

LEVERAGING QUANTUM-BASED ARCHITECTURES FOR ROBUST DIAGNOSTICS

Shabnam Sodagari Tommy Long

Computer Engineering & Computer Science
California State University, Long Beach, CA 90840, USA
shabnam@csulb.edu tommy.long@student.csulb.edu

ABSTRACT

The objective of this study is to diagnose and differentiate kidney stones, cysts, and tumors using Computed Tomography (CT) images of the kidney. This study leverages a hybrid quantum–classical framework in this regard. We combine a pretrained ResNet50 encoder, with a Quantum Convolutional Neural Network (QCNN) to explore quantum-assisted diagnosis. We pre-process the kidney images using denoising and contrast limited adaptive histogram equalization to enhance feature extraction. We address class imbalance through data augmentation and weighted sampling. Latent features extracted by the encoder are transformed into qubits via angle encoding and processed by a QCNN. The model is evaluated on both 8-qubit and 12-qubit configurations. Both architectures achieved rapid convergence with stable learning curves and high consistency between training and validation performance. The models reached a test accuracy of 0.99, with the 12-qubit configuration providing improvements in overall recall and precision, particularly for Cyst and Tumor detection, where it achieved perfect recall for Cysts and a tumor F1-score of 0.9956. Confusion matrix analysis further confirmed reliable classification behavior across all classes, with very few misclassifications. Results demonstrate that integrating classical pre-processing and deep feature extraction with quantum circuits enhances medical diagnostic performance.

Index Terms— biomedical imaging, quantum machine learning, quantum convolutional neural network (QCNN), medical image classification, kidney CT imaging, hybrid quantum–classical model, feature extraction, data augmentation, computed tomography.

1. INTRODUCTION

Recent advances in quantum machine learning have accelerated research on Quantum Convolutional Neural Networks (QCNNs) for medical image analysis. Various studies have demonstrated that QCNNs can enhance diagnostic accuracy and computational efficiency. One study proposed a QCNN-based classification model for brain MRI images, consisting of preprocessing and QCNN-based categorization stages [1]. The approach showed that quantum models can better capture

complex spatial relationships than traditional convolutional networks. In contrast to that work, which focused on MRI-based brain tumor detection, the present study applies a different type of QCNN targeting a different dataset, i.e., computed tomography (CT) images of the kidney, addressing a more diverse four-class problem involving normal, stone, cyst, and tumor categories.

A separate investigation introduced a hybrid QCNN for diabetic retinopathy detection, integrating quantum layers with conventional convolutional ones [2]. The model leveraged quantum superposition and entanglement to enhance feature extraction, outperforming a purely classical CNN. Moreover, in breast cancer diagnosis, a quantum–classical hybrid convolutional neural network (QCCNN) was developed [3]. The hybrid model improved robustness and generalization compared with CNN and logistic regression models. Another study investigated QCNNs for pneumonia classification from chest X-ray images under constrained feature conditions [4]. The results demonstrated the robustness of QCNNs with limited data dimensionality. In contrast, the present QCNN is applied to a four-class CT dataset requiring more sophisticated hierarchical feature extraction, supported by quantum interaction mechanisms that enable richer internal representations.

A multiqubit QCNN was also proposed for AI-driven healthcare analytics, utilizing parameterized quantum circuits for diagnostic tasks such as breast cancer, diabetes, and heart failure prediction [5]. That architecture incorporated convolutional, pooling, and interaction layers to promote inter-qubit communication. Moreover, a hybrid quantum–classical framework using quantum feature extraction was applied to autism image classification [6]. The system combined a fixed four-qubit circuit with a classical classifier, showing modest accuracy gains over classical baselines. In contrast, the QCNN used in this study uses interaction layers [7], allowing adaptive learning of quantum feature correlations specific to medical images. Additionally, a medical image classification model combining a quantum convolutional network with ResNet50 [8] was applied to the Medical MNIST dataset [9]. That model demonstrated the benefits of embedding quantum operations within a classical backbone.

Prior research consistently reports that QCNN and hybrid quantum–classical designs improve diagnostic performance

and parameter efficiency. The present work differentiates itself by employing explicit interaction layers [7] within the QCNN for complex multi-class kidney CT classification.

Thus, the main contributions of this paper include: (i) We pre-process kidney CT images from four classes—normal, stone, cyst, and tumor—to enhance image quality and enable effective latent feature extraction using the encoder of a pretrained ResNet50 autoencoder. (ii) To address dataset imbalance, we apply data augmentation and weighted sampling techniques, improving model generalization and classification consistency across minority classes. (iii) The extracted latent features are encoded into qubits via angle encoding and processed using a quantum CNN with interaction layers following the architecture of [7]. (iv) We evaluate the leveraged hybrid model on 8-qubit and 12-qubit systems, demonstrating robust classification performance.

The organization of this paper is as follows. In Section 2 we identify the problem and the approach. Section 3 contains simulation results, while Section 4 concludes the paper.

2. PROBLEM AND APPROACH

The objective of this study is to diagnose and differentiate kidney stones, cysts, and tumors using CT images of the kidney. Our dataset contains four classes: Normal with 5077 images, Tumor with 2283 images, Cyst with 3709 images, and Stone with 1377 images [10].

The model that we used was built with PyTorch and PennyLane. For the preprocessing step, we split each class into their own respective training, testing, and validation sets using scikit-learn. We set the size of the training set to 80%, with the testing and validation set being 10% each. Each class split was then instantiated as a PyTorch Dataset class object.

The PyTorch Dataset class object loads each image using OpenCV and converts it to grayscale. For our image preprocessing function we applied denoising and Contrast Limited Adaptive Histogram Equalization (CLAHE) [11]. We observed that this helps to improve feature extraction. For instance, in Fig. 1 we compare the original input image with either CLAHE or denoising or both methods applied.

We used data augmentation to help with class imbalance. Because the Tumor and Stone classes have the least number of images, the training data for both classes were given separate data augmentation, where images would undergo various geometric transformations such as `RandomZoomOut()`, `RandomResizedCrop()`, `RandomRotation()`, and `RandomHorizontalFlip()` from the PyTorch transforms library to increase dataset variety.

We passed our datasets into a PyTorch DataLoader where `WeightedRandomSampler()` was used for the training dataset to help with sampling minority classes such as Tumor and Stone. The sample size parameter in `WeightedRandomSampler()` was determined by multiplying the number of classes, 4, with an adjustable number of samples per class,

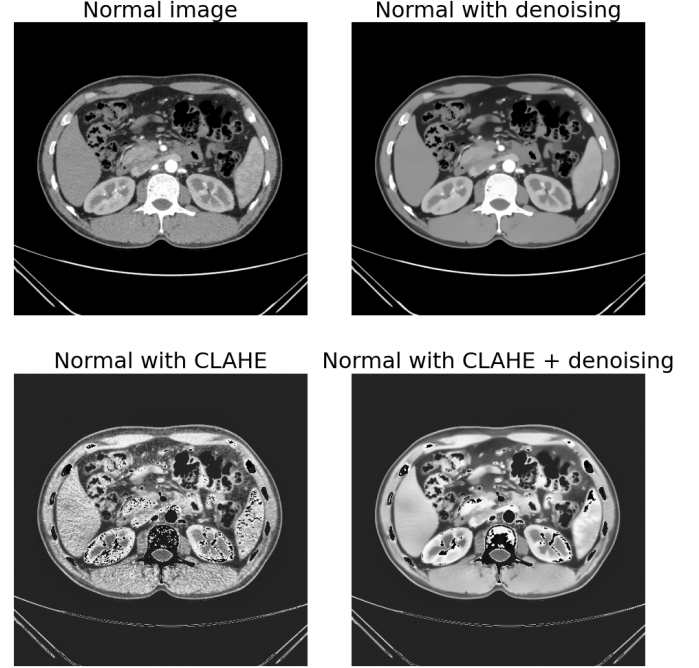


Fig. 1: An example of our pre-processing of a kidney image using denoising and CLAHE.

$n_{\max}$, to form $n_{\max} \times 4$ samples drawn per epoch with replacement. Other class data like Normal and Cyst were given a basic transform where we only applied grayscale.

We then concatenated all separate class dataset splits together to form one train, test, and validation dataset.

At the end of both transforms, we applied a ResNet50 preprocessing function. This helped prepare the images for input to an encoder part of an autoencoder where the ResNet50 pretrained model was used for feature extractions and inference of latent variables. For feature extraction we froze the ResNet50 model's weights during training, except that we adjusted the final layer of the pretrained model to have two fully connected linear layers that reduced from 2048 neurons (to ReLU) to 256 and then to a tunable number of latent variables.

We fed the latent variables that we obtained at the output of our ResNet50 encoder to an angle encoder to represent classical variables as qubits. The output of the angle encoding scheme was the input to a particular QCNN architecture with interaction [7] layers.

More specifically, after angle encoding, the qubits are passed into the first quantum convolutional layer composed of an architecture with 15 parameters utilizing $U_3(\theta, \phi, \lambda) = R_z(\phi) R_x(-\frac{\pi}{2}) R_z(\theta) R_x(\frac{\pi}{2}) R_z(\lambda)$, as well as R_y and R_z rotation gates (similar to the convolutional circuit 9 of [12]).

After two layers of the first convolutional layer (the ansatz described above), the output is fed into a quantum pooling layer [12], with parameterized Z and X rotation gates, where the number of qubits is reduced by half, keeping only the

Algorithm 1: Kidney CT Classification (12 qubits:
8 data $q_0 : q_7$ + 4 ancilla $a_0 : a_3$)

Input: Image x ; class set $\{0, 1, 2, 3\}$

Output: $\hat{y} \in \mathbb{R}^4$

- 1 **Denoise & Enhance:** grayscale $\rightarrow$ fast NLM (10, 7, 21) $\rightarrow$ CLAHE (clip= 5) +30 shift $\rightarrow$ clip to $[0, 255]$.
 - 2 **Encode & Embed:** ResNet50: $z \leftarrow \text{Enc}(x) \in \mathbb{R}^8$; rescale to $[0, \pi]^8$;
 - 3 Apply $R_Y(z_j)$ on each q_j .
 - 4 **Quantum Convolution Layer 1 (2 passes; 15 params each):** apply the convolutional circuit 9 of [12] on (q_0, q_7) , then $(q_0, q_1), (q_2, q_3), (q_4, q_5), (q_6, q_7)$, then $(q_1, q_2), (q_3, q_4), (q_5, q_6)$; repeat with new 15 parameters.
 - 5 **Quantum Pooling (2 params; [12]):** for $i \in \{0, 2, 4, 6\}$ apply $\text{CRZ}(\phi_0) - X - \text{CRX}(\phi_1)$ on $(q_{i+1} \rightarrow q_i)$; retain q_0, q_2, q_4, q_6 .
 - 6 **Quantum Interaction Layer 1 (paramless; [7]):** Toffoli cascade among $\{q_0, q_2, q_4, q_6\}$ to establish entanglement.
 - 7 **Quantum Convolution Layer 2 (2 passes; 15 params each):** on the 4 active data wires (reindexed locally 0..3), apply the same ansatz on $(0, 3), (0, 1), (2, 3), (1, 2)$; repeat with new parameters (*equivalently*: physical pairs $(0, 6), (0, 2), (4, 6), (2, 4)$).
 - 8 **Quantum Interaction Layer 2 (12 params; [7]):** on $\{q_0, q_2, q_4, q_6\}$ apply Toffoli- R_X -Toffoli- R_Y -Toffoli- R_Z .
 - 9 **Classifier Interaction (12 params; [7]):** CNOT ring on data $q_6 \rightarrow q_0 \rightarrow q_2 \rightarrow q_4 \rightarrow q_6$; for $k=0..3$: $\text{CNOT}(q_{2k} \rightarrow a_k)$ then $R_Z(\beta_{3k}) R_Y(\beta_{3k+1}) R_Z(\beta_{3k+2})$ on a_k
 - 10 Readout $\mathbf{q} = (\langle Z \rangle_{a_0}, \langle Z \rangle_{a_1}, \langle Z \rangle_{a_2}, \langle Z \rangle_{a_3})$.
 - 11 **Classical Head:** MLP($\mathbf{q}$) $\rightarrow \hat{y}$.
-

even-numbered qubits. After pooling, the remaining qubits go into the first interaction layer [7], which uses Toffoli gates to establish three-qubit interactions and create entanglement between the qubits.

A second quantum convolutional layer is then applied, the output of which is fed into the second interaction layer. In Convolutional Layer 2 (two passes), the same two-qubit ansatz used in the first convolutional layer is employed, maintaining 15 trainable parameters per pass. After pooling reduces the number of active data qubits to $\{q_0, q_2, q_4, q_6\}$, the ansatz is applied to a new connectivity pattern among these qubits—specifically, on local pairs $(0, 3), (0, 1), (2, 3)$, and $(1, 2)$ (corresponding to physical pairs $(0, 6), (0, 2), (4, 6)$, and $(2, 4)$). The second interaction layer with 12 parameters is similar to the first interaction layer, but uses R_X , R_Y , and R_Z rotation gates along with trainable parameters between Toffoli gates [7]. This layer helps the model learn more specific features of the training data. The last interaction layer takes the remaining qubits and uses CNOT gates to entangle these qubits with the ancilla qubits. The ancilla qubits represent the number of classes which is 4 for the kidney CT images under study. The 4 ancilla qubits are then passed into a classical network with 3 fully connected Linear layers with ReLU activations in between each Linear layer. Algorithm 1 summarizes the steps.

Table 1: Performance metrics comparison between 8-qubit and 12-qubit models.

Metric	8Q	12Q
Normal Precision	0.9921	0.9980
Normal Recall	0.9921	0.9902
Normal F1 Score	0.9921	0.9941
Cyst Precision	0.9893	0.9973
Cyst Recall	0.9946	1.0000
Cyst F1 Score	0.9919	0.9987
Stone Precision	0.9856	0.9716
Stone Recall	0.9928	0.9928
Stone F1 Score	0.9892	0.9821

Metric	8Q	12Q
Tumor Precision	0.9956	0.9956
Tumor Recall	0.9825	0.9956
Tumor F1 Score	0.9890	0.9956
Macro Precision	0.9906	0.9906
Micro Precision	0.9912	0.9944
Macro Recall	0.9905	0.9946
Micro Recall	0.9912	0.9944
Test Accuracy	0.99	0.99

3. SIMULATION RESULTS

We tested the method on the CT kidney dataset [10] using various hyperparameters, such as learning rate, batch size, sample size, number of qubits, etc. We used the Adam optimizer and a class-weighted cross-entropy loss. Each training epoch drew 4,000 samples, 1,000 per class across the four diagnostic categories (normal, cyst, stone, tumor), using a weighted random sampler to maintain class balance despite dataset size differences. Sampling with replacement ensured equal representation of underrepresented classes, while the choice of 1,000 samples per class balanced computational efficiency with statistical representativeness. We considered latent dimensions of 8 and 4, which, combined with 4 ancilla qubits for class encoding, yield 12- and 8-qubit systems, respectively.

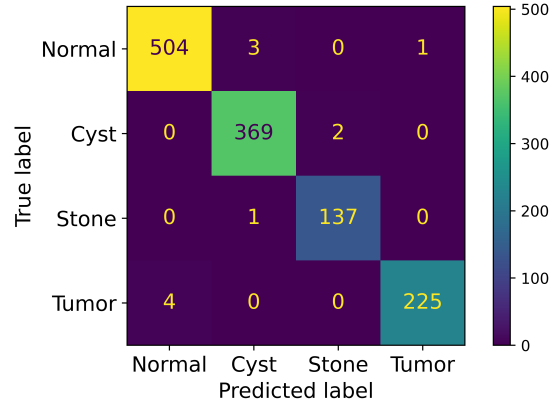
As shown in Table 1, the 12-qubit model shows stronger performance on several key metrics, particularly in Cyst and Tumor classification. It achieves perfect recall for Cyst and the highest F1-score. It also delivers higher micro-precision and micro-recall compared to the 8-qubit model.

Fig. 2 compares the confusion matrices for 8 qubits and 12 qubits both with a learning rate of 0.01, batch size of 8, and sample size of 4000.

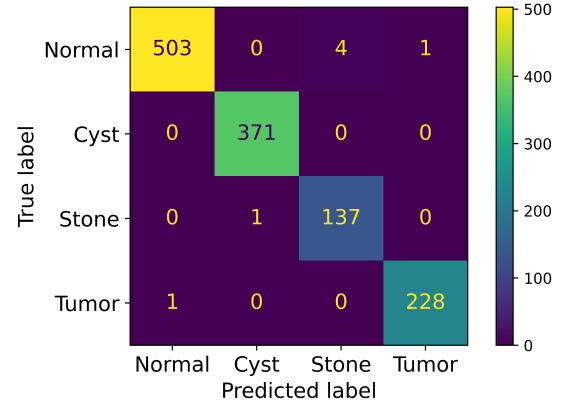
As illustrated in Fig. 2, both models show excellent accuracy for the Normal and Stone classes, with only isolated misclassifications. However, minor performance differences are observed in the Cyst and Tumor categories. The 12-qubit model achieves slightly better separation, with fewer instances of Cyst and Tumor cases being misclassified. While both configurations perform well, the 12-qubit model demonstrates higher robustness, particularly in clinically sensitive tumor differentiation.

Fig. 3 compares the training and validation accuracy and loss for the hyperparameters in Fig. 2.

As shown in Fig. 3 the loss and accuracy curves for both models demonstrate stable convergence and strong generalization. This indicates effective learning with no clear signs of overfitting. The validation accuracy closely tracks the training accuracy throughout. The 12-qubit model achieves marginally lower final loss and slightly more stable validation accuracy near convergence.

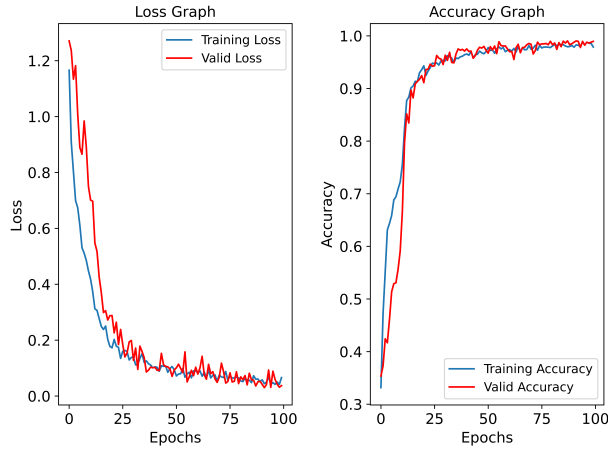


(a)

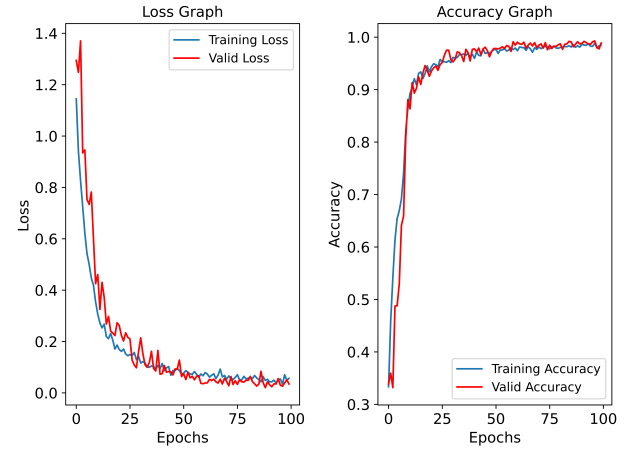


(b)

Fig. 2: Comparison of confusion matrices for (a) 8 qubits and (b) 12 qubits, both trained with a learning rate of 0.001, batch size of 8, and sample size of 4000.



(a)



(b)

Fig. 3: Comparison of training and validation accuracy and loss for (a) 8 qubits and (b) 12 qubits, both trained over 100 epochs with a learning rate of 0.001, batch size of 8, and sample size of 4000.

4. CONCLUSION

In this work, we leveraged a hybrid quantum–classical model to investigate the potential of Quantum Convolutional Neural Networks for kidney CT image diagnosis. By integrating advanced pre-processing, data augmentation, and the encoder of a pretrained autoencoder (ResNet50), we enhanced the model’s ability to extract discriminative latent features prior to quantum processing. The experimental results achieved strong and stable convergence, with closely aligned training and validation loss and accuracy curves. Each model exhibited highly reliable classification across all categories. While their overall performance was comparable, the 12-qubit model showed superior recall and F1-scores for Cyst and Tumor detection, including

perfect Cyst recall, indicating better generalization in clinically sensitive cases. These findings suggest that coupling robust classical feature extraction with quantum neural architectures provide diagnostic capability in complex medical datasets. Future research will focus on optimizing quantum circuit depth and extending to larger qubit systems.

5. REFERENCES

- [1] B. Tantawi, H. K. Ahmed, M. Magdy, M. Adel, and G. I. Sayed, “Exploring the power of quantum convolutional neural networks for brain MRI image classification,” in *2023 Intelligent Methods, Systems, and Applications (IMSA)*, 2023, pp. 549–584.

- [2] G. S. Babu, M. Lakshmi Sathwika, V. Yamuna, K. Sahithi, N. Rohith, and P. Abhishek, "Hybrid quantum convolutional neural networks for enhanced diabetic retinopathy detection," in *4th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0*. IEEE, 2025, pp. 1–6.
- [3] Q. Xiang, D. Li, Z. Hu, Y. Yuan, Y. Sun, Y. Zhu, Y. Fu, Y. Jiang, and X. Hua, "Quantum-classical hybrid convolutional neural networks for breast cancer diagnosis," *Scientific Reports*, vol. 14, p. 24699, 2024.
- [4] R. Khatoniar, V. Vyas, V. Acharya, A. Saxena, A. Saxena, R. Neiwal, and K. Korgaonkar, "Quantum convolutional neural network for medical image classification: A hybrid model," in *2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies*. IEEE, 2024, pp. 1–5.
- [5] T. B. Ovi, N. Bashree, A. S. Alam, R. Tanzim, M. A. Wahed, and H. Nyeem, "Multiqubit quantum convolutional neural networks for efficient AI-driven healthcare analytics," in *2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)*. IEEE, 2025, pp. 1–6.
- [6] M. Vimal Cruz, G. Usha, S. Dhar, K. Gautam, and S. Hajra, "Quantum image classification for sustainable medical image diagnosis," in *IEEE International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2025.
- [7] J. Mahmud, R. Mashtura, S. A. Fattah, and M. Saquib, "Quantum convolutional neural networks with interaction layers for classification of classical data," *Quantum Machine Intelligence*, vol. 6, no. 1, p. 11, 2024.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [9] E. Hassan, M. S. Hossain, A. Saber, S. Elmougy, A. Ghoneim, and G. Muhammad, "A quantum convolutional network and ResNet (50)-based classification architecture for the MNIST medical dataset," vol. 87, 2024, p. 105560. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S174680942300993X>
- [10] M. N. Islam, "CT kidney dataset: Normal-cyst-tumor and stone," 2021, [Online]. Available: <https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone>.
- [11] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics gems IV*, 1994, pp. 474–485.
- [12] T. Hur, L. Kim, and D. K. Park, "Quantum convolutional neural network for classical data classification," *Quantum Machine Intelligence*, vol. 4, no. 1, p. 3, 2022. [Online]. Available: <https://doi.org/10.1007/s42484-021-00061-x>