The impact of a quantum layer on classification performance

POZNAN UNIVERSITY OF TECHNOLOGY

Dominika Nowak 151583 [1]

[1]Computer Science, Artificial Intelligence Specialization

Abstract

This project researches the impact of integrating a quantum layer into the classical convolutional ResNet18 model for biomedical image classification. Experiments were conducted on the imbalanced BreastMNIST dataset from the MedMNIST collection[5]. I compare the performance of the standard ResNet18 with two quantum-enhanced variants: one replaces the final fully connected (FC) layer with a quantum layer followed by a linear output layer; the other inserts a quantum layer after the third residual block, followed by dimensionality reduction and classification. Results suggest that the variant with the quantum layer replacing the FC head achieves comparable or slightly improved performance. These findings highlight the potential benefits of incorporating quantum components into classical deep learning architectures for biomedical tasks.

Introduction and Related Work

Biomedical image classification remains a challenging task due to small, imbalanced datasets and the limited generalizability of classical models. Recent studies suggest that hybrid quantum-classical neural networks can enhance feature representation and optimization, even on near-term quantum hardware [1, 2, 3].

Prior work has typically focused on integrating quantum components as final classifiers. For instance, Rifat et al. [1] and Sagingalieva et al. [2] both use variational quantum circuits at the end of classical pipelines to replace the fully connected classification head. Sobrinho et al. [3], on the other hand, introduced a quanvolutional layer early in the model to transform input features prior to classical processing.

In this project, I extend these approaches by evaluating two integration strategies within a full ResNet18 architecture: replacing the final fully connected layer (Quantum-End) and inserting a quantum layer mid-network after the third residual block (Quantum-Mid). ResNet18 was selected for its top AUC performance on the BreastMNIST dataset in the original MedMNIST benchmark [5].

Data

The experiments were conducted on the BreastMNIST dataset, part of the MedMNIST v2 collection [5]. It contains 702 grayscale images (28×28 pixels) of breast ultrasound scans, labeled as malignant (0) or normal/benign (1).

Due to significant class imbalance (malignant cases make up only 27% of the training data), I undersampled normal/benign data to create a balanced training set with 147 samples per class.

- **Training set:** 147 malignant, 399 normal/benign → balanced to 147+147.
- **Validation set:** 42 malignant, 114 normal/benign.
- **Test set:** 42 malignant, 114 normal/benign.

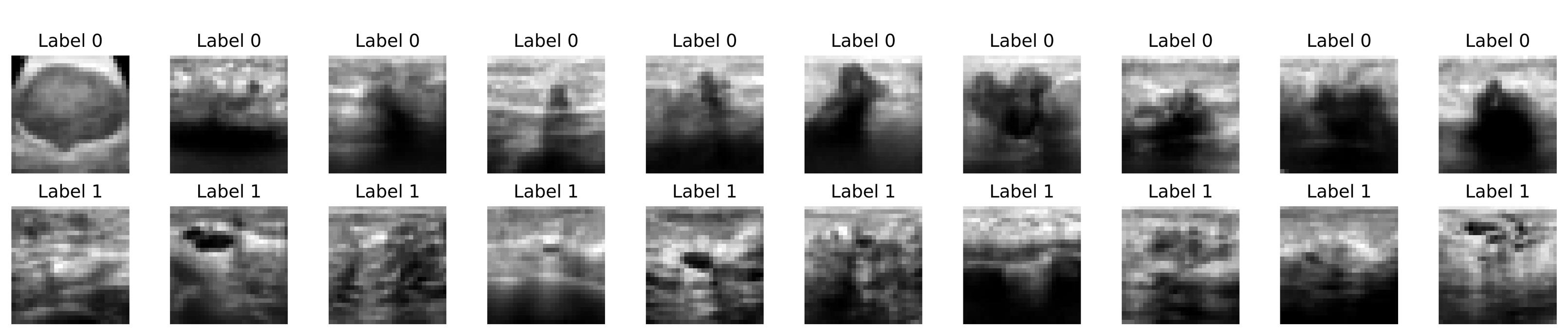

Figure 1. Examples of malignant (top) and benign (bottom) samples from BreastMNIST.

Methods

I use ResNet18 as a baseline convolutional model, modified to accept single-channel 28×28 input images. Two quantum-enhanced variants were developed for comparison:

- **Quantum-End ResNet18:** A variational quantum layer replaces the fully connected head. Features from the penultimate layer are linearly projected to match the 4-qubit input size, passed through a quantum circuit, and then classified.
- **Quantum-Mid ResNet18:** A quantum layer is inserted after the third residual block. Intermediate feature maps are pooled and projected to 4 dimensions before entering the quantum circuit.

Both quantum models use 4 qubits, angle embedding, and two layers of `BasicEntanglerLayers` implemented via PennyLane's `TorchLayer`.

All models were trained using the Adam optimizer (learning rate: $1e{-}4$ for the classical model, $1e{-}5$ for the quantum variants), with cross-entropy loss. For quantum models, early stopping was applied if the validation G-mean exceeded that of the classical ResNet18.

The implementation was based on PyTorch and PennyLane. Training was performed on a GPU for the classical model and on a CPU for the quantum-enhanced models due to quantum simulator constraints.

Experiments

To evaluate the impact of incorporating quantum layers into a classical convolutional architecture, I conducted a series of experiments comparing the three models described above.

Performance was assessed using the following metrics:

- Accuracy and Balanced Accuracy
- F1-score
- Precision and Recall
- Geometric Mean (G-mean)
- Area Under the ROC Curve (AUC)

Metric	Baseline	Quantum-End	Quantum-Mid
Accuracy	0.7436	**0.7692**	0.5897
F1-score	0.6364	**0.6604**	0.5429
Precision	0.9205	**0.9239**	0.9310
Recall	0.7105	**0.7456**	0.4737
G-mean	0.7695	**0.7883**	0.6546
AUC	**0.8590**	0.8285	0.8262

Table 1. Comparison of model performance on the BreastMNIST test set.

The original MedMNIST benchmark reported higher accuracy (0.901) and a slightly higher AUC (0.863) for ResNet18 on BreastMNIST[5]. However, these results were obtained using the full, imbalanced dataset, where class imbalance can inflate accuracy while obscuring performance on the minority class. The AUC achieved by the baseline model in this study (0.859) is comparable to the benchmark despite being trained on a smaller, balanced subset.

To compute relevant metrics such as G-mean and ensure a fair comparison across models, it was necessary to retrain and evaluate the classical ResNet18 from scratch on the same balanced dataset used for the quantum models.

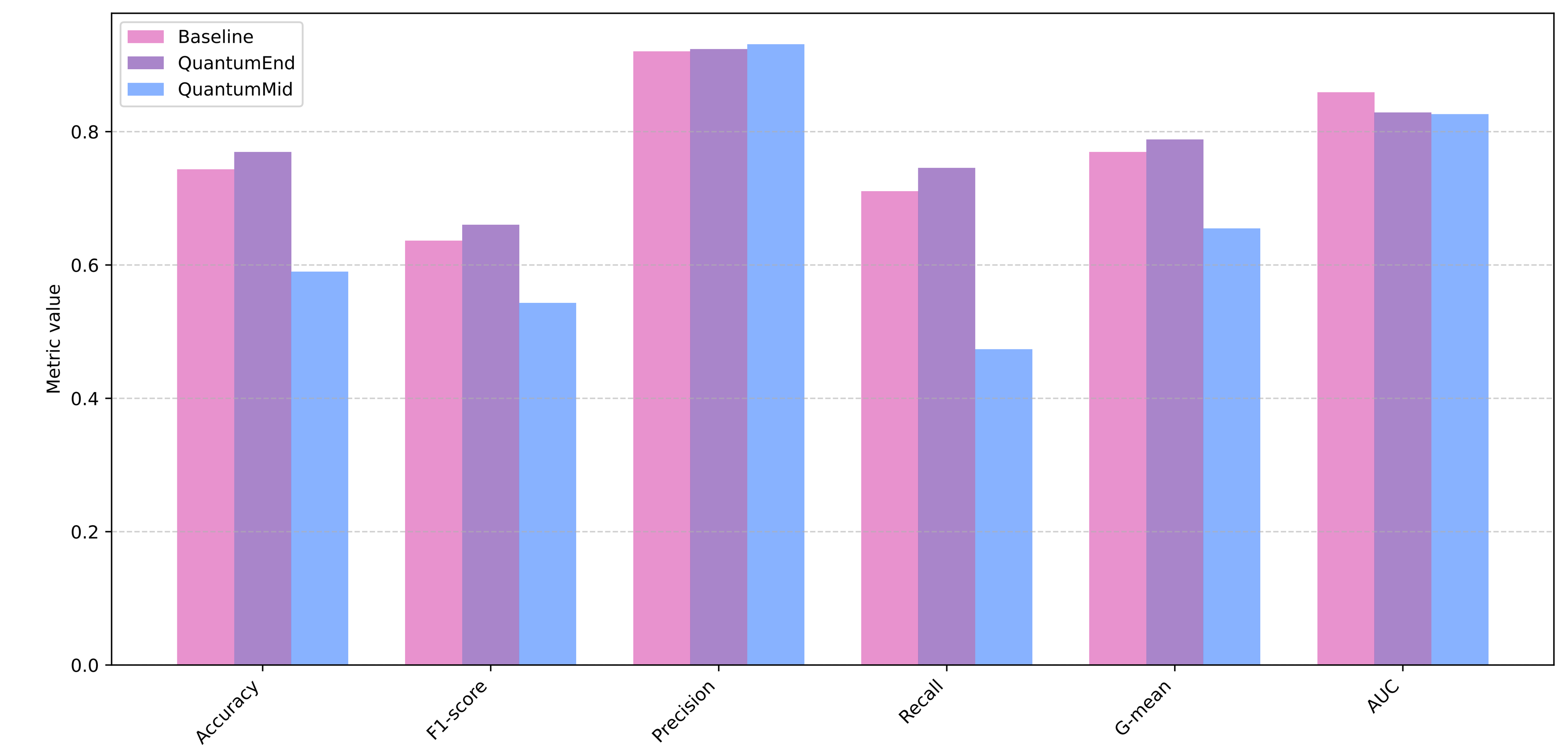

Figure 2. Comparison of model performance across metrics.

Conclusions

This study explored the integration of quantum layers into a classical ResNet18 architecture for binary biomedical image classification using the BreastMNIST dataset. Two hybrid configurations were evaluated: one replacing the final fully connected layer (Quantum-End) and another inserting a quantum layer mid-network (Quantum-Mid).

The results indicate that the Quantum-End model slightly outperformed the classical baseline in most metrics, including AUC, F1-score, and G-mean. In contrast, the Quantum-Mid model underperformed, particularly in recall and balanced accuracy, suggesting that the location of quantum integration significantly affects model performance.

These findings support the potential benefits of hybrid quantum-classical architectures in medical imaging tasks, especially when quantum components are carefully integrated at the classifier level. Future work could explore more advanced quantum circuit designs, adaptive embedding techniques, or fine-tuned hyperparameter settings to improve robustness across architectures.

References

[1] Rezuan Chowdhury Rifat, Md. Tahmid Ul Islam Tonmoy, Rifa Tasniya Aziz, and Mohammad Abdul Qayum. Enhanced hybrid quantum neural network for breast cancer detection. In *2024 27th International Conference on Computer and Information Technology (ICCIT)*, pages 1464–1469, 2024. doi: 10.1109/ICCIT64611.2024.11021989.

[2] Asel Sagingalieva, Mo Kordzanganeh, Andrii Kurkin, Artem Melnikov, Daniil Kuhmistrov, Michael Perelshtein, Alexey Melnikov, Andrea Skolik, and David Von Dollen. Hybrid quantum ResNet for car classification and its hyperparameter optimization. *Quantum Machine Intelligence*, 5(2):38, 2023. ISSN 2524-4906, 2524-4914.

[3] Yasmin Sobrinho, Enzo Soares, João Ribeiro Manesco, Jawaher Al-Tuweity, Rafael Pires, and João Papa. A hybrid quantum-classical model for breast cancer diagnosis with quanvolutions. 05 2025.

[4] Xanadu AI. Pennylane documentation – torchlayer. `https://docs.pennylane.ai/en/stable/code/api/pennylane.qnn.TorchLayer.html`. Accessed: 2025-06-20.

[5] Jiancheng Yang, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister, and Bingbing Ni. Medmnist v2-a large-scale lightweight benchmark for 2d and 3d biomedical image classification. *Scientific Data*, 10(1):41, 2023.

dominika.nowak@student.put.poznan.pl