Hybrid Quantum-Classical Convolutional Neural Networks for Image Classification in Multiple Color Spaces

Kwok-Ho Ng[a], Tingting Song[a], Zhiquan Liu[a]

[a] *College of Cyber Security, Jinan University, 601 Huangpu Avenue West, Guangzhou, 510632, Guangdong, China*

Abstract

The growing complexity and scale of image processing tasks challenge classical convolutional neural networks (CNNs) with high computational costs. Hybrid quantum-classical convolutional neural networks (HQCNNs) show potential to improve performance by accelerating processing speed, enhancing classification accuracy, and reducing model parameters, though studies have primarily focused on the RGB color space. However, the effectiveness of HQCNNs in non-RGB color spaces, such as Lab, YCrCb, and HSV, remains largely unexplored. We propose an HQCNN to evaluate image classification across diverse color spaces. The HQCNN integrates parameterized quantum circuits (PQCs) with a classical CNN, leveraging quantum entanglement and trainable gates to enhance expressiveness across varied color representations. We assess performance on MNIST, CIFAR-10, EuroSAT, and SAT-4 datasets. Experimental results demonstrate that the HQCNN outperforms the classical CNN across all tested color spaces for the ten-class MNIST task, achieving a best accuracy of 94.3% in Lab compared to 92.8% in RGB for the CNN, with superior performance on other datasets in various color spaces. These findings highlight the potential of non-RGB color spaces and optimized PQC designs to improve classification performance. We provide new insights for advancing hybrid quantum-classical computer vision through optimized PQC architectures and diverse color space applications.

Keywords: Quantum convolutional neural network, Parameterized quantum circuit, Image classification

1. Introduction

Computer image classification is one of the core research areas in computer vision, with widespread applications across multiple fields. Notable examples include object classification in autonomous driving [1, 2], disease classification using computed tomography images in medicine [3, 4], defect detection in industrial settings [5, 6], and product identification and classification in agriculture [7, 8]. Advances in classification accuracy and algorithmic development continue to drive rapid and efficient progress across these disciplines. Convolutional neural networks (CNNs) have been a cornerstone of computer vision tasks, particularly excelling in image recognition over the past decade [9, 10]. Their hierarchical structure is designed to extract image features efficiently. The evolution of model scale and the emergence of innovative architectures, such as ResNet [11] and DenseNet [12], suggest that CNNs remain at the forefront of image classification methodologies.

However, classical algorithms face significant computational challenges when processing high-dimensional data, a bottleneck that becomes especially apparent in more complex tasks. Quantum information and computing technologies are advancing rapidly, with quantum algorithms promising exponential performance improvements [13, 14, 15]. These advancements have the potential to substantially enhance performance in computer vision and other applications requiring high-dimensional data processing. Additionally, quantum neural networks (QNNs), a subset of quantum algorithms, are being actively developed to integrate quantum computing with machine learning, potential offering advantages [16]. Research by Gil-Fuster et al. [17] indicates that quantum machine learning may achieve effective generalization with minimal training data. Recent studies [18] further demonstrate that classical functions can be approximated using parameterized quantum circuits (PQCs), consistent with the universal approximation theorem.

Quantum convolutional neural networks (QCNNs) are quantum algorithms inspired by classical CNNs. Proposed by Cong et al. [19], their structure mirrors that of CNNs, incorporating convolutional, pooling, and fully connected layers to extract features and process information from quantum states. By combining the layered architecture of classical convolutional networks with the high-dimensional capabilities of quantum computing, QCNNs may offer notable advantages in handling complex data. Notably, QCNNs can utilize a small number of variational parameters for efficient training, suggesting their potential for feature extraction and classification in image processing tasks [17]. Moreover, QCNNs have been shown to avoid barren plateaus [20], distinguishing them from other QNNs by remaining trainable even under random initialization, thus avoiding the vanishing gradient issues prevalent in other quantum models.

Currently, QCNN applications in image classification primarily focus on single-channel datasets, such as MNIST, where they achieve high classification accuracy with significantly fewer parameters than traditional models [21]. However, their application to color image datasets remains limited, with most research focusing on the RGB color space. The exploration of alternative color spaces, such as Lab, YCrCb, and HSV, has been relatively underexplored.

In contrast, classical CNN research has extensively investigated various color spaces. For example, the Lab color space excels in identifying plant pests [22], while YCrCb outperforms RGB in image forgery detection [23], and the HSV color model improves tomato ripeness identification in complex environments for automated harvesting [24].

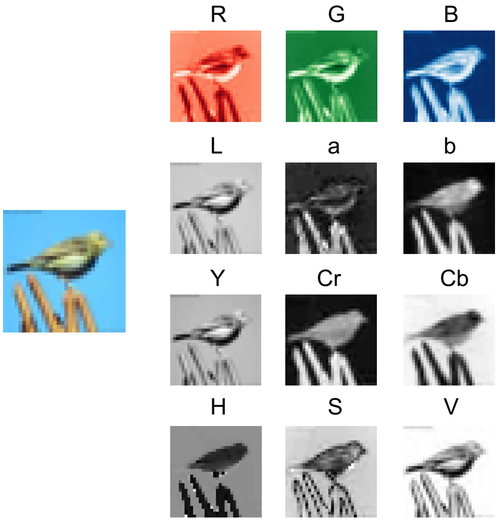

Figure 1: Plotted each channel of the image in four color spaces: RGB, Lab, YCrCb, and HSV.

Studies suggest that the selection of color space significantly affects the accuracy of classification model [25]. Research on hybrid color spaces in CNNs has demonstrated improved accuracy, with a model utilizing 1.75M parameters and multiple color spaces performing comparably to one with 27M parameters. Converting images in different color spaces and visualizing their channel information, as shown in Figure 1, reveals noticeable differences in brightness and edge details, motivating the exploration of various color spaces to enhance model robustness and generalization. Further extending QCNNs to these color spaces could provide valuable insights into performance across various color characteristics and highlight potential advantages in diverse image datasets.

To investigate the potential of QCNNs in various color spaces, we propose a hybrid quantum-classical convolutional neural network (HQCNN) model, the architecture of which is shown in Figure 2. The HQCNN combines the strengths of quantum computing and classical neural networks: the quantum component extracts complex features, while the classical component handles fully connected layer operations and loss computation to optimize classification performance. Initially, the image dataset is scaled and converted into multiple color spaces, including Lab, YCrCb, and HSV. Each image is processed using a fixed 2×2 quantum convolutional kernel, encoding four pixel values from each channel into quantum states. After encoding, PQCs evolve the quantum states, followed by measurement to obtain feature values corresponding to the three channels.

The main contributions of this paper:

(1) This paper proposes a HQCNN model that is among the first to demonstrate image classification accuracy across various color spaces, offering a novel perspective on harnessing quantum advantages in computer vision. Given the limited exploration of this area in current quantum computing research, this study aims to contribute significantly to the field. For each color channel, after encoding and applying the first quanvolutional circuits, we introduce an optimized additional two-qubit parameterized unitary operator before the pooling layer to extract channel information and assess its impact on model performance.

(2) We analyze the types of quantum rotation gates used in the encoding layer and reduce the number of two-qubit gates in the quantum circuit, resulting in shallower circuits with fewer parameters. For each color channel, after encoding and applying the first quanvolutional circuits, we introduce an optimized additional two-qubit parameterized unitary operator before the pooling layer to extract channel feature and assess its impact on model performance. This flexibility facili-

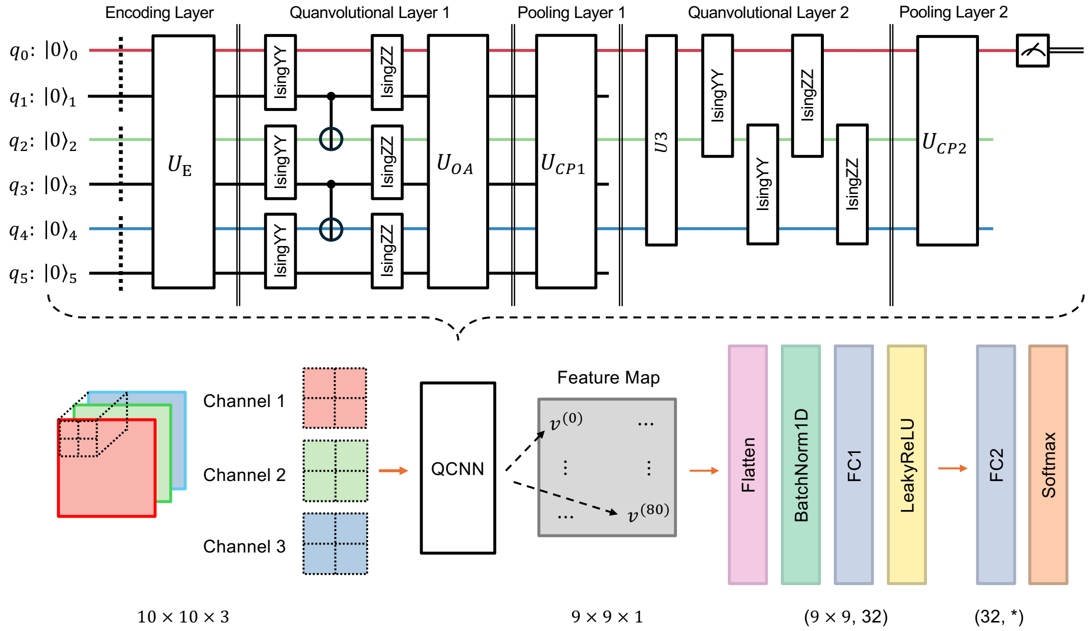

Figure 2: Details of our proposed HQCNN architecture, showing the main components of the QCNN. This includes the relationship between image channel encoding and qubit positions within the QCNN, and the output feature map size of the quantum layer.

tates exploration of quantum structures in specific tasks.

The remainder of this paper is organized as follows: Section 2 reviews current HQCNN applications in image classification. Section 3 details the composition and implementation of the proposed HQCNN model. Section 4 presents experimental results from classification tasks in various color spaces. Finally, Section 5 summarizes the study and proposes future research directions.

2. Related work

This section provides briefly overview of PQCs and reviews recent advancements in HQCNN models for color image classification.

The current era of quantum computing is characterized by noisy intermediate-scale quantum (NISQ) devices, where quantum hardware is prone to noise. Consequently, designing quantum algorithms that function effectively on these noisy devices is essential. PQCs address this challenge by providing a flexible framework for quantum computing [26]. Researchers have utilized PQCs to develop various quantum algorithms, such as the variational quantum eigensolver (VQE) [27] and the quantum approximate optimization algorithm (QAOA) [28]. These algorithms have demonstrated potential in practical applications, including electronic structure research [29] and solving combinatorial optimization problems, such as the MaxCut problem [30]. The flexibility of PQC makes them an ideal choice for exploring potential quantum advantages in NISQ devices. It can approximate target values for specific tasks by adjusting the variational parameters. The QCNN model takes full advantage of this flexibility, showing promising performance in particular quantum computing tasks. Building on this foundation, HQCNN further integrates the robustness of classical algorithms, enhancing the reliability and stability of the final classification results.

Early work by Henderson et al. [31] compared the classification accuracy of their proposed QCNN with that of a classical CNN and a random model on the MNIST dataset. Their results demonstrated that, compared to the classical CNN, the accuracy of QCNN improved with more training iterations, and adding a certain number of quantum filters enhanced the model performance. In deeper hybrid models, quantum convolutional layers boosted network accuracy. Ovalle-Magallanes et al. [32] conducted ablation studies on the QCNN model for the MNIST classification task, revealing that increasing the size and number of quanvolutional kernels in the PQC does not always result in optimal classification accuracy. Liu et al [33] proposed an early HQCNN for image classification tasks.

This model uses a global phase measurement method to measure all qubits and generate multi-channel feature maps. The classification results on the Tetris dataset have shown that it could even surpass the accuracy of classical CNN.

Jing et al. [34] proposed two quantum convolutional circuits that compare different methods for extracting inter-channel and intra-channel information. Experimental results show that these newly proposed quantum convolutional circuits achieve higher test accuracy than classical CNN in RGB image classification tasks. Moreover, their work demonstrates that the design of larger PQCs can improve the performance of multi-class classification tasks.

Riaz et al. [35] proposed a QCNN requiring no parameter updates, which showed outperforming QCNN with randomly generated PQCs on the MNIST and CIFAR-10 datasets. However, its accuracy decreased on the German Traffic Sign Recognition Benchmark (GTSRB) dataset, suggesting that the application of quantum algorithm-based classifiers requires further investigation for colored and complex datasets. Smaldone et al. [36] developed a QCNN for multi-channel image processing, which achieved higher accuracy on the CIFAR-10 dataset compared to prior work [34], with reduced quantum circuit complexity.

Yang and Sun [37] proposed a hybrid quantum-classical model for semiconductor wafer defect detection, conducted an ablation study on various PQC [38]. The experimental results showed that angle encoding outperformed basis state and amplitude encoding, with stronger entangled and expressive circuits showing clear advantages. Their hybrid model achieved higher accuracy with limited quantum resources than classical CNN on the LIT Hotspot (ICCAD-2012) and Defect Wafer Map (WM-811K) datasets. Senokosov et al. [21] proposed two hybrid neural networks. The first model, using parallelized quantum circuits, achieved over 99% accuracy on MNIST and Medical MNIST and over 82% on CIFAR-10 with eight times fewer parameters than classical models. The second model featuring quantum convolutional layers achieved performance comparable to classical CNNs while using four times fewer parameters.

Research on QCNN-based color image classification has predominantly focused on the RGB color spaces, while broader studies have mainly addressed grayscale image processing [39]. However, the application of QCNNs or hybrid quantum models to image classification tasks in alternative color spaces remains limited.

3. Methods

The HQCNN model proposed in this study integrates quantum and classical components to utilize the strengths of both paradigms. The quantum component encodes classical image data into quantum states, utilizing the high-dimensional Hilbert spaces of quantum computing. These states evolve through PQC, and local measurements are performed on qubits to extract image features while mitigating vanishing gradients associated with the barren plateau problem [40]. This approach enables direct quantum processing of raw input data without requiring additional preprocessing through a classical CNN.

3.1. Quantum component

The quantum component consists of three primary layers: encoding, quanvolutional, and pooling layers, executed sequentially, following the quantum circuit structure proposed by Barenco et al. [41].

3.1.1. Encoding layer

In the quantum encoding (embedding) layer, single-qubit parameterized rotation gates $R_x(\theta)$, $R_y(\theta)$, and $R_z(\theta)$ are employed for encoding.

$$R_x(\theta) = e^{-i\theta\sigma_x/2} = \begin{bmatrix} \cos(\theta/2) & -i\sin(\theta/2) \\ -i\sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \tag{1}$$

$$R_y(\theta) = e^{-i\theta\sigma_y/2} = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \tag{2}$$

$$R_z(\theta) = e^{-i\theta\sigma_z/2} = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix} \tag{3}$$

The structure of the unitary operators U_{RY} and U_{RZ} in the quantum circuit, as depicted in Figure 3, utilizes two distinct single-qubit rotation gates per qubit to provide local degrees of freedom, preparing the quantum states of the image pixels.

For the R channel of an RGB image, the quanvolutional kernel extracts four pixel values from the channel. These values are encoded into quantum states as follows: the first two values are processed using angle encoding via the $R_y(\theta)$ gate, while the remaining two are encoded with the $R_z(\theta)$ gate, as depicted in Figure 3 (d). This encoding process is applied to each of the three channels, resulting in twelve encoded pixel values for the selected kernel. Each set of twelve pixel values corresponds to the image pixels selected by the quanvolution kernel in three channels.

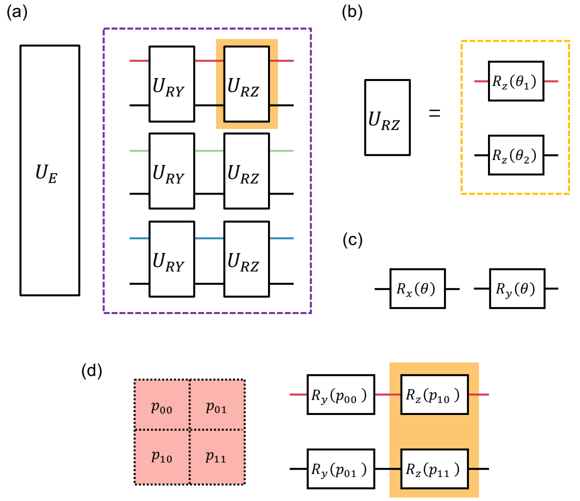

Figure 3: (a) The structure of the quantum encoding layer in QCNN, exemplifying the R_{YZ} rotation gate combination (along with R_{XZ} and R_{XY} combinations). (b) The structure of the U_{RZ} unitary operator module, applying an $R_z(\theta)$ operator on each circuit. (c) The representation of $R_x(\theta)$ and $R_y(\theta)$ operators in the quantum circuit. (d) An example of selecting the R channel of an RGB image using a 2×2 kernel, where the four pixels are sequentially encoded into corresponding rotation gates at specific positions.

3.1.2. Quanvolutional layer

Following encoding, the first quanvolutional layer employs two distinct parameterized two-qubit operators, IsingYY and IsingZZ, each implemented using two CNOT gates for entanglement and a parameterized rotation gate, as shown in Figure 4 (c). For an RGB image, as an illustrative example, the four pixel values of each color channel are mapped to two qubits via angle encoding, with a total of six qubits representing the red, blue, and green channels.

Initially, the IsingYY unitary operator is applied to the two qubits within each channel, using trainable parameters to provide non-local degrees of freedom. This generates and optimizes entangled states, enhancing the expressiveness of the variational quantum circuit (VQC), also referred to as the PQC, for complex data. Subsequently, CNOT gates are applied to create cross-channel entanglement, specifically between the R and G channels (q_1-control, q_2-target) and the G and B channels (q_3-control, q_4-target), resulting in a globally entangled state across all six qubits. Finally, the IsingZZ unitary operator is similarly applied. It introduces relative phases through trainable parameters to optimize the phase structure of the entangled state, further improving the PQC's expressiveness for image data.

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{4}$$

To further enhance the model, an additional optimized unitary operator module U_{OA} is introduced between the first quanvolutional and pooling layers, as shown in Figure 4 (b). The U_{OA} operator comprises three parameterized controlled rotation gates, which adjust the entanglement structure across channels to capture non-local correlations of pixel features, enabling the evaluation of the impact of different parameterized gates on the expressiveness of the QCNN.

$$CR_x(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\theta/2) & -i\sin(\theta/2) \\ 0 & 0 & -i\sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \tag{5}$$

Sim et al. [38] found that the $CR_x(\theta)$ gate outperforms the $CR_z(\theta)$ gate in expressiveness and entanglement capability. Accordingly, ablation experiments are conducted using three types of parameterized two-qubit gates, $CR_x(\theta)$, $CR_y(\theta)$, and $CR_z(\theta)$, to assess their effects.

$$U3(\theta,\phi,\delta) = \begin{bmatrix} \cos(\theta/2) & -e^{i\delta}\sin(\theta/2) \\ e^{i\phi}\sin(\theta/2) & e^{i(\phi+\delta)}\cos(\theta/2) \end{bmatrix} \tag{6}$$

Following the first pooling layer, the second quanvolutional layer, as shown in Figure 2, applies the parameterized $U3$ gate to qubits q_0, q_2, and q_4, using trainable parameters (θ, ϕ, δ) to adjust local amplitudes and phases, thereby improving the expressiveness of the main features. Subsequently, IsingYY and IsingZZ unitary operators are applied to adjust interactions between color channels, strengthening non-local feature correlations among the main feature qubits across the three channels.

3.1.3. Pooling layer

The first pooling layer employs the unitary operator U_{CP1}, as shown in Figure 5 (a), consisting of three controlled-phase gates $CP(\theta)$. These gates utilize trainable parameters to provide non-local degrees of freedom, adjusting the phase structure across the red, green, and blue channels. Specifically, q_0, q_2, and q_4 serve as target qubits, representing the main features of each channel and are retained for subsequent processing,

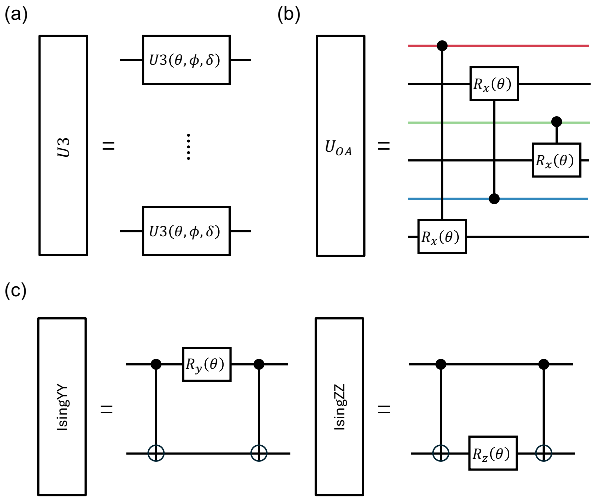

Figure 4: (a) The unitary operator $U3$ consists of independently applying a single-qubit parameterized gate $U3(\theta, \phi, \delta)$ to each qubit. (b) An example of the composition of the optimized additional unitary operator module U_{OA}, utilizing type of $CR_x(\theta)$ gate, with optional controlled $R_y(\theta)$ and $R_z(\theta)$ gates. (c) The structure of the IsingYY and IsingZZ unitary operators in the quantum circuit.

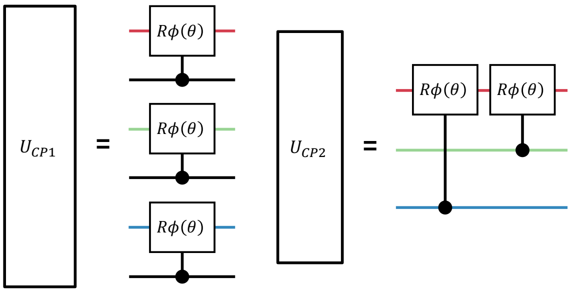

Figure 5: The structure of the U_{CP1} and U_{CP2} unitary operators, each comprising a distinct quantum pooling layer implemented with $CP(\theta)$.

while other qubits are disregarded to emulate classical pooling functionality.

The second pooling layer uses the unitary operator U_{CP2}, as shown in Figure 5 (b), where two parameterized controlled gates compress the features of the blue and green channels into the red channel.

$$CP(\theta) = CR_\phi(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{bmatrix} \quad (7)$$

3.1.4. Measurement

The quantum state is processed by a parameterized unitary $U_{Ansatz}(\theta)$ and an angle encoding unitary $U_E(\theta)$, acting on six qubits. The circuit consists of angle encoding, two quanvolutional layers, and two pooling layers. Finally, the Pauli-Z expectation value is computed for qubit q_0, resulting in

$$v^{(i)} = \langle 0|U_E^\dagger(x_i)U_{Ansatz}(\theta)^\dagger Z^0 U_{Ansatz}(\theta)U_E(x_i)|0\rangle, \quad (8)$$

where Z^0 denotes the Pauli-Z operator on q_0, $U_{Ansatz}(\theta)$ encompasses all quanvolutional and pooling operations, and x denotes the 12 pixels extracted by the quanvolutional kernel, with i indicating the index, corresponding to the feature map ($v^{(i)}$) as shown in Figure 2. The expectation value $v^{(i)}$ is calculated using the reduced density matrix of q_0,

$$\rho_0 = Tr_{1,2,3,4,5}(|\psi_i\rangle\langle\psi_i|) \quad (9)$$

with $|\psi_i\rangle = U_{Ansatz}(\theta)U_E(x_i)|0\rangle^{\otimes 6}$, thus $v^{(i)} = Tr_i(\rho_0 Z)$. This expectation value represents the feature encoded in q_0, serving as the output for cross-channel pixel characteristics in quantum image processing tasks.

3.2. Classical component

The classical component of the HQCNN model primarily consists of fully connected layers and classical optimization algorithms.

After processing through the QCNN, each image generates a feature map derived from the Pauli-Z expectation values of the output qubits, as depicted in Figure 2. This feature map is flattened into a one-dimensional vector, normalized by a BatchNorm1D layer, and passed to the first fully connected classical layer (FC1). A leaky rectified linear unit (LeakyReLU) activation function [42] is applied to ensure effective backpropagation, even for negative inputs. The output is then fed into the second fully connected layer (FC2), with its output size matching the number of classes in the classification task. A softmax function [43] converts the FC2 outputs into a probability distribution. Finally, the cross-entropy loss function [44] computes the error between predicted and true labels, enabling backpropagation to update the parameters of both quantum and classical components.

3.3. Simple CNN

To evaluate the performance of the HQCNN model, a simple classical CNN model is introduced as a baseline for comparison. This classical model adopts a structure similar to that of the HQCNN, as shown in Figure 6.

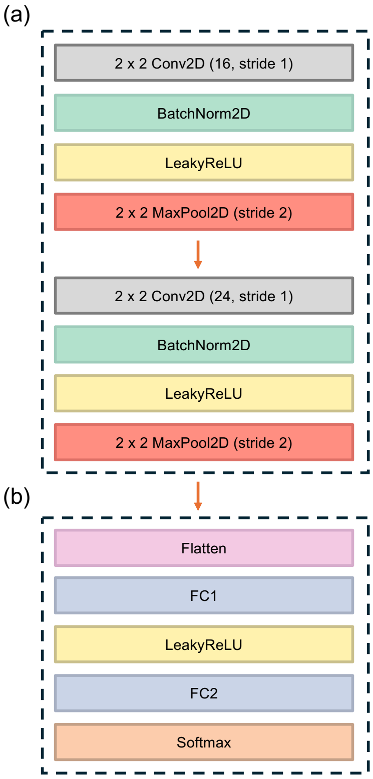
(a)
2 x 2 Conv2D (16, stride 1)
BatchNorm2D
LeakyReLU
2 x 2 MaxPool2D (stride 2)
2 x 2 Conv2D (24, stride 1)
BatchNorm2D
LeakyReLU
2 x 2 MaxPool2D (stride 2)
(b)
Flatten
FC1
LeakyReLU
FC2
Softmax

Figure 6: (a) Two convolutional and pooling layers. (b) Similar to the classical part of the HQCNN, omitting the BatchNorm1D layer.

3.4. Model scale comparison

The quantum circuit complexity of the HQCNN is analyzed using the gate set $\mathcal{G} = \{CNOT, R_x, R_y, R_z\}$ as we consider. The QCNN decomposition results in 48 CNOT gates, a circuit depth of 20, and 27 trainable parameters. For binary classification tasks, the classical component of the HQCNN includes 2852 trainable parameters, compared to 2762 parameters for the simple CNN. In the ten-class classification task, only the output dimension of the FC2 differs, adding 264 parameters to the binary classification baseline. Consequently, the HQCNN has 117 more parameters than the classical CNN.

4. Experiments

4.1. Datasets

To evaluate the performance of the HQCNN and the classical CNN, four datasets are selected: CIFAR-10 [45], EuroSAT [46], and SAT-4 [47] for binary classification tasks, and MNIST [48] for a ten-class classification task. CIFAR-10 comprises complex color images that challenge the expressive capabilities of models. EuroSAT and SAT-4, both satellite image datasets, assess the generalization in domain-specific tasks. MNIST, due to its simplicity and widespread use, serves as a benchmark for evaluating basic deep learning models. Compared to the large-scale ImageNet dataset, these smaller datasets are better suited for evaluating lightweight HQCNN and CNN architectures, which face challenges with ImageNet's computational and data complexity. The single-channel grayscale images of MNIST are transformed into three-channel images by replicating the grayscale channel to align with the three-channel format of color datasets.

All datasets are preprocessed consistently, with images resized to $(10\times10\times3)$ $(\mathcal{H}\times\mathcal{W}\times C)$ using bilinear interpolation, and RGB images are converted into Lab, YCrCb, and HSV color spaces using OpenCV [49]. To address varying color space ranges, the L channel of Lab, the Y channel of YCrCb, HSV, and RGB are normalized to $[0, 1]$, while the a and b channels of Lab and the Cr and Cb channels of YCrCb are normalized to $[-1, 1]$. Each dataset is randomly sampled without replacement and split into a training set of 500 images, a validation set of 100 images, and a test set of 100 images.

4.2. Experimental setup

A quantum convolutional kernel of size 2×2 is applied across all three channels of the input image with a stride of 1, producing an output of size $2 \times 2 \times 3$. The HQCNN and CNN models are trained using the AdamW optimizer [50] to update parameters, with the ReduceLROnPlateau scheduler employed for dynamic learning rate adjustment. The initial learning rate is set to 0.01, reduced to 10% of its current value (with a minimum of 10^{-4}) if the validation loss does not decrease for three consecutive epochs over a total of 20 epochs. In the classical component of both models, the FC1 has an output size of 32, while the FC2 matches the number of classes in the classification task. The QCNN is constructed and simulated using the PennyLane framework [51], while the HQCNN and CNN models are implemented and trained using the PyTorch library [52].

Parameters of both quantum and classical components are optimized via the backpropagation algorithm, implemented in PyTorch to compute gradients of the cross-entropy loss function, enabling optimization through gradient descent. In HQCNN simulation, the backpropagation method provides faster gradient computation compared to the parameter shift rule [53], which is typically required for quantum hardware. Due to hardware limitations, direct application of backpropagation is not feasible on quantum hardware, necessitating the use of the parameter shift rule for gradient computation. All quantum simulations and computations for both HQCNN and classical CNN models on an Apple M4 CPU.

4.3. Evaluation metrics

For both binary and ten-class classification tasks, the performance of the HQCNN and classical CNN models is assessed using macro-averaged precision, recall, F1 score, and accuracy. These metrics are computed as arithmetic averages across all classes to account for balanced class distributions in the selected datasets. The definitions of the metrics are as follows:

- True Positives (TP_i): The number of correctly predicted instances for class i.
- False Positives (FP_i): The number of instances incorrectly predicted as class i.
- False Negatives (FN_i): The number of instances of class i incorrectly predicted as another class.

The precision (P_i) and recall (R_i) for each class i are defined as:

$$P_i = \frac{TP_i}{TP_i + FP_i}, \tag{10}$$

$$R_i = \frac{TP_i}{TP_i + FN_i}. \tag{11}$$

Given the balanced class sizes in the datasets, accuracy is calculated as the proportion of correctly classified instances across all samples, where N denotes the total number of samples:

$$Accuracy = \frac{1}{N}\sum_{i=0}^{N-1} TP_i. \tag{12}$$

Macro-averaged metrics are computed as follows, where n represents the number of classes:

$$Macro - Precision = \frac{1}{n}\sum_{i=1}^{n} P_i, \tag{13}$$

$$Macro - Recall = \frac{1}{n}\sum_{i=1}^{n} R_i, \tag{14}$$

$$Macro - F1 = \frac{1}{n}\sum_{i=1}^{n} 2 \cdot \frac{P_i \cdot R_i}{P_i + R_i}. \tag{15}$$

These metrics provide a comprehensive evaluation of model performance across different classification tasks.

4.4. CIFAR-10

Two groups of classes, C_1 (airplane-frog) and C_2 (frog-ship) are selected from the CIFAR-10 dataset to evaluate the performance of the CNN and HQCNN models. As shown in Table 1, the CNN model achieves an accuracy of 95% in the YCrCb on the C_1 test set, outperforming other color spaces. For C_2, accuracies of 96% are obtained in both Lab and YCrCb, surpassing the RGB performance of 94%. In the HQCNN model, the default encoding employs the R_{XZ} operator. When the U_{OA} operator employs the R_x gate, accuracies of 94% are achieved in both RGB and Lab for C_1, outperforming configurations using R_y and R_z gates. For C_2, the HQCNN in HSV achieves the highest accuracy of 97%, compared to 94.5% in RGB and the CNN accuracies across all four color spaces.

Average accuracy and standard deviation from five training runs, as shown in Table 2. For the CNN, an average accuracy of 94.1% is achieved in the YCrCb color space for the C_1 task, outperforming RGB 91.7%. In C_2 task, the CNN achieves an accuracy of 94.5% in Lab and YCrCb, while RGB reaches 93.3% and HSV 92.5%. For the HQCNN, no encoding combination surpasses the highest CNN accuracy. However, the U_{OA} operator with the R_x gate ($U_E : R_{XZ}$) in the Lab color space achieves an average accuracy of 93% for C_1, outperforming HQCNN's best RGB performance of 90.3% ($R_{XY} - R_y$). Similar accuracies are obtained with the R_x and R_y gates, while the R_z gate, though generally less effective, remains competitive in HSV.

4.5. EuroSAT

Performance metrics for the EuroSAT dataset, as shown in Table 3. For the E_1 (annualcrop-fores) classification task, comparable performance is observed between the CNN and HQCNN models For the E_2 (industrial-river) task, the HQCNN using R_x gate achieves an accuracy of 98.5% in Lab, while HSV reaches 97.5%, matching the CNN's HSV performance and surpassing RGB at 96%.

Average accuracy are summarized in Table 4. The CNN model attains an average accuracy of 98.9% in Lab for E_1 and 98.6% in YCrCb for E_2, both surpassing RGB's 98.5%. In the HQCNN model, the $R_{YZ} - R_x$ configuration achieves an accuracy of 99% in RGB for E_1, outperforming the CNN. The R_{XY}-R_y configuration in YCrCb achieves the highest accuracy of 99.6% for E_1. For E_2, the HQCNN does not outperform the CNN. However, the U_E-U_{OA} configurations in YCrCb and Lab generally outperform RGB for E_1, while YCrCb and HSV show higher average performance than RGB in E_2 with the R_{XY} encoding.

Table 1: The best performance metrics results (%) of the CNN and HQCNN models on the CIFAR-10 dataset for two binary classification tasks.

Model	U_{OA}		Dataset	CIFAR-10						
			C_1^{RGB}	C_1^{Lab}	C_1^{YCrCb}	C_1^{HSV}	C_2^{RGB}	C_2^{Lab}	C_2^{YCrCb}	C_2^{HSV}
CNN		Acc.	93.5	94.5	**95.0**	92.0	94.0	**96.0**	**96.0**	94.5
		F1	93.4	94.4	95.0	91.9	93.9	95.9	96.0	94.4
		Rec.	93.5	94.5	95.0	92.0	94.0	96.0	96.0	94.5
		Pre.	93.5	94.5	95.0	92.1	94.0	96.0	96.0	94.5
HQCNN	R_x	Acc.	**94.0**	**94.0**	92.5	93.0	94.5	94.0	95.0	**97.0**
		F1	93.9	93.9	92.4	92.9	94.4	93.9	94.9	96.9
		Rec.	94.0	94.0	92.5	93.0	94.5	94.0	95.0	97.0
		Pre.	94.1	94.1	92.8	93.0	94.8	94.1	95.1	97.0
	R_y	Acc.	91.0	91.5	**93.0**	91.0	94.5	**95.5**	95.0	93.0
		F1	90.9	91.4	92.9	90.9	94.4	95.4	94.9	92.9
		Rec.	91.0	91.5	93.0	91.0	94.5	95.5	95.0	93.0
		Pre.	91.0	91.8	93.0	91.0	94.5	95.5	95.0	93.0
	R_z	Acc.	87.0	86.0	88.0	**91.5**	91.5	88.5	**92.5**	91.5
		F1	87.0	85.9	87.9	91.4	91.4	88.4	92.4	91.4
		Rec.	87.0	86.0	88.0	91.5	91.5	88.5	92.5	91.5
		Pre.	87.0	86.2	88.1	91.5	91.8	89.1	92.5	91.6

Table 2: The average accuracy and standard deviation of CNN and HQCNN models across five training runs on the CIFAR-10 test set for two binary classification tasks, comparing performance across RGB, Lab, YCrCb, and HSV color spaces and HQCNN configurations.

Model	U_E	U_{OA}	Dataset	CIFAR-10						
			C_1^{RGB}	C_1^{Lab}	C_1^{YCrCb}	C_1^{HSV}	C_2^{RGB}	C_2^{Lab}	C_2^{YCrCb}	C_2^{HSV}
CNN			91.7 ± 1.5	93.0 ± 0.9	**94.1 ± 0.9**	91.1 ± 0.9	93.3 ± 0.5	**94.5 ± 0.9**	94.5 ± 1.0	92.5 ± 1.3
HQCNN	R_{YZ}	R_x	89.6 ± 2.0	89.9 ± 1.9	**91.1 ± 3.0**	89.7 ± 1.4	92.1 ± 1.9	91.1 ± 1.6	**93.1 ± 2.0**	90.8 ± 2.2
		R_y	90.4 ± 1.9	**91.7 ± 1.4**	90.7 ± 1.2	90.1 ± 1.5	91.6 ± 4.9	92.1 ± 1.0	**92.5 ± 1.3**	91.2 ± 2.5
		R_z	86.9 ± 4.0	81.1 ± 3.5	85.4 ± 2.4	**88.5 ± 1.9**	**90.0 ± 1.3**	82.8 ± 0.9	87.4 ± 1.4	89.7 ± 3.0
	R_{XZ}	R_x	89.6 ± 3.0	**93.0 ± 0.7**	90.4 ± 1.8	89.4 ± 2.5	91.3 ± 2.4	**92.9 ± 1.0**	92.5 ± 1.5	92.8 ± 2.2
		R_y	88.6 ± 1.7	89.2 ± 1.4	**91.4 ± 1.0**	89.9 ± 1.0	93.1 ± 1.1	**93.3 ± 1.5**	92.7 ± 1.6	91.2 ± 1.3
		R_z	83.6 ± 3.8	81.7 ± 2.4	85.2 ± 2.3	**88.9 ± 1.4**	88.6 ± 4.5	85.8 ± 3.4	87.2 ± 3.0	**89.9 ± 2.2**
	R_{XY}	R_x	90.2 ± 2.0	**91.6 ± 1.8**	89.5 ± 1.6	88.9 ± 1.8	91.2 ± 2.1	92.1 ± 1.1	**92.3 ± 1.7**	89.2 ± 1.2
		R_y	90.3 ± 1.5	**91.6 ± 1.5**	91.0 ± 1.3	90.5 ± 1.5	**93.2 ± 2.0**	92.9 ± 1.5	92.4 ± 1.2	90.8 ± 2.6
		R_z	86.7 ± 2.3	80.9 ± 3.7	84.8 ± 1.5	**88.1 ± 2.0**	87.8 ± 5.2	82.6 ± 3.3	88.9 ± 1.5	**90.6 ± 1.3**

4.6. SAT-4

Similar to the EuroSAT dataset, both the CNN and HQCNN models achieve high accuracy in the SAT-4 dataset for the S_1 (barrenland-grassland) and S_2 (barrenland-trees) classification tasks. For the CNN model, an accuracy of 98% is achieved in Lab for S_1. In the HQCNN model, the R_y gate achieves an accuracy of 98.5% in both RGB and YCrCb, surpassing the CNN accuracies of 95% and 97.5% in these color spaces. Both models exhibit comparable performance across various color spaces in the S_2 task, as shown in Table 5.

Average performance as shown in Table 6. In S_1, the CNN model achieves an average accuracy of 96.7% in the Lab and HSV, while the HQCNN model achieves accuracies of 96.3% ($R_{YZ} - R_x$) in RGB and 97.2% ($R_{XZ} - R_y$) in YCrCb, compared to the CNN's corresponding results. In S_2, the CNN model achieves an average accuracy of 99.7% in the Lab. The HQCNN, across various configurations, achieves an average accuracy of 99.8% in each color space, outperforming the CNN. Unlike in the CIFAR-10 dataset, the R_z gate does not exhibit significant performance differences compared to other rotation gates in these tasks.

4.7. MNIST

To ensure consistency with color datasets, a three-channel version of the MNIST dataset is employed for both training and testing, unlike prior studies that utilized single-channel grayscale images. The results are shown in Table 7. The CNN model achieves an average accuracy of 91.1% in RGB, with a best accuracy of

Table 3: The best accuracy (%) from five training runs on the EuroSAT dataset.

Model	U_{OA}		Dataset			EuroSAT				
			E_1^{RGB}	E_1^{Lab}	E_1^{YCrCb}	E_1^{HSV}	E_2^{RGB}	E_2^{Lab}	E_2^{YCrCb}	E_2^{HSV}
CNN		Acc.	**99.5**	**99.5**	99.0	**99.5**	98.5	**99.5**	**99.5**	97.5
		F1	99.4	99.4	98.9	99.4	98.4	99.4	99.4	97.4
		Rec.	99.5	99.5	99.0	99.5	98.5	99.5	99.5	97.5
		Pre.	99.5	99.5	99.0	99.5	98.5	99.5	99.5	97.5
HQCNN	R_x	Acc.	99.0	99.5	**100**	99.5	96.0	**98.5**	96.5	97.5
		F1	99.0	99.4	100	99.4	96.0	98.4	96.4	97.4
		Rec.	99.0	99.5	100	99.5	96.0	98.5	96.5	97.5
		Pre.	99.0	99.5	100	99.5	96.0	98.5	96.5	97.5
	R_y	Acc.	**99.5**	**99.5**	99.0	99.0	97.0	**97.5**	**97.5**	96.5
		F1	99.4	99.4	98.9	98.9	96.9	97.4	97.4	96.4
		Rec.	99.5	99.5	99.0	99.0	97.0	97.5	97.5	96.5
		Pre.	99.5	99.5	99.0	99.0	97.0	97.5	97.5	96.5
	R_z	Acc.	**99.5**	99.0	99.0	99.0	**97.5**	96.5	96.5	96.5
		F1	99.4	98.9	99.0	98.9	97.4	96.4	96.4	96.4
		Rec.	99.5	99.0	99.0	99.0	97.5	96.5	96.5	96.5
		Pre.	99.5	99.0	99.0	99.0	97.5	96.5	96.5	96.5

Table 4: The average and standard deviation of results (%) from five test set evaluations.

Model	U_E	U_{OA}	Dataset			EuroSAT				
			E_1^{RGB}	E_1^{Lab}	E_1^{YCrCb}	E_1^{HSV}	E_2^{RGB}	E_2^{Lab}	E_2^{YCrCb}	E_2^{HSV}
CNN			98.5 ± 1.0	**98.9** ± **0.5**	98.6 ± 0.3	98.4 ± 0.7	97.8 ± 0.7	97.9 ± 0.8	**98.6** ± **0.4**	96.9 ± 0.3
HQCNN	R_{YZ}	R_x	**99.0** ± **0.4**	98.6 ± 0.7	98.8 ± 0.5	98.0 ± 0.5	94.7 ± 0.8	**95.3** ± **0.9**	95.1 ± 0.8	94.8 ± 1.9
		R_y	98.2 ± 0.6	**99.1** ± **0.5**	98.6 ± 0.5	97.5 ± 1.5	94.1 ± 0.2	94.9 ± 0.9	94.6 ± 1.5	**95.6** ± **1.1**
		R_z	98.5 ± 1.1	99.0 ± 1.0	98.6 ± 1.1	**99.0** ± **0.6**	**94.5** ± **1.3**	94.0 ± 1.4	94.0 ± 1.1	94.2 ± 1.9
	R_{XZ}	R_x	98.2 ± 0.8	98.8 ± 0.9	**99.3** ± **0.6**	99.0 ± 0.5	94.6 ± 1.2	95.5 ± 1.7	**95.8** ± **0.5**	95.4 ± 1.3
		R_y	98.6 ± 1.1	**98.8** ± **0.4**	98.7 ± 0.2	97.5 ± 1.5	95.0 ± 1.1	94.9 ± 1.4	94.8 ± 1.8	**95.3** ± **1.0**
		R_z	98.3 ± 0.8	**98.4** ± **0.3**	**98.4** ± **0.3**	98.3 ± 0.6	**96.6** ± **0.6**	93.8 ± 1.6	95.2 ± 0.8	95.0 ± 1.0
	R_{XY}	R_x	98.3 ± 0.7	**98.6** ± **0.6**	98.1 ± 0.5	97.9 ± 1.2	94.0 ± 0.9	93.3 ± 1.1	94.6 ± 1.9	**95.5** ± **0.8**
		R_y	98.3 ± 0.7	98.2 ± 0.9	**99.6** ± **0.2**	98.6 ± 0.8	93.7 ± 1.6	93.4 ± 0.9	**95.7** ± **0.6**	95.7 ± 1.6
		R_z	98.2 ± 1.2	98.7 ± 0.5	**98.8** ± **0.5**	98.0 ± 1.0	93.6 ± 0.5	94.7 ± 0.8	94.6 ± 1.8	**95.5** ± **1.1**

92.8%. The HQCNN model achieves the highest average accuracy of 93.4% in Lab, with a best accuracy of 94.3%, and 93.9% in RGB. Furthermore, the HQCNN model consistently outperforms the CNN across all four color spaces evaluated.

Notably, the HQCNN demonstrates faster convergence to optimal performance, achieving stable accuracy and minimal loss within fewer epochs compared to the CNN, as shown in Figure 7.

5. Conclusion

This study presents an HQCNN model that integrates quantum and classical components to enhance image classification performance across multiple color spaces. The HQCNN demonstrates superior performance compared to a classical simple CNN baseline in several tasks, particularly in the ten-class classification task on a three-channel MNIST dataset, where it achieves the best accuracy of 94.3% in the Lab color space, outperforming the CNN's 92.8% in RGB. Additionally, the HQCNN exhibits faster convergence, stabilizing within approximately 10 epochs of a 20-epoch training period, suggesting that quantum circuits contribute to improved efficiency in multi-channel image classification.

Across the CIFAR-10, EuroSAT, and SAT-4 datasets, the HQCNN consistently achieves competitive or superior performance in non-RGB color spaces such as Lab, YCrCb, and HSV. For instance, in the CIFAR-10 C_2 task, the HQCNN attains 97% accuracy in HSV, surpassing the CNN's 94.5%. In the EuroSAT E_1 task, the HQCNN reaches an average accuracy of 99.6% in YCrCb with the R_{XY}-R_y configuration, outperforming the CNN's 98.5% in RGB. Similarly, in the SAT-4 S_1

Table 5: The highest accuracy (%) was achieved across five training sessions on the SAT-4 dataset.

Model	U_{OA}		Dataset			SAT-4				
			$\mathrm{S}_1^{\mathrm{RGB}}$	$\mathrm{S}_1^{\mathrm{Lab}}$	$\mathrm{S}_1^{\mathrm{YCrCb}}$	$\mathrm{S}_1^{\mathrm{HSV}}$	$\mathrm{S}_2^{\mathrm{RGB}}$	$\mathrm{S}_2^{\mathrm{Lab}}$	$\mathrm{S}_2^{\mathrm{YCrCb}}$	$\mathrm{S}_2^{\mathrm{HSV}}$
CNN		Acc.	95.0	**98.0**	97.5	97.5	100	100	100	100
		F1	94.9	97.9	97.4	97.4	100	100	100	100
		Rec.	95.0	98.0	97.5	97.5	100	100	100	100
		Pre.	95.0	98.0	97.5	97.5	100	100	100	100
HQCNN	R_x	Acc.	96.5	**97.5**	96.0	96.5	100	99.5	99.5	99.5
		F1	96.4	97.4	95.9	96.4	100	99.4	99.4	99.4
		Rec.	96.5	97.5	96.0	96.5	100	99.5	99.5	99.5
		Pre.	96.6	97.5	96.0	96.5	100	99.5	99.5	99.5
	R_y	Acc.	**98.5**	98.0	**98.5**	97.0	100	100	100	100
		F1	98.4	97.9	98.4	96.9	100	100	100	100
		Rec.	98.5	98.0	98.5	97.0	100	100	100	100
		Pre.	98.5	98.0	98.5	97.0	100	100	100	100
	R_z	Acc.	95.5	95.5	**98.0**	93.5	100	100	100	99.0
		F1	95.4	95.4	97.9	93.4	100	100	100	98.9
		Rec.	95.5	95.5	98.0	93.5	100	100	100	99.0
		Pre.	95.5	95.7	98.0	93.5	100	100	100	99.0

Table 6: The average accuracy (%) and standard deviation of the results of the five test sets.

Model	U_E	U_{OA}	Dataset			SAT-4				
			$\mathrm{S}_1^{\mathrm{RGB}}$	$\mathrm{S}_1^{\mathrm{Lab}}$	$\mathrm{S}_1^{\mathrm{YCrCb}}$	$\mathrm{S}_1^{\mathrm{HSV}}$	$\mathrm{S}_2^{\mathrm{RGB}}$	$\mathrm{S}_2^{\mathrm{Lab}}$	$\mathrm{S}_2^{\mathrm{YCrCb}}$	$\mathrm{S}_2^{\mathrm{HSV}}$
CNN			94.0 ± 0.9	96.7 ± 1.0	96.1 ± 1.1	**96.7 ± 0.6**	99.4 ± 0.6	**99.7 ± 0.2**	99.4 ± 0.7	99.5 ± 0.3
HQCNN	R_{YZ}	R_x	**96.3 ± 0.9**	95.1 ± 1.5	95.9 ± 0.6	94.4 ± 2.2	99.5 ± 0.5	**99.8 ± 0.2**	99.4 ± 0.3	**99.8 ± 0.2**
		R_y	95.2 ± 1.0	94.9 ± 0.7	**96.2 ± 1.6**	96.1 ± 0.2	99.5 ± 0.3	**99.7 ± 0.2**	99.2 ± 0.2	99.7 ± 0.4
		R_z	95.1 ± 1.1	93.6 ± 0.7	**95.3 ± 2.1**	92.3 ± 1.1	**99.5 ± 0.4**	**99.5 ± 0.4**	99.1 ± 0.2	97.9 ± 1.2
	R_{XZ}	R_x	95.5 ± 1.2	**96.8 ± 0.5**	95.2 ± 0.6	95.1 ± 0.8	**99.4 ± 0.4**	99.2 ± 0.6	99.2 ± 0.4	98.6 ± 0.8
		R_y	95.9 ± 2.1	95.0 ± 2.0	**97.2 ± 0.9**	94.5 ± 1.7	**99.8 ± 0.2**	99.6 ± 0.3	99.4 ± 0.3	99.3 ± 0.9
		R_z	94.1 ± 1.2	93.5 ± 1.1	**96.1 ± 1.1**	90.4 ± 3.9	99.6 ± 0.3	99.4 ± 0.9	**99.8 ± 0.2**	97.6 ± 1.5
	R_{XY}	R_x	95.3 ± 0.8	**95.9 ± 1.2**	**95.9 ± 1.2**	94.4 ± 2.3	99.4 ± 0.4	**99.8 ± 0.2**	99.5 ± 0.4	99.3 ± 0.4
		R_y	95.2 ± 1.9	94.6 ± 0.4	**96.0 ± 1.0**	93.9 ± 0.8	99.4 ± 0.4	**99.5 ± 0.5**	99.4 ± 0.3	99.3 ± 0.6
		R_z	94.6 ± 1.1	94.4 ± 2.2	**94.9 ± 1.0**	93.0 ± 1.5	98.7 ± 0.6	99.1 ± 0.5	**99.3 ± 0.7**	98.4 ± 1.3

task, the HQCNN achieves 98.5% in RGB and YCrCb with the R_y gate, exceeding the CNN's 95% in RGB. These results suggest that non-RGB color spaces often outperform RGB, highlighting the importance of color space selection in image classification.

The experimental results demonstrate that the design of PQCs significantly influences performance. Configurations using R_x and R_y gates generally outperform those with R_z, though R_z remains competitive in specific tasks, such as those in the HSV color space. For example, the R_{XZ}-R_z configuration achieves competitive accuracy on CIFAR-10, demonstrating the expressiveness of optimized quantum circuits.

The hybrid quantum-classical approach enables exploration of diverse color space representations and their impact on classification performance. However, the current study is limited by its reliance on quantum simulations rather than NISQ hardware, which may introduce practical challenges. The varying performance of PQC designs also indicates that further optimization is needed to fully harness quantum advantages.

Future research will focus on optimizing PQC designs for hybrid systems to enhance performance. Evaluating the HQCNN on NISQ hardware will be essential to validate its practical applicability. Additionally, exploring larger datasets and other color spaces may further reveal the potential of quantum-enhanced models. These findings are expected to offer valuable insights for advancing quantum computing applications in computer vision, particularly in color image classification.

In summary, this study demonstrates the potential of the HQCNN to leverage quantum advantages in image classification tasks across various color spaces. By identifying effective color space representations and

Table 7: The performance metrics (%) for the ten-class classification task on the MNIST dataset. The HQCNN applies the R_{XZ} encoding configuration.

Model	U_{OA}		Dataset	MNIST		
			RGB	Lab	YCrCb	HSV
CNN		Acc.	**91.1 ± 1.0**	90.4 ± 0.8	90.3 ± 0.7	90.8 ± 0.5
		Best	**92.8**	91.3	91.6	91.5
		F1	92.8	91.2	91.5	91.4
		Rec.	92.8	91.3	91.6	91.5
		Pre.	92.9	91.3	91.7	91.5
HQCNN	R_x	Acc.	93.0 ± 0.8	**93.4 ± 0.6**	93.2 ± 1.1	93.0 ± 0.6
		Best	93.9	**94.3**	94.2	93.9
		F1	93.9	94.3	94.2	93.8
		Rec.	93.9	94.3	94.2	93.9
		Pre.	93.9	94.4	94.2	93.9

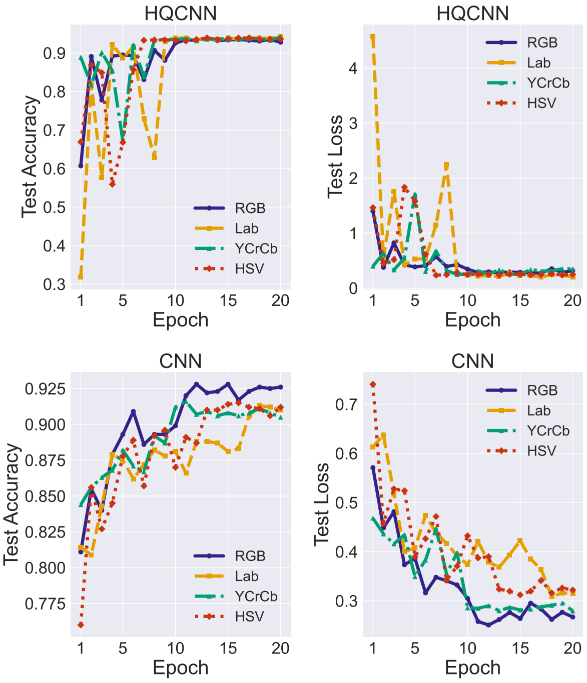

Figure 7: The best accuracy and loss curves of the HQCNN and the classical CNN on the MNIST test set. Data are sourced from Table 7.

PQC configurations, the approach highlights opportunities for quantum-enhanced models in specific application scenarios, while acknowledging the need for further development to address current limitations.

6. Acknowledgement

This work is supported by the National Natural Science Foundation of China (Grant No. 62271234), Guangdong Basic and Applied Basic Research Foundation (Grant No. 2023A1515030290), Guangzhou Basic and Applied Basic Research Foundation (Grant No. 2025A04J3585).

References

[1] A. De La Escalera, L. E. Moreno, M. A. Salichs, J. M. Armingol, Road traffic sign detection and classification, IEEE transactions on industrial electronics 44 (6) (1997) 848–859.

[2] T. Turay, T. Vladimirova, Toward performing image classification and object detection with convolutional neural networks in autonomous driving systems: A survey, IEEE Access 10 (2022) 14076–14119.

[3] Q. Song, L. Zhao, X. Luo, X. Dou, Using deep learning for classification of lung nodules on computed tomography images, Journal of healthcare engineering 2017 (1) (2017) 8314740.

[4] T. D. Pham, A comprehensive study on classification of covid-19 on computed tomography with pretrained convolutional neural networks, Scientific reports 10 (1) (2020) 16942.

[5] Z. Ren, F. Fang, N. Yan, Y. Wu, State of the art in defect detection based on machine vision, International Journal of Precision Engineering and Manufacturing-Green Technology 9 (2) (2022) 661–691.

[6] S. A. Singh, K. A. Desai, Automated surface defect detection framework using machine vision and convolutional neural networks, Journal of Intelligent Manufacturing 34 (4) (2023) 1995–2011.

[7] M. Yang, P. Kumar, J. Bhola, M. Shabaz, Development of image recognition software based

on artificial intelligence algorithm for the efficient sorting of apple fruit, International Journal of System Assurance Engineering and Management 13 (Suppl 1) (2022) 322–330.

[8] A. S. Paymode, V. B. Malode, Transfer learning for multi-crop leaf disease image classification using convolutional neural network vgg, Artificial Intelligence in Agriculture 6 (2022) 23–33.

[9] W. Rawat, Z. Wang, Deep convolutional neural networks for image classification: A comprehensive review, Neural computation 29 (9) (2017) 2352–2449.

[10] A. Voulodimos, N. Doulamis, A. Doulamis, E. Protopapadakis, Deep learning for computer vision: A brief review, Computational intelligence and neuroscience 2018 (1) (2018) 7068349.

[11] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[12] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708.

[13] R. Jozsa, N. Linden, On the role of entanglement in quantum-computational speed-up, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 459 (2036) (2003) 2011–2032.

[14] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, D. A. Spielman, Exponential algorithmic speedup by a quantum walk, in: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, 2003, pp. 59–68.

[15] R. Babbush, J. R. McClean, M. Newman, C. Gidney, S. Boixo, H. Neven, Focus beyond quadratic speedups for error-corrected quantum advantage, PRX quantum 2 (1) (2021) 010103.

[16] M. Schuld, I. Sinayskiy, F. Petruccione, The quest for a quantum neural network, Quantum Information Processing 13 (2014) 2567–2586.

[17] E. Gil-Fuster, J. Eisert, C. Bravo-Prieto, Understanding quantum machine learning also requires rethinking generalization, Nature Communications 15 (1) (2024) 2277.

[18] L. Gonon, A. Jacquier, Universal approximation theorem and error bounds for quantum neural networks and quantum reservoirs, IEEE Transactions on Neural Networks and Learning Systems (2025) 1–14doi:10.1109/TNNLS.2025.3552223.

[19] I. Cong, S. Choi, M. D. Lukin, Quantum convolutional neural networks, Nature Physics 15 (12) (2019) 1273–1278.

[20] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, P. J. Coles, Absence of barren plateaus in quantum convolutional neural networks, Physical Review X 11 (4) (2021) 041011.

[21] A. Senokosov, A. Sedykh, A. Sagingalieva, B. Kyriacou, A. Melnikov, Quantum machine learning for image classification, Machine Learning: Science and Technology 5 (1) (2024) 015040.

[22] M. Khanramaki, E. A. Asli-Ardeh, E. Kozegar, Citrus pests classification using an ensemble of deep learning models, Computers and Electronics in Agriculture 186 (2021) 106192.

[23] H. Phan-Xuan, T. Le-Tien, T. Nguyen-Chinh, T. Do-Tieu, Q. Nguyen-Van, T. Nguyen-Thanh, Preserving spatial information to enhance performance of image forgery classification, in: 2019 International Conference on Advanced Technologies for Communications (ATC), IEEE, 2019, pp. 50–55.

[24] G. Moreira, S. A. Magalhães, T. Pinho, F. N. dos Santos, M. Cunha, Benchmark of deep learning and a proposed hsv colour space models for the detection and classification of greenhouse tomato, Agronomy 12 (2) (2022) 356.

[25] S. N. Gowda, C. Yuan, Colornet: Investigating the importance of color spaces for image classification, in: Asian conference on computer vision, Springer, 2018, pp. 581–596.

[26] M. Benedetti, E. Lloyd, S. Sack, M. Fiorentini, Parameterized quantum circuits as machine learning models, Quantum Science and Technology 4 (4) (2019) 043001.

[27] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. O'brien, A variational eigenvalue solver on a photonic quantum processor, Nature communications 5 (1) (2014) 4213.

[28] E. Farhi, J. Goldstone, S. Gutmann, A quantum approximate optimization algorithm, arXiv preprint arXiv:1411.4028 (2014).

[29] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, nature 549 (7671) (2017) 242–246.

[30] Z. Wang, S. Hadfield, Z. Jiang, E. G. Rieffel, Quantum approximate optimization algorithm for maxcut: A fermionic view, Physical Review A 97 (2) (2018) 022304.

[31] M. Henderson, S. Shakya, S. Pradhan, T. Cook, Quanvolutional neural networks: powering image recognition with quantum circuits, Quantum Machine Intelligence 2 (1) (2020) 2.

[32] E. Ovalle-Magallanes, D. E. Alvarado-Carrillo, J. G. Avina-Cervantes, I. Cruz-Aceves, J. Ruiz-Pinales, Quantum angle encoding with learnable rotation applied to quantum–classical convolutional neural networks, Applied Soft Computing 141 (2023) 110307.

[33] J. Liu, K. H. Lim, K. L. Wood, W. Huang, C. Guo, H.-L. Huang, Hybrid quantum-classical convolutional neural networks, Science China Physics, Mechanics & Astronomy 64 (9) (2021) 290311.

[34] Y. Jing, X. Li, Y. Yang, C. Wu, W. Fu, W. Hu, Y. Li, H. Xu, Rgb image classification with quantum convolutional ansatz, Quantum Information Processing 21 (3) (2022) 101.

[35] F. Riaz, S. Abdulla, H. Suzuki, S. Ganguly, R. C. Deo, S. Hopkins, Accurate image multi-class classification neural network model with quantum entanglement approach, Sensors 23 (5) (2023) 2753.

[36] A. M. Smaldone, G. W. Kyro, V. S. Batista, Quantum convolutional neural networks for multichannel supervised learning, Quantum Machine Intelligence 5 (2) (2023) 41.

[37] Y.-F. Yang, M. Sun, Semiconductor defect detection by hybrid classical-quantum deep learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 2323–2332.

[38] S. Sim, P. D. Johnson, A. Aspuru-Guzik, Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms, Advanced Quantum Technologies 2 (12) (2019) 1900070.

[39] R. Kharsa, A. Bouridane, A. Amira, Advances in quantum machine learning and deep learning for image classification: a survey, Neurocomputing 560 (2023) 126843.

[40] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nature communications 12 (1) (2021) 1791.

[41] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, H. Weinfurter, Elementary gates for quantum computation, Physical review A 52 (5) (1995) 3457.

[42] A. L. Maas, A. Y. Hannun, A. Y. Ng, et al., Rectifier nonlinearities improve neural network acoustic models, in: Proc. icml, Vol. 30, Atlanta, GA, 2013, p. 3.

[43] R. J. Williams, D. Zipser, A learning algorithm for continually running fully recurrent neural networks, Neural computation 1 (2) (1989) 270–280.

[44] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning internal representations by error propagation, parallel distributed processing, explorations in the microstructure of cognition, ed. de rumelhart and j. mcclelland. vol. 1. 1986, Biometrika 71 (599-607) (1986) 6.

[45] A. Krizhevsky, G. Hinton, et al., Learning multiple layers of features from tiny images [dataset] (2009).

[46] P. Helber, B. Bischke, A. Dengel, D. Borth, Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification [dataset], IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12 (7) (2019) 2217–2226.

[47] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, R. Nemani, Deepsat: a learning framework for satellite imagery [dataset], in: Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems, 2015, pp. 1–10.

[48] L. Deng, The mnist database of handwritten digit images for machine learning research [best of the web] [dataset], IEEE signal processing magazine 29 (6) (2012) 141–142.

[49] G. Bradski, A. Kaehler, et al., Opencv, Dr. Dobb's journal of software tools 3 (2) (2000).

[50] I. Loshchilov, Decoupled weight decay regularization, arXiv preprint arXiv:1711.05101 (2017).

[51] V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. AkashNarayanan, A. Asadi, et al., Pennylane: Automatic differentiation of hybrid quantum-classical computations, arXiv preprint arXiv:1811.04968 (2018).

[52] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett (Eds.), Advances in Neural Information Processing Systems, Vol. 32, Curran Associates, Inc., 2019.

[53] D. Wierichs, J. Izaac, C. Wang, C. Y.-Y. Lin, General parameter-shift rules for quantum gradients, Quantum 6 (2022) 677.