Article

Uncovering Emotion Correlates to Transitions in EEG Energy Landscapes

Anubhav * and Kantaro Fujiwara

Department of Mathematical Informatics, The University of Tokyo, Tokyo 113-8654, Japan
* Correspondence: anubhav2901@g.ecc.u-tokyo.ac.jp

Abstract: Wearable brain-computer interfaces (BCIs) have made it feasible to monitor brain activity for emotion recognition in real-world settings. While deep learning models achieve high classification accuracy on electroencephalography (EEG) data, they often lack interpretability, limiting their neuroscientific relevance. In this study, we present an interpretable framework for EEG-based emotion analysis rooted in energy landscape analysis. EEG signals from the DEAP dataset were standardized and binarized prior to quantification of neural state transitions. We found significant subject-specific correlations between the number of state transitions and emotional ratings of valence and arousal. Further analysis revealed that certain binary brain states, particularly complementary pairs, were among the most frequently observed and showed emotion-dependent frequency differences. Transitions between these state pairs varied across subjects, suggesting their role as local minima in the brain's dynamic landscape. Our findings demonstrate that energy landscape analysis provides an interpretable alternative to black-box models, offering insights into how brain dynamics relate to emotional experiences. This approach contributes toward building explainable affective computing systems and supports the use of neural state modeling in emotion-aware BCIs.

Keywords: EEG,; energy landscape analysis; emotion recognition

1. Introduction

Interpreting human emotions through neurophysiological signals has become a central challenge in affective computing and cognitive neuroscience. Among various modalities, electroencephalography (EEG) has emerged as a valuable tool for real-time monitoring of emotional states due to its high temporal resolution and non-invasive nature. With recent advances in brain-computer interface (BCI) technology, particularly in developing portable and wearable EEG systems, it has become increasingly feasible to design low-cost, low-power devices capable of continuous emotion monitoring. In the post-COVID era, where mental health disorders such as anxiety and depression have become more prevalent, such technologies can enable early detection and intervention through real-time tracking of emotional fluctuations [1,2].

EEG has been extensively used for studying affective states and uncovering neural correlates of emotions. Existing research indicates that emotional responses are reflected in distinct spatiotemporal patterns of brain activity, which can be captured using machine learning techniques. Numerous studies have employed EEG-based features such as power spectral density, differential entropy, and connectivity measures to classify emotions, typically within the valence-arousal dimensional framework [3,4]. In parallel, other works focus on learning spatiotemporal representations of EEG signals directly, without relying on hand-crafted features, aiming to develop computationally efficient and real-time emotion recognition frameworks [5,6].

Recent studies have significantly improved EEG-based emotion recognition through deep learning methods. For instance, ref. [7] propose a graph convolutional neural network (GCNN) combined with long-short term memory (LSTM) to capture both spatial and temporal characteristics of EEG signals, attaining classification accuracies exceeding 90% for

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, *2*(1), 15–25

valence and arousal. Similarly, a residual LSTM model for multimodal emotion recognition on the DEAP dataset [8], demonstrating strong performance on EEG signals alone was introduced by ref. [9]. Other approaches employing convolutional neural networks (CNNs), LSTMs, attention mechanisms, or hybrid models have also reported high classification accuracies, often exceeding 90% under subject-dependent settings [10–14]. However, these models function primarily as black boxes and offer limited insights into the underlying neural mechanisms of emotional state transitions.

Despite the advances in classification accuracy, existing models fail to provide an interpretable framework for understanding how emotional states emerge and evolve over time in the brain. Specifically, no prior work has linked transitions between emotional states to the energy landscape topography derived from EEG activity. While energy landscape analysis has been applied to model brain dynamics in other neuroimaging modalities such as functional magnetic resonance imaging (fMRI) and functional near-infrared spectroscopy (fNIRS) [1,2,15,16], its potential for uncovering interpretable transitions in EEG-based emotion recognition remains unexplored. Given the importance of explainable models in mental health applications, it is critical to investigate alternative frameworks that can offer insight into the neurophysiological basis of emotions.

This study proposes an interpretable framework for EEG-based emotion recognition by utilizing energy landscape analysis to model state transitions corresponding to different emotional responses (Figure 1). At each time point, multichannel EEG activity is mapped to a binary pattern of above-baseline versus baseline activation across channels, and a pairwise maximum entropy model is used to construct an energy landscape over these binary states. Local minima in this landscape correspond to metastable brain states, and transitions between them trace trajectories of neural dynamics during emotional stimulation. Using binarized EEG activity from the DEAP dataset, we investigate how the number and structure of these state transitions relate to individual valence and arousal ratings. Unlike traditional deep learning approaches that primarily optimize prediction accuracy, our framework provides mechanistically interpretable representations of brain activity that can serve as candidate neural biomarkers for emotion-related brain dynamics.
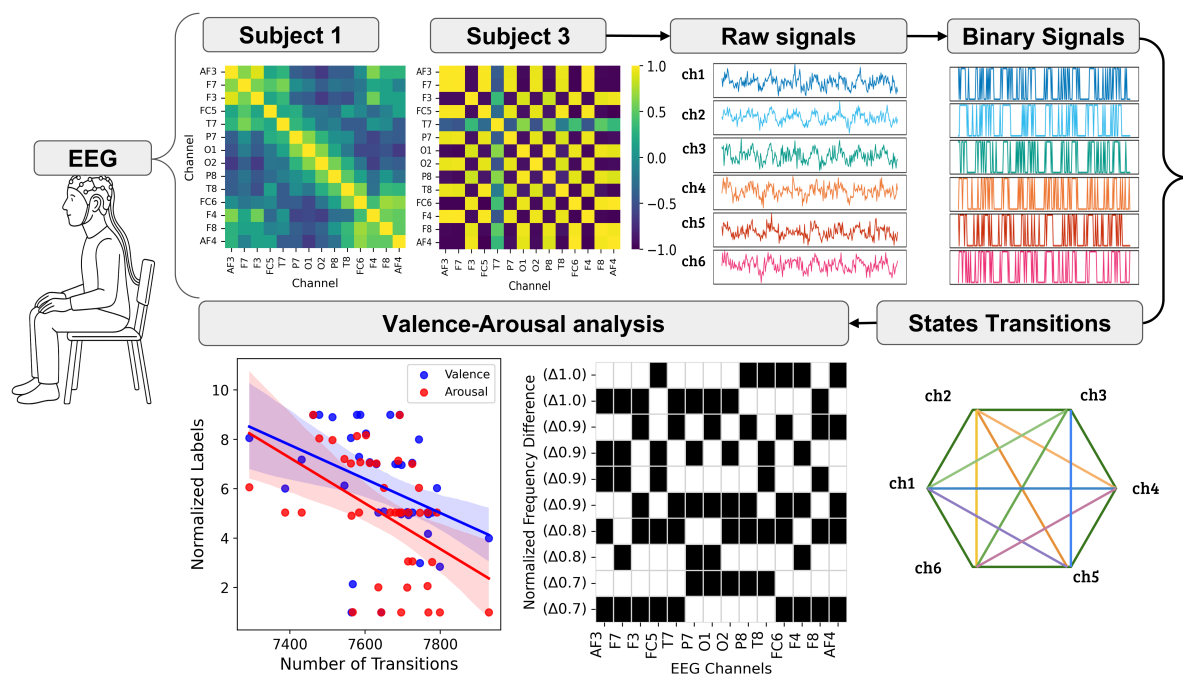

Figure 1. Overview of the proposed framework for interpretable EEG-based emotion analysis. EEG signals are recorded from multiple subjects and preprocessed to obtain raw time-series data. The signals are standardized and binarized to represent neural activity in binary state space. State transitions are computed and analyzed with valence and arousal labels. Subject-specific analyses highlight individual differences in transition-emotion correlations and dominant emotion-specific brain states.

The remainder of this paper is organized as follows: We present a literature review in Section 2 followed by Section 3 detailing the methodological procedures, including data preprocessing, binarization, and analysis of state transitions. Section 4 presents the experimental results, including identified metastable states and transition dynamics. Section 5 discusses the implications of the findings, limitations, and potential avenues for future work with conclusion in Section 6.

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, 2(1), 15–25

2. Related Works

2.1. Deep Learning Approaches for EEG-Based Emotion Recognition

EEG-based emotion recognition has attracted considerable research attention due to its characteristics of non-invasively capturing high-temporal-resolution brain activity. The DEAP dataset [8] has become a benchmark for evaluating classification models of emotional dimensions such as valence and arousal. Early studies employed hand-crafted features and conventional classifiers, but recent works have demonstrated that deep-learning models can substantially improve classification accuracy. CNNs and recurrent neural networks (RNNs), such as LSTMs, have been widely used for modeling spatiotemporal EEG patterns. While ref. [7] introduced a hybrid model combining GCNN with LSTM to capture spatial connectivity and temporal dynamics in EEG data, achieving over 90% accuracy on DEAP for valence and arousal, ref. [9] proposed a multimodal residual LSTM (MMResLSTM) network that yielded competitive results using EEG alone.

Other studies have further integrated attention mechanisms or hybrid optimization methods to enhance performance. In ref. [13], authors incorporated an attention module into an LSTM framework, improving the model's ability to focus on temporally salient EEG segments, while ref. [14] applied a meta-heuristic hybrid model combining artificial bee colony and grey wolf optimization to tune CNNs, reporting near-perfect accuracy under subject-dependent conditions. While these methods report high classification performance, they often lack interpretability, treating EEG signals as opaque inputs to black-box models.

Several studies acknowledge the trade-off between classification accuracy and explainability in affective computing systems. Works such as, refs. [3,12], underscore this issue, noting that although deep models outperform classical approaches, their internal representations offer limited insight into the underlying neurophysiological processes. As a result, there is increasing interest in methods that maintain high accuracy while providing interpretable mappings between brain activity and emotional states.

2.2. EEG-Microstate and Energy Landscape Analysis

To address the need for interpretability in emotion recognition systems, several researchers have explored state-transition analyses of neural activity. These approaches assume that brain dynamics can be represented as transitions among discrete neural states and that the structure of these transitions carries meaningful information about cognitive or affective processes. In the EEG domain, ref. [17] proposed a microstate-based framework to analyze cognitive effort during a Stroop task. By computing transition costs between EEG microstates using optimal transport theory, they demonstrated that increased task difficulty corresponds to higher cognitive reconfiguration effort. While they quantify cognitive effort through EEG microstate transition costs, our approach focuses on emotional dynamics by modeling transitions between binarized EEG states. Both frameworks emphasize interpretable neural dynamics, but our method highlights emotion-specific attractor states in an energy landscape, thus extending microstate-based interpretability into the affective domain.

Beyond EEG, energy landscape analysis has been extensively applied to fMRI and fNIRS data [15]. In ref. [18] used a pairwise maximum entropy model to estimate energy landscapes from resting-state fMRI signals. They found that brain state transitions are organized around hub configurations, with the default mode network (DMN) mediating between distinct cognitive states. Similarly, ref. [2] applied this framework to study Alzheimer's disease, finding that patients exhibited altered dwell times and transition frequencies, which correlated with cognitive impairment. In ref. [1] extended this methodology to fNIRS data in clinical populations. By constructing energy landscapes from cognitive task recordings, they showed that individuals with depression displayed more metastable states and shallower attractor basins. These quantitative features served as both interpretable markers of neurodynamics and effective discriminators between clinical and control groups.

Although energy landscape models have demonstrated success in characterizing brain dynamics across modalities, their application to EEG-based emotion recognition remains largely unexplored. To our knowledge, no existing study has linked emotional state transitions in EEG to energy landscape topology. This gap highlights the need for frameworks combining the predictive power of state-of-the-art classification algorithms with principled analyses of state-based affect behaviors.

3. Methods

3.1. Dataset Description

This study utilizes the DEAP dataset [8], a benchmark resource for emotion analysis using physiological signals. The dataset consists of 32 participants undergoing 40 trials wherein one-minute music videos were presented to evoke emotional responses. During each trial, 32-channel EEG signals were recorded, which

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, *2*(1), 15–25

were later downsampled to a sampling rate of 128 Hz. The dataset authors preprocessed all EEG data to remove ocular artifacts and line noise through independent component analysis and bandpass filtering (0.5–45 Hz) (http://eecs.qmul.ac.uk/mmv/datasets/deap/readme.html (accessed on 11 July 2024)).

3.2. Proposed Framework

We propose an interpretable analysis of EEG signals based on state transitions derived from binarized EEG activity and modeled using energy landscape analysis. The following methodological steps were followed to ensure consistency and computational tractability.

3.2.1. Data Preparation

Following prior energy landscape analyses of brain activity, which represent regional activation as binary on/off states, we binarized each EEG channel after subject-wise standardization. Specifically, values with positive z scores were set to 1 and non-positive values to 0. Because DEAP provides artifact-corrected, bandpass-filtered (0.5–45 Hz) EEG, the resulting binary states can be interpreted as above-baseline versus baseline activity for each subject and channel. This transformation converts continuous EEG data into discrete binary state vectors suitable for energy landscape modeling.

Due to the exponential increase in computational complexity with the number of EEG channels in pairwise maximum entropy models, we restricted the analysis to a subset of 14 channels. These correspond to the configuration of the Emotiv Epoc headset, a commercially available wearable EEG device. This coarse coding focuses on the configuration of active channels rather than precise amplitudes and keeps the pairwise maximum entropy model tractable in the 2^{14} state space considered here. This subset selection allows our framework to remain computationally efficient while remaining applicable in mobile and real-world settings, where portable EEG hardware is typically constrained in channel count.

We computed pairwise correlations between channel activity across trials to understand the subject-wise consistency and variability in EEG dynamics. The analysis revealed moderate to strong intra-subject consistency but significant inter-subject variability, highlighting the necessity to evaluate the energy landscape features under subject-dependent and subject-independent protocols. This comparative strategy enables investigation of the trade-off between personalized emotion modeling and generalizable affective decoding.

3.2.2. State-Transition Analysis

We then quantified the number of state transitions for each trial and grouped them by subject. Figure 2 shows the distribution of transition counts across subjects. For some subjects, transitions are tightly concentrated across all trials, indicating consistent dynamic reconfiguration of brain states in response to emotional stimuli. In contrast, other subjects exhibit a wider spread in transition counts, which may reflect heterogeneous affective responses or the stabilization of neural dynamics into quasi-attractor states with limited transitions.
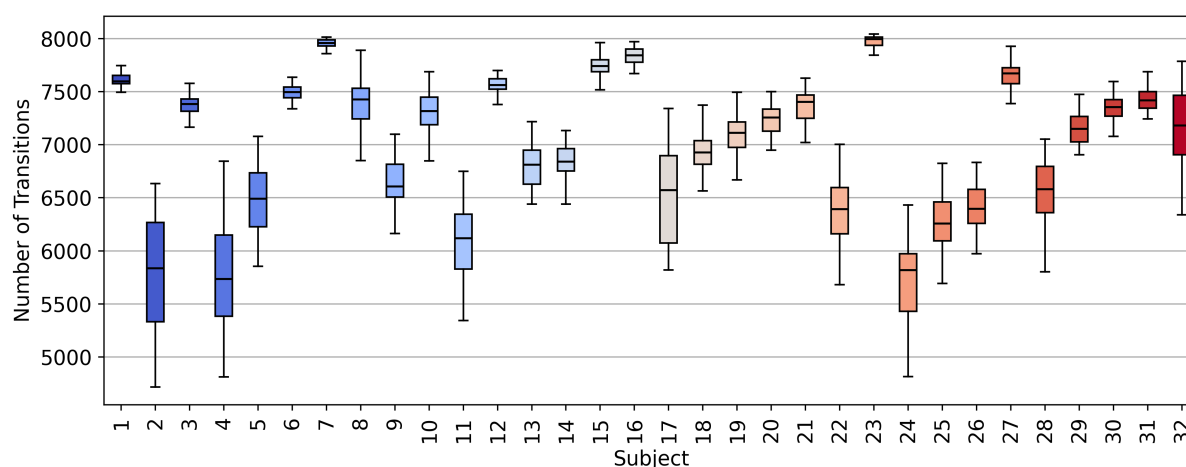

Figure 2. Distribution of number of state transitions across trials for each subject. A narrower spread suggests consistent state-switching behavior across emotional conditions, whereas a broader spread may indicate increased state stability.

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, 2(1), 15–25

Each trial consists of approximately 63 s of continuous EEG sampled at 128 Hz (8064 time points) after downsampling by the dataset authors. After binarization, we obtain for every trial a sequence of 8064 binary state vectors across 14 channels. We define a state transition whenever the binary pattern at time t differs from that at time $t + 1$ in at least one channel. We deliberately do not apply additional temporal windowing, since our aim is to characterize the full evolution of brain states during each emotional clip rather than extract short windowed features. Although transitions are counted at the native 128 Hz sampling rate, this choice only sets the temporal resolution at which state changes are observed; exploring coarser temporal sampling (e.g., downsampled trajectories) is an interesting direction for future work.

3.2.3. Analysis

In addition to Spearman rank correlations, which capture monotonic associations without assuming linearity, we also fitted subject-wise logistic regression models. For each participant, we binarized valence and arousal ratings into high versus low labels based on the within-subject z scores and modeled the probability of a high rating as a function of standardized transition counts using a logistic link. We further included a quadratic term in transitions to probe for potential non-monotonic relationships.

For completeness, we also fitted a linear mixed-effects model to the pooled data, with a random intercept per subject. For each emotion dimension y_{ij} (valence or arousal) of subject i in trial j, we specified

$$y_{ij} = \beta_0 + \beta_1 \, \mathrm{tnm}_{ij}^{(z)} + u_i + \varepsilon_{ij}, \tag{1}$$

where $\mathrm{tnm}_{ij}^{(z)}$ is the standardized number of transitions for that trial, $u_i \sim \mathcal{N}(0, \sigma_u^2)$ is a subject-specific random intercept, and $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ is the residual error. Because ratings were standardized within subject, most between-subject mean differences are absorbed by the standardization, so the model primarily estimates a common slope β_1 across subjects.

In the next section, we explore the correlations between the number of transitions and subjective valence and arousal ratings. This analysis examines whether the dynamic properties of EEG state transitions are predictive of emotional dimensions and can offer interpretable insights into affective brain dynamics.

4. Results

The DEAP dataset provides emotion annotations on a scale of 1 to 9 along two continuous dimensions: valence and arousal. To ensure that inter-subject differences in rating tendencies do not bias the analysis, we independently standardized the valence and arousal labels for each subject. This normalization step maps the emotional ratings for each subject to a zero-centered distribution, facilitating fair comparison across trials and individuals.

To evaluate the association between EEG transition dynamics and emotional states, we performed subject-wise Spearman correlation analyses between the number of state transitions and normalized valence and arousal scores across the 40 trials. The rationale behind this approach is that the frequency of transitions between metastable states may reflect the brain's affective reactivity or flexibility under emotional stimulation.

Figures 3a,b illustrate representative regression plots for Subject 10 and Subject 27, respectively. For Subject 10, a strong positive correlation was observed between valence and number of transitions ($\rho = 0.50$, $p = 0.001$), suggesting that higher valence (positive emotions) are associated with greater neural state switching. In contrast, for Subject 27, both valence and arousal show significant negative correlations with the number of transitions ($\rho = -0.50$, $p = 0.001$ for valence; $\rho = -0.53$, $p < 0.001$ for arousal), indicating that lower affective ratings coincide with more frequent state changes.

Pooling all 32 subjects and 40 trials each ($N = 1280$), we observed essentially no group-level monotonic association between transitions and valence (Spearman's $\rho = -0.002$, $p = 0.936$). For arousal, the pooled association was small but statistically significant (Spearman's $\rho = 0.060$, $p = 0.032$) (see Figure 4). These tiny effect sizes indicate that most of the interesting structure lies in individual transition–emotion relationships rather than a strong global trend.

Table 1 summarizes Spearman correlation coefficients and p values for subjects that showed significant correlations in at least one emotion dimension. Subjects 10, 16, 17, and 27 exhibit statistically significant correlations in at least one emotional dimension. Subject 3 and Subject 5 also showed a significant correlation with arousal. However, the direction and strength of these correlations vary substantially across subjects, further supporting the notion of individualized neural-affective coupling.
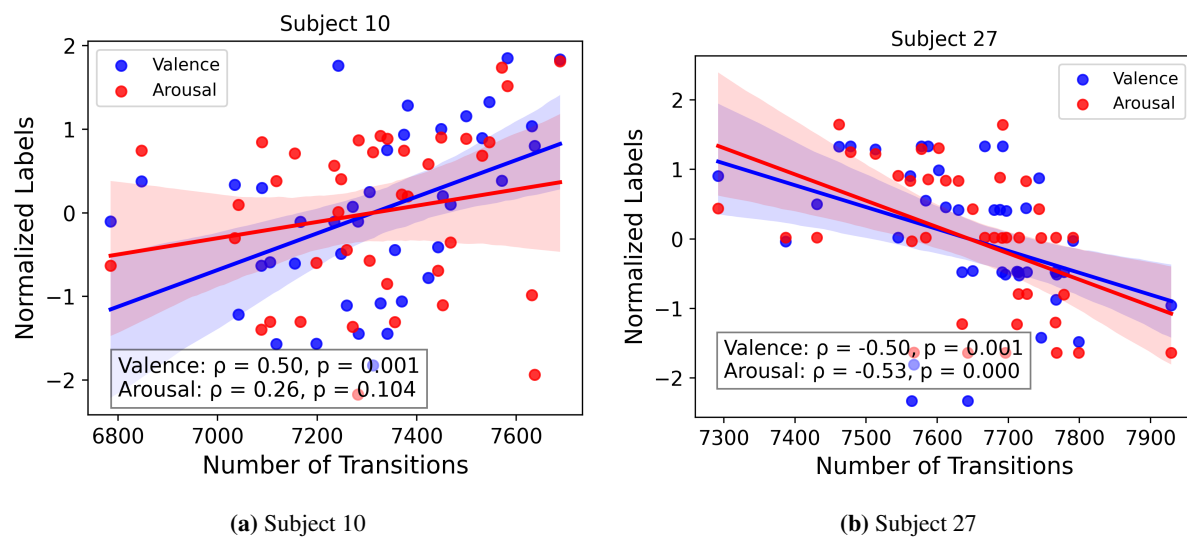
(a) Subject 10   **(b)** Subject 27

Figure 3. Regression between number of transitions and normalized valence/arousal ratings.
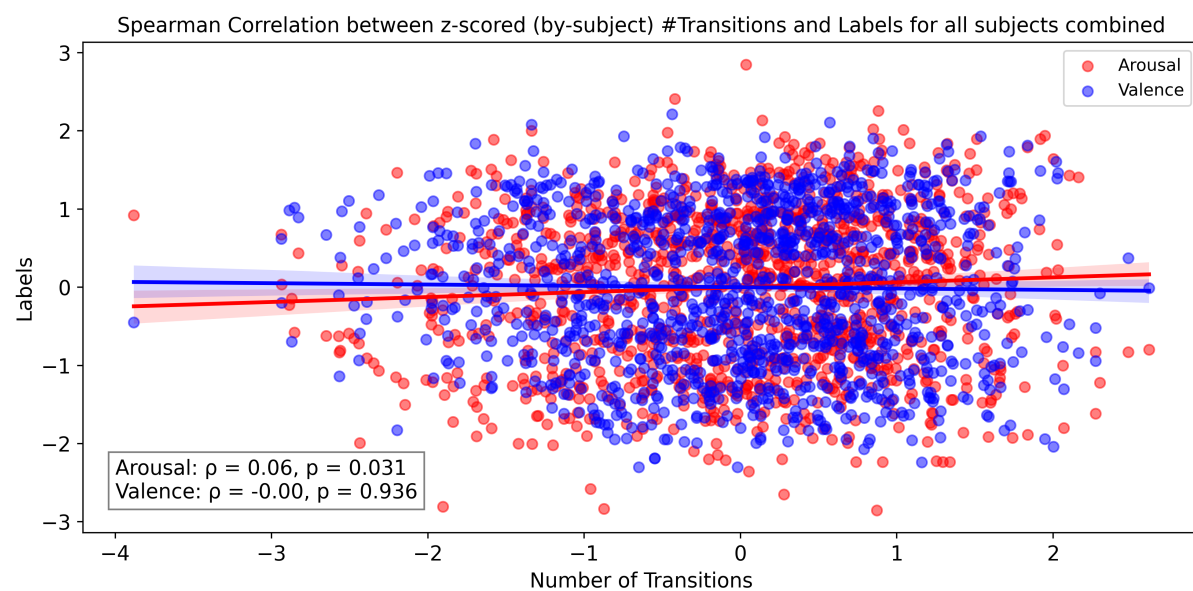

Figure 4. Group-level correlation analysis between the number of transitions and normalized valence/arousal ratings. Valence shows essentially no global trend, while arousal exhibits a small but significant positive association. The large spread of points reflects pronounced individual variability, underscoring the importance of personalized modeling approaches in emotion recognition systems.

Table 1. Selected subject-level Spearman's ρ and p-values between standardized transition counts and emotion ratings, highlighting participants with significant or near-significant associations.

Subject	Valence ρ (*p*)	Arousal ρ (*p*)	Significance
1	0.17 (0.289)	0.18 (0.268)	–
3	−0.10 (0.528)	**0.32 (0.047)**	Arousal
5	0.19 (0.243)	**0.38 (0.015)**	Arousal
6	**−0.31 (0.048)**	0.06 (0.707)	Valence
10	**0.50 (0.001)**	0.26 (0.104)	Valence
11	**−0.44 (0.004)**	0.20 (0.214)	Valence
16	0.23 (0.145)	**0.42 (0.007)**	Arousal
17	0.23 (0.149)	**0.54 (0.000)**	Arousal
27	**−0.50 (0.001)**	**−0.53 (0.000)**	**Both**
32	**0.42 (0.006)**	−0.05 (0.736)	Valence

The mixed effects model yielded a slope close to zero for valence ($\beta_1 = -0.017$, $p = 0.54$) and a small but significant positive slope for arousal ($\beta_1 = 0.063$, $p = 0.024$), mirroring the pooled Spearman results and confirming that any group-level association between transitions and emotion is very weak.

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, 2(1), 15–25

The subject-wise logistic models corroborated this picture. For valence, 17 of 32 participants exhibited positive slopes relating higher transition counts to a higher probability of high valence, and 15 showed negative slopes, but only 4 slopes reached $p < 0.05$ and the median odds ratio (OR) was close to one ($OR = 1.05$). Thus, at the individual level valence associations are heterogeneous and generally weak. For arousal, 19 participants had positive slopes and 13 had negative slopes (median $OR = 1.09$), with three significantly positive and one significantly negative slope, consistent with the small positive pooled correlation. When we added a quadratic term in transitions, only 3 subjects (valence) and 2 subjects (arousal) had significant curvature, close to the number expected by chance. We therefore conclude that simple monotonic models are adequate and that more complex non-monotonic functions are not systematically required.

These results demonstrate that the number of metastable state transitions derived from energy landscape analysis carries information relevant to emotional experience, particularly at the individual level. However, the heterogeneity of correlation direction across subjects suggests that emotional correlates of brain dynamics are likely subject-specific. Therefore, in future prediction models, we should incorporate transition metrics for affective computing and devise mechanisms for personalization to account for inter-subject variability.

5. Discussion

This study proposes interpretable correlates of EEG-based emotion dynamics, specifically the number and structure of state transitions. Our subject-wise analyses revealed that, for a subset of participants, the number of transitions in brain states is significantly associated with emotional ratings, while overall patterns are heterogeneous and generally weak at the group-level. Notably, both positive and negative correlations were observed, highlighting individual-specific trajectories of emotional dynamics in the energy landscape. These findings reinforce the notion that emotional experiences are reflected in brain activity and are dynamically constructed through ongoing transitions. This dynamic perspective complements traditional static feature-based emotion recognition and highlights the potential of modeling emotion as a process embedded within evolving brain states.

To interpret these results further, we investigated whether specific brain states are more prevalent for certain emotions. We compared the most frequent states in trials labeled with high vs low valence and similarly for arousal. The resulting heatmaps (Figure 5a,b) depict the top 10 most discriminative brain states ranked by their normalized frequency difference. This prevalence suggests that emotional information may be encoded not just in isolated brain states but in the transitions between them, supporting the idea that emotion is inherently tied to the neurodynamics.
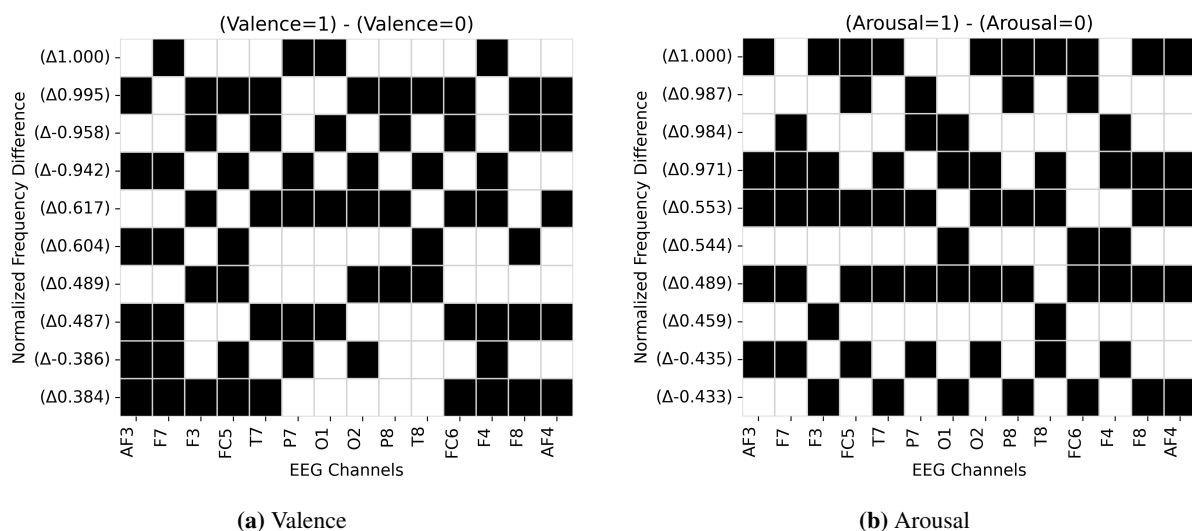

(a) Valence **(b)** Arousal

Figure 5. Top 10 most discriminative brain states for valence and arousal. Each row corresponds to a state, black indicating an active EEG channel, and the normalized frequency difference is annotated on the y-axis.

These visualizations show that many prominent states occur in complementary pairs, i.e., state configurations that are binary inverses of one another. This observation suggests the existence of opposing or mutually exclusive neural activation patterns tied to emotional polarity, possibly reflecting transitions between antagonistic neural regimes (e.g., approach vs. withdrawal circuits). The recurring presence of these complementary states implies that they represent local minima in the brain's energy landscape, attractor-like states that are recurrent across different emotions.

To investigate this further, we analyzed the frequency of transitions between these complementary state pairs

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, 2(1), 15–25

across all subjects and compared them for high vs. low valence and arousal conditions. As shown in Figure 6, the difference in mean complement transitions varies considerably across subjects. Positive values denote more frequent transitions during positive (valence = 1 or arousal = 1) trials, while negative values indicate predominance during low-value emotion trials (valence = 0 or arousal = 0). Statistical significance markers indicate subjects where these differences were found significant using a *t*-test.
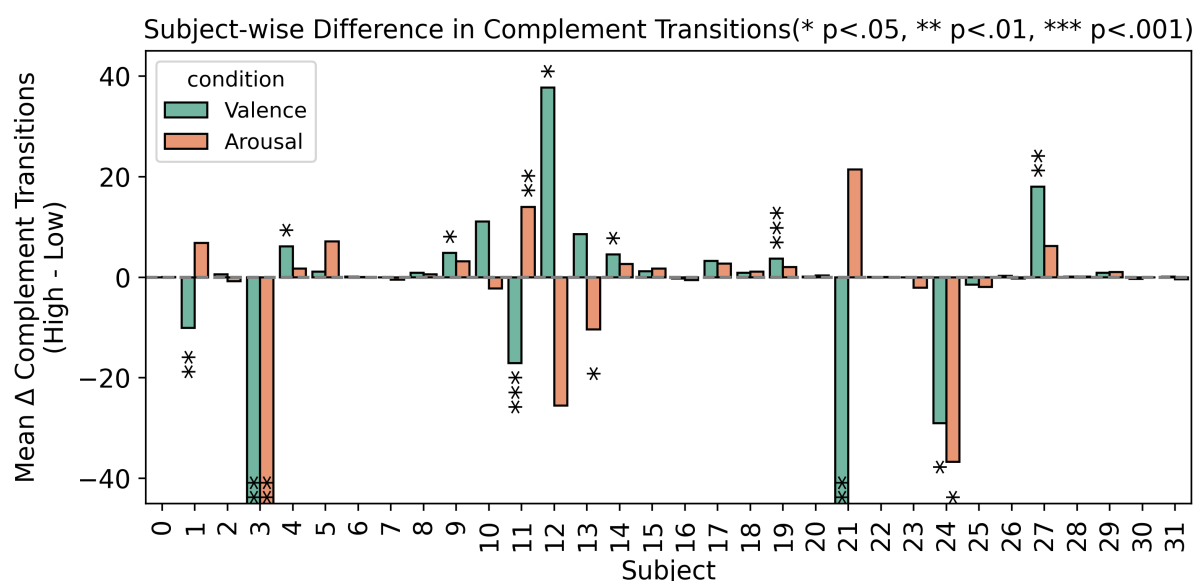

Figure 6. Subject-wise difference in mean transitions between complementary states under high vs. low valence/arousal conditions. Asterisks indicate statistical significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

These findings emphasize that complementary state transitions are among the most common transitions observed across trials, suggesting their structural relevance within the energy landscape. The variability in their association with emotion labels across subjects highlights the need for personalized interpretation. Nonetheless, the observation that emotion-relevant trials consistently recruit such state pairs supports their interpretive value as potential neurodynamic biomarkers. Future work involving detailed topological analysis of the energy surface (e.g., basin stability, saddle paths) may uncover the latent neural mechanisms underlying these recurrent transitions.

To assess whether these group-level trends depended on the exact binarization threshold, we recomputed transition counts using a range of thresholds applied to the standardized EEG signals (from $z = -1.0$ to $z = 1.0$ in steps of 0.1) and repeated the pooled Spearman correlations, see Figure 7. Valence correlations remained close to zero and non-significant for all thresholds (Spearman's ρ between -0.02 and 0.01), whereas arousal correlations stayed small but consistently positive (Spearman's $\rho = 0.035$–0.075). This pattern indicates that the absence of a valence effect and the weak positive arousal effect are robust to moderate changes in the binarization rule.

Existing EEG-based emotion recognition models primarily employ deep learning architectures that optimize for predictive accuracy but offer limited interpretability [7, 9, 13]. These models do not explain how dynamic brain activity patterns evolve across emotional conditions. By contrast, our study adopts an energy landscape perspective rooted in statistical physics, offering a mechanistic interpretation of emotional states as stable or metastable configurations and transitions as quantifiable dynamical events. Prior work has applied energy landscapes to fMRI and fNIRS [2, 18], but not to EEG-based emotion recognition.

By linking the number of neural state transitions and the structure of energy basins to emotion dimensions, our work provides a pathway for interpretable emotion decoding grounded in neurodynamics. This reinforces the emerging view that emotion is a dynamic phenomenon shaped by transient neural states. Our approach bridges this perspective with empirical EEG data, showing that changes in emotional state correspond to measurable shifts in brain dynamics across energy landscapes. Moreover, our identification of emotion-specific and complementary states aligns with findings from microstate and attractor landscape theories in cognitive neuroscience, where specific brain states reflect functional processing.
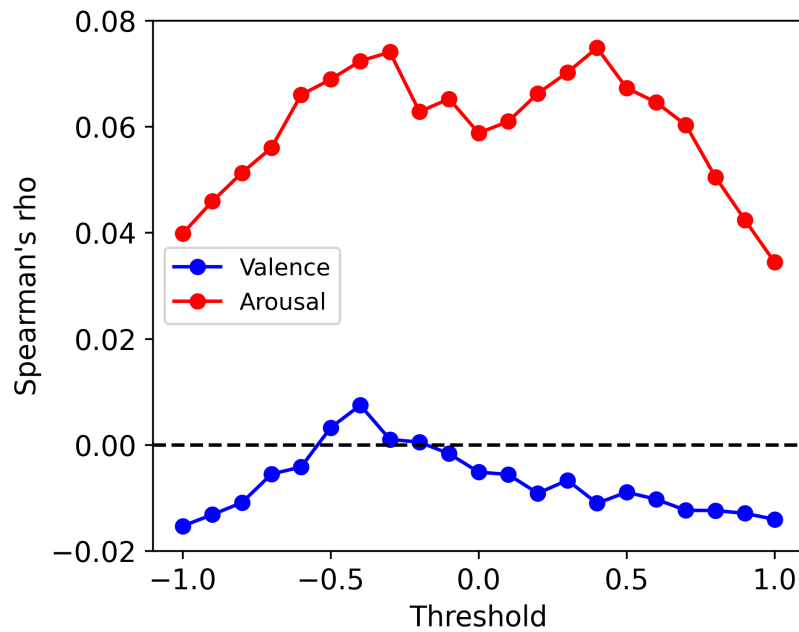
Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, 2(1), 15–25

Figure 7. Spearman correlation between standardized transition counts and emotion ratings as a function of the binarization threshold applied to the z-scored EEG signals. For each threshold, transitions were recomputed and correlations were evaluated on the normalized labels. Valence correlations remain close to zero across all thresholds, whereas arousal correlations stay small but consistently positive, indicating that the main group-level conclusions are robust to the exact threshold choice.

5.1. Limitations

Despite the promising findings, this study has several limitations. First, the results depend on the binarization method applied to EEG signals. However, for future applications such as online emotion estimation, reducing computational complexity is crucial, and binarization can contribute positively to this goal. A related challenge is that artifacts or signal noise may distort the binary configurations and the inferred energy landscapes. To address this concern, our threshold sensitivity analysis suggests that the main group-level trends remain stable across a range of thresholds; nonetheless, more adaptive binarization techniques should be investigated to improve reliability. Second, the trial duration in DEAP is limited to one minute, potentially restricting the range of state transitions that can be empirically observed. Therefore, exploring longer EEG recordings would further enable observing slower or less frequent transitions that may be crucial for emotional processing. Third, we restricted our analysis to a subset of 14 EEG channels matching the Emotiv Epoc layout in order to keep the pairwise maximum entropy model tractable. While this choice aligns the framework with realistic wearable EEG hardware, it inevitably discards information from the remaining DEAP channels and may limit the representational capacity of the inferred energy landscapes.

5.2. Future Work

Future studies can address these limitations by employing signal processing pipelines that compute spectral band power features (e.g., alpha, beta bands) before binarization. Band-specific features are widely used in cognitive neuroscience and may enhance both interpretability and robustness to noise. Furthermore, detailed energy landscape metrics such as basin depth, transition entropy, or attractor proximity could be computed to model individual differences in emotion dynamics. In addition, extending the framework to higher-density EEG and systematically comparing landscapes inferred from different channel subsets would clarify the trade-off between computational tractability and representational richness. Finally, coupling the energy landscape framework with deep learning classifiers could yield hybrid models that retain interpretability while achieving high classification accuracy.

6. Conclusions

With the increasing adoption of wearable BCIs, interpretable models of brain activity have vast potential for reliable and personalized emotion monitoring and prediction. This study presents an interpretable approach to EEG-based emotion recognition rooted in energy landscape analysis to characterize transitions among brain states associated with different emotions. By quantifying the frequency and structure of these state transitions, we demonstrated how individual differences in neural dynamics relate to subjective experiences of valence and arousal.

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, *2*(1), 15–25

The subject-wise correlation analysis revealed that the number of transitions between neural states is significantly associated with emotional ratings in several individuals. However, no consistent trend was observed at the population level. By analyzing high-frequency emotion-specific states and their complementary counterparts, we identified candidate brain states that recur across emotional trials. The prevalence of transitions between complementary states and their emotion-specific dependency further supports the hypothesis that these states represent meaningful structures in the brain's energy landscape.

By integrating empirical EEG data with a physics-based modeling approach, this research addresses a critical gap in the literature: the lack of interpretability in deep learning-based emotion recognition systems. Our findings suggest that energy landscape dynamics offer a viable path toward identifying neural biomarkers of affective processing, particularly when analyzed in a subject-specific manner.

Nevertheless, the current analysis may suffer from limitations in the use of binarization methods and trial duration. Thus, in future work, we will explore more robust spectral feature-based transformations and investigate topological properties of energy landscapes in greater depth. By doing so, we aim to bridge the divide between high-performance emotion classification and interpretable EEG modeling, thereby advancing the development of explainable affective brain-computer interfaces.

Author Contributions

A.: conceptualization, methodology, software, formal analysis, investigation, visualization, writing—original draft preparation; K.F.: supervision, validation, resources, funding acquisition, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding

Institutional Review Board Statement

Ethical review and approval were waived for this study, as it involved the analysis of a publicly available dataset with no personally identifiable information, and no new data were collected from human participants. The original data collection procedures and ethical approvals are described in the DEAP dataset documentation.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

During the preparation of this work, the authors used ChatGPT (OpenAI) to assist with language polishing and improving the clarity and flow of the text. After using this tool, the authors reviewed and edited all content as needed and take full responsibility for the content of the published article.

References

1. Wu, Y.; Qiao, S.; Zhong, J.; et al. FNIRS-based energy landscape analysis to signify brain activity dynamics of individuals with depression. *CNS Neurosci. Ther.* **2024**, *30*, e70139.
2. Xing, L.; Guo, Z.; Long, Z. Energy landscape analysis of brain network dynamics in Alzheimer's disease. *Front. Aging Neurosci.* **2024**, *16*, 1375091.
3. Liu, J.; Wu, G.; Luo, Y.; et al. EEG-Based Emotion Classification Using a Deep Neural Network and Sparse Autoencoder. *Front. Syst. Neurosci.* **2020**, *14*, 43.
4. Anubhav.; Nath, D.; Singh, M.; et al. An Efficient Approach to EEG-Based Emotion Recognition using LSTM Network. In Proceedings of the 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), Langkawi, Malaysia, 28–29 February 2020; pp. 88–92.

Anubhav and Fujiwara

Trans. Artif. Intell. **2026**, 2(1), 15–25

5. Anubhav; Fujiwara, K. Reservoir Splitting Method for EEG-based Emotion Recognition. In Proceedings of the 11th International Winter Conference on Brain-Computer Interface (BCI), Gangwon, Republic of Korea, 20–22 February 2023; pp. 1–5.

6. Anubhav.; Fujiwara, K. Across Trials vs. Subjects vs. Contexts: A Multi-Reservoir Computing Approach for EEG Variations in Emotion Recognition. In Proceedings of the 26th International Conference on Multimodal Interaction, San Jose, Costa Rica, 4–8 November 2024; pp. 518–525.

7. Yin, Y.; Zheng, X.; Hu, B.; et al. EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. *Appl. Soft Comput.* **2021**, *100*, 106954.

8. Koelstra, S.; Muhl, C.; Soleymani, M.; et al. DEAP: A Database for Emotion Analysis ;Using Physiological Signals. *IEEE Trans. Affect. Comput.* **2012**, *3*, 18–31.

9. Ma, J.; Tang, H.; Zheng, W.L.; et al. Emotion Recognition Using Multimodal Residual LSTM Network. In Proceedings of the 27th ACM International Conference on Multimedia (MM'19), Nice, France, 21–25 October 2019; pp. 1776–1782.

10. Hasan, M.; Anzum, R.N.; Yasmin, S.; et al. Fine-Grained Emotion Recognition from EEG Signal Using Fast Fourier Transformation and CNN. In Proceedings of the Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) and 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 16–20 August 2021; pp. 1–9.

11. Nath, D.; Anubhav.; Singh, M.; Sethia, D.; Kalra, D.; Indu, S. A Comparative Study of Subject-Dependent and Subject-Independent Strategies for EEG-Based Emotion Recognition Using LSTM Network. In Proceedings of the 4th International Conference on Compute and Data Analysis, Silicon Valley, CA, USA, 9–12 March 2020; pp. 142–147.

12. Alhalaseh, R.; Alasasfeh, S. Machine-Learning-Based Emotion Recognition System Using EEG Signals. *Computers* **2020**, *9*, 95.

13. Kim, Y.; Choi, A. EEG-Based Emotion Classification Using Long Short-Term Memory Network with Attention Mechanism. *Sensors* **2020**, *20*, 6727.

14. Karthiga, M.; Suganya, E.; Sountharrajan, S.; et al. EEG based smart emotion recognition using meta heuristic optimization and hybrid deep learning techniques. *Sci. Rep.* **2024**, *14*, 30251.

15. Watanabe, T.; Hirose, S.; Wada, H.; et al. A Pairwise Maximum Entropy Model Accurately Describes Resting-State Human Brain Networks. *Nat. Commun.* **2013**, *4*, 1370.

16. Watanabe, T.; Masuda, N.; Megumi, F.; et al. Energy Landscape and Dynamics of Brain Activity during Human Bistable Perception. *Nat. Commun.* **2014**, *5*, 4765.

17. Barzon, G.; Ambrosini, E.; Vallesi, A.; et al. EEG microstate transition cost correlates with task demands. *PLoS Comput. Biol.* **2024**, *20*, e1012521.

18. Kang, J.; Pae, C.; Park, H.J. Graph-theoretical analysis for energy landscape reveals the organization of state transitions in the resting-state human cerebral cortex. *PLoS ONE* **2019**, *14*, e0222161.