deep learning has become one of the most widely used mechanisms for recognizing data derived from EEG scans. Earlier studies relied on convolutional neural networks (CNNs), which were effective for identifying localized patterns in brainwave signals (Zheng & Lu, 2015). Later investigations, however, employed recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, which better captured how brain activity changes over time (Hizlisoy et al., 2021). While both models offered promising preliminary insights, they struggled to demonstrate the relationship between signals separated by large durations of time. This all changed, however, when Vaswani et al. (2017) introduced the Transformer architecture, which uses a mechanism known as self-attention to simultaneously examine connections across an entire sequence of data. Given that emotions are reflected as a coordinated activity occurring in multipls areas of the brain at various points in time, the broader perspective offered by the Transformer system can better distinguish between emotional states when compared to its earlier counterparts.

Innovations in this area continued, as Song et al. (2023) introduced the EEG-Conformer, an AI model capable of recognizing patterns in brainwave recordings associated with different emotional states. It does this through an integrated dual approach model, wherein it (1) identifies significant local patterns in EEG signals, and (2) examines how those patterns relate to one another across the entire recording. Through such integration, EEG-Conformer has consistently performed well on EEG emotion-recognition datasets, such as BCI Competition IV and SEED (Song et al., 2023)—two widely used collections of EEG recordings gathered while participants viewed emotion-eliciting videos.

Another primary drawback associated with previous studies is that they often focus on technical performance metrics, such classification accuracy and confusion matrices, that can prove difficult for non-specialists to interpret. As a result, there remains a marked need for approaches that translate emotion-recognition outputs in such a way that they are accessible to the average user. However, recent advances in artificial intelligence have enabled users to generate music from simple text descriptions, and in doing so, opened up new opportunities for transforming computational assessments of emotion into engaging—and even healing—auditory experiences (Hung et al., 2021). Building upon these developments, the present study uses EEG-Conformer as the foundation for connecting emotion recognition with customized music generation. The EEG-Conformer system uses EEG data to identify a user's emotional state to then generate music recommendations tailored to said emotional state.

Building on these developments, the present study uses EEG-Conformer as the foundation for a system that connects emotion recognition with music generation. More specifically, the system identifies a user's emotional state from EEG data and then generates music recommendations tailored to that emotional state. Through this interactive platform, users can upload EEG recordings, receive emotion predictions, and explore music generated in response to those predictions.

To assess the feasibility of this approach, this study tested the proposed system using the EEG-Conformer architecture and evaluated the publicly available SEED dataset, which is widely used as a collection of EEG recordings against which new research is emotion recognition is frequently benchmarked.

REVIEW OF THE LITERATURE

EEG-Based Emotion Recognition

EEG-based emotion recognition involves three stages, the first of which is signal preprocessing. Signal preprocessing necessitates cleaning and preparing raw EEG recordings so that meaningful neural activity can be separated from background noise or other ambient acoustic interference. After this stage is complete, the second stage, feature extraction**,** commences. Feature extraction involves detecting measurable patterns within the EEG signals that may be associated with different emotional states. The features extracted from these signals can provide information about varying degrees of intensity exhibited over time, activity witnessed within different

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

frequency bands, or patterns that fluctuate in terms of both time and frequency (Zheng & Lu, 2015). The final stage, classification, uses computational models to determine which emotional categories best correspond with the EEG patterns noted. In most emotion-recognition studies, these categories indicate positive, negative, or neutral emotional states. Table I (introduced below) summarizes the major EEG frequency bands and the mental states most commonly associated with each.

Table I
EEG Frequency Bands and Their Associations with Emotion/Cognition

EEG Band	Frequency Range (Hz)	Commonly Associated Mental State
Delta	0.5–4	Deep sleep and unconscious states
Theta	4–8	Relaxation, meditation, and drowsiness
Alpha	8–13	Relaxed wakefulness and resting states
Beta	13–30	Alertness, attention, and active thinking
Gamma	>30	Higher-order cognition and perceptual integration

However, it is worth pointing out that not all EEG signals are indicative of a particular emotional state. In fact, research has shown that certain frequency bands are more useful for emotion recognition than others. Zheng and Lu (2015), for example, found that the beta (13- 30 Hz) and gamma (greater than 30 Hz) frequency bands were especially useful for distinguishing emotional states. The authors noted that the location of the brain where the activity is observed likewise has an impact on distinguishing emotional states, with the frontal and temporal regions of the brain contributing the most information regarding the user's emotional state.

In addition to identifying informative EEG features, researchers have worked to improve machine learning techniques to improve these model's capacity to recognize emotional states. Towards this end, Li et al. (2018) developed a bi-hemispheric domain adversarial neural network (BiDANN) that draws on the emotional asymmetry between the left and right hemispheres of the brain, mapping the EEG features of each hemisphere into separate discriminative feature spaces to better distinguish emotional states across participants. Along these same lines, Chen et al. (2021) suggested a well-known method called Multi-Source marginal Distribution Adaptation (MS-MDA) so that the model's performance would remain consistent across participants and

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

recording sessions—an objective they were able to achieve when the model was tested on the above-referenced SEED dataset.

When considered alongside one another, these studies highlight that EEG-based emotion recognition is dependent not only on identifying meaningful neural signals, but also on developing computational models capable of capturing complex patterns across various spatial and temporal contexts, and with differing users and recording conditions. To address these challenges, in 2023, researchers introduced an EEG-Conformer model that combines deep learning with Transformer architecture to more reliably ascertain emotional state from EEG data. The section that follows details the developments offered by the EEG-Conformer in greater detail.

EEG-Conformer, Transformer Decoding, and Music Generation

Transformer-based models have demonstrated promising results for use in other applications, including language processing and computer vision, which has consequently prompted researchers to explore how its use might similarly benefit EEG analysis. The appeal of this particular type of AI architecture lies in its ability to simultaneously examine relationships across entire sequences of information, thereby enabling them to identify patterns that may otherwise be overlooked when data are analyzed in individual segments. Earlier approaches relied heavily on interpreting manually selected signal characteristics, which introduces the pitfalls traditionally associated with selection bias. Song et al. (2023), however, showed how an EEG-Conformer model can overcome this limitation by learning relevant patterns from the data itself, thereby precluding the need for researchers to preemptively designate the neural features that are the most salient to analysis.

But correctly identifying a user's emotional state only solves half the problem. The other half lays in designing a meaningful intervention, raising the question of how such technologies can be utilized to improve user well-being. One potential application for EEG data for such purposes pertains to music therapy. Music therapy has long been used to foster both physical and mental health, and recent development in AI only furthered the ability to tailor musical experiences to the specific user (Li et al., 2024). To illustrate, Hung et al. (2021) created the EMOPIA dataset—a collection of piano recordings annotated according to the emotional characteristics often ascribed to them. The authors used this dataset to study how data regarding emotional state can inform the creation of AI-generated music. In a similar vein, large-scale commercial platforms like Suno and Udio have demonstrated similar advances in that they illustrate how high-quality music can be generated from simple text prompts. While these signify remarkable progress in the integration of AI and music, systems that bridge physiological signals, emotion recognition, and music generation remain relatively uncommon. The present study therefore seeks to address this gap by creating a Brain-Computer Music Interface (BCMI) that links EEG-based emotion recognition with AI-assisted music generation, and in doing so, render neural activity a highly personalized auditory experience.

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

METHODS

Data Source

To perform its analyses, this study used data from the publicly available SEED dataset (Zheng & Lu, 2015). To review, this dataset is composed of electroencephalogram (EEG) recordings collected from participants as they viewed video clips that were designed to elicit a response that could be categorized as positive, negative, or neutral in terms of emotional state. For each participant, EEG activity was recorded across 62 channels (i.e., recording locations on the scalp) using the standardized international placement system of approx. 10-20 electrodes. Because this study relied exclusively on a previously collected and publicly available dataset, no new human participants were recruited and no additional institutional review board approval was required.

EEG-Conformer Architecture

The classification of emotions was performed using the EEG-Conformer architecture originally proposed by Song et al. (2023). In the present study, this architecture was implemented and trained on the SEED dataset to evaluate its ability to distinguish among positive, negative, and neutral emotional states. The convolutional part allows researchers to extract local spatiotemporal patterns from the EEG recordings. The Transformer part, on the other hand, offers complementary data speaking to the relationships inherent within signal sequences. The TransformerEncoder consisted of six encoder blocks, each of which itself contained multi-head self-attention and feed-forward network layers. In other words, the EEG-Conformer model was able to analyze EEG recordings by using these six sequential processing layers that learned increasingly complex patterns associated with different emotional states. The ClassificationHead then transformed the feature representations the system learned, designating them as belonging in one of three emotional categories: positive, negative, or neutral. Figure 1 below depicts the EEG-Conformer architecture used in this study.

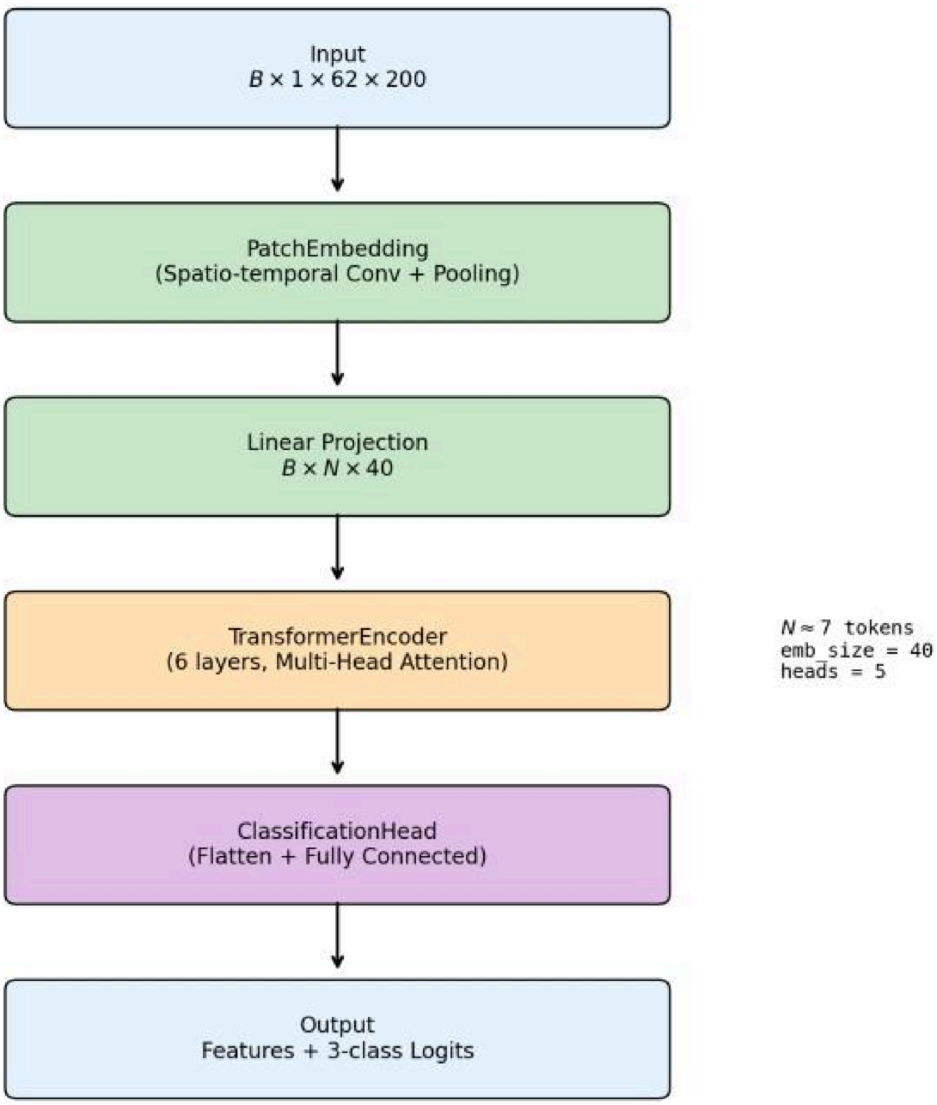
SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Figure 1
EEG-Conformer Architecture



Note. Raw EEG recordings are first processed to identify local spatiotemporal patterns through convolutional operations. The resulting features are then analyzed by a Transformer encoder, which learns relationships across the entire recording before classifying the EEG signal as representing a positive, negative, or neutral emotional state.

MODEL TRAINING AND PERFORMANCE EVALUATION

The researcher trained the EEG-Conformer model on the SEED dataset and evaluated the model with respect to this established dataset. To assess the model's performance and draw conclusions regarding reliability, a five-fold cross-validation was employed. This procedure involves repeatedly dividing a dataset into training and testing subsets so that the model's overall performance can be evaluated across multiple iterations. More detailed analyses of the SEED dataset focused on Subject 15 because the participant's first recording session contained approximately 3,375 one-second EEG segments, providing a sufficiently large sample for evaluating model performance and generating visualization analyses. These datapoints help assess model performance along the lines of classification accuracy, confusion matrices, feature visualizations, and attention-mapping techniques—all of which are designed to identify the neural signals that most strongly contributed to emotion classification. These provide insight into not only how accurately (or inaccurately) the model's performance was, but how it made its decisions.

System Architecture

After training the model, the researcher developed a prototype web-based application to demonstrate how EEG-based emotion recognition could be integrated with AI-assisted music generation in a real-world system. The system used a front-end/back-end architectural design. The front end was developed using Streamlit, and allowed for EEG file upload, visualization of the emotion results, and playback of audio recordings. The back end, by contrast, was implemented using FastAPI and enabled emotion classification, music generation, and interactions among large language models. When a user uploads an EEG segment, the system first attempts to classify the emotion using the trained EEG-Conformer model. The emotional category that results is then translated into a text prompt that then guides music generation. Additional functionality was incorporated into the platform to allow future users to refine subsequent music recommendations through either questionnaire responses or interactions with a large language model.

RESULTS

Analysis of the Model's Classification

The model performed well when it came to distinguishing among positive, negative, and neutral emotional states. Negative emotions were identified with an accuracy rate of 97.0%, whereas neutral and positive emotions

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

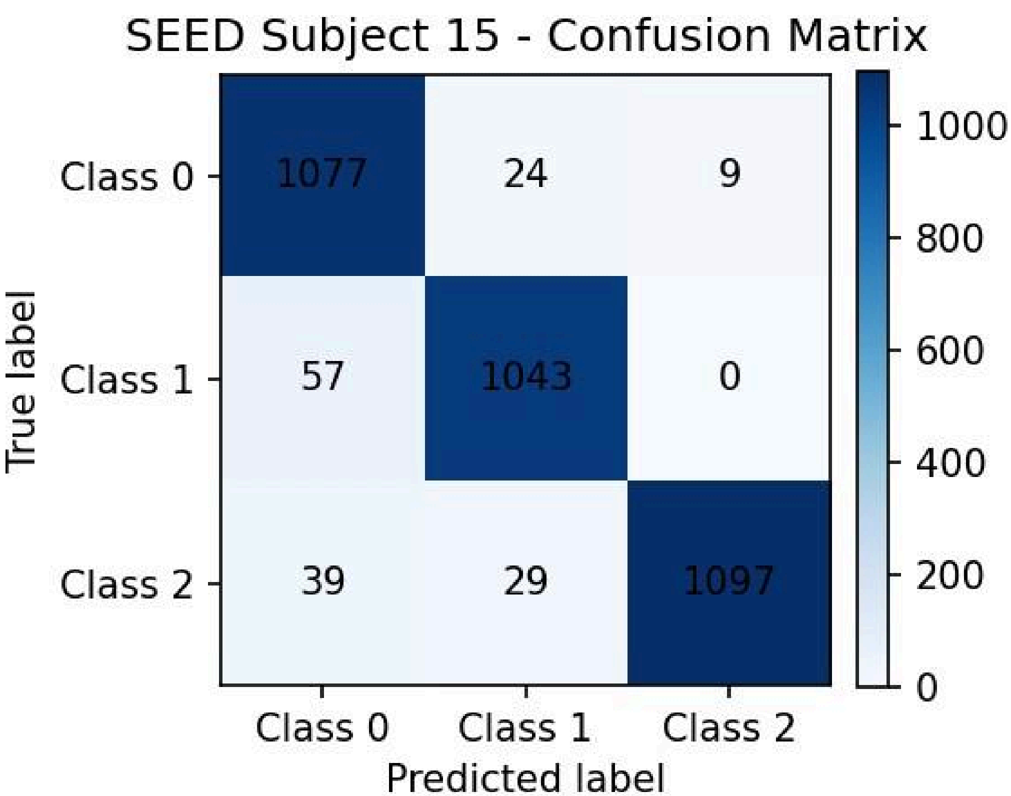
E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

were identified with accuracy rates of 94.8% and 94.2%, respectively. The average classification accuracy across all five validation folds was at 95.32%. Figure 2 below presents the confusion matrix based on the model's classifications. The figure provides a detailed breakdown of correct and incorrect predictions across the three emotional categories and illustrates where classification errors were most likely to occur.

Figure 2

Confusion Matrix for EEG-Conformer Emotion Classification



Note. The confusion matrix summarizes the model's classification performance across the three aforementioned emotional categories (that is, positive, negative, and neutral). The values that appear along the diagonal plane represent correct classifications, whereas those off-diagonal values indicate instances in which the model misclassified one emotional state as another. The concentration of observations along the diagonal demonstrates the model's strong overall performance, with an average classification accuracy of 95.32%.

While overall performance for the model was high, as indicated by these figures, some errors in classification were nonetheless observed. Most of these errors were witnessed in the classification of negative or neutral states; positive and negative emotional states witnessed errors to a comparatively lesser extent. What this suggests is that certain emotional states may exhibit more (or more overt) neural characteristics that make them more difficult to distinguish using just EEG data.

To investigate how the model performed over time, the researcher examined classification accuracy throughout the training process. The model's performance improved rapidly during the initial stages of training as it increased from a baseline of 35% accuracy to an end-stage accuracy of greater than 90% over the course of 150 training cycles. The model's accuracy continued to improve before eventually stabilizing, and ultimately reached around 97% accuracy during the fifth validation fold. This pattern suggests that the model was indeed capable of efficiently learning and exhibiting a stable performance over the course of the training period.

Feature Visualization Analysis

To gauge whether the model learned how to make meaningful distinctions among different emotional states, a t-SNE visualization was generated. This particular technique is useful for distilling complex, high-dimensional data into an easy-to-understand two-dimensional representation. The resulting visualization demonstrated that the different classifications of emotional sates (i.e., positive, negative, and neutral) formed distinct clusters. While there was some overlap noted at the boundaries between categories, their general separation indicates that instead of relying on arbitrary statistical relationships, the model was actually able to discern and learn meaningful patterns associated with different emotional experiences. The following figure, Figure 3, is a t-SNE visualization of the feature representations that the EEG-Conformer successfully learned. The clustering of observations according to emotional category provides further evidence for the conclusion that the model was able to successfully distinguish among positive, negative, and neutral emotional states.

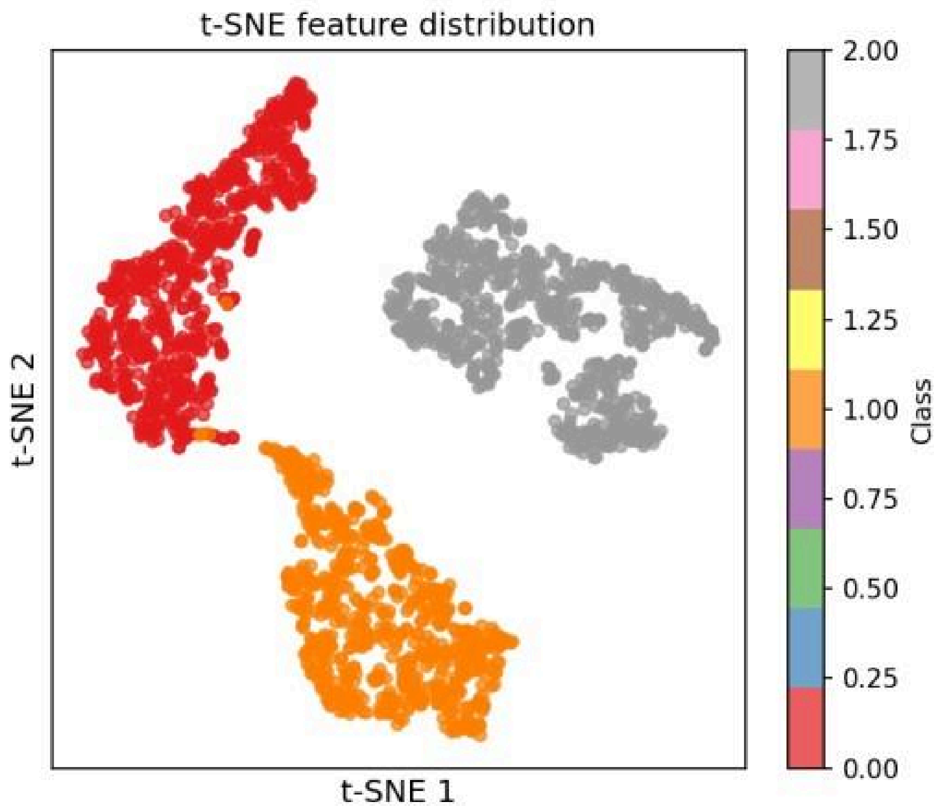
SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Figure 3
t-SNE Visualization of Learned Emotional Representations



Note. t-distributed Stochastic Neighbor Embedding (t-SNE) is a method of visualization that distills complex, high-dimensional data into a two-dimensional mapping. Each data point represents an EEG recording. The points located closer together signify recordings that the model interpreted as having similar characteristics. The fact that the emotional states were divided into the categories of positive, negative, and neutral as distinct clusters means that the model was able to make meaningful distinctions among these emotional categories.

Further analyses were performed using Gradient-weighted Class Activation Mapping, also known as Grad-CAM. This technique helps the analyst to identify which segments of the EEG recordings contributed most significantly to the model's decisions. The results of this analysis suggested that negative emotions were primarily associated with activity that occurred in the frontal and temporal regions of the brains. Positive emotions, by contrast, were primarily observed in the temporal and parietal regions. Finally, neutral emotions were much more distributed in terms of the regions of the brain they activated. Collectively, these findings

SCHOLARLY REVIEW

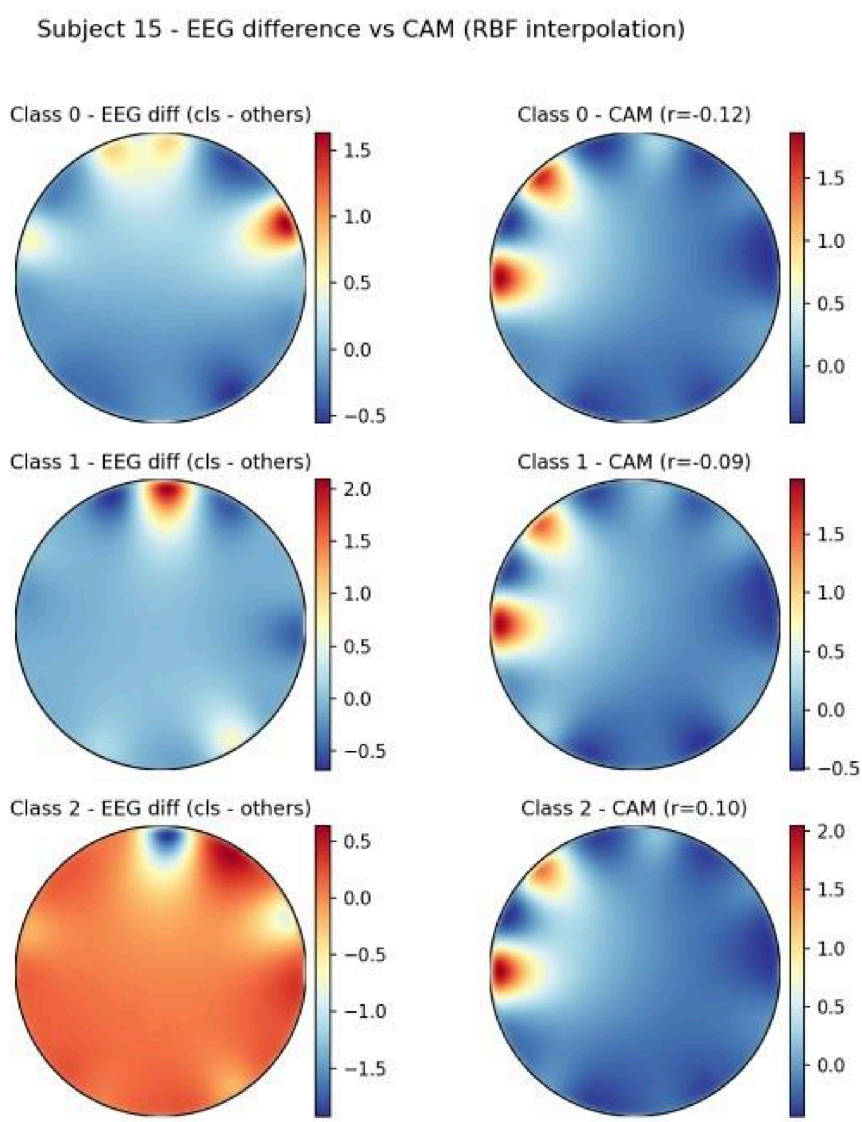
SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

support previous neuroscience on emotional processing and provide evidence that the model's classifications are biologically plausible. Figure 4 illustrates the areas of the brain that were observed as having contributed most significantly to the model's classification of the different emotions. By emphasizing the fact that neural activity influenced each prediction, the visualization offers researchers insight into how the model arrived at its conclusions as well as the extent to which those conclusions align with existing neuroscience research.

Figure 4
Grad-CAM Visualization of Neural Regions Associated with Emotional Classification



Note. This study used Gradient-weighted Class Activation Mapping (Grad-CAM) to identify the regions of the EEG recordings (i.e., the regions of the brain) that contributed to the model's classifications the most. Warmer colors indicate greater influence, whereas cooler colors indicate a lesser degree of influence. Negative emotions were most strongly associated with activity in the frontal and temporal regions, positive emotions with the temporal and parietal regions. Neutral emotions were more broadly distributed by comparison.

Implications for Real-World System Functionality

These results point to important takeaways regarding system functionality offline. Outside of emotion classification, the system demonstrated feasibility when it came to rendering neural signals in terms of the average user's experience. So, once an emotional state was identified, the platform would automatically generate a music prompt tailored to that specific emotional category. (See Table II below for a detailed description of negative, neutral, and positive emotional states.) The platform was also designed to support additional personalization through questionnaire responses or large language model interactions, enabling future users to refine subsequent music recommendations based on their preferences and experiences. In this way, the model effectively moved beyond simple detection to produce intelligent outputs.

The three prompts were not selected arbitrarily. Their musical parameters were derived from two established frameworks in music psychology and music therapy. The first is the iso principle, a long-standing clinical method holding that a music intervention is most effective when it begins with music matching the listener's present affective state and then moves gradually toward the desired state, rather than immediately presenting contrasting music (Heiderscheit & Madson, 2015). This accounts for why the prompt associated with negative affect specifies slow, gentle piano and strings rather than cheerful music, since matching precedes redirection. The second is the circumplex model of affect, which represents emotional states along the two dimensions of valence and arousal (Russell, 1980). Tempo was mapped primarily onto the arousal dimension, and mode and melodic contour onto the valence dimension. This mapping is supported by evidence that musical tempo drives cardiovascular and respiratory arousal more strongly than musical style does (Bernardi et al., 2006), and by mechanistic accounts of how structural features of music induce emotional responses (Juslin & Västfjäll, 2008; Koelsch, 2014). Table 2 presents each prompt alongside the psychological rationale for its musical parameters.

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Table 2
Emotion Categories and Corresponding Music Generation Prompts

Emotion Category	Music Generation Prompt	Psychological Rationale
Negative	A gentle and slow piano and string composition with a calm, soothing atmosphere designed to reduce tension, comfort sadness, and promote relaxation.	Follows the iso principle: the prompt first matches the listener's low-valence, low-arousal state rather than opposing it, then supports gradual mood change (Heiderscheit & Madson, 2015). Slow tempo and sustained strings correspond to low arousal in the circumplex model (Russell, 1980) and are associated with reduced heart rate and respiratory rate (Bernardi et al., 2006).
Neutral	A light and peaceful instrumental composition with a moderate tempo, simple melody, and balanced rhythm intended to support concentration while maintaining a relaxed emotional state.	Moderate tempo and simple melodic structure place the stimulus near the center of the valence-arousal space (Russell, 1980), sustaining a stable affective baseline. Predictable structure limits strong expectancy violations, one of the mechanisms through which music induces emotion (Juslin & Västfjäll, 2008).
Positive	An upbeat but not overly stimulating composition featuring a pleasant melody, clear rhythm, and optimistic tone intended to maintain positive emotions and psychological well-being.	Faster tempo, major mode, and clear rhythm correspond to high valence and moderate arousal (Russell, 1980) and engage reward-related processing associated with pleasurable music listening (Koelsch, 2014). Stimulation is deliberately bounded so that arousal is raised without overshooting (Bernardi et al., 2006).

Note. Following emotion classification, the system translated each emotional category into a text prompt that could be used by an AI music-generation platform to create music aligned with the user's detected emotional state. The rightmost column reports the theoretical basis for the musical parameters specified in each prompt. Prompts were constructed a priori from these frameworks and were not empirically validated in the present study.

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Proposed Extension: Continuous Modulation of Prompt Parameters

The prompt scheme described above remains categorical, in that each of the three emotional classes maps to one fixed prompt. A natural extension, proposed here for future development rather than tested in the present study, would use information the model already produces to modulate prompt parameters continuously. Two such signals are available without any change to the trained architecture. The first is the softmax probability assigned to the predicted class, which can be read as a measure of how strongly that emotional state is expressed and used to scale the intensity of the requested musical features. The second is the relative power in the beta and gamma bands, which the Grad-CAM analysis identified as contributing substantially to classification and which prior work has linked to alertness and active processing (Zheng & Lu, 2015). Because tempo maps onto the arousal dimension of the circumplex model (Russell, 1980) and exerts a stronger physiological effect than musical style does (Bernardi et al., 2006), these quantities could be used to specify a tempo in beats per minute rather than selecting among three preset descriptions. Under the iso principle, the generated sequence could then shift tempo gradually across successive segments, beginning near the listener's estimated arousal level and moving toward a target level (Heiderscheit & Madson, 2015). Such a design would link the sensing and generation components more directly than a fixed set of category prompts allows, and it represents the most immediate direction for extending the present system.

LIMITATIONS

The results of the present study support the promise of EEG-based emotion recognition to create a brain-computer music interface. As is detailed in the previous section, this study's analysis points to the feasibility of an end-to-end pipeline that bgins with EEG-based emotion recognition and ends with customized music recommendations. That said, it did not evaluate the therapeutic effects of said recommendations. It can therefore not make claims as to emotional effects. Future research may wish to design a control experiment, wherein users are presented with different musical recommendations (based on the emotional state detected) to see how effective these interventions are in ameliorating a negative emotion or promoting a positive one.

A further limitation concerns the music prompts themselves. Although the prompts presented in Table 2 were derived from established principles in music psychology and music therapy, they were not empirically validated in the present study. No listening experiment was conducted, so it is not known whether listeners perceived the generated music as carrying the intended valence and arousal characteristics, nor whether the prompts produced any measurable change in affective state. The prompts should therefore be understood as theoretically motivated starting points rather than as validated therapeutic stimuli. Future work should evaluate them directly, for example by collecting listener ratings of valence and arousal for the generated excerpts and pairing those ratings with physiological measures such as heart rate and respiratory rate.

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Another limitation worth mentioning here is that the study relied entirely on secondary data in the form of the publicly available SEED dataset rather than introduce primary data collected by way of humans-subject research. The SEED dataset admittedly offers several advantages, the first of which is that it has demonstrated reliability and validity across multiple study contexts. Yet, while the SEED dataset is a widely used reference point in research on emotion recognition, the experiments took place under controlled experimental conditions that may not accurately reflect the complexities of everyday life.

An additional limitation lays in how the study focused on a single participant from the dataset (Subject 15). This approach was intentional as it allowed for a more in-depth examination of the model's performance, but future studies should make these evaluations based on a larger participant pool. This way, they can eliminate the possibility that any conclusion drawn is not an individual expression or anomaly. Emotional responses can vary dramatically from person to person, so broadening the research sample holds important implications for the study's overall generalizability. Future research should also consider employing cross-subject validation techniques for similar purposes.

CONCLUSION

This study investigated the utility of EEG-Conformer, a deep-learning model designed to classify emotional states based on electroencephalogram (EEG) recordings. Using the publicly available SEED dataset, the model was able to successfully distinguish among positive, negative, and neutral emotional states, exhibiting an average accuracy classification rate in excess of 95%. Additional analyses further confirmed the model's ability to make meaningful distinctions among emotional categories and the fact that the model relied on neural patterns already established within the existing scholarship in this area.

More than correctly classify emotions, however, the current study demonstrated how EEG-based emotion recognition could be translated into a practical user experience. As it connects emotion predictions to AI-assisted music generation, the system featured here provides evidence speaking to how brain-computer interfaces may support personalized interventions aimed at promoting individual psychological well-being. Rather than treating emotion recognition as purely within the purview of computer science, this project represents an interdisciplinary endeavor, one that brings together the principles of artificial intelligence and neuroscience. As mentioned in the discussion on limitations, future research should examine the system's performance across diverse populations and investigate whether EEG-guided music recommendations can influence mood, moderate stress levels, or regulate emotions. Through intelligent innovations such as this, new avenues may be identified and pursued to better support human emotional well-being.

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

REFERENCES

Bernardi, L., Porta, C., & Sleight, P. (2006). Cardiovascular, cerebrovascular, and respiratory changes induced by different types of music in musicians and non-musicians: The importance of silence. *Heart, 92*(4), 445–452. https://doi.org/10.1136/hrt.2005.064600

Chen, H., Jin, M., Li, Z., Fan, C., Li, J., & He, H. (2021). MS-MDA: Multisource marginal distribution adaptation for cross-subject and cross-session EEG emotion recognition. *Frontiers in Neuroscience, 15*, Article 778488. https://doi.org/10.3389/fnins.2021.778488

Grekow, J. (2021). Music emotion recognition using recurrent neural networks and pretrained models. *Journal of Intelligent Information Systems, 57*, 531–546. https://doi.org/10.1007/s10844-021-00658-5

Heiderscheit, A., & Madson, A. (2015). Use of the iso principle as a central method in mood management: A music psychotherapy clinical case study. *Music Therapy Perspectives, 33*(1), 45–52. https://doi.org/10.1093/mtp/miu042

Hizlisoy, S., Yildirim, S., & Tufekci, Z. (2021). Music emotion recognition using convolutional long short-term memory deep neural networks. *Engineering Science and Technology, an International Journal, 24*(3), 760–767. https://doi.org/10.1016/j.jestch.2020.10.009

Hung, H. T., Ching, J., Doh, S., Kim, N., Nam, J., & Yang, Y. H. (2021). EMOPIA: A multi-modal pop piano dataset for emotion recognition and emotion-based music generation. In *Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR)* (pp. 318–325). ISMIR.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*(5), 559–575. https://doi.org/10.1017/S0140525X08005293

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience, 15*(3), 170–180. https://doi.org/10.1038/nrn3666

Koelstra, S., Muhl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2012). DEAP: A database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing, 3*(1), 18–31. https://doi.org/10.1109/T-AFFC.2011.15

Li, Y., Zheng, W., Cui, Z., Zhang, T., & Zong, Y. (2018). A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)* (pp. 1561–1567). International Joint Conferences on Artificial Intelligence. https://doi.org/10.24963/ijcai.2018/216

Picard, R. W. (1997). *Affective computing*. MIT Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. https://doi.org/10.1037/h0077714

SCHOLARLY REVIEW

SCHOLARLY REVIEW JOURNAL

E-ISSN: 2996-8380
Published by Leadership & Innovation Lab

FALL 2026

Song, Y., Zheng, Q., Liu, B., & Gao, X. (2023). EEG-Conformer: Convolutional transformer for EEG decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31*, 710–719. https://doi.org/10.1109/TNSRE.2022.3230250

Steptoe, A., & Kivimäki, M. (2012). Stress and cardiovascular disease. *Nature Reviews Cardiology, 9*(6), 360–370. https://doi.org/10.1038/nrcardio.2012.45

Su, Y., Liu, Y., Xiao, Y., Ma, J., & Li, D. (2024). A review of artificial intelligence methods enabled music-evoked EEG emotion recognition and their applications. *Frontiers in Neuroscience, 18*, Article 1400444. https://doi.org/10.3389/fnins.2024.1400444

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998–6008). Curran Associates.

Zheng, W. L., & Lu, B. L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development, 7*(3), 162–175. https://doi.org/10.1109/TAMD.2015.2431497