Journal of
Quantum Computing

Tech Science Press

https://doi.org/10.32604/jqc.2026.086882

ARTICLE

EEG-Based Emotion Recognition Using Deep Quantum Features

Reon Yoshida[1], **Keiko Ono**[2,*], **Kentaro Ohki**[3] **and Takuya Futagami**[2]

[1]Graduate School of Science and Engineering, Doshisha University, Kyoto, Japan
[2]Faculty of Science and Engineering, Doshisha University, Kyoto, Japan
[3]School of Information Science and Technology, Department of Applied Computer Engineering, Tokai University, Kanagawa, Japan

*Corresponding Author: Keiko Ono. Email: kono@mail.doshisha.ac.jp

Received: 06 June 2026; Accepted: 05 August 2026; Published: 07 September 2026

ABSTRACT: Quantum machine learning (QML) has attracted significant attention for its potential to accelerate computation and improve efficiency, particularly through quantum feature maps that may enable the separation of data not linearly separable in classical spaces. Although this capability remains largely theoretical, it represents a promising direction for addressing complex learning tasks. However, current quantum devices suffer from low error tolerance and a limited number of qubits, which has spurred interest in hybrid quantum–classical approaches. One such application is EEG-based emotion recognition, which involves complex, nonlinear signals and substantial inter-subject variability. While prior studies have applied QML to EEG analysis—using QSVMs on handcrafted features, sequential quantum-then-classical pipelines on band-power attributes, or fully quantum hybrid models for brain–computer interfacing—none of these approaches integrate high-level convolutional representations with entanglement-based quantum feature maps in a parallel fusion architecture, nor do they investigate subject-wise adaptive contribution of quantum features. To address these gaps, this study proposes a hybrid architecture that integrates classical and quantum computing by leveraging *deep quantum features*. These features are synthesized by combining deep representations extracted via a Convolutional Neural Network (CNN) with quantum features generated through a ZZFeatureMap-based quantum circuit using angle encoding, in which each of four qubits is semantically aligned with a canonical EEG frequency band (θ, α, β, γ). The two branches are concatenated and processed by a two-layer multilayer perceptron for binary classification of Valence and Arousal on the DEAP dataset. Across ten independent training runs, the proposed model achieved average classification accuracies of 79.6% for Valence and 81.5% for Arousal. Based on the normality of the paired subject-level differences, two-sided paired t-tests or Wilcoxon signed-rank tests showed significant improvements over the standalone classical and quantum models ($p < 0.05$) under the same segment-level subject-dependent protocol. Additional robustness and ablation analyses further showed that the same ordering, Proposed > Classical > Quantum, was preserved under repeated-seed evaluation, and that replacing the quantum branch with shape-matched random noise substantially degraded performance. Furthermore, interpretability analysis using LIME showed subject-dependent use of both feature branches, with the quantum-derived branch receiving greater importance in many cases. This result supports the usefulness of the quantum-derived representation as complementary predictive information, although it should not be interpreted as causal evidence of a uniquely quantum advantage. Together, the parallel CNN–quantum fusion design, the band-aligned qubit encoding, and the subject-wise interpretability analysis distinguish this work from existing QML-EEG studies, and highlight the efficacy of classical–quantum hybrid approaches for complex biomedical signal analysis.

KEYWORDS: Quantum machine learning; emotion recognition; hybrid architecture; EEG data

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

In recent years, research on quantum computing has garnered significant attention. However, with current technology, the number of available qubits remains insufficient for solving many complex problems, and quantum systems are highly susceptible to noise. As a result, existing quantum computers often produce outputs with non-negligible errors. A computational paradigm designed to leverage quantum computers while accounting for this noise is referred to as Noisy Intermediate-Scale Quantum (NISQ) technology [1,2].

One promising application of NISQ devices is Quantum Machine Learning (QML). QML is expected to introduce new possibilities to conventional machine learning methods (hereinafter referred to as classical machine learning) through advantages such as novel feature representations and algorithmic diversity. In particular, quantum feature maps can incorporate entanglement and non-linear transformations, enabling expressive representations that differ fundamentally from those in classical kernels [3–5]. This structural complexity may allow QML models to capture intricate relationships in data without relying on extremely deep architectures.

In classical deep learning for classification, effective feature extraction typically requires numerous convolution operations and very deep network architectures. In practice, it is common to increase the depth of the network until sufficiently discriminative features are obtained, which often results in substantial computational cost and a need for large-scale training data. QML, on the other hand, offers an alternative approach by generating expressive features through quantum circuits, potentially reducing the reliance on deep networks. Consequently, research applying QML to various domains is gaining momentum. While much of the existing research in QML focuses on benchmark datasets such as MNIST or Digits, efforts to apply QML to real-world problems have recently begun [6–8]. For example, studies have demonstrated that QML-based models can outperform classical approaches in practical tasks such as COVID-19 detection from chest X-ray images [9], MRI classification [10], vehicle recognition [11], and analysis of electronic health records [12].

Inspired by these advances, this study focuses on applying QML to electroencephalogram (EEG) data and proposes a novel approach for EEG-based emotion recognition. We target this problem as a practical application that can leverage QML's strength in capturing complex patterns through entanglement and non-linear transformations. We hypothesize that QML can identify intricate relationships within EEG data with greater accuracy than classical machine learning.

Recent years have seen growing interest in applying QML to EEG analysis. Garg et al. [13] applied a Quantum Support Vector Machine (QSVM) to DEAP-based emotion recognition, reporting accuracies of 65.6% and 75.0% for valence and arousal, respectively, using handcrafted statistical features. Chandanwala et al. [14] proposed a hybrid pipeline in which band-power features (δ, θ, α, β) are encoded into a quantum circuit with rotation and entanglement gates and then passed to a classical deep classifier, reporting up to 95% accuracy on a custom EEG test set. Aishwarya et al. [15] compared Variational Quantum Classifiers (VQC), quantum annealing, and hybrid quantum–classical neural networks for cognitive-state prediction, achieving up to 61.53% validation accuracy. Behera et al. [16] introduced QSVM-QNN, a fully quantum hybrid combining QSVM decision boundaries with QNN expressive power, evaluated on motor-imagery BCI benchmarks with accuracies of 0.990 and 0.950.

Despite this progress, three gaps remain. First, none of these works integrate high-level CNN-derived representations—which have been shown to capture spatio-temporal EEG topography effectively—with entanglement-based quantum feature maps in a parallel two-branch fusion architecture. Second, the choice of qubit encoding has so far been treated as an abstract dimensionality-reduction step; a band-aligned and interpretable mapping in which each qubit corresponds to a canonical EEG frequency band (θ, α, β, γ) has

not been investigated. We use this mapping as an engineering design that preserves the semantic identity of EEG frequency bands during quantum feature extraction, rather than as a neuroscientific theory of band interactions. Third, the question of *when* and *for which subjects* quantum features provide complementary information beyond classical models has not been quantitatively analyzed. The present work addresses all three gaps.

In classical computing approaches to EEG-based emotion recognition, large-scale models are often employed to capture complex temporal and spatial patterns in the data. Research in this area has been active since around 2016, with early methods relying on handcrafted features such as Power Spectral Density (PSD) [17]. For example, Marjit et al. [18] combined PSD with a Genetic Algorithm-optimized MLP, achieving 91% accuracy for binary and 83.5% for four-class classification. More recent studies have shifted toward advanced deep learning architectures, including CNNs, GNNs, and 3D CNNs, to automatically learn spatial and temporal features from raw EEG data. Deng et al. [19] reported binary accuracies above 92% and four-class accuracy of 74.2%, while Lin et al. [20] employed graph neural networks to model functional connectivity, achieving over 90% accuracy on multiple datasets. Other works explored lightweight models, self-supervised learning, and subject-dependent strategies, with some reporting near-perfect binary classification performance. For instance, Oka et al. [21] utilized an LSTM-based model and demonstrated high accuracy in four-class classification. However, such approaches often lead to a significant increase in the number of parameters and model complexity, posing challenges in terms of computational cost and interpretability.

Given this background, this study focuses on EEG data, which exhibits complex and diverse patterns, and investigates EEG-based emotion recognition using a novel QML approach. Specifically, we aim to construct a compact yet effective model by leveraging QML's ability to extract high-quality features. To achieve this, we propose a hybrid architecture that utilizes deep quantum features, combining deep features generated by classical machine learning with quantum features extracted via QML. The rationale for this hybrid design stems from preliminary observations that a quantum model achieved superior accuracy for specific subjects where conventional methods underperformed. This suggests that optimal feature representations vary across individuals. Consequently, our approach integrates both feature types, enabling the model to adaptively select the most informative representation for each subject.

Our contributions are as follows:

- **Parallel hybrid architecture combining CNN deep features with ZZFeatureMap quantum features**, distinct from prior works that either (i) use quantum models alone on handcrafted features [13,15,16], or (ii) use quantum circuits only as a sequential preprocessing stage on raw band-power features without integrating CNN-extracted high-level representations [14].
- **Compact yet high-performing model design**, reducing reliance on large-scale deep networks while maintaining competitive accuracy.
- **LIME-based interpretability analysis revealing subject-wise adaptive weighting** between classical and quantum features. To our knowledge, no prior QML-EEG study has empirically analyzed how a trained hybrid classifier uses quantum-derived features across subjects.
- **Empirical validation on the DEAP dataset**, demonstrating statistically significant improvement (normality-dependent paired t-test or Wilcoxon signed-rank test, $p < 0.05$) over both standalone classical CNNs and standalone quantum classifiers under the same ten-run, segment-level subject-dependent protocol—a comparison absent from prior QML-EEG studies.

2 Related Work

QML has attracted attention for its ability to exploit quantum phenomena such as superposition and entanglement to process complex and nonlinear data efficiently. QML is expected to achieve computational speedups and enable expressive feature transformations by leveraging these quantum properties. A key concept is the quantum feature map, which embeds classical data into a quantum Hilbert space, potentially enabling the separation of data that is not linearly separable in classical spaces. It has motivated the development of various QML models and hybrid architectures [22]. QML models are typically implemented using quantum circuits, which require encoding schemes to map classical data onto quantum states suitable for processing. Specifically, QML employs angle and amplitude encoding, where data values are represented as parameters for rotational gates and as amplitudes of quantum states. Additionally, Hadamard gates (H) are often used to induce superposition, enabling quantum parallelism. This quantum phenomenon allows for the simultaneous representation of multiple states, enabling information to be distributed across the qubit register and facilitating the parallel processing of diverse possibilities [23,24]. Recent QML research has also expanded beyond direct classification toward quantum-assisted feature selection, robust learning, neighborhood-based classification, and multi-view fusion. These studies are relevant to the present work because they treat quantum circuits not only as final classifiers but also as representation modules that can reshape the feature space before classical decision making. In this context, our study is positioned as an EEG-specific feature-fusion framework: the quantum circuit is used to generate an auxiliary representation from band-power features, while the final decision is made after fusion with CNN-derived features.

2.1 QML Approaches for EEG Analysis

A growing body of work has begun to apply QML directly to EEG signals for emotion recognition and related brain-state classification tasks. We summarize the four most relevant studies and contrast them with our proposed approach.

Garg et al. [13] were among the first to apply quantum machine learning to DEAP-based emotion recognition. Their approach extracts handcrafted statistical features from preprocessed EEG signals and feeds them to a Quantum Support Vector Machine (QSVM) trained on a D-Wave quantum annealer. They reported accuracies of 65.6% and 75.0% for valence and arousal, respectively. While this work demonstrates the feasibility of QML for emotion recognition, it does not exploit deep representations, and the quantum component operates as a standalone classifier rather than being fused with classical deep features.

Chandanwala et al. [14] proposed a hybrid quantum deep learning model in which δ, θ, α, and β band-power attributes—obtained via bandpass filtering and the Welch method—are mapped onto a quantum circuit composed of rotation and entanglement gates. The quantum-encoded representations are then passed to a classical deep classifier, achieving up to 95% accuracy. However, their evaluation is conducted on a custom EEG dataset (the specific dataset is not identified by name in the original paper, and the reported sampling frequency of 250 Hz differs from the standard DEAP configuration) rather than on a publicly available benchmark such as DEAP, which is the standard dataset adopted in the present study; therefore, a direct numerical comparison with their reported accuracy is not feasible. Moreover, in their architecture, the quantum circuit is used *only* as a sequential preprocessing stage on raw band-power features: high-level CNN representations of the EEG topography are not extracted, and there is no parallel fusion between classical and quantum branches.

Aishwarya et al. [15] compared three quantum techniques—Variational Quantum Classifiers (VQC), quantum annealing classifiers, and hybrid quantum–classical neural networks (HNN)—for predicting cognitive states from EEG signals, reaching up to 61.53% validation accuracy. Their study targets cognitive-state

classification rather than affective dimensions (valence/arousal), and the quantum and classical components are arranged sequentially rather than in a parallel fusion architecture.

Most recently, Behera et al. [16] proposed QSVM-QNN, a fully quantum hybrid model that integrates a Quantum Support Vector Machine with a Quantum Neural Network. The model is evaluated on two motor-imagery BCI benchmarks, achieving accuracies of 0.990 and 0.950, with additional analysis of robustness to six quantum noise models. While this work demonstrates the strength of quantum–quantum integration, the target task is BCI classification rather than emotion recognition, and the architecture does not involve any classical deep representation branch. Compared with these studies, the present work does not claim to outperform strong purely classical EEG emotion-recognition systems. Instead, it asks a narrower question: whether quantum-circuit-derived features can provide useful complementary information when fused with a conventional CNN branch under the same training protocol.

2.2 Positioning of the Present Work

The most closely related study is Garg et al. [13], which also targets DEAP-based valence/arousal classification; however, their approach relies on a single QSVM applied to handcrafted features, whereas our model fuses high-level CNN representations with quantum-encoded band-power features in a parallel architecture on the same dataset. Crucially, none of the four prior works adopts (i) a parallel two-branch CNN–quantum fusion design, (ii) a frequency-band-aligned qubit encoding tied to canonical EEG bands, or (iii) a subject-wise interpretability analysis of the classical–quantum complementarity. These three elements jointly constitute the contribution of the present work, and a quantitative literature-level comparison with these prior QML-EEG studies is provided later in the Performance Evaluation section.

Several studies have explored QML for classification tasks. Havlicek et al. [6] extended the concept of classical SVM by proposing a Quantum Support Vector Machine (QSVM) that encodes input data into quantum states to construct a quantum kernel matrix, enabling classification in complex feature spaces through quantum superposition and entanglement. Kariri et al. [7] constructed a Quantum Neural Network for binary classification by connecting a single-qubit quantum circuit to the final layer of a CNN, demonstrating its effectiveness on the MNIST handwritten digit dataset. Houssein et al. [9] proposed a hybrid approach that performs feature extraction on chest X-ray images using quantum circuits and then feeds the output to a CNN for training, achieving improved COVID-19 detection accuracy. Furthermore, Senokosov et al. [25] and Alam et al. [8] verified the effectiveness of QML models in image data classification, highlighting the potential of quantum-enhanced feature spaces for complex visual tasks. As such, numerous applications of QML to image data have been reported, suggesting a strong affinity between image features and QML. For example, Dong et al. [10] proposed an improved hybrid quantum-classical CNN for multi-class brain tumor MRI classification, demonstrating enhanced accuracy compared to classical CNNs. Sagingalieva et al. [11] introduced a hybrid quantum ResNet for vehicle recognition, leveraging quantum layers to optimize feature extraction and improve classification performance. Krunic et al. [12] applied quantum kernels to electronic health records for real-world prediction tasks, showing that quantum-enhanced feature spaces can outperform traditional kernel methods. These studies collectively highlight QML's potential in practical applications beyond benchmark datasets.

On the other hand, early studies on EEG-based emotion recognition primarily relied on handcrafted features such as PSD, which captures signal power within specific frequency bands and has been widely used for its robustness and interpretability. For example, Qiao et al. [26] divided EEG signals into emotion-related frequency bands and calculated PSD, while Marjit et al. [18] enhanced this approach by employing a Multi-Layer Perceptron (MLP) optimized with a genetic algorithm, improving classification accuracy through feature selection and nonlinear modeling. Building on these foundations, Li et al. [27] proposed

arranging PSD values for each electrode into a 2D array to reflect spatial electrode layout, enabling CNN-based processing similar to image recognition. This marked a shift toward leveraging deep learning models for automatic feature learning. Recent works have further advanced this trend by introducing architectures capable of capturing spatiotemporal dynamics. Wang et al. [28] utilized 3D CNNs to jointly model spatial and temporal dependencies, while Yang et al. [29] integrated 2D PSD maps with time-series features using a hybrid CNN-LSTM framework, effectively modeling dynamic emotional patterns. Because the present study evaluates DEAP trials after converting them into short temporal windows, the data-generation and evaluation protocols used in prior DEAP studies are particularly relevant. Several studies have expanded each video trial into multiple segment-level or window-level samples: Li et al. [27] extracted multidimensional PSD features using window lengths as short as 1 s, Salama et al. [30] removed the 3-s baseline and represented 60-s trials as 1-s frames and 6-frame chunks for 3D-CNN learning, Alhagry et al. [31] divided each video into twelve 5-s EEG segments for LSTM-based classification, Liu et al. [32] generated multiple 8-s or 12-s windows from each DEAP trial and assigned the original trial label to the derived segments, Moon et al. [33] generated overlapping 3-s segments and randomly divided the EEG data into five clusters, and Moctezuma et al. [34] evaluated 2-, 5-, and 10-s EEG segments with subject-wise training, validation, and test partitions. More recent preprint studies also use short-window FFT features followed by sample-level train/test splitting [35,36]. These studies support the view that converting DEAP trials into short temporal samples is a common data-generation strategy. However, they do not eliminate the methodological concern that segments from the same trial may be temporally correlated. Indeed, TSception explicitly adopted trial-wise validation and noted that random segment-level splitting can place adjacent segments into both training and testing sets [37]. Therefore, the present evaluation is described as segment-level subject-dependent evaluation rather than strict unseen-trial or subject-independent generalization.

Lin et al. [20] emphasize that this evolution reflects a broader movement from handcrafted features toward end-to-end deep learning approaches. Transforming EEG signals into structured representations such as 2D or 3D feature maps allows the application of CNNs originally designed for computer vision. Moreover, hybrid models combining CNNs and RNNs (e.g., LSTM, GRU) have become popular for capturing both spatial correlations among electrodes and temporal dynamics of emotional states. Emerging architectures such as Capsule Networks (CapsNet) [38] and Graph Neural Networks (GNN) have also been introduced to overcome limitations of CNNs in modeling complex electrode relationships and functional connectivity. These developments demonstrate a clear trajectory toward models that exploit spatial, spectral, and temporal characteristics of EEG signals, significantly improving recognition performance compared to traditional handcrafted approaches.

3 Proposed Method

Based on the above background, we propose a fusion model that integrates classical machine learning and QML for emotion prediction from EEG signals.

3.1 Overview

The proposed model, illustrated in Fig. 1, is a hybrid architecture that leverages "deep quantum features." These features combine deep representations, extracted by feeding preprocessed EEG data into CNN, with quantum features obtained through a quantum circuit. The rationale for this hybrid approach arises from preliminary observations indicating that a purely quantum model achieved high accuracy for certain subjects that conventional methods struggled to classify. This finding suggests that the most effective feature representations may vary across individuals. Consequently, we adopted a strategy to design a hybrid model that integrates both feature types, enabling the model to adaptively select the optimal representations during training.

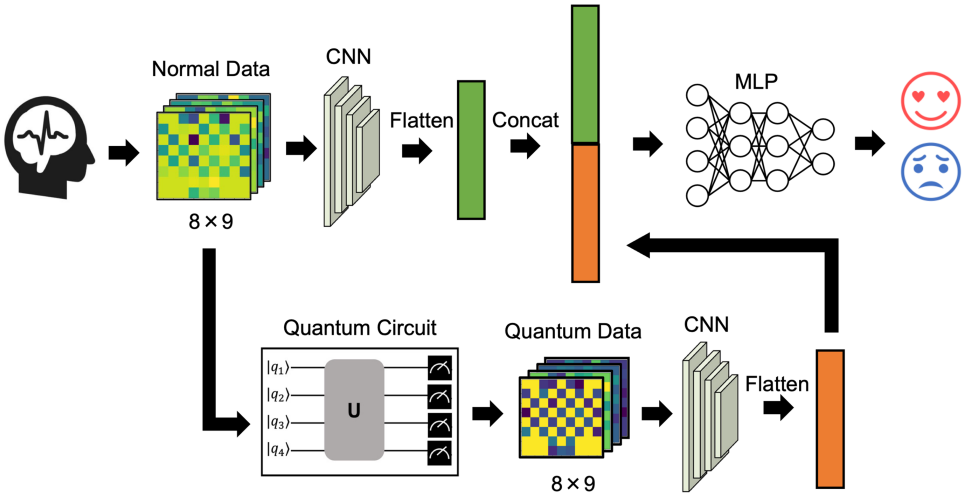

Figure 1: Overview of the proposed hybrid architecture. The model integrates deep features extracted by the CNN and quantum features extracted by the quantum circuit. These are concatenated to form "deep quantum features," which are then input into an MLP for classification.

EEG signals are recorded from 32 channels as raw time-series data. From these signals, PSD values for four frequency bands are computed, and a 2D representation is constructed (details are provided in Section 3.2). The 2D-arranged EEG data is then flattened into a 1D vector, and specific data points corresponding to the four frequency bands are selected for encoding into the quantum circuit using angle encoding. The CNN architecture used in the proposed model is shown in Fig. 2. Deep representations are obtained by applying an average pooling layer to the CNN output, reducing the dimensionality to $1 \times 1 \times 128$. These features are concatenated with quantum features to form the combined "deep quantum features." For the quantum component, a 4-qubit circuit is employed. The fused features are then processed by a two-layer MLP for binary classification. The quantum branch should not be interpreted as a single four-dimensional input replacing the full EEG map. The 4-qubit circuit is applied patch-wise to local four-band vectors, and the circuit outputs are reconstructed into an $8 \times 9 \times 4$ feature map before entering the quantum-feature branch. Thus, the classical and quantum branches use spatially aligned feature maps of the same reconstructed size, while differing in how the four-band information is represented.

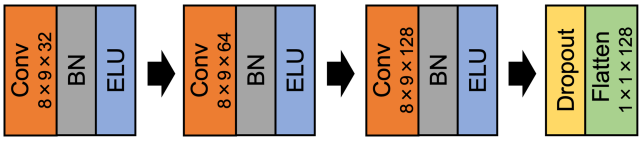

Figure 2: Architecture of the CNN for deep feature extraction. The network processes the 3D EEG feature representation ($8 \times 9 \times 4$) and outputs a 128-dimensional deep feature vector via average pooling.

3.2 EEG Data Preprocessing

EEG-based emotion estimation requires careful data preprocessing. Following the DEAP preprocessing and segment-generation practices identified in prior EEG emotion-recognition studies, we use the following five-step preprocessing pipeline. Among widely used datasets for this task, DEAP and SEED are common benchmarks; in this study, we adopt the more challenging DEAP dataset. Details of the dataset are provided in Section 4.1.

1. **EEG Data Segmentation**. Each trial consists of 63 s of EEG data, divided into a 3-s pre-stimulus baseline and a 60-s viewing period. The EEG data from the viewing period is segmented into 1-s non-overlapping windows. Following prior research on DEAP-based EEG emotion recognition and short-window spectral feature extraction [26,27,30,32], each segment is decomposed into four frequency bands: Theta (θ: 4–8 Hz), Alpha (α: 8–14 Hz), Beta (β: 14–31 Hz), and Gamma (γ: 31–45 Hz).
2. **Feature Extraction**. PSD is computed for each electrode within the segmented data to extract frequency-domain features.
3. **Baseline Subtraction**. To account for frequency-domain changes caused by emotional fluctuations, the baseline PSD (from the pre-stimulus period) is subtracted from the PSD obtained during the video-viewing period.
4. **Standardization**. Due to inter-subject variability in EEG signals, standardization is applied across all subjects.
5. **Feature Arrangement Based on Electrode Position**. Following prior research [27], an 8×9 2D array is constructed to preserve the spatial relationships among EEG electrodes before CNN processing. The extracted features are mapped onto this array to create a 2D feature map. By combining the feature maps for the four frequency bands, a 3D representation ($8 \times 9 \times 4$) is generated.

3.3 Feature Extraction using the Quantum Circuit

The quantum circuit, called the `ZZFeatureMap`, employed in this research is illustrated in Fig. 3. This feature map is widely used in quantum machine learning because it effectively captures correlations between input features through entanglement. By applying parameterized rotations and controlled-Z interactions, the ZZFeatureMap enables the representation of complex relationships in the data, which is particularly beneficial for tasks involving non-linear patterns such as EEG-based emotion recognition. Furthermore, we selected the ZZFeatureMap due to its reported stability and robustness in practical implementations, as highlighted in recent studies [39]. This stability makes it a suitable choice for NISQ devices, where noise and limited qubit resources pose significant challenges. Fig. 4 illustrates the procedure used to generate the quantum-derived feature map from the preprocessed EEG representation.

1. **Encoding Data into Quantum States**. An encoding circuit is employed to map classical data onto quantum states. In this study, we utilize the ZZFeatureMap provided by Qiskit, which introduces entanglement between qubits to capture pairwise feature interactions. Qiskit is a Python framework that facilitates cloud access to IBM's quantum processors (IBM Quantum) and enables the simulation of quantum computations on classical computers [40].
2. **Decoding via Pauli Z Expectation**. Decoding is performed by calculating the expectation value of the Pauli Z operator [41]. This process converts quantum state information into classical probability values corresponding to the basis states (i.e., $|0\rangle$ and $|1\rangle$), resulting in four final output values.
3. **Feature Extraction and Reconstruction**. Since the quantum circuit outputs four values and consists of 4 qubits, where each qubit corresponds to one of the four EEG frequency bands, the input EEG feature map (Normal data in Fig. 2) is divided into patches using a sliding window approach. Each patch is processed by the quantum circuit to generate quantum features. Using these probabilities, a total of four feature channels are extracted for each patch. Finally, the features output from the quantum circuit are reconstructed (reshaped) into a 2D array structure for integration with deep features. Consequently, the quantum-only baseline and the quantum branch of the proposed model do not operate on only four scalar values for an entire trial. They operate on patch-wise four-band vectors and reconstruct the resulting outputs into the same $8 \times 9 \times 4$ spatial-band format used by the CNN-based processing stream.

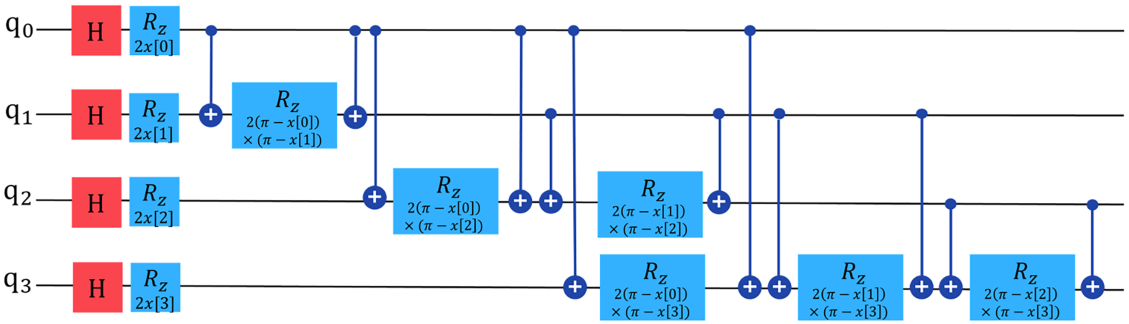

Figure 3: Quantum circuit structure for feature extraction. The circuit consists of 4 qubits, where each qubit corresponds to one of the four EEG frequency bands. Classical input data is mapped onto quantum states using the ZZFeatureMap (angle encoding).

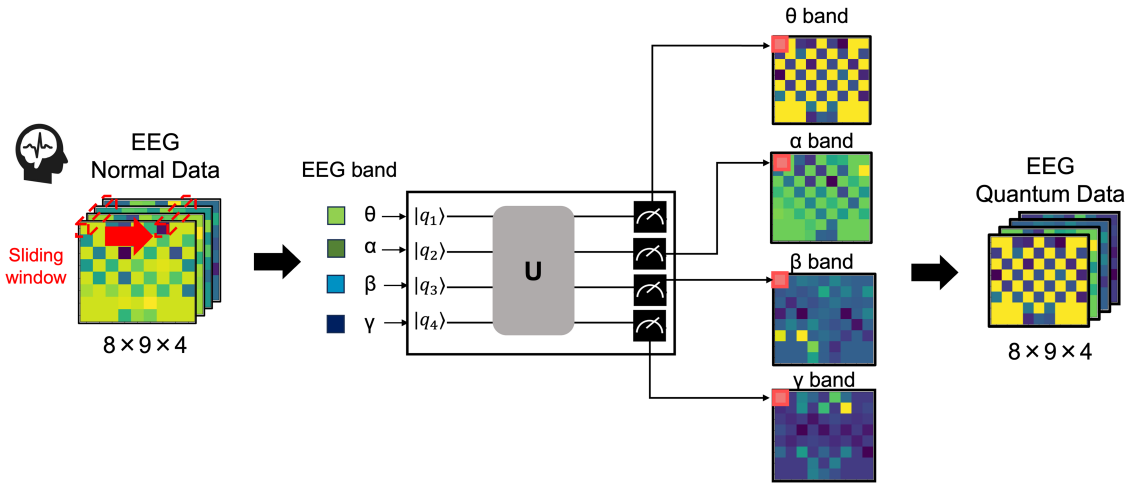

Figure 4: Procedure for generating the quantum-derived EEG feature map. The preprocessed EEG representation is arranged as an 8 × 9 × 4 spatial-band feature map. For each spatial location or local patch, a four-dimensional vector corresponding to the theta, alpha, beta, and gamma bands is extracted and encoded into a 4-qubit ZZFeatureMap using angle encoding. The Pauli-Z expectation values obtained from the quantum circuit are then reconstructed into an 8 × 9 × 4 quantum-derived feature map.

4 Performance Evaluation

We perform binary classification for both Valence and Arousal using a threshold of 5 to compare the proposed model against baseline models. The comparison models are defined as follows: (1) a classical machine learning model that uses only deep features, and (2) a quantum machine learning model that uses only quantum features. This study evaluates the performance of the proposed model against these baselines to verify its effectiveness for EEG-based emotion recognition. In addition, to clarify the positioning of the proposed model with respect to the existing QML-EEG literature, we compare it with prior QML-based EEG studies along five axes: target task, quantum component, classical component, feature-fusion strategy, and dataset/accuracy.

4.1 DEAP Dataset

In this experiment, we use the DEAP (A Database for Emotion Analysis using Physiological Signals) dataset [42]. DEAP is a dataset containing physiological signals, including EEG, recorded from 32 subjects

while they watched 40 one-minute-long music videos. Classification is performed using the Arousal (intensity of emotion) and Valence (degree of pleasure or displeasure) labels. Valence and Arousal are indices in Russell's circumplex model, as shown in Fig. 5. Subjects rated these labels on a scale of 1 to 9 after each viewing. Each trial consists of a 3-s baseline (pre-stimulus) period and the 60-s music video. The EEG data was recorded using 32 electrodes placed according to the international 10–20 system.

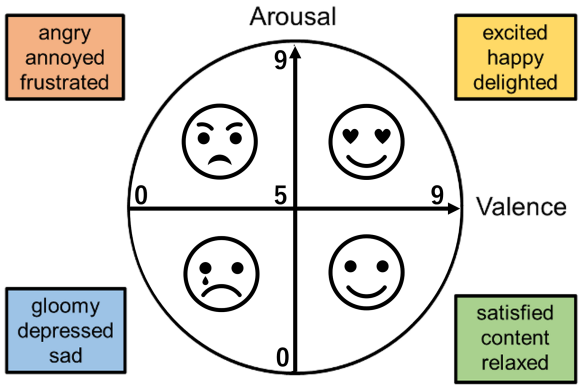

Figure 5: Russell's circumplex model of affect. The horizontal axis represents Valence (degree of pleasure or displeasure), and the vertical axis represents Arousal (intensity of emotion). Emotions are mapped within this two-dimensional space.

4.2 Experimental Settings

All experiments were conducted using the DEAP dataset (details provided in Section 4.1). For each subject, the 40 trials were first segmented into 60 non-overlapping 1-s windows, producing 2400 segment-level samples per subject. The trial-level valence/arousal labels were assigned to all 1-s segments derived from the corresponding trial. The resulting segment pool for each subject was then divided into training (60%), validation (20%), and test (20%) sets. Thus, the present evaluation is a subject-dependent segment-level protocol. This protocol is consistent with a common evaluation style in DEAP-based EEG emotion-recognition studies that expand trials into short temporal windows or segment-level feature samples before model training and evaluation [27,30–36]. The proposed hybrid model was implemented in Python, utilizing PyTorch for the classical components and Qiskit for quantum circuit simulation. To ensure a controlled comparison, the baseline models were constructed by removing one branch from the proposed architecture while keeping the remaining classifier structure unchanged as much as possible. To clarify the architectural differences among the comparison models, Fig. 6 illustrates the classical and quantum baseline. The classical baseline uses only the CNN-derived deep feature branch. Specifically, the $8 \times 9 \times 4$ EEG feature map is processed by the CNN shown in Fig. 2, followed by average pooling and the same two-layer MLP classifier used in the proposed model. The quantum baseline uses only the quantum-derived feature branch. In this model, local four-band vectors are processed patch-wise by the 4-qubit circuit, reconstructed into an $8 \times 9 \times 4$ feature map, and then passed to the same downstream feature extraction and MLP classification structure. For the CNN, the architecture shown in Fig. 2 was employed, consisting of convolutional layers followed by an average pooling layer to produce a 128-dimensional feature vector. The number of CNN layers and kernel sizes are specified in Fig. 2. The MLP classifier comprises two fully connected layers with ELU activation functions and a final sigmoid output for binary classification. The model was trained using the Adam optimizer with a learning rate of 1×10^{-4} and a batch size of 64. Cross-entropy was used as the loss function. To mitigate overfitting, early stopping was applied based on validation loss with a patience of 10 epochs. All experiments were conducted on a hybrid quantum-classical computing setup. Classical computations

were accelerated using an NVIDIA Tesla T4 GPU, while quantum circuits were simulated using Qiskit Aer's statevector method. Accordingly, the reported results evaluate the proposed feature-generation framework in a noiseless simulator environment. They do not constitute evidence of quantum hardware advantage, and hardware-noise evaluation remains outside the present experimental scope. This scope is also consistent with a common practice in early-stage applied QML studies, where simulator-based, framework-based, or small-scale quantum implementations are first used to validate the representation design on practical datasets before full hardware-oriented deployment. Examples include hybrid QML studies on COVID-19 chest X-ray classification, general image classification, quanvolutional image-classification frameworks, vehicle recognition, and electronic health-record prediction tasks [8,9,11,12,25]. These studies support the view that early-stage QML experiments can be useful for evaluating the potential of quantum-enhanced feature spaces in real-world data domains.

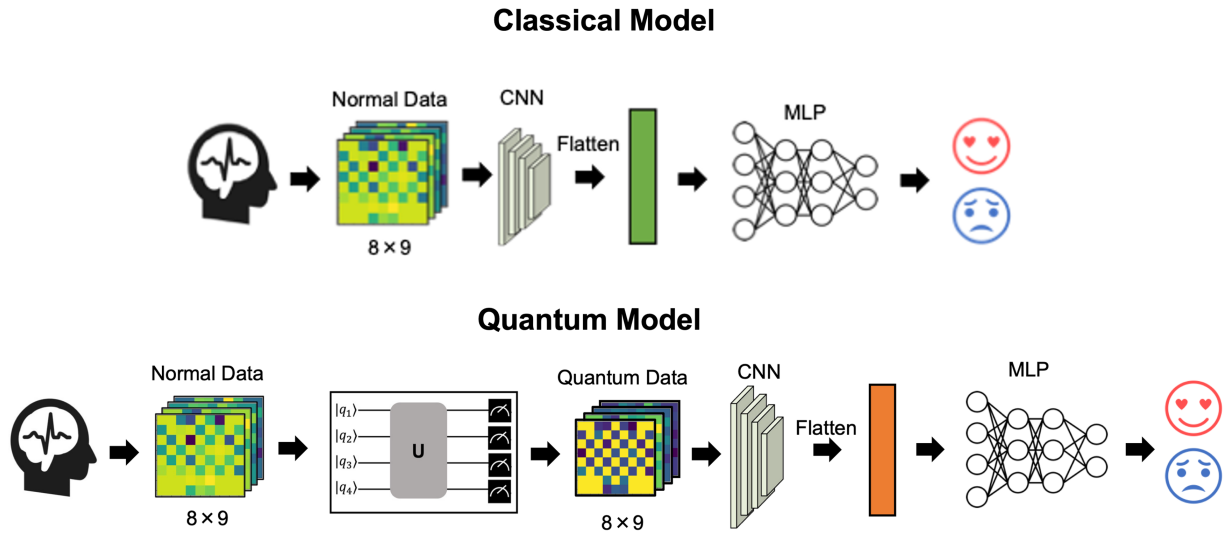

Figure 6: Architectures of the baseline model. The classical baseline uses only the CNN-derived deep feature branch, whereas the quantum baseline uses only the quantum-derived feature branch generated by applying the 4-qubit ZZFeatureMap circuit patch-wise to local four-band EEG vectors.

In the experiments, we evaluate performance using two metrics: classification accuracy and Cross-Entropy Difference (CE Difference). Here, CE Difference is defined as follows. Let $H(Q, P)$ denote the cross-entropy, which quantifies the divergence of the predicted distribution P from the true distribution Q; a smaller value implies higher prediction accuracy,

$$\text{CE Difference} = H(Q, P_{\text{uniform}}) - H(Q, P_{\text{model}}), \tag{1}$$

where P_{uniform} represents a uniform distribution (i.e., the probability distribution for random classification). In the case of binary classification, the cross-entropy for a uniform distribution is approximately 0.693. P_{model} represents the probability distribution generated by the proposed model (specifically, corresponding to the final validation loss). Consequently, the CE Difference serves as a metric where a larger positive value indicates that the proposed model has learned more effectively compared to random guessing.

To improve robustness, all experiments were repeated ten times with different random seeds, and the reported results correspond to the mean subject-level performance across these runs.

5 Results

Table 1 summarizes the average classification accuracy (ACC) and Cross-Entropy (CE) Difference for Valence and Arousal classification, while Table 2 presents the two-class accuracy of Valence and Arousal for all 32 subjects in the proposed and comparison models. All accuracy values reported in Tables 1 and 2 are subject-level means over ten independent training runs. The overall average ACC results are as follows: For Valence classification, the Classical Model achieved 77.0%, and the Quantum Model achieved 74.2%. In contrast, the Proposed Model achieved the highest average ACC under this protocol with an average ACC of 79.6%. For Arousal classification, the Classical Model reached 79.0%, and the Quantum Model reached 76.7%, whereas the Proposed Model achieved the highest accuracy of 81.5%.

Table 1: Summary of average classification accuracy (ACC) and cross-entropy (CE) Difference for Valence and Arousal classification. ACC values are averages across 32 subjects after averaging ten independent training runs for each subject.

Metric	Classical Model		Quantum Model		Proposed Model	
	Valence	Arousal	Valence	Arousal	Valence	Arousal
Average ACC (%)	77.0	79.0	74.2	76.7	**79.6***	**81.5***
Average CE Difference	–	–	–	–	0.252	0.281

Note: Bold values indicate that the Proposed Model achieved the highest average accuracy among the three models. * Statistically significant difference ($p < 0.05$) vs. both comparison models; see Table 3.

Table 2: 2-class accuracy of Valence and Arousal for 32 subjects in the proposed and comparison models. Each value is the mean of ten independent training runs. The (+) marker identifies the more accurate of the classical and quantum comparison models for each subject and label. Bold values indicate that the Proposed Model outperformed both comparison models.

Subject	Classical Model		Quantum Model		Proposed Model	
	Valence	Arousal	Valence	Arousal	Valence	Arousal
1	0.922(+)	0.929(+)	0.599	0.618	0.878	0.880
2	0.622	0.614	0.732(+)	0.770(+)	**0.744**	**0.775**
3	0.820	0.877	0.907(+)	0.946(+)	**0.926**	**0.946**
4	0.684	0.675	0.732(+)	0.769(+)	**0.764**	**0.780**
5	0.699	0.696	0.760(+)	0.749(+)	**0.765**	**0.776**
6	0.899(+)	0.838(+)	0.749	0.602	0.843	0.768
7	0.910(+)	0.907(+)	0.734	0.629	0.859	0.842
8	0.817(+)	0.855(+)	0.724	0.750	0.772	0.789
9	0.621	0.639	0.676(+)	0.690(+)	0.670	0.689
10	0.808	0.851(+)	0.818(+)	0.812	**0.835**	0.828
11	0.631	0.637	0.691(+)	0.664(+)	**0.703**	**0.693**
12	0.789(+)	0.874(+)	0.688	0.864	0.739	0.866
13	0.740(+)	0.856	0.574	0.859(+)	0.678	0.856
14	0.613	0.719	0.631(+)	0.748(+)	**0.686**	**0.765**
15	0.843(+)	0.826	0.836	0.836(+)	**0.858**	**0.868**
16	0.867(+)	0.880(+)	0.801	0.804	**0.910**	**0.884**
17	0.553	0.641(+)	0.609(+)	0.640	0.608	**0.653**

(Continued)

Table 2 (continued)

Subject	Classical Model		Quantum Model		Proposed Model	
	Valence	Arousal	Valence	Arousal	Valence	Arousal
18	0.854(+)	0.854(+)	0.814	0.832	**0.869**	**0.897**
19	0.799(+)	0.769	0.798	0.803(+)	**0.800**	**0.811**
20	0.885(+)	0.931(+)	0.665	0.797	0.817	0.883
21	0.772(+)	0.841	0.751	0.853(+)	**0.793**	**0.886**
22	0.728	0.680	0.909(+)	0.883(+)	**0.915**	0.872
23	0.852(+)	0.846(+)	0.739	0.803	**0.906**	**0.889**
24	0.711	0.826	0.824(+)	0.883(+)	0.824	**0.896**
25	0.733(+)	0.769	0.723	0.824(+)	**0.763**	**0.847**
26	0.730	0.659	0.786(+)	0.703(+)	0.781	**0.729**
27	0.924(+)	0.898(+)	0.810	0.765	0.885	0.844
28	0.778(+)	0.787(+)	0.686	0.640	0.736	0.691
29	0.815(+)	0.871(+)	0.707	0.746	0.779	0.814
30	0.735	0.721	0.831(+)	0.807(+)	**0.838**	0.806
31	0.774(+)	0.759(+)	0.723	0.702	0.755	**0.760**
32	0.726(+)	0.760(+)	0.721	0.751	**0.766**	**0.787**

To visualize the variations in accuracy among subjects, the distributions of ACC for Valence and Arousal are presented as boxplots in Figs. 7 and 8, respectively. These figures illustrate that the proposed model consistently maintains high accuracy across the subject pool compared to the comparison models.

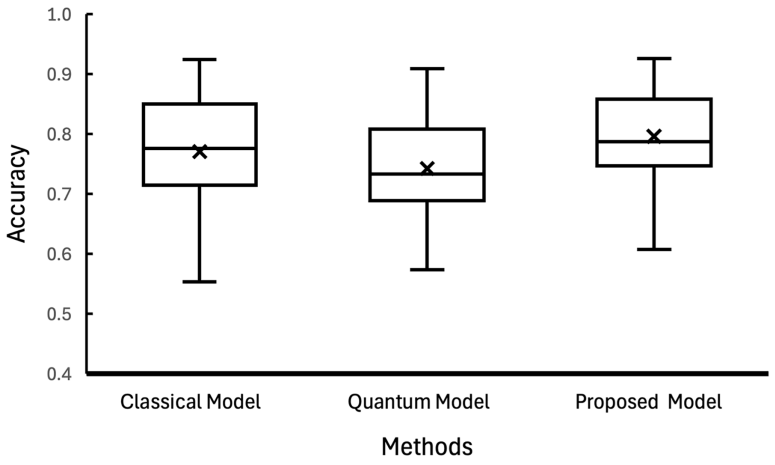

Figure 7: Classification accuracy for Valence. The distribution of accuracy across 32 subjects is shown for the classical, quantum, and proposed models.

To assess the statistical significance of these results, we conducted tests on the subject-level ten-run mean accuracy differences between the proposed model and the comparison models (classical and quantum). First, the Shapiro–Wilk test was applied to the paired accuracy differences for each comparison. When normality was not rejected ($p \geq 0.05$), a two-sided paired t-test was used; when normality was rejected ($p < 0.05$), a two-sided Wilcoxon signed-rank test was used. The complete results are reported in Table 3. At

an unadjusted significance level of $\alpha = 0.05$, the selected tests indicated statistically significant differences in all four evaluated comparisons:

Scenario 1 Valence—Proposed vs. Classical

Scenario 2 Valence—Proposed vs. Quantum

Scenario 3 Arousal—Proposed vs. Classical

Scenario 4 Arousal—Proposed vs. Quantum

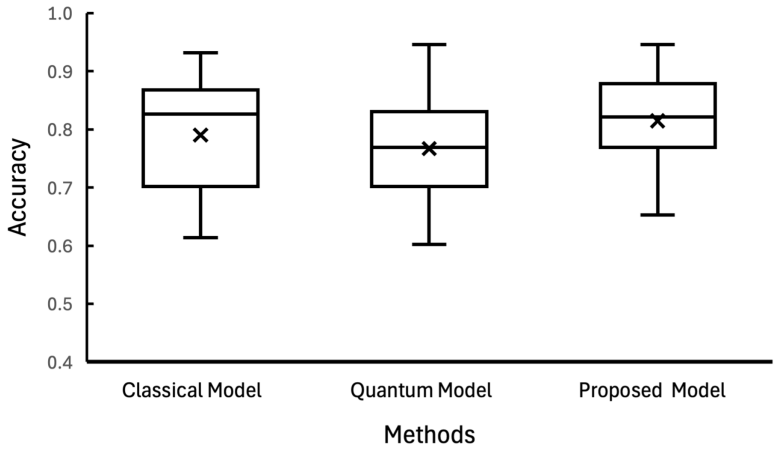

Figure 8: Classification accuracy for arousal. The distribution of accuracy across 32 subjects is shown for the classical, quantum, and proposed models.

Table 3: Statistical comparisons between the proposed model and the comparison models using subject-level ten-run mean accuracies. Normality was assessed for the paired accuracy differences. A two-sided paired t-test was used when normality was not rejected; otherwise, a two-sided Wilcoxon signed-rank test was used.

Label	Comparison	Selected Test	p
Valence	Proposed vs. Classical	Paired t-test	0.033
Valence	Proposed vs. Quantum	Wilcoxon	3.08×10^{-6}
Arousal	Proposed vs. Classical	Paired t-test	0.049
Arousal	Proposed vs. Quantum	Wilcoxon	1.43×10^{-7}

These findings show that the proposed hybrid model achieved significantly higher subject-level mean accuracy than both the classical model (utilizing only deep features) and the quantum model (utilizing only quantum features) under the present ten-run, segment-level subject-dependent evaluation protocol.

Furthermore, we calculated the Cross-Entropy Difference (CE Difference) to evaluate the learning performance of the model. While the cross-entropy for a uniform (random) binary classification is approximately 0.693, the proposed model achieved an average CE Difference of 0.252 for Valence classification and 0.281 for Arousal classification, based on the ten-run mean test losses, across all subjects. The positive values of CE Difference indicate that the model effectively learned discriminative features, performing better than random guessing. Conversely, a negative value would suggest that the model failed to generalize or was overfitting; however, no such trend was observed in this study. Unlike accuracy, which measures overall correctness, CE Difference reflects the model's learning effectiveness compared to random guessing, providing a complementary perspective on performance.

In addition to predictive performance, the model complexity of each architecture was also examined. The standalone comparison model and the proposed hybrid model contain 102,370 and 204,546 trainable parameters, respectively, with the additional parameters originating from the quantum-feature branch. For context, recently reported EEG emotion-recognition models operate on a considerably larger scale, ranging from approximately 0.38 M parameters in a CNN–LSTM hybrid [43] and 0.8 M in the EEG Conformer [44] to 15.25 M in a temporal-difference minimizing neural network [45]. Relative to these architectures, the proposed model remains in a low-parameter regime. As parameter counts are not consistently reported across previous studies, this comparison is intended to characterize model complexity rather than to establish a ranking.

5.1 Additional Robustness and Ablation Analyses

To address whether the improvement of the proposed model is attributable to the quantum-derived branch itself or merely to the existence of a second branch, we conducted additional robustness and ablation analyses under the same training protocol. The additional evaluation included a noise-fusion control in which the quantum feature map was replaced with random values of the same shape and distribution scale, two band-input interventions, and an entanglement-free circuit.

For the band-input interventions, the *band-shuffle* condition randomly permuted the correspondence between the four EEG bands and the four qubits before quantum feature generation. The *orthogonal-rotation* condition applied a random orthogonal transformation to the four-dimensional band-power vector before the quantum circuit, preserving vector norms while removing the direct interpretability of the original band axes. For the circuit-structure intervention, the *no-entanglement* condition replaced the ZZFeatureMap with four independent single-qubit R_y rotations, thereby removing explicit entangling operations while preserving a four-qubit, four-output feature-generation structure.

Table 4 summarizes the mean accuracies and the paired statistical comparisons between the original quantum-feature condition and each ablation condition.

Table 4: Additional robustness and ablation results. Values are average accuracies across 32 subjects. For each original-vs.-ablation comparison, normality was assessed using the Shapiro–Wilk test; a two-sided paired t-test was used when normality was not rejected, and a two-sided Wilcoxon signed-rank test was used otherwise.

Model/Condition	Valence ACC (%)	p	Arousal ACC (%)	p
Proposed model (original quantum features)	**79.6**	–	**81.5**	–
Proposed + noise branch	66.0	10^{-6}*	71.0	10^{-6}*
Proposed + band-shuffled quantum features	78.8	0.012*	80.7	0.020*
Proposed + orthogonally rotated band inputs	78.3	0.002*	81.3	0.588
Proposed + no-entanglement circuit	71.2	8.21×10^{-5}*	75.4	0.001*

Note: Bold values indicate that the Proposed Model with the original quantum features achieved the highest accuracy among the compared conditions. * Each ablation is compared with the Proposed original condition. Arousal band shuffle used the Wilcoxon test; the other seven comparisons used paired t-tests.

The noise-fusion control showed a large decrease relative to the proposed model, suggesting that the improvement is not explained solely by adding a second branch or injecting arbitrary auxiliary inputs. The no-entanglement condition also degraded performance, indicating that the ZZFeatureMap structure contributes meaningfully under the present simulator-based setting. Following the same normality-based test-selection procedure, all differences were significant except original vs. orthogonal rotation for Arousal. In contrast to the larger noise and no-entanglement effects, band-shuffle and orthogonal-rotation produced

results close to the original band-aligned condition. Therefore, the band-aligned encoding should be understood primarily as an interpretable and structured design choice; the present results do not support a strong claim that assigning each specific EEG band to a fixed qubit is itself the dominant source of the performance gain.

Finally, to position the proposed model within the broader QML-EEG literature, we compare it with prior QML-based EEG studies along five axes: target task, quantum component, classical component, feature-fusion strategy, and dataset/accuracy. The results are summarized in Table 5. Among the existing studies, Garg et al. [13] target the same DEAP-based valence/arousal classification task as the present work and therefore provide the closest QML-EEG reference point, although differences in preprocessing and evaluation protocol should be considered; their QSVM-based approach reports 65.6% for valence and 75.0% for arousal, whereas the proposed model reports numerically higher ten-run mean accuracies of 79.6% and 81.5%, respectively. Chandanwala et al. [14] report up to 95% accuracy on a custom EEG test set, but their evaluation is conducted on a different dataset and the protocol is not directly comparable to DEAP-based valence/arousal classification. Aishwarya et al. [15] target cognitive-state prediction rather than affective dimensions, achieving 61.53% validation accuracy, while Behera et al. [16] address motor-imagery BCI classification with accuracies of 0.990 and 0.950 on benchmark datasets that differ from DEAP in both task and label structure. Although these numerical values cannot be directly compared due to differences in target tasks and datasets, the proposed model is the only one that simultaneously combines a parallel CNN–quantum fusion architecture, a band-aligned and interpretable qubit encoding, and subject-wise interpretability analysis. Within the present DEAP-based valence/arousal setting, these results support the effectiveness of the proposed hybrid design under a segment-level subject-dependent protocol.

Table 5: Comparison with prior QML-based EEG studies. "V" and "A" denote valence and arousal, respectively.

Study	Task	Quantum Component	Classical Component	Feature Fusion	Dataset/Accuracy
Garg et al. [13]	Emotion (V/A)	QSVM (D-Wave)	Handcrafted features	None	DEAP/65.6% (V), 75.0% (A)
Chandanwala et al. [14]	Emotion	Rotation + entanglement circuit	DL classifier (post-quantum)	Sequential	Custom EEG/up to 95%
Aishwarya et al. [15]	Cognitive state	VQC/QA/HNN	MLP	None/sequential	EEG (cognitive)/61.53%
Behera et al. [16]	BCI (motor imagery)	QSVM + QNN	None (fully quantum hybrid)	Quantum–Quantum	EEG benchmarks/0.990, 0.950
Ours	**Emotion (V/A)**	**ZZFeatureMap (band-aligned 4-qubit)**	**CNN (3D EEG map)**	**Parallel concat → MLP**	**DEAP/79.6% (V), 81.5% (A)**

Note: Bold entries indicate the proposed method presented in this study.

6 Discussion

The experimental results indicate that the proposed hybrid model, which incorporates deep quantum features, achieved a statistically significant improvement in accuracy compared to both the classical model (utilizing only deep features) and the quantum model (utilizing only quantum features). This discussion analyzes how the trained hybrid classifier used the two feature branches, based on Local Interpretable

Model-agnostic Explanations (LIME). LIME indicates which input features influenced the trained MLP's predictions; it does not, by itself, establish a causal quantum-mechanical mechanism.

As noted in the introduction, a primary motivation for this study was the observation that a quantum-only model produced promising results for specific subjects where conventional classical models underperformed. To validate our hypothesis that optimal feature representations vary across subjects, we analyzed the predictive basis of the proposed hybrid model using LIME. The updated LIME analysis showed subject-dependent use of both feature branches. Across the 64 subject–label cases, the branch with greater LIME importance agreed with the better standalone model in 42 cases. Quantum-derived features had greater LIME importance in 25 of 32 Valence cases and 26 of 32 Arousal cases, indicating that the trained MLP frequently used the quantum-derived branch, but not always in direct correspondence with standalone-model accuracy. To quantitatively verify this trend and test the hypothesis that the model adaptively selects effective features, we conducted a detailed comparison between the classification accuracy of the individual models (Classical vs. Quantum) and the feature importance scores derived from LIME. The results for all 32 subjects are summarized in Table 6. In this table, the higher accuracy between the two single models is underlined, and the feature type with the greater LIME contribution in the proposed model is highlighted in bold. We analyzed the relationship based on the absolute difference in accuracy ("Diff") between the Classical and Quantum models, using a threshold of 0.10 (10%) to distinguish between significant and negligible performance gaps. This analysis revealed two key trends:

- **Significant Performance Gap (Diff ≥ 0.10):** Among the 22 subject–label cases where the standalone-model accuracy difference was at least 10%, LIME importance agreed with the better standalone model in 16 cases: 9 of 12 Valence cases and 7 of 10 Arousal cases. Subjects 1 and 22 for Valence are representative aligned cases, whereas Subjects 12, 20, and 29 for Valence and Subjects 8, 28, and 29 for Arousal were exceptions.
- **Small Performance Gap (Diff < 0.10):** Among the remaining 42 subject–label cases with a standalone-model accuracy difference below 10%, accuracy–LIME agreement was observed in 26 cases. Because the standalone accuracies were close in this group, disagreement should be interpreted as evidence that standalone accuracy and within-fusion feature usage measure different properties, rather than as a strict feature-selection error.

The updated exceptions show that the hybrid model does not simply select the branch associated with the higher standalone accuracy. For example, the Classical Model was more accurate for Subjects 20 and 29 in Valence and Subjects 8, 28, and 29 in Arousal by at least 10%, while the quantum-derived features received greater LIME importance within the fused model. This result provides an important insight: the hybrid model does not simply rely on accuracy metrics to select features. Even when the standalone Quantum model yields lower overall accuracy, its extracted features (via angle encoding with the `ZZFeatureMap`) may still provide predictive variables that are used differently after fusion. However, LIME alone cannot determine whether this usage arises from uniquely quantum structure, generic nonlinear representation, or interactions learned by the fusion head. EEG signals are characterized by complexity, non-linearity, and significant inter-subject variability. While deep features captured by CNNs are effective for pattern recognition in image-like PSD maps, they may not fully capture the spectrum of brain activity for all individuals. The observed exceptions suggest that the hybrid model can use quantum-derived features as additional predictive variables even when the standalone quantum model is not superior. In conclusion, the improved performance under the present protocol of the proposed hybrid model is not merely described by the standalone performance of its two branches. The LIME results indicate subject-dependent integration of deep and quantum-derived representations, but the moderate accuracy–LIME agreement rate requires a cautious interpretation of complementarity.

Table 6: Comparison of ten-run mean single-model accuracy and updated LIME feature importance for the proposed model across 32 subjects. The "Accuracy" columns present the ten-run mean performance of the classical and quantum models. The "LIME" columns report the updated contribution scores of classical and quantum features within the proposed model. Underlined values indicate the higher accuracy between the two single models, while bold values in the LIME columns denote the feature type with the greater contribution. "Diff" represents the absolute difference in accuracy between the classical and quantum models.

Subject	Valence					Arousal				
	Accuracy		Diff	LIME (Proposed)		Accuracy		Diff	LIME (Proposed)	
	Classical Model	Quantum Model		Classical Feature	Quantum Feature	Classical Model	Quantum Model		Classical Feature	Quantum Feature
1	0.922	0.599	0.324	**0.983**	0.018	0.929	0.618	0.310	**0.606**	0.028
2	0.622	0.732	0.110	−0.006	**0.308**	0.614	0.770	0.156	0.007	**0.340**
3	0.820	0.907	0.087	0.045	**0.907**	0.877	0.946	0.069	−0.002	**0.047**
4	0.684	0.732	0.049	0.013	**0.332**	0.675	0.769	0.094	0.058	**0.214**
5	0.699	0.760	0.062	0.022	**0.307**	0.696	0.749	0.053	0.071	**0.342**
6	0.899	0.749	0.150	**0.097**	0.022	0.838	0.602	0.236	**0.331**	0.041
7	0.910	0.734	0.177	**0.209**	0.027	0.907	0.629	0.278	**0.473**	0.023
8	0.817	0.724	0.092	0.073	**0.309**	0.855	0.750	0.104	0.097	**0.332**
9	0.621	0.676	0.055	0.012	**0.198**	0.639	0.690	0.051	0.007	**0.064**
10	0.808	0.818	0.010	0.025	**0.693**	0.851	0.812	0.038	0.003	**0.673**
11	0.631	0.691	0.060	0.010	**0.248**	0.637	0.664	0.027	0.017	**0.075**
12	0.789	0.688	0.101	0.069	**0.315**	0.874	0.864	0.010	−0.003	**0.053**
13	0.740	0.574	0.167	**0.141**	0.027	0.856	0.859	0.003	0.001	**0.014**
14	0.613	0.631	0.018	0.044	**0.249**	0.719	0.748	0.029	0.029	**0.132**
15	0.843	0.836	0.007	0.049	**0.455**	0.826	0.836	0.010	0.078	**0.537**
16	0.867	0.801	0.065	**0.179**	0.150	0.880	0.804	0.077	0.335	**0.422**
17	0.553	0.609	0.055	0.021	**0.087**	0.641	0.640	0.001	0.011	**0.117**
18	0.854	0.814	0.040	0.128	**0.520**	0.854	0.832	0.022	0.051	**0.380**
19	0.799	0.798	0.000	0.022	**0.486**	0.769	0.803	0.034	0.044	**0.241**
20	0.885	0.665	0.219	0.114	**0.201**	0.931	0.797	0.134	**0.067**	0.035
21	0.772	0.751	0.021	0.098	**0.367**	0.841	0.853	0.012	−0.003	**0.035**
22	0.728	0.909	0.181	0.008	**0.770**	0.680	0.883	0.203	−0.006	**0.497**
23	0.852	0.739	0.113	**0.209**	0.163	0.846	0.803	0.043	**0.038**	0.035
24	0.711	0.824	0.113	0.013	**0.729**	0.826	0.883	0.056	0.005	**0.076**
25	0.733	0.723	0.010	0.030	**0.303**	0.769	0.824	0.055	0.004	**0.099**
26	0.730	0.786	0.055	0.016	**0.185**	0.659	0.703	0.045	−0.017	**0.284**
27	0.924	0.810	0.114	**0.057**	0.038	0.898	0.765	0.133	**0.084**	0.081
28	0.778	0.686	0.091	0.035	**0.130**	0.787	0.640	0.148	0.101	**0.196**
29	0.815	0.707	0.108	0.105	**0.346**	0.871	0.746	0.125	0.104	**0.253**
30	0.735	0.831	0.096	0.002	**0.200**	0.721	0.807	0.086	0.004	**0.475**
31	0.774	0.723	0.052	0.057	**0.229**	0.759	0.702	0.057	0.066	**0.250**
32	0.726	0.721	0.005	0.035	**0.290**	0.760	0.751	0.010	0.034	**0.078**

7 Conclusion and Future Work

The primary objective of this study was to propose and validate a novel approach for EEG-based emotion recognition—a task inherently challenged by high-dimensional and complex data structures—by integrating classical machine learning and Quantum Machine Learning. To achieve this, we have proposed a hybrid architecture utilizing deep quantum features, synthesized by combining deep features extracted via CNN with quantum features derived from a quantum circuit (specifically, the `ZZFeatureMap` from Qiskit).

The efficacy of the proposed model was evaluated on the binary classification tasks of Valence and Arousal using the DEAP dataset under the segment-level subject-dependent protocol described above. Experimental results demonstrated that the proposed hybrid model achieved an average classification accuracy of 79.6% for Valence and 81.5% for Arousal across ten independent training runs. Statistical analysis using paired t-tests when the subject-level differences were normally distributed and Wilcoxon signed-rank tests otherwise indicated significant improvements over both the baseline classical model (utilizing only deep features) and the quantum model (utilizing only quantum features) at the unadjusted $\alpha = 0.05$ level.

Furthermore, interpretability analysis using LIME provided crucial insights into how the trained hybrid model used the classical and quantum-derived feature branches. The updated results showed that the branch with greater LIME importance agreed with the better standalone model in 42 of 64 subject–label cases. This suggests that deep and quantum-derived features can play complementary predictive roles, although LIME should be interpreted as a model-usage analysis rather than causal evidence of a quantum-mechanical contribution. Additional ablation analyses supported this interpretation. Replacing the quantum branch with random noise reduced performance. Removing entangling operations from the feature map also degraded performance. By contrast, band-shuffle and orthogonal-rotation interventions yielded results close to the original band-aligned encoding, so the specific fixed assignment of each EEG band to each qubit should be treated as an interpretable design choice rather than as a confirmed physiological mechanism.

In addition, we positioned the proposed model within the broader QML-EEG literature by comparing it with representative prior studies, including QSVM-based emotion recognition [13], sequential quantum-then-classical pipelines on band-power features [14], variational quantum classifiers for cognitive-state prediction [15], and fully quantum hybrid models for brain–computer interfacing [16]. Among QML-EEG studies that also report DEAP-based valence/arousal classification, Garg et al. [13] provide the closest reference point, reporting 65.6%/75.0%, while the proposed model reports ten-run mean accuracies of 79.6%/81.5% under the present protocol. Beyond numerical performance, the proposed model is the only one among the surveyed studies that simultaneously combines a parallel CNN–quantum fusion architecture, a band-aligned and interpretable qubit encoding, and subject-wise interpretability analysis, supporting the effectiveness of the proposed hybrid design within the present segment-level evaluation setting.

Despite these promising results, several limitations should be acknowledged. Although the proposed approach was evaluated on a challenging benchmark dataset (DEAP), which is widely used in EEG-based emotion recognition research, this evaluation alone does not guarantee practical applicability. Real-world deployment will require validation on actual hardware and under realistic conditions, where factors such as noise, latency, and device constraints may significantly impact performance. Second, all quantum-circuit outputs in this study were generated using Qiskit Aer statevector simulation. This setup is appropriate for evaluating the proposed representation design, but it does not demonstrate performance on noisy quantum hardware or prove quantum advantage. Third, the proposed method has so far been validated mainly on DEAP. Preliminary investigation on SEED suggests that the same fusion idea can be applied to another EEG dataset, but the quantum-derived branch did not show as clear a benefit as in DEAP; broader dataset-level validation is therefore required.

Furthermore, the current method relies on a fixed quantum feature map (ZZFeatureMap), which may not fully exploit the expressive power of parameterized quantum circuits for diverse problem settings. In light of this limitation, future research will focus on exploring Variational Quantum Circuit (VQC) architectures [8,46,47], which offer adaptive parameterization and potentially richer feature representations tailored to specific tasks. We also plan to investigate hybrid optimization strategies that jointly tune CNN and quantum circuit parameters, evaluate the approach on larger and multi-class emotion recognition datasets, conduct trial-wise and subject-independent validation, compare against stronger classical parallel-fusion

baselines, and assess the feasibility of deploying such hybrid models on near-term quantum hardware under realistic conditions. These directions aim to bridge the gap between benchmark performance and practical implementation, paving the way for robust EEG-based emotion recognition in real-world settings.

Acknowledgement: Not applicable.

Funding Statement: This work was supported by JSPS KAKENHI Grant Number 21K12097.

Author Contributions: Conceptualization, Reon Yoshida and Keiko Ono; methodology, Reon Yoshida, Keiko Ono and Kentaro Ohki; validation, Reon Yoshida; writing—original draft preparation, Reon Yoshida; writing—review and editing, Keiko Ono, Kentaro Ohki and Takuya Futagami; supervision, Keiko Ono. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The DEAP dataset analyzed during the current study is a publicly available dataset [42].

Ethics Approval: Not applicable. This study used a publicly available anonymized dataset and did not involve new human or animal experiments by the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Preskill J. Quantum computing in the NISQ era and beyond. Quantum. 2018;2:79. doi:10.22331/q-2018-08-06-79.
2. Corcoles AD, Kandala A, Javadi-Abhari A, McClure DT, Cross AW, Temme K, et al. Challenges and opportunities of near-term quantum computing systems. Proc IEEE. 2020;108(8):1338–52. doi:10.1109/jproc.2019.2954005.
3. Power L, Guha K. Feature importance and explainability in quantum machine learning. arXiv:2405.08917. 2024. doi:10.48550/arXiv.2405.08917.
4. Toufah A, Kadim MA, El Hafidi MY. Overcoming SVM limitations in lung cancer classification with a quantum feature map. BMC Artif Intell. 2025;1(1):15. doi:10.1186/s44398-025-00016-3.
5. Alvarez-Estevez D. Benchmarking quantum machine learning kernel training for classification tasks. IEEE Trans Quantum Eng. 2025;6:1–15. doi:10.1109/tqe.2025.3541882.
6. Havlíček V, Córcoles AD, Temme K, Harrow AW, Kandala A, Chow JM, et al. Supervised learning with quantum-enhanced feature spaces. Nature. 2019;567(7747):209–12. doi:10.1038/s41586-019-0980-2.
7. Kariri A, Aljamal M, Al Aridhee J, Albukhari I, Matondo-Mvula N, Elleithy K. Enhancement of CNNHQ with an activation function. In: Proceedings of the 2023 IEEE Long Island Systems, Applications and Technology Conference (LISAT); 2023 May 5; Old Westbury, NY, USA. p. 1–5. doi:10.1109/lisat58403.2023.10179557.
8. Alam M, Kundu S, Topaloglu RO, Ghosh S. Quantum-classical hybrid machine learning for image classification (ICCAD special session paper). In: Proceedings of the 2021 IEEE/ACM International Conference on Computer Aided Design (ICCAD); 2021 Nov 1–4; Munich, Germany. p. 1–7. doi:10.1109/iccad51958.2021.9643516.
9. Houssein EH, Abohashima Z, Elhoseny M, Mohamed WM. Hybrid quantum-classical convolutional neural network model for COVID-19 prediction using chest X-ray images. J Comput Des Eng. 2022;9(2):343–63. doi:10.1093/jcde/qwac003.
10. Dong Y, Fu Y, Liu H, Che X, Sun L, Luo Y. An improved hybrid quantum-classical convolutional neural network for multi-class brain tumor MRI classification. J Appl Phys. 2023;133(6):064401. doi:10.1063/5.0138021.
11. Sagingalieva A, Kordzanganeh M, Kurkin A, Melnikov A, Kuhmistrov D, Perelshtein M, et al. Hybrid quantum ResNet for car classification and its hyperparameter optimization. Quantum Mach Intell. 2023;5(2):38. doi:10.1007/s42484-023-00123-2.
12. Krunic Z, Flother F, Seegan G, Earnest-Noble N, Omar S. Quantum kernels for real-world predictions based on electronic health records. IEEE Trans Quantum Eng. 2022;3(1):1–11. doi:10.1109/tqe.2022.3176806.

13. Garg D, Verma GK, Singh AK. EEG-based emotion recognition using quantum machine learning. SN Comput Sci. 2023;4(5):480. doi:10.1007/s42979-023-01943-6.
14. Chandanwala AA, Bhowmik S, Chaudhury P, Pravin SC. Hybrid quantum deep learning model for emotion detection using Raw EEG signal analysis. arXiv:2411.17715. 2024. doi:10.48550/arXiv.241117715.
15. Aishwarya S, Abeer V, Sathish BB, Subramanya KN. Quantum computational techniques for prediction of cognitive state of human mind from EEG signals. J Quantum Comput. 2020;2(4):157–70. doi:10.32604/jqc.2020.015018.
16. Behera BK, Al-Kuwari S, Farouk A. QSVM-QNN: quantum support vector machine based quantum neural network learning algorithm for brain–computer interfacing systems. IEEE Trans Artif Intell. 2026;7(1):308–20. doi:10.1109/tai.2025.3572852.
17. Li X, Zhang Y, Tiwari P, Song D, Hu B, Yang M, et al. EEG based emotion recognition: a tutorial and review. ACM Comput Surv. 2023;55(4):1–57. doi:10.1145/3524499.
18. Marjit S, Talukdar U, Hazarika SM. EEG-based emotion recognition using genetic algorithm optimized multi-layer perceptron. In: Proceedings of the 2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA); 2021 Sep 20–22; Goa, India. p. 304–9. doi:10.1109/iria53009.2021.9588702.
19. Deng X, Zhu J, Yang S. SFE-net: EEG-based emotion recognition with symmetrical spatial feature extraction. In: Proceedings of the 29th ACM International Conference on Multimedia; 2021 Oct 20–24; Chengdu, China. p. 2391–400. doi:10.1145/3474085.3475403.
20. Lin X, Chen J, Ma W, Tang W, Wang Y. EEG emotion recognition using improved graph neural network with channel selection. Comput Methods Programs Biomed. 2023;231(5):107380. doi:10.1016/j.cmpb.2023.107380.
21. Oka H, Ono K, Panagiotis A. Attention-based PSO-LSTM for emotion estimation using EEG. Sensors. 2024;24(24):8174. doi:10.3390/s24248174.
22. Schuld M, Killoran N. Quantum machine learning in feature Hilbert spaces. Phys Rev Lett. 2019;122(4):040504. doi:10.1103/physrevlett.122.040504.
23. Tudisco A, Marchesin A, Zamboni M, Graziano M, Turvani G. Evaluating angle and amplitude encoding strategies for variational quantum machine learning: their impact on model's accuracy. arXiv:2508.00768. 2025. doi:10.48550/arXiv.2508.00768.
24. Zaman K, Ahmed T, Hanif MA, Marchisio A, Shafique M. A comparative analysis of hybrid-quantum classical neural networks. arXiv:2402.10540. 2024. doi:10.48550/arXiv.2402.10540v1.
25. Senokosov A, Sedykh A, Sagingalieva A, Kyriacou B, Melnikov A. Quantum machine learning for image classification. Mach Learn Sci Technol. 2024;5(1):015040. doi:10.1088/2632-2153/ad2aef.
26. Qiao R, Qing C, Zhang T, Xing X, Xu X. A novel deep-learning based framework for multi-subject emotion recognition. In: Proceedings of the 2017 4th International Conference on Information, Cybernetics and Computational Social Systems (ICCSS); 2017 Jul 24–26; Dalian, China. p. 181–5. doi:10.1109/iccss.2017.8091408.
27. Li Y, Huang J, Zhou H, Zhong N. Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks. Appl Sci. 2017;7(10):1060. doi:10.3390/app7101060.
28. Wang Y, McCane B, McNaughton N, Huang Z, Shadli H, Neo P. AnxietyDecoder: an EEG-based anxiety predictor using a 3-D convolutional neural network. In: Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN); 2019 Jul 14–19; Budapest, Hungary. p. 1–8. doi:10.1109/ijcnn.2019.8851782.
29. Yang Y, Wu Q, Qiu M, Wang Y, Chen X. Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network. In: Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN); 2018 Jul 8–13; Rio de Janeiro. p. 1–7. doi:10.1109/ijcnn.2018.8489331.
30. Salama ES, El-Khoribi RA, Shoman ME, Wahby MA. EEG-based emotion recognition using 3D convolutional neural networks. Int J Adv Comput Sci Appl. 2018;9(8):329–37. doi:10.14569/ijacsa.2018.090843.
31. Alhagry S, Aly A, Reda A. Emotion recognition based on EEG using LSTM recurrent neural network. Int J Adv Comput Sci Appl. 2017;8(10):355–8. doi:10.14569/ijacsa.2017.081046.
32. Liu J, Wu G, Luo Y, Qiu S, Yang S, Li W, et al. EEG-based emotion classification using a deep neural network and sparse autoencoder. Front Syst Neurosci. 2020;14:43. doi:10.3389/fnsys.2020.00043.

33. Moon SE, Jang S, Lee JS. Convolutional neural network approach for EEG-based emotion recognition using brain connectivity and its spatial information. arXiv:1809.04208. 2018. doi:10.48550/arXiv.1809.04208.
34. Moctezuma LA, Abe T, Molinas M. Two-dimensional CNN-based distinction of human emotions from EEG channels selected by multi-objective evolutionary algorithm. Sci Rep. 2022;12(1):3523. doi:10.1038/s41598-022-07517-5.
35. Ahmed SMM, Sabur ET. Emotion analysis on EEG signal using machine learning and neural network. arXiv:2307.05375. 2023. doi:10.48550/arXiv.2307.05375.
36. Sateesh SK, Sparsh BK, Uma D. Decoding human emotions: analyzing multi-channel EEG data using LSTM networks. arXiv:2408.10328. 2024. doi:10.48550/arXiv.2408.10328.
37. Ding Y, Robinson N, Zhang S, Zeng Q, Guan C. TSception: capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition. IEEE Trans Affect Comput. 2023;14(3):2238–50. doi:10.1109/taffc.2022.3169001.
38. Mazzia V, Salvetti F, Chiaberge M. Efficient-CapsNet: capsule network with self-attention routing. Sci Rep. 2021;11(1):14634. doi:10.1038/s41598-021-93977-0.
39. Singh N, Pokhrel SR. Modeling feature maps for quantum machine learning. arXiv:2501.08205. 2025. doi:10.48550/arXiv.2501.08205.
40. Javadi-Abhari A, Treinish M, Krsulich K, Wood CJ, Lishman J, Gacon J, et al. Quantum computing with Qiskit. arXiv:2405.08810. 2024. doi:10.48550/arXiv.2405.08810.
41. Takeuchi Y, Morimae T, Hayashi M. Quantum computational universality of hypergraph states with Pauli-X and Z basis measurements. Sci Rep. 2019;9(1):13585. doi:10.1038/s41598-019-49968-3.
42. Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, et al. DEAP: a database for emotion analysis; using physiological signals. IEEE Trans Affect Comput. 2012;3(1):18–31. doi:10.1109/t-affc.2011.15.
43. Mahmood MA, Alsalem K, Elbashir MK, El-Ghany SA, El-Aziz AAA. Segmentation-enhanced approach for emotion detection from EEG signals using the fuzzy C-mean and SVM. Sci Rep. 2025;15(1):31956. doi:10.1038/s41598-025-17220-w.
44. Song Y, Zheng Q, Liu B, Gao X. EEG conformer: convolutional transformer for EEG decoding and visualization. IEEE Trans Neural Syst Rehabil Eng. 2023;31:710–9. doi:10.1109/TNSRE.2022.3230250.
45. Ju X, Li M, Tian W, Hu D. EEG-based emotion recognition using a temporal-difference minimizing neural network. Cogn Neurodyn. 2024;18(2):405–16. doi:10.1007/s11571-023-10004-w.
46. Maheshwari D, Sierra-Sosa D, Garcia-Zapirain B. Variational quantum classifier for binary classification: real vs synthetic dataset. IEEE Access. 2022;10(5):3705–15. doi:10.1109/access.2021.3139323.
47. Park G, Huh J, Park DK. Variational quantum one-class classifier. Mach Learn Sci Technol. 2023;4(1):015006. doi:10.1088/2632-2153/acafd5.