RSF Conference Proceeding Series: Engineering and Technology Vol. 4 No. 1 (2025) https://doi.org/10.31098/cset.v4i1.1014

Research Paper

Hybrid-Quantum CNN for Enhanced Facial Emotion Recognition: A Comparative Study with VGG16 on the RAF-DB Dataset

Mangaras Yanu Florestiyanto[1*], Herman Dwi Surjono[2], Handaru Jati[3], Wilis Kaswidjanti[4], Revta Faritzy[5]

[1,2,3] Engineering Science, Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia

[1,4] Informatics, Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia

[5] Communication Studies, Universitas Pembangunan Nasional Veteran Yogyakarta, Indonesia

Received : September 4, 2025	Revised : September 30, 2025	Accepted : October 2, 2025	Online : October 15, 2025

Abstract

Facial expression recognition (FER) underpins applications in affective computing but remains challenged by computational cost and the ambiguity of compound emotions. We introduce a Hybrid-Quantum Convolutional Neural Network (HQ-CNN) that integrates quantum principles (superposition, entanglement) into a classical CNN pipeline to enhance representational power and efficiency. Evaluated on the Real-World Affective Faces Database (RAF-DB), the HQ-CNN improves accuracy by 4.60% on basic emotions and 4.47% on compound emotions, while reducing computation time by up to 22.11% and 6.20%, respectively, relative to a VGG16 baseline. Confusion-matrix analysis shows fewer misclassifications on challenging compound categories, indicating better separation of overlapping affective cues. These results support the use of quantum-enhanced architectures as a viable path toward robust, real-time FER systems.

Keywords *Facial Expression Recognition, Quantum Machine Learning, Hybrid-Quantum CNN, VGG16, RAF-DB, Compound Emotions*

INTRODUCTION

Facial expression recognition (FER) is a vital component of affective computing, with applications in mental health monitoring, security, human-computer interaction, and social robotics. (Florestiyanto et al., 2024; Hassan et al., 2021; Zhao et al., 2022). As the integration of technology into human-centric applications increases, the demand for FER systems that accurately identify both basic and compound emotional states has grown (Mengoni et al., 2021). Convolutional Neural Networks (CNNs), especially VGG16, have become the preferred approach to FER due to their capability to capture intricate spatial hierarchies in facial expressions (Wei et al., 2022). However, CNNs like VGG16 face challenges in processing large datasets (Tran & Liu, 2021) and distinguishing subtle nuances in complex emotions, which can lead to misclassifications (Kheirandish et al., 2021; Wang & Hu, 2021).

To address these limitations, researchers are exploring the integration of quantum computing into CNN architectures (Henderson et al., 2020). Quantum principles such as superposition and entanglement present opportunities to reduce computational complexity and enhance feature extraction (Kumbhakar et al., 2021). Hybrid-Quantum CNN models combine the strengths of traditional CNNs with the efficiencies of quantum circuits (Mengoni et al., 2021), which are particularly advantageous for recognising compound emotions that involve subtle emotional distinctions (Rengasamy et al., 2024).

This study proposes a Hybrid-Quantum CNN and compares its performance with the

Corresponding author's email: mangaras.yanu@upnyk.ac.id

established VGG16 model using the Real-World Affective Faces Database (RAF-DB) (Li et al., 2017). By evaluating various hyperparameter settings and employing confusion matrix analysis, the research aims to assess the impact of quantum computing in improving both accuracy and computational efficiency in FER tasks. The findings aim to contribute to the understanding of quantum machine learning's potential in facial emotion recognition, ultimately advancing the development of real-time emotion recognition systems in the field of affective computing.

LITERATURE REVIEW

Facial expression recognition (FER) “in the wild” remains difficult due to pose, illumination, occlusion, and the ambiguity of compound emotions. The Real-World Affective Faces Database (RAF-DB) is a widely used benchmark addressing these issues: ~30,000 images annotated by around 40 crowd workers per image, with both basic and compound categories, enabling rigorous evaluation under real-world variability (Li & Deng, 2019). This design exposes label noise and overlapping affective cues, motivating methods that better separate subtly mixed expressions. In that context, VGG16 remains a transparent and reproducible baseline for FER studies and is often retained to contextualise newer architectures. Recent surveys further highlight RAF-DB’s central role and the need for models that handle label ambiguity and class imbalance (Li & Deng, 2019; Sajjad et al., 2023).

Beyond canonical CNNs, contemporary FER increasingly exploits attention and Transformer designs to fuse local and global cues, as well as long-range dependencies, on RAF-DB and related datasets. Vision-Transformer-based models (e.g., self-supervised or hybrid local-attention variants) report gains by combining fine-grained facial regions with global context, underscoring the importance of representational bias rather than depth alone (Chen et al., 2023; Tian et al., 2024). Parallel work specifically addresses compound emotions through soft/label-distribution learning and joint recognition of basic and compound categories, thereby improving robustness to label ambiguity and cross-class overlap (Jiang et al., 2024). These trends justify using VGG16 as a canonical reference while acknowledging the stronger contemporary baselines that anchor today’s FER literature.

In parallel, quantum–classical hybrids have emerged as a complementary approach to compact and expressive feature mappings. The Quantum Convolutional Neural Network (QCNN) exhibits hierarchical circuits with favorable parameter scaling ($\mathbf{O}(\mathbf{log}\,\boldsymbol{N})$ variational parameters for $\boldsymbol{N}$ qubits), indicating efficient trainability on near-term devices (Cong et al., 2019). Crucially, the parameter-shift rule yields analytic gradients for variational circuits, enabling end-to-end differentiable training of hybrid models (Schuld et al., 2019). Recent journal work on quantum data encodings (including amplitude encoding and its efficient/approximate realizations) clarifies practical trade-offs among qubit count, circuit depth, and fidelity on noisy hardware (Daimon & Matsushita, 2024; Mitsuda et al., 2024; Rath & Date, 2024). Positioned within this literature, our Hybrid-Quantum CNN (HQ-CNN) amplitude-encodes low-resolution grayscale inputs into an entangling circuit, fuses quantum measurements with a lightweight classical branch, and benchmarks against VGG16 on RAF-DB to test whether quantum-derived embeddings sharpen class boundaries, particularly for compound emotions.

RESEARCH METHOD

This section details the preprocessing pipeline, architecture design, training strategies, and comparative models used in this study. Our objective is to benchmark the proposed Hybrid Quantum-Classical CNN (HQ-CNN) against well-established deep learning baselines using the compound emotion subset of the RAF-DB dataset.

1. Comparative Deep Learning Baselines

The selection of baseline models in this study was guided by relevance and fairness in facial expression recognition (FER). VGG16 was chosen as a canonical convolutional architecture that remains widely used in FER literature, providing a historical benchmark. To ensure a fair comparison, we used ImageNet-pretrained weights and trained the model on 224×224 RGB images with appropriate augmentation. A grayscale, down-sampled variant (VGG16-Tiny) was also included to control for input-resolution effects relative to our hybrid model.

Our proposed HQ-CNN offers a lightweight hybrid design that integrates quantum feature extraction into a classical CNN pipeline using low-dimensional (16×16 grayscale) inputs. By contrasting HQ-CNN with both VGG16 (full resolution) and VGG16-Tiny (matched resolution), we disentangle the influence of input resolution from the contribution of the quantum branch, allowing for a fair and controlled assessment of compound emotion classification.

2. Data Collection and Preprocessing

The study utilises the Real-World Affective Faces Database (RAF-DB) (Li et al., 2017), which comprises two subsets: Basic (5,755 images across seven emotions—anger, disgust, fear, happiness, neutral, sadness, and surprise) and Compound (1,485 images across nine compound classes). Together, they provide a solid basis for analyzing both fundamental and nuanced facial emotions.

All images were re-labelled into class folders. For HQ-CNN, images were converted to grayscale, resized to 16×16, normalized to [0, 1], and flattened for amplitude embedding. For the VGG16 baseline, images were resized to 224×224 RGB and processed using the standard function. Training utilized data augmentation (random flips, 20° rotations, affine shear, and brightness/contrast jitter), along with a stratified split of the training (64%), validation (16%), and test (20%) sets. A summary table of the preprocessing steps is provided below:

Table 1. Data Preprocessing Overview

Model	Image Size	Colour Space	Normalisation / Scaling	Augmentation Applied	Notes
HQ-CNN	16 × 16	Grayscale	Pixel values scaled to [0, 1]	Yes (flip, crop, contrast, affine, shear)	Input is flattened and amplitude-encoded for the quantum feature map
VGG16	224 × 224	RGB	Using preprocess_input from Keras	Yes (same augmenters as HQ-CNN)	Features extracted using frozen convolutional layers

3. Hybrid Quantum–Classical CNN (HQ-CNN)

HQ-CNN integrates a shallow classical CNN with a quantum feature branch to combine efficient classical processing with quantum-derived representations. The pipeline is illustrated in Figure 1.

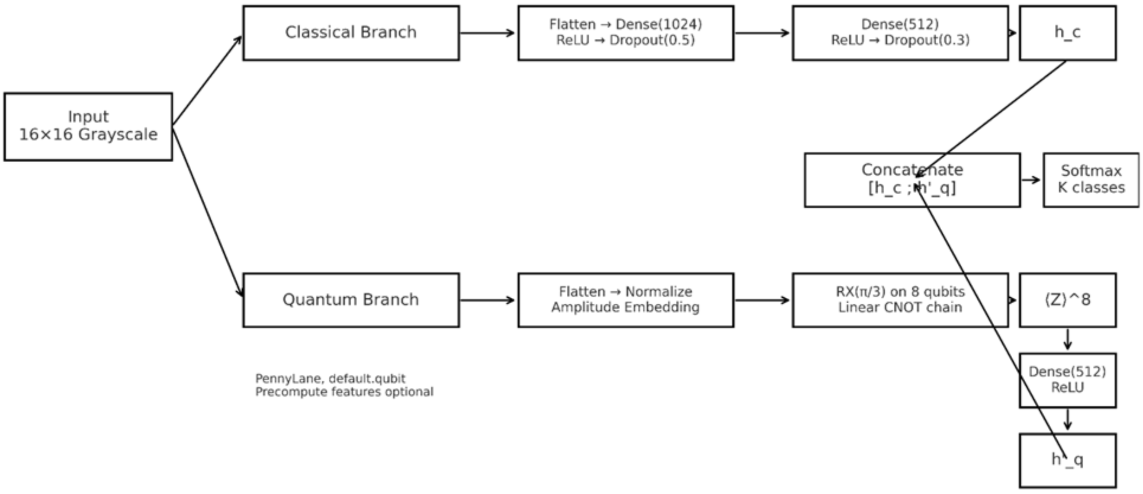

Figure 1. HQ-CNN architecture. A 16×16 grayscale image feeds two branches: (i) a classical pathway (Flatten → Dense(1024) → Dropout(0.5) → Dense(512) → Dropout(0.3) → $\boldsymbol{h_c}$); and (ii) a quantum pathway that normalises the flattened vector, performs amplitude embedding into 8 qubits, applies fixed RX(π/3) rotations with linear CNOT entanglement, and measures Pauli-Z expectations to produce an 8-D vector projected by Dense(512) to $\boldsymbol{h'_q}$. The features are concatenated and fed to a softmax classifier.

a. Classical branch

16×16 grayscale inputs are flattened and passed to two fully connected layers (1024, 512; ReLU) with Dropout (0.5, 0.3). This lightweight pathway provides a compact, nonlinear embedding of the image.

b. Quantum feature extraction

Flattened, normalized 256-D vectors are amplitude-encoded into 8 qubits. Each qubit receives a fixed RX(π/3) rotation; linear entanglement is introduced via chained CNOTs. Expectation values of Pauli-Z on all qubits yield an 8-D quantum feature vector, which is projected with a Dense(512) layer and concatenated with the classical features before the final softmax.

c. Training

Quantum features for train/validation/test are precomputed. The hybrid model uses Adam (learning rate = 1e-4), 100 epochs, batch size = 64, and validation accuracy monitoring; labels are one-hot encoded. Although quantum simulation is offline, the pipeline supports end-to-end differentiability via PennyLane's parameter-shift rule (e.g., with TorchLayer) if integrated during training.

4. VGG16

A comprehensive transfer learning pipeline was established utilizing the VGG16 architecture, which was pretrained on the extensive ImageNet dataset, deliberately excluding the top classification layers. To ensure consistency in input dimensions, all images were meticulously resized to a standard 224×224 pixels in RGB format. The convolutional base of the VGG16 model was intentionally frozen during the training process, allowing for the extraction of rich and relevant features from the images, which were then flattened for further processing.

The classifier architecture was thoughtfully designed to enhance performance and mitigate overfitting, incorporating the following layers:

a. A Dense layer with 1024 neurons, followed by a Dropout layer with a drop rate of 0.5 to promote regularization
b. A subsequent Dense layer with 512 neurons, again accompanied by a Dropout layer, this time set to 0.3

The output was produced by a softmax layer, facilitating multi-class classification. Training was conducted over 100 epochs using the Adam optimizer, configured with a learning rate of 0.0001 and a batch size of 64. Before the training of the classifier, the relevant features were meticulously extracted from the frozen convolutional base, thereby setting the stage for practical model training and performance evaluation.

5. Evaluation Protocol

All models were evaluated on the same test set to ensure consistency and comparability in performance assessments. The evaluation metrics included several key indicators: accuracy, precision, recall, and F1-score, both in macro and weighted forms. Additionally, per-class accuracy was calculated to understand how well the models performed on each class. At the same time, the confusion matrix provided a detailed breakdown of the predictions, highlighting true positives, false positives, and false negatives.

To gain insights into the computational costs associated with the HQ-CNN model specifically, training time, quantum feature engineering time, and prediction time were also meticulously recorded. This comprehensive approach not only enables a robust evaluation of model performance but also facilitates an understanding of the models' efficiency and scalability in practical applications.

FINDINGS AND DISCUSSION

This section reports the experimental evaluation of the Hybrid-Quantum CNN (HQ-CNN) against VGG16 on the RAF-DB Basic and Compound subsets. We assess classification accuracy, computational time, and error patterns via confusion matrices and classification reports, while testing robustness by varying epochs, learning rates, and batch sizes. Overall, the quantum-enhanced model demonstrates measurable gains in accuracy and efficiency, and it better captures nuanced expressions.

1. The Experimental Results

Table 2 summarises results across hyperparameter settings for both subsets: each row lists (epochs, learning rate, batch size), and columns report each model's accuracy and runtime. The table reveals sensitivity trends and the relative advantages of HQ-CNN over VGG16.

Table 2. Performance Comparison

No	Subset of the Dataset	Epoch	Learning Rate	Batch Size	Accuracy of VGG16	Accuracy of Hybrid-Quantum	Time Model VGG16	Time Model Hybrid-Quantum
1	Basic	70	0,001	16	23.78%	38.87%	150.05 s	113.3 s
2		70	0,001	64	41.69%	45.06%	93.58 s	74.42 s
3		70	0,0001	16	47.23%	49.62%	168.20 s	114.96 s
4		70	0,0001	64	45.82%	47.01%	95.69 s	74.14 s
5		100	0,001	16	29.97%	39.74%	201.08 s	143.94 s

No	Subset of the Dataset	Epoch	Learning Rate	Batch Size	Accuracy of VGG16	Accuracy of Hybrid-Quantum	Time Model VGG16	Time Model Hybrid-Quantum
6		100	0,001	64	39.52%	46.25%	102.68 s	81.85 s
7		100	0,0001	16	48.43%	49.73%	212.27 s	138.76 s
8		100	0,0001	64	44.73%	45.39%	110.21 s	82.01 s
9	Compound	70	0,001	16	17.65%	31.09%	45.86 s	39.50 s
10		70	0,001	64	26.05%	31.09%	40.44 s	26.48 s
11		70	0,0001	16	34.45%	31.93%	60.81 s	36.37 s
12		70	0,0001	64	26.89%	28.99%	44.23 s	25.50 s
13		100	0,001	16	18.49%	32.35%	60.69 s	44.31 s
14		100	0,001	64	21.01%	31.51%	33.32 s	28.51 s
15		100	0,0001	16	30.67%	32.35%	68.15 s	44.30 s
16		100	0,0001	64	30.25%	32.35%	43.28 s	28.92 s

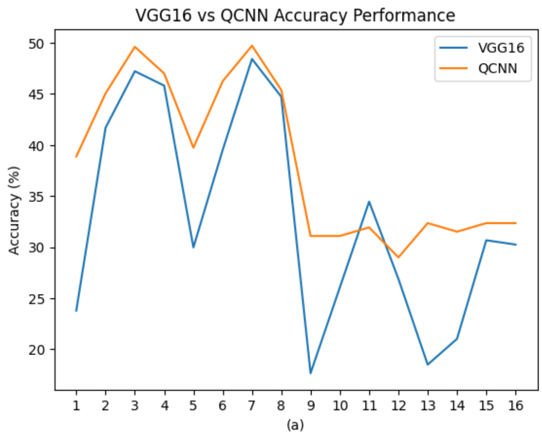

Figure 2a. VGG16 vs QCNN Accuracy Performance

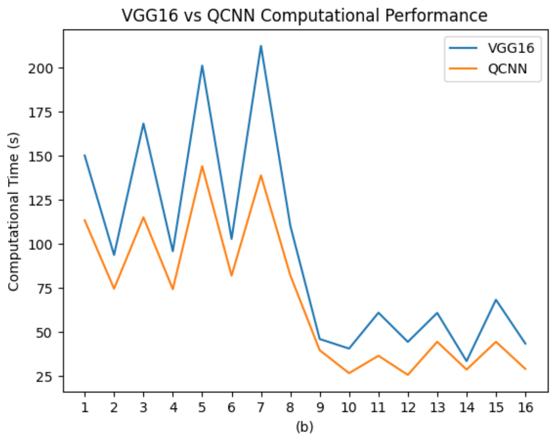

Figure 2b. VGG16 vs QCNN Computational Performance.

The experimental results demonstrate significant performance differences between the VGG16 and Hybrid-Quantum CNN (QCNN) models on both the Basic and Compound subsets. For instance, with a configuration of 70 epochs, a learning rate of 0.001, and a batch size of 16, the Hybrid-Quantum CNN achieved an accuracy of 38.87% on the Basic subset, outperforming VGG16's 23.78% by approximately 15 percentage points. Similarly, in the Compound subset, the Hybrid-Quantum model reached 31.09% accuracy, compared to VGG16's 17.65%, indicating gains of about 13–14%.

While increasing the number of epochs generally led to longer computation times for both models, the accuracy improvements were not always proportional to the increase in epochs (Figure 2). Lower learning rates (0.0001) improved training stability and accuracy, and smaller batch sizes (16) slightly enhanced accuracy but increased training duration. Overall, the Hybrid-Quantum CNN consistently surpassed VGG16 in accuracy and computational efficiency, particularly in recognising complex compound emotions. The accompanying accuracy and computational time graph further reinforces these findings, showcasing the Hybrid-Quantum approach's potential for real-time facial emotion recognition.

2. The Comparison of the Confusion Matrix

The confusion matrix (Figures 3 and 4) shows the performance of the VGG16 model and the Hybrid-Quantum CNN (QCNN) on the basic expressions subset and compound facial expressions. Each row represents the actual emotion, while each column shows the predicted emotion. Comparing these matrices reveals the accuracy and misclassification patterns of each model in interpreting facial expressions.

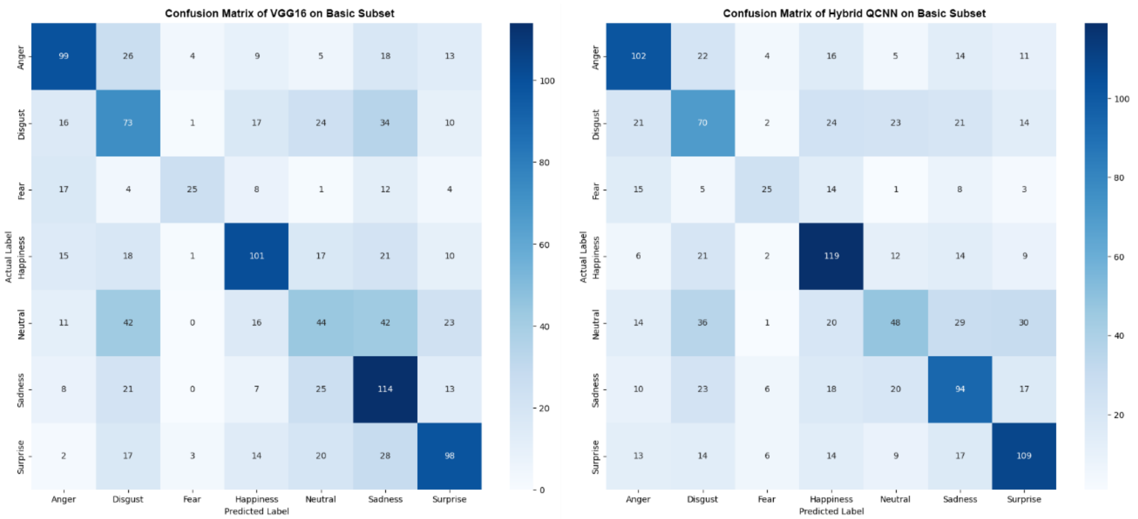

Figure 3. Confusion Matrix of VGG16 and Hybrid-Quantum CNN on The Basic Expressions Subset

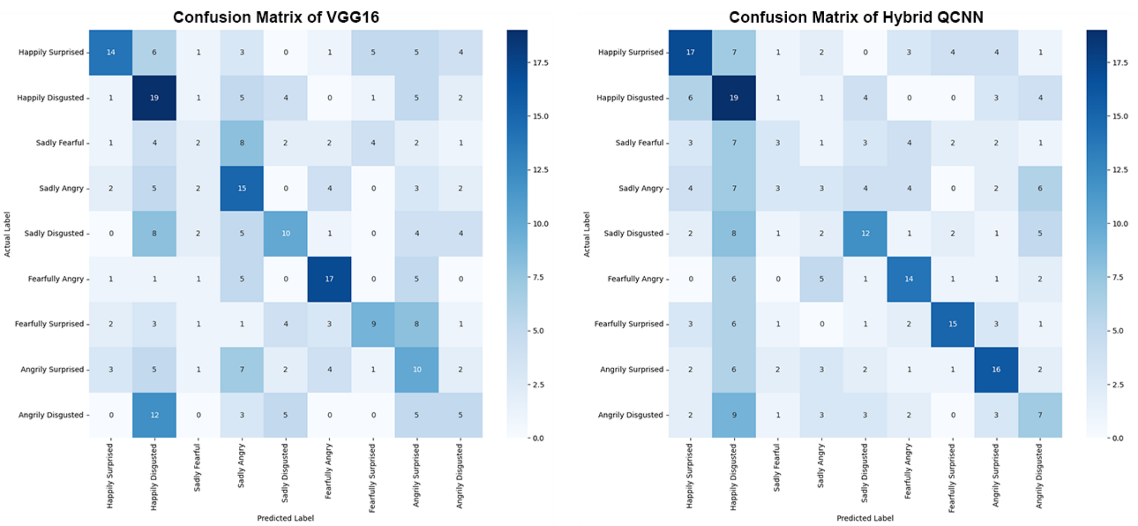

Figure 4. Confusion Matrix of VGG16 and Hybrid-Quantum CNN on The Compound Expressions Subset

a. Basic Subset

HQ-CNN reduces key confusions and increases true positives for Anger (102 vs. 99), Happiness (119 vs. 101), and Surprise (109 vs. 98), with comparable results for Fear (25 vs. 25), but a cleaner error profile. Disgust shows fewer spillovers into Sadness despite a slight drop in corrects, while Neutral improves modestly. The principal regression is Sadness (94 vs 114 for VGG16), indicating class-specific tuning needs.

b. Compound Subset

HQ-CNN improves several overlap-heavy classes ("Happily Surprised" (17 vs 14), "Angrily Disgusted" (7 vs 5), and "Fearfully Surprised" (15 vs 9)) and slightly betters "Sadly Fearful" (3 vs 2) with a more balanced error spread. "Sadly Angry" remains a VGG16 strength. Overall, HQ-CNN sharpens boundaries in surprise-involving and nuanced categories, while targeted refinements are needed for specific negative blends.

3. The Classification Report Comparison

Figures 5 and 6 compare precision, recall, and F1 scores for each emotion class on the basic expressions subset. This graph succinctly highlights the strengths and weaknesses of VGG16 and the Hybrid-Quantum CNN, making it clear where quantum enhancements yield improvements and where traditional methods still excel.

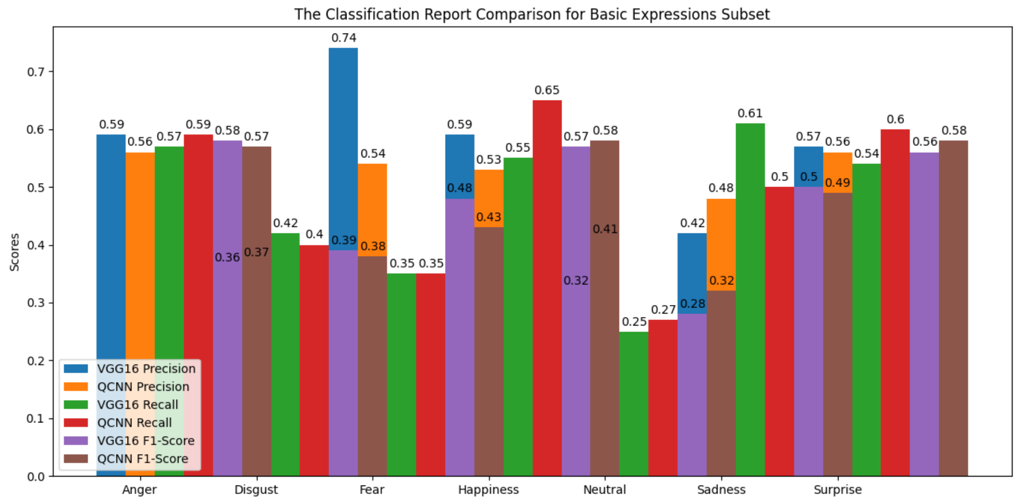

Figure 5. The Classification Report Comparison for Basic Expressions Subset

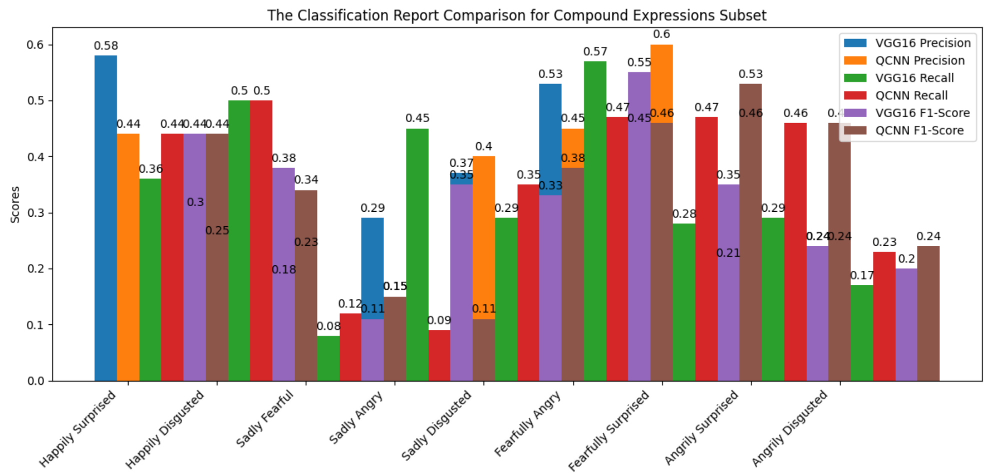

Figure 6. The Classification Report Comparison for Compound Expressions Subset

a. Basic Subset
VGG16 and HQ-CNN perform similarly on Anger and Disgust; VGG16 is stronger on Fear (notably higher precision), while HQ-CNN improves Happiness, Neutral, and Surprise via higher recall and net F1 gains. Sadness is effectively a tie (precision–recall trade-off yields near-equal F1).

b. Compound Subset
HQ-CNN shows more evident advantages in “Fearfully Surprised” and “Angrily Surprised” (higher F1), offers balanced precision–recall on “Happily Surprised” (vs VGG16’s high-precision/low-recall), and is comparable on “Happily Disgusted” and “Angrily Disgusted”. Both models struggle with “Sadly Fearful” (HQ-CNN performs modestly better), while “Sadly Angry” remains a weakness for HQ-CNN (VGG16 achieves a higher F1).

4. Discussion

This study compares a traditional VGG16 baseline with a Hybrid-Quantum CNN (QCNN) on RAF-DB (Basic and Compound subsets). Across metrics (accuracy, runtime, and error profiles), QCNN consistently outperforms VGG16. On the Basic subset, it achieves higher accuracy while reducing computation time by more than 20%. On the Compound subset, under 70 epochs, learning rate = 0.001, batch = 16, QCNN achieves 31.09% accuracy versus 17.65% for VGG16, with more apparent separation in difficult classes (e.g., “Fearfully Surprised”, “Angrily Surprised”).

Performance gains are not uniform: VGG16 is stronger for Fear in Basic, and QCNN underperforms on “Sadly Angry”. Both models are sensitive to hyperparameters; lower learning rates and smaller batches improve accuracy but increase training time. Overall, quantum-enhanced feature mapping yields robust improvements, particularly for compound emotions, suggesting QCNN as a promising and resource-conscious option. Future work will focus on circuit optimisation and integrating attention or multi-scale features to address remaining class-specific weaknesses.

CONCLUSIONS

This study demonstrates that integrating quantum computing principles into convolutional neural networks can significantly enhance facial emotion recognition performance. The Hybrid-Quantum CNN achieved higher accuracy, especially for complex compound expressions, and reduced computational time compared to the traditional VGG16 model. Notably, improvements in capturing nuanced expressions such as Fearfully Surprised and Angrily Surprised underscore the potential of quantum-enhanced architectures in addressing the inherent challenges of FER.

However, specific categories, such as Sadly Angry, still pose significant challenges, suggesting that further refinements are needed. Future research should optimise quantum circuits, integrate advanced feature extraction techniques (e.g., attention mechanisms and multi-scale processing), and expand the dataset to include more diverse and balanced samples. These directions will help bridge the remaining gaps and pave the way for robust, real-time emotion recognition systems in practical applications. Overall, the findings provide a strong foundation for the continued exploration of quantum-enhanced deep learning in affective computing, with promising implications for improving accuracy and efficiency in facial emotion recognition.

LIMITATIONS & FURTHER RESEARCH

Limitations

This study has several constraints:

1. The comparative scope is narrow: results are reported against VGG16 (and a resolution-matched VGG16-Tiny control), without stronger modern baselines (e.g., ResNet, EfficientNet, or transformer models), which limits external validity
2. There is an input-modality mismatch between models: HQ-CNN operates on 16×16 grayscale, while VGG16 uses 224×224 RGB; although VGG16-Tiny helps control for resolution, residual modality effects may remain
3. The absolute accuracies—particularly on compound emotions—remain modest for deployment-grade systems, indicating headroom for methodological improvement
4. Reported scores are based on single-run training without multi-seed repeats, dispersion (mean ± SD), or statistical significance testing, so robustness under re-runs is undetermined.
5. The quantum branch is evaluated via offline simulation with precomputed quantum features; fully end-to-end differentiable training through the circuit and execution on real hardware are not demonstrated.
6. Runtime accounting may be confounded if quantum feature extraction is not included uniformly in wall-clock comparisons.
7. The evaluation is single-dataset (RAF-DB), whose compound subset is relatively small and imbalanced, and class-specific weaknesses persist (e.g., certain negative/overlapping emotions), suggesting sensitivity to label ambiguity and skew.
8. The hyperparameter search is limited, and reproducibility could be strengthened with fuller reporting of software/hardware environments.

Future Work

We see four priority directions, as presented below:

1. Broaden baselines and controls: include ResNet/EfficientNet and transformer models; maintain strictly resolution- and modality-matched classical controls to isolate quantum effects.
2. Strengthen experimental rigour: train with ≥ 5 random seeds, report mean ± SD and paired significance tests (e.g., Wilcoxon); expand the search over optimizers, learning-rate schedules, batch sizes, and regularization.
3. Advance the hybrid pipeline: enable end-to-end training via parameter-shift gradients (e.g., PennyLane TorchLayer), study convergence/gradient variance, and benchmark on noisy intermediate-scale hardware to assess practicality; report fair efficiency metrics (end-to-end wall clock including preprocessing) plus FLOPs/parameter counts.
4. Improve modelling and evaluation: explore richer quantum ansätze (deeper layers, alternative entanglement topologies, learnable encodings), attention/multi-scale modules, and ablations to attribute gains; extend to additional FER datasets and apply imbalance-aware learning (class weights, focal loss, calibrated thresholds), emphasizing macro-F1 and per-class analyses. To facilitate reuse and verification, future releases should include code, circuit diagrams, training scripts, and environment specs, alongside a lightweight inference demo for real-time feasibility.

REFERENCES

Chen, X., Zheng, X., Sun, K., Liu, W., & Zhang, Y. (2023). Self-supervised vision transformer-based few-shot learning for facial expression recognition. *Information Sciences*, *634*, 206–226. https://doi.org/10.1016/j.ins.2023.03.105

Cong, I., Choi, S., & Lukin, M. D. (2019). Quantum convolutional neural networks. *Nature Physics*, *15*(12), 1273–1278. https://doi.org/10.1038/s41567-019-0648-8

Daimon, S., & Matsushita, Y. (2024). Quantum circuit generation for amplitude encoding using a transformer decoder. *Physical Review Applied*, *22*(4), L041001. https://doi.org/10.1103/PhysRevApplied.22.L041001

Florestiyanto, M. Y., Surjono, H. D., & Jati, H. (2024). Harnessing Emotions for Effective Online Learning: A Systematic Review of Emotion Recognition Techniques. *Proceedings of the 2023 1st International Conference on Advanced Informatics and Intelligent Information Systems (ICAI3S 2023)*, 71–89. https://doi.org/10.2991/978-94-6463-366-5_8

Hassan, T., Seus, D., Wollenberg, J., Weitz, K., Kunz, M., Lautenbacher, S., Garbas, J. U., & Schmid, U. (2021). Automatic Detection of Pain from Facial Expressions: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(6), 1815–1831. https://doi.org/10.1109/TPAMI.2019.2958341

Henderson, M., Shakya, S., Pradhan, S., & Cook, T. (2020). Quanvolutional neural networks: powering image recognition with quantum circuits. *Quantum Machine Intelligence*, *2*(1), 2. https://doi.org/10.1007/s42484-020-00012-y

Jiang, J., Wang, M., Xiao, B., Hu, J., & Deng, W. (2024). Joint recognition of basic and compound facial expressions by mining latent soft labels. *Pattern Recognition*, *148*, 110173. https://doi.org/10.1016/j.patcog.2023.110173

Kheirandish, D., Haghparast, M., Reshadi, M., & Hosseinzadeh, M. (2021). Efficient techniques for fault detection and location of multiple controlled Toffoli-based reversible circuit. *Quantum Information Processing*, *20*(11), 370. https://doi.org/10.1007/s11128-021-03292-w

Kumbhakar, P., Shanmugam, A., Sharma, C. H., Reno, J. L., & Thalakulam, M. (2021). Quantum point contact galvanically coupled to planar superconducting resonator: a shot-noise-limited broad-band electrical amplifier. *Quantum Science and Technology*, *6*(4), 045006. https://doi.org/10.1088/2058-9565/ac107f

Li, S., & Deng, W. (2019). Reliable Crowdsourcing and Deep Locality-Preserving Learning for Unconstrained Facial Expression Recognition. *IEEE Transactions on Image Processing*, *28*(1), 356–370. https://doi.org/10.1109/TIP.2018.2868382

Li, S., Deng, W., & Du, J. (2017). Reliable Crowdsourcing and Deep Locality-Preserving Learning for Expression Recognition in the Wild. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2584–2593. https://doi.org/10.1109/CVPR.2017.277

Mengoni, R., Incudini, M., & Di Pierro, A. (2021). Facial expression recognition on a quantum computer. *Quantum Machine Intelligence*, *3*(1), 8. https://doi.org/10.1007/s42484-020-00035-5

Mitsuda, N., Ichimura, T., Nakaji, K., Suzuki, Y., Tanaka, T., Raymond, R., Tezuka, H., Onodera, T., & Yamamoto, N. (2024). Approximate complex amplitude encoding algorithm and its application to data classification problems. *Physical Review A*, *109*(5), 052423. https://doi.org/10.1103/PhysRevA.109.052423

Rath, M., & Date, H. (2024). Quantum data encoding: a comparative analysis of classical-to-quantum mapping techniques and their impact on machine learning accuracy. *EPJ Quantum Technology*, *11*(1), 72. https://doi.org/10.1140/epjqt/s40507-024-00285-3

Rengasamy, K., Joshi, P., & Raveendra, V. V. S. (2024). Real-time hybrid facial expression analysis model using quantum distance-based classifier and classical artificial neural networks: QUADCANN. *Quantum Information Processing*, *23*(6), 220. https://doi.org/10.1007/s11128-024-04437-3

Sajjad, M., Ullah, F. U. M., Ullah, M., Christodoulou, G., Alaya Cheikh, F., Hijji, M., Muhammad, K., & Rodrigues, J. J. P. C. (2023). A comprehensive survey on deep facial expression recognition: challenges, applications, and future guidelines. *Alexandria Engineering Journal*, *68*, 817–840. https://doi.org/10.1016/j.aej.2023.01.017

Schuld, M., Bergholm, V., Gogolin, C., Izaac, J., & Killoran, N. (2019). Evaluating analytic gradients on quantum hardware. *Physical Review A*, *99*(3), 032331. https://doi.org/10.1103/PhysRevA.99.032331

Tian, Y., Zhu, J., Yao, H., & Chen, D. (2024). Facial Expression Recognition Based on Vision Transformer with Hybrid Local Attention. *Applied Sciences*, *14*(15), 6471. https://doi.org/10.3390/app14156471

Tran, L., & Liu, X. (2021). On Learning 3D Face Morphable Model from In-the-Wild Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(1), 157–171. https://doi.org/10.1109/TPAMI.2019.2927975

Wang, J., & Hu, X. (2021). Convolutional Neural Networks with Gated Recurrent Connections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. https://doi.org/10.1109/TPAMI.2021.3054614

Wei, X.-S., Song, Y.-Z., Aodha, O. Mac, Wu, J., Peng, Y., Tang, J., Yang, J., & Belongie, S. (2022). Fine-Grained Image Analysis With Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(12), 8927–8948. https://doi.org/10.1109/TPAMI.2021.3126648

Zhao, S., Yao, X., Yang, J., Jia, G., Ding, G., Chua, T. S., Schuller, B. W., & Keutzer, K. (2022). Affective Image Content Analysis: Two Decades Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(10), 6729–6751. https://doi.org/10.1109/TPAMI.2021.3094362.