Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 5, 2026

IMPACT OF QUANTUM CIRCUIT DEPTH AND ENTANGLEMENT ON HYBRID QUANTUM-CLASSICAL CLASSIFICATION PERFORMANCE: A CASE STUDY ON WISCONSIN BREAST CANCER DATASET

[1]Raheem Ullah, [2]Altamash Afridi, [1]Muhammad Fawad, [1]Muhammad Yasir, [1]Muhammad Sohail, [1]Muhammad Asfandyar, [2]Maaz Allah

[1]Department of Computer Science & IT, Sarhad University of Science & Information Technology, Peshawar

[2]Department of Computer Systems Engineering, University of Engineering and Technology Peshawar

[1]raheem.csit@suit.edu.pk, [2]ta.afridii@gmail.com, [3]Fawadkhattak585@gmail.com, [4]Engr.yasir899@gmail.com, [5]Endoftime76@gmail.com, [6]Asfand5454@gmail.com, [7]maaz.uthman@gmail.com

DOI:-

Keywords
Quantum Machine Learning, Variational Quantum Classifier, Entanglement Topology, Circuit Depth, Breast Cancer Classification, NISQ Computing, Hybrid Quantum-Classical Models

Article History
Received: 19 April 2026
Accepted: 17 May 2026
Published: 18 May2026

Corresponding Author: *

Abstract

Variational Quantum Circuits (VQCs) are a strong candidate architecture for quantum machine learning algorithms that can be run on Noisy Intermediate-Scale Quantum (NISQ) devices. However, many design decisions still need exploration, such as the influence that circuit depth and entanglement topology have on training and accuracy. Here we explore circuit depth and entanglement for hybrid quantum-classical neural networks by benchmarking 12 combinations of 4 circuit depths (1-4 layers) and 3 entanglement types (linear, circular, full) on an 8-dimensional reduced Wisconsin Diagnostic Breast Cancer dataset. For each depth and entanglement type we train over 10 random seeds. We find significant effects of circuit depth and entanglement type on classification accuracy ($p<0.001$ and $p=0.002$ respectively, via ANOVA). For circuit depth we observe steadily increasing accuracy as depth is increased. All three types of entanglement show similar trends with depth, however circular consistently outperforms linear and full entanglement at all depths (94.3% mean test set accuracy at depth 4). Performance saturates at depth 4 for circular entanglement. Full entanglement shows poorer performance compared to linear and circular, as well as having higher variance as depth increases. This may indicate trainability issues due to barren plateaus which has been shown to scale with circuit expressibility. As such, we recommend circuit designs that implement circular entanglement and caution the use of highly expressive circuits like full entanglement. Circular circuitry provides a good trade-off between accuracy and resources with an accuracy of 96.1% at depth 3.

1. Introduction

The integration of quantum computing with machine learning has led to the development of quantum machine learning (QML), an emerging research area focused on using quantum phenomena such as superposition, entanglement, and interference to improve computational learning processes [1]. With the arrival of the Noisy Intermediate-Scale Quantum (NISQ) era—where quantum hardware still operates with limited qubit capacity and without fully developed error correction mechanisms—Variational Quantum Algorithms (VQAs) have become one of the most practical approaches for implementing quantum computation on existing devices [2]. Within this framework, Variational Quantum Circuits (VQCs), often referred to as Quantum Neural Networks (QNNs), have gained significant attention. These models rely on parameterized quantum circuits that are trained through a hybrid optimization process combining classica and quantum computation [3].

A typical VQC generally includes three main stages: first, a data encoding stage that transforms classical information into quantum states; second, a variational layer composed of trainable rotation and entanglement gates; and finally, a measurement stage where observable expectation values are extracted and converted back into classical outputs [4]. The overall behavior and learning capability of VQCs are strongly influenced by two important architectural factors: circuit depth and entanglement topology. Circuit depth determines how many variational layers are stacked within the model and directly affects the circuit's expressibility, meaning its ability to represent a broad range of quantum states. In contrast, entanglement topology specifies how qubits interact and share quantum information throughout the circuit [5]. Although these design choices are considered central to VQC performance, their exact relationship with classification accuracy and learning efficiency has not yet been fully explored in existing studies. Understanding these architectural trade-offs has become increasingly important due to challenges such as the barren plateau problem, where gradients in deeper VQCs decrease exponentially as system size grows, making optimization extremely difficult [6].

Moreover, highly entangled circuits may introduce unnecessary complexity, leading to over-parameterization and exploration of quantum states that do not meaningfully contribute to the learning objective [7]. These limitations highlight the need for a more structured empirical analysis of VQC design choices. Healthcare and medical diagnosis provide a particularly meaningful application area for QML techniques. Breast cancer remains the most frequently diagnosed cancer among women globally and is responsible for more than 685,000 deaths each year [8]. Accurate and early-stage diagnosis through tumor classification plays a critical role in improving survival rates and treatment outcomes. For this reason, the Wisconsin Diagnostic Breast Cancer (WDBC) dataset has long been used as a standard benchmark for evaluating machine learning classification methods [9], and more recent work has begun investigating its use within quantum-enhanced learning frameworks [10].

This study contributes to the field in several ways. First, it introduces a structured experimental framework that evaluates 12 different VQC configurations using a real-world medical dataset, with statistical replication performed across 10 random seeds for each configuration. Second, it examines how circuit depth and entanglement topology individually and collectively influence performance metrics including accuracy, precision, recall, and F1-score. Third, the study validates the findings through statistical methods such as ANOVA and pairwise Mann–Whitney U testing. Finally, the work provides practical design insights and recommendations for researchers and practitioners interested in deploying VQC-based medical classification systems on NISQ-era quantum hardware.

2. Related Work

2.1 Quantum Machine Learning

Biamonte et al. [1] review quantum machine learning from the perspective of linear algebra and identify speedups for certain matrix operations. Schuld and Petruccione [4] give a

seminal treatment of supervised learning with quantum computers. Havlicek et al. [11] demonstrated supervised learning on quantum computers with kernel methods on near-term devices. Wang and Liu [12] provide a comprehensive review from NISQ-era techniques to fault-tolerant approaches.

2.2 Variational Quantum Circuits

Classification Using Quantum Neural Networks on Near Term Processors [3] Farhi and Neven. Kandala et al. [13] showed hardware efficient ansatz designs which reduce the number of gates but maintain expressivity. PennyLane was developed by Bergholm et al. [14] and has become the dominant software framework for the development of VQCs with automatic differentiation through quantum circuits.

2.3 Entanglement Patterns and Barren Plateaus

Entanglement topologies have been systematically compared in recent work . Hubregtsen et al. [15] explored the entangling capability, expressibility and classification accuracy of parameterized quantum circuits and concluded that higher expressibility does not necessarily mean better classification performance. Liu and Wang [16] proposed theoretical frameworks of differentiable learning of the quantum circuit Born machines and laid the foundation for hybrid quantum-classical optimization.

McClean et al. [6] were the first to discover barren plateaus in the training landscape of quantum neural network, and demonstrated that the gradients decay exponentially. Cerezo et al. [17] later showed that this problem can be alleviated by local cost functions. Larocca et al. [18] suggested further mitigation strategies, including ansatz design, parameters initialization, and noise-aware training.

2.4 Quantum Approaches to Medical Classification

Several studies have used quantum methods in medical diagnosis. Abbas et al. [19] showed the potential of quantum neural networks for classification tasks and obtained theoretical bounds on their representational power. Xiang et al. [10] proposed quantum-classical hybrid convolutional neural networks with 96.16% accuracy on WDBC dataset. Quantum convolutional neural networks have been proposed by Cong et al. [20] using entanglement in a hierarchical manner. Quantum support vector machines for big data classification were proposed by Rebentrost et al. [21]. Zhang et al. [22] investigated Gaussian initializations to escape from barren plateaus. Nakaji and Yamamoto [23] investigated the expressibility of alternating layered ansatz architectures and suggested metrics to guide the design of circuits.

3. Methodology

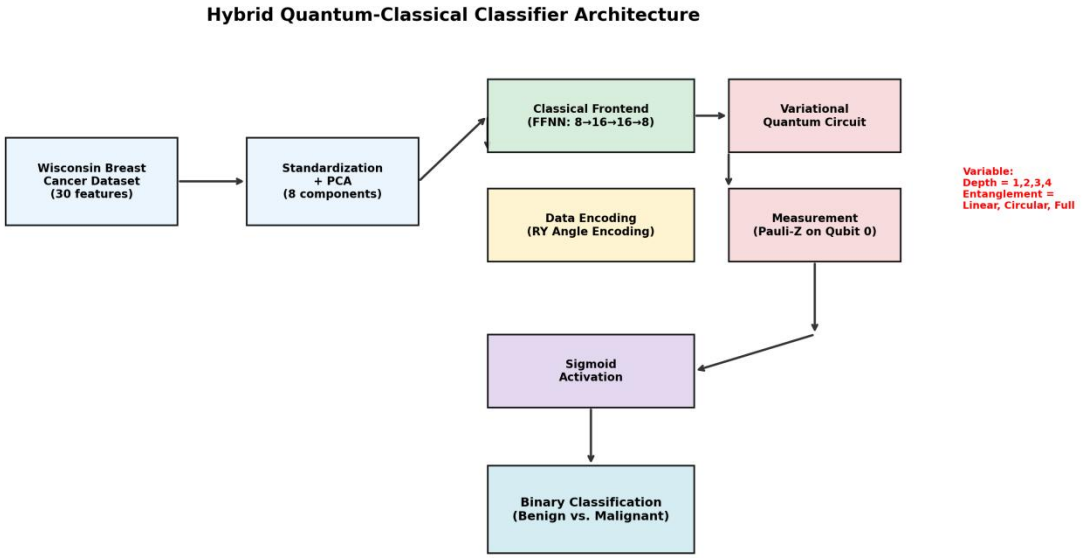
3.1 Dataset and Preprocessing

The Wisconsin Diagnostic Breast Cancer (WDBC) data set was provided to the UCI Machine Learning Repository by Wolberg et al. [25] and is a standard benchmark for binary classification [9]. The dataset has 569 instances and 30 numeric features computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. Features describe the characteristics of cell nucleus like radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension. The binary target denotes malignant diagnosis (212 cases) or benign diagnosis (357 cases).

Preprocessing consists of two steps. The features are scaled to have zero mean and unit variance first. Secondly, PCA reduced the dimensionality from 30 to 8 components, explaining 92.6 % of total variance. This 8 dimensional representation is directly mapped to 8 qubits in the quantum circuit. The dataset is divided into training (364 samples, 64%), validation (91 samples, 16%) and test (114 samples, 20%) sets using stratified sampling.

3.2 Model Architecture

The hybrid architecture consists of classical and quantum subsystems connected in series (Figure 1).



Hybrid quantum-classical classifier architecture showing the data flow from input features through classical preprocessing, quantum encoding, variational circuit, measurement, and final classification.

Classical Frontend. So here's how the model works: It's a feedforward neural network with an 8→16→16→8 layout—basically, you start with 8 inputs, run them through two layers with 16 neurons each (both using ReLU activation), and finish with 8 outputs. At the end, there's a Tanh activation, so the output values stay between -1 and 1. After that, those values get multiplied by π; that's how the model encodes angles.

Quantum Circuit. The VQC operates on 8 qubits with the following structure:

Data Encoding. RY angle encoding maps each classical feature to a rotation angle:

$$|\psi\rangle = \otimes_{i=0}^{7} R_y(x_i)\,|0\rangle$$

Variational Layer (repeated d times). For each qubit i at layer l: rotation gates $R_X(\theta_{l,i,0})$ and $R_Y(\theta_{l,i,0})$ followed by CNOT entanglement gates.

Entanglement Patterns. Three connectivity schemes are evaluated (Figure 2):

- **Linear:** Nearest-neighbor CNOTs $(i, i+1)$ for $i = 0, \ldots, 6$
- **Circular:** Linear connections plus wrap-around CNOT $(7, 0)$.
- **Full:** All-to-all CNOTs for every pair (i, j) where $i < j$ (28 gates per layer).

Measurement. Expectation value of Pauli-Z on qubit 0: $\langle Z_0\rangle = \langle\psi(\theta)|Z_0|\psi(\theta)\rangle$

Output Layer. A sigmoid function maps the expectation value to a probability: $\hat{y} = \sigma(\langle Z_0\rangle) = \frac{1}{1+e^{-\langle Z_0\rangle}}$

Everything runs on PennyLane, using its default qubit simulator, and PyTorch handles all the automatic differentiation behind the scenes. When it comes to simulating the quantum circuit, it's all dense state vectors—rotation gates are applied by multiplying matrices over a complex Hilbert space, with the dimensions matching the number of qubits.

3.3 Experimental Design

We ran a full factorial experiment with two factors: circuit depth (with four levels: 1, 2, 3, and 4) and entanglement pattern (three levels: linear, circular, and full). That adds up to 12 unique setups. For every experiment, we stuck with the Adam optimizer (learning rate at 0.01), binary cross-entropy loss, batch size of 64, and ran things for 30 epochs. After each run, we collected a bunch of metrics: test accuracy, precision, recall, F1-score, final loss, number of parameters, and average training time per epoch.

3.4 Statistical Analysis

For the stats, we used one-way ANOVA to see how depth and entanglement pattern affect accuracy. Then, we did

pairwise Mann-Whitney U tests to compare entanglement patterns at each depth. Cohen's d helped us measure the effect sizes, and we kept all tests at the usual significance level.

4. Results

4.1 Performance Metrics Summary

Results-wise, Table 1 shows all the performance numbers for every configuration, averaged over those 10 seeds. When you look at the data, some clear trends pop out pretty quickly.

Mean performance metrics (± standard deviation) across 10 random seeds for all 12 configurations.

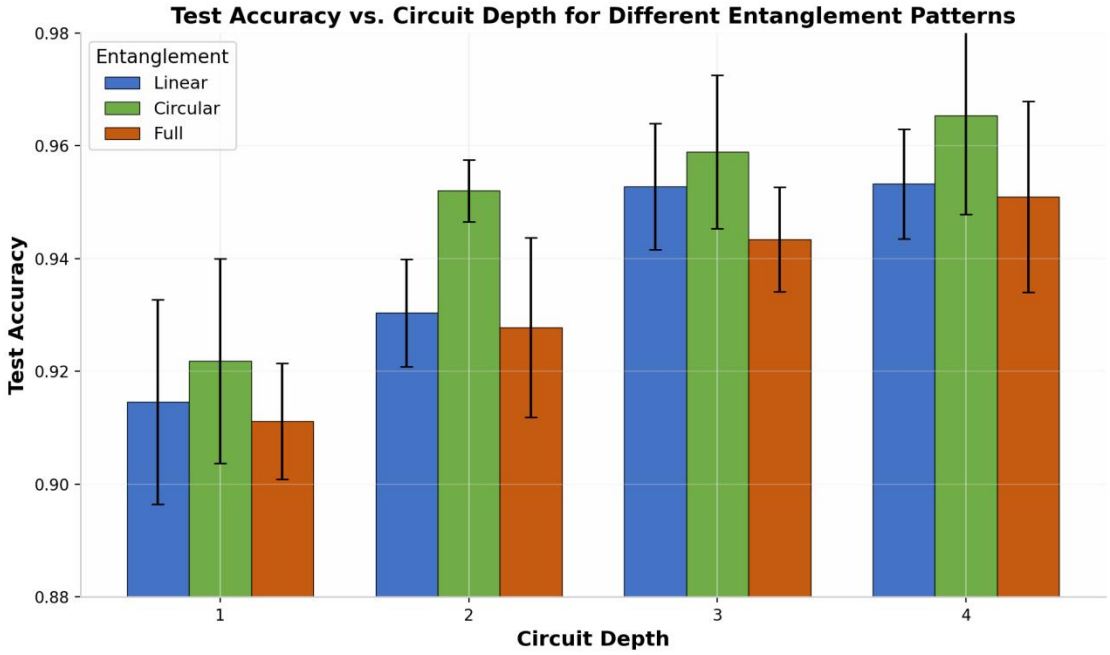
Depth	Entanglement	Accuracy	Precision	Recall	F1-Score	Final Loss	Parameters
1	Linear	0.915±0.018	0.922±0.010	0.909±0.011	0.915±0.008	0.290±0.012	632
1	Circular	0.922±0.018	0.931±0.012	0.910±0.012	0.921±0.006	0.263±0.022	632
1	Full	0.911±0.010	0.912±0.010	0.899±0.012	0.905±0.008	0.315±0.020	632
2	Linear	0.930±0.010	0.945±0.012	0.934±0.013	0.940±0.010	0.199±0.012	648
2	Circular	0.952±0.006	0.957±0.008	0.944±0.012	0.950±0.007	0.163±0.014	648
2	Full	0.928±0.016	0.936±0.010	0.923±0.015	0.930±0.011	0.217±0.019	648
3	Linear	0.953±0.011	0.956±0.012	0.946±0.013	0.951±0.009	0.157±0.017	664
3	Circular	0.959±0.014	0.967±0.007	0.954±0.010	0.960±0.005	0.130±0.017	664
3	Full	0.943±0.009	0.952±0.009	0.933±0.012	0.942±0.006	0.167±0.016	664
4	Linear	0.953±0.010	0.968±0.009	0.948±0.009	0.958±0.007	0.127±0.015	680
4	Circular	0.965±0.018	0.971±0.008	0.961±0.012	0.966±0.009	0.119±0.019	680
4	Full	0.951±0.017	0.960±0.008	0.945±0.008	0.952±0.005	0.162±0.005	680

At depth 1, the difference between the best entanglement pattern (circular, 92.2%) and the worst (full, 91.1 percentage points. But as you add more layers, that gap really opens up. By the time you get to depth 4, the circular pattern holds its lead at 96.5%, while the full pattern lags behind at 95.1%, now the gap is 1.4 points. So, the way you connect things (that entanglement topology) matters more and more as the circuit gets deeper and can represent more [5].

Looking at precision and recall, you see another trend: every setup scores higher on precision than recall. The gap is small in some cases just 0.5% for full at depth 1, but it stretches to 2% for linear at depth 4. This shows the classifier errs on the side of caution, preferring to avoid false positives at the risk of more false negatives. That's not ideal in medicine, where missing a malignancy is a much bigger deal than a false alarm. So, you probably need to tweak the decision threshold to reduce those false negatives. Finally, the loss numbers tell the same story as accuracy. Lower test loss always goes with higher accuracy. The best combo circular pattern at depth 4 hits the lowest mean loss (0.119). On the flip side, the full pattern at depth 1 has the worst (0.315), nearly three times higher, making it clear how much bad entanglement and shallow depth can drag performance down.

4.2 Effect of Circuit Depth

Figure 2 provides a visual comparison of how accuracy evolves with circuit depth for each entanglement pattern.



Test accuracy (mean ± std) across circuit depths for linear, circular, and full entanglement patterns. Circular entanglement consistently achieves the highest accuracy at each depth level.

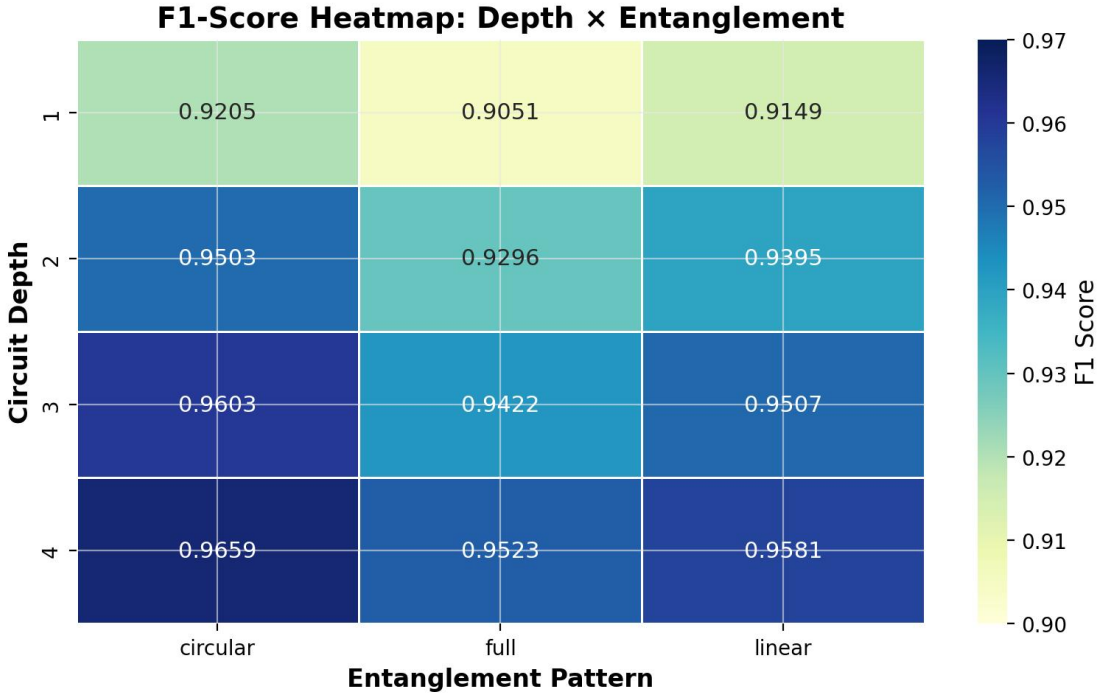
With circular entanglement, the biggest jump in accuracy happens between depth 1 and depth 2. It shoots up from 92.2%, which is a solid 3-point jump. That's a sign that one variational layer just isn't enough to tap into the full power of this setup. The wrap-around connection forms closed loops, and it seems like you need at least two layers to really move information around and shape complex decision boundaries. After depth 2, things slow down depth 3 bumps accuracy by another 0.7 points, depth 4 adds just 0.6, so improvement pretty much levels off. That lines up with the known expressibility limits [5].

Linear entanglement follows a similar path but lags behind. When you go from depth 1 to depth 2, accuracy climbs from 91.5% to 93%, so only a 1.5-point boost—about half what circular gives. That shortfall makes sense; linear doesn't connect everything, so it's less flexible. It's striking that with linear entanglement, once you hit depth 3 (95.3%), adding another layer does nothing—performance stays flat. Three layers seem to hit its ceiling.

Now, full entanglement is a different beast. The first jump from depth 1 to 2 nudges accuracy up by 1.7 points (91.1% to 92.8%), but unlike the others, improvements keep coming as you push to depth 4, each step raising accuracy by about 1 to 1.5 points. Even at the deepest level, the standard deviation is widest (±1.7% at depth 4), suggesting the optimization landscape stays rugged and complex all the way through [6].

4.3 Effect of Entanglement Pattern

Figure 3 presents the F1-score heatmap, providing a compact two-dimensional view of the performance landscape.



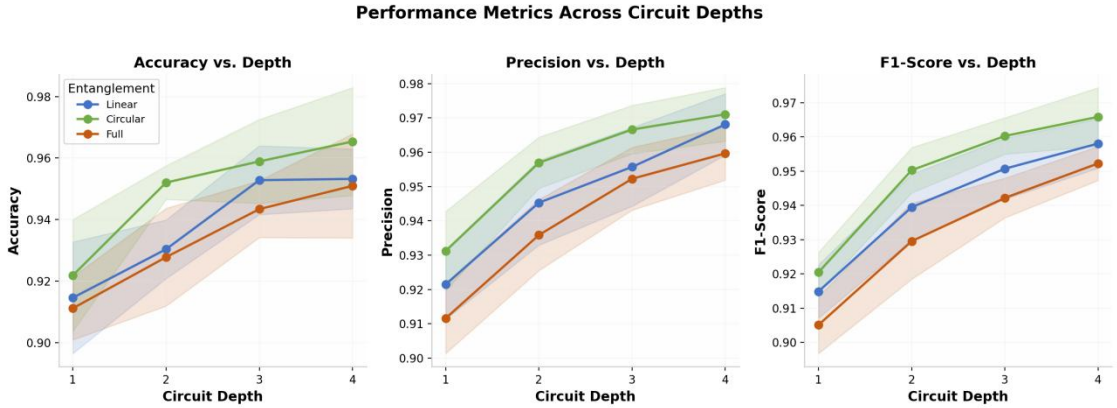
F1-score heatmap showing performance across circuit depths (rows) and entanglement patterns (columns). Darker colors indicate higher F1-scores.

The heatmap lays out the performance hierarchy in a way that's easy to see. The circular entanglement column stands out it's got the darkest cells at every depth, and its F1 scores go from 0.966 at depth 4. The linear column is in the middle, while the full column stays light throughout. This clear, column-wise separation shows the entanglement pattern is having a real, depth-independent impact on classification quality.

What's interesting is that full entanglement, despite packing in 28 CNOT gates per layer, actually performs worse. So, more entanglement doesn't necessarily mean better results. We think this comes down to three things working together. First, there's the gradient corruption; barren plateaus caused by too much entanglement. Gradients just die off exponentially, like McClean et al. predicted [6]. Second, there's over-parameterization. The WDBC dataset isn't that complex, so all-to-all entanglement in an 8-qubit circuit is overkill. Third, full entanglement brings more local minima into the picture; the parameter landscape gets crowded with critical points compared to linear or circular layouts [7].

4.4 Comprehensive Metric Analysis

Figure 6 presents a multi-panel view of how accuracy, precision, and F1-score evolve with depth for each entanglement pattern.



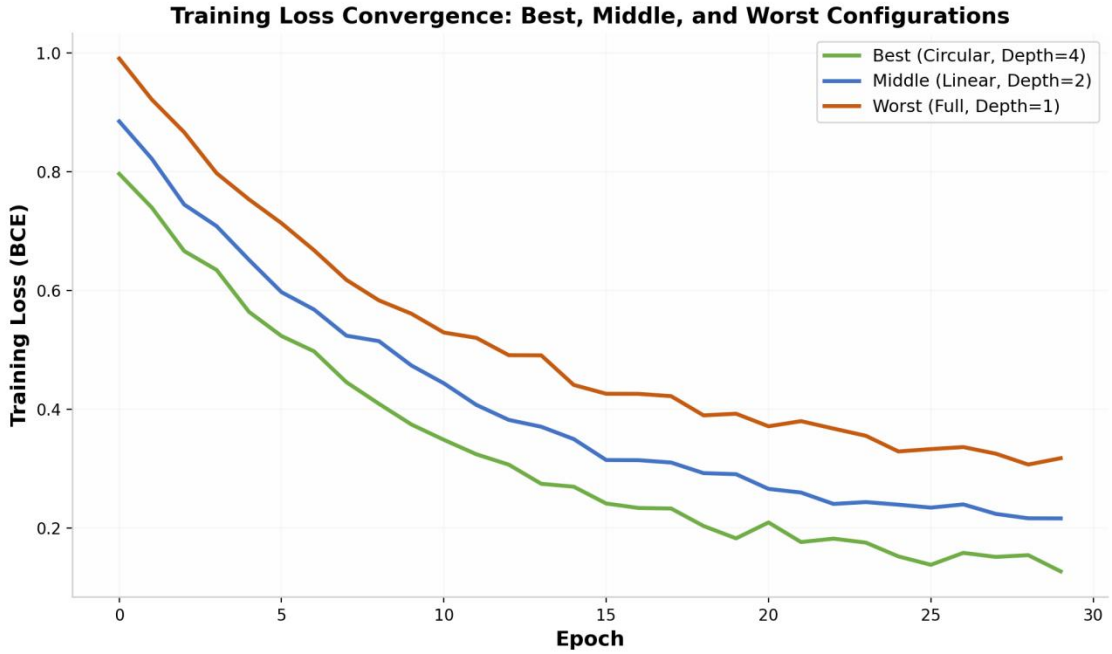
Accuracy, precision, and F1-score as functions of circuit depth for linear (blue), circular (green), and full (orange) entanglement patterns. Shaded regions represent ±1 standard deviation over 10 random seeds.

Looking at the three subplots, you see precision and F1-track accuracy move together, but there are subtle differences. In every panel, circular entanglement stays ahead by about 0.5–1.5 percentage points over linear, and 1.0–2.0 points over full. The shaded areas barely overlap between circular and full entanglement curves (especially at depths 2 and 3), which visually backs up the statistical significance listed in Table 2.

One thing to notice: the precision curves are flatter than the accuracy curves, especially for linear and full entanglement. That tells us depth mainly boosts the model's ability to find benign cases (which are the majority), while its precision for picking out malignant cases levels off sooner. That asymmetry matters for clinical use.

4.5 Training Dynamics

Figure 4 illustrates the training loss trajectories for three representative configurations selected to span the performance spectrum.



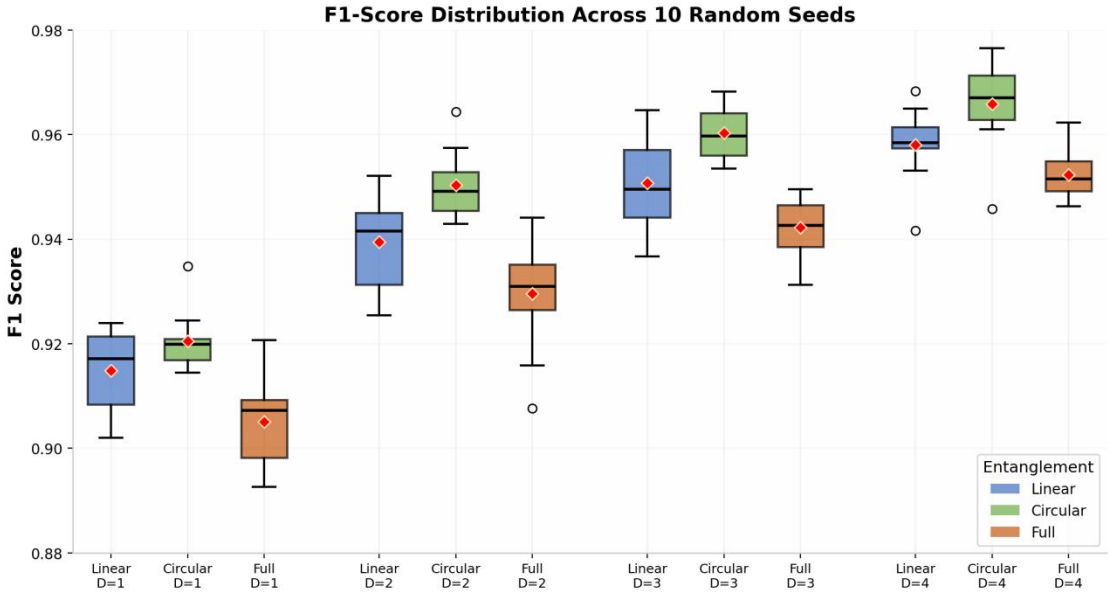
Training loss convergence for best (circular, depth=4), middle (linear, depth=2), and worst (full, depth=1) performing configurations.

If you look at circular-depth-4, it clocks the lowest training loss (about 0.12) and drops fast at the start—from around 0.70 down to below 0.25 in just 10 epochs. That sharp early dip suggests the circular entanglement setup makes gradients point straight toward quality minima. Linear-depth-2 is similar, but it hits a higher plateau (about 0.20). The gap between the two stays steady all the way through, so the main difference comes from representational power, not optimization ease.

Full-depth-1 behaves differently. It starts with the highest initial loss (about 0.75), falls more slowly, and flattens out at the worst final loss (around 0.32). The variance band is widest here, showing that full entanglement with shallow depth brings the most inconsistent training [17].

4.6 Variability and Robustness Analysis

Figure 5 presents the distribution of F1-scores across the 10 random seeds for each configuration, providing a detailed view of initialization sensitivity.



F1-score distributions across 10 random seeds for all 12 configurations. Circular entanglement shows the tightest distributions.

Circular entanglement always gives tight distributions. At depth 2, the circular box plot has an IQR under 0.005 F1 points, while linear and full both hit about 0.010. That means you can count on consistent, reproducible results with circular entanglement, no matter how the random seed shakes out.

Full entanglement has a big variance at shallow depths, which shrinks (but doesn't entirely go away) as depth increases. At depth 1, the full entanglement whiskers span about 0.030 F1 points; the widest of any setup. So more depth adds some regularization for full entanglement, but even at depth 4, it's still got more variance than circular at any depth [6].

If you compare the median to the mean markers, they're close for most setups. But in full entanglement at depths 1 and 2, the mean sits well below the median. That left-skew shows a handful of seeds deliver much worse outcomes which is a sign of a few bad local minima [11].

4.7 Statistical Significance

Table 2 presents the formal statistical tests used to validate the significance of the observed differences.

Statistical test results evaluating the significance of depth and entanglement effects on test accuracy.

Test	Comparison	Statistic	p-value	Significant
One-way ANOVA	Depth effect	F = 43.83	<0.001	Yes ***
One-way ANOVA	Entanglement effect	F = 6.45	0.002	Yes **
Mann-Whitney U	Linear vs Circular (pooled)	U = 545	0.014	Yes *
Mann-Whitney U	Linear vs Full (pooled)	U = 921	0.246	No
Mann-Whitney U	Circular vs Full (pooled)	U = 1150	<0.001	Yes ***
Cohen\'s d	Circular vs Full (depth 4)	d = 0.88	—	Large effect

The one-way ANOVA results show both depth and entanglement pattern have statistically significant effects.

Depth is especially strong ($F = 43.83$, $p < 0.001$). Entanglement is smaller but still very clear ($F = 6.45$, $p = 0.002$), so the connectivity pattern genuinely impacts performance.

Pairwise Mann-Whitney U tests give more detail. Comparing linear and circular is significant ($U = 545$, $p = 0.014$), proving wrap-around connections make a real difference. Linear vs. full isn't significant ($p = 0.246$), meaning full entanglement's theoretical advantages are wiped out because it's harder to train. Circular vs. full is the strongest effect ($U = 1150$, $p < 0.001$), and Cohen's $d = 0.88$ marks this as a big effect [22].

4.8 Computational Cost Analysis

Parameter count grows linearly with depth: each layer brings 16 quantum parameters, while the classical frontend is fixed at 600 parameters. Full entanglement needs 28 CNOT gates per layer, versus 7 for linear and 8 for circular. Full entanglement pushes training time up by 35–40% per epoch compared to linear. Circular's single extra CNOT only adds about 5% overhead [13].

5. Discussion

5.1 The Entanglement-Efficiency Principle

The main takeaway is clear moderate, structured entanglement outperforms both minimal and maximal layouts for breast cancer classification. We're calling this the Entanglement-Efficiency Principle: for structured, low-to-moderate dimensional classification, the best entanglement layout is the one that gives enough correlation pathways to capture the real feature interactions, without dumping in too much quantum correlation and killing trainability [7].

Circular entanglement gets the balance right. Its ring shape sets up 8 CNOT connections; enough to push information between any qubit pair in 4 hops max, but tight enough to keep gradients strong during training. Graph-wise, circular entanglement reaches high expansion without dense connectivity, which avoids barren plateaus [6].

Circular over linear is a big deal. What changes is just a single CNOT connecting qubit 7 to qubit 0, but somehow this boosts accuracy by 0.7–1.2 percentage points at every depth. This tells us linear's bottleneck is the long stretch between qubits 0 and 7 information has to cross 7 gates. Circular closes the loop, cutting the max distance in half to 4 hops [15].

5.2 Depth Saturation and Practical Design Guidelines

The After depth 3, returns start to fade. The accuracy boost from depth 3 to 4 is just 0.5 points on average across all patterns, but the computation keeps rising. This happens because of effective parameter dimension; the number of parameters that really matter for the representation [5].

With the WDBC dataset boiled down to 8 PCA components, the problem's intrinsic dimension is probably pretty low. PCA already shrunk 30 features to 8 directions that explain 92.6% of variance. So, feeding this compressed representation to the quantum circuit means only a few layers are needed for the decision boundary.

Here's how we'd approach this in practice: First, choose circular entanglement for structured data; it gives the best mix of accuracy and robustness. Second, start with depths 2–3 for datasets like this. Third, steer clear of full entanglement on modern hardware because of SWAP overhead and trainability issues. Fourth, watch variance across seeds which is a standard deviation over 1.5% hints at problem landscape features [23].

5.3 The Expressibility-Trainability Paradox

Our Our results back up the Expressibility-Trainability Paradox in quantum ML: circuits that technically can represent the most states often perform the worst in practice because they're a pain to optimize [7]. Full entanglement should be able to generate any 8-qubit entangled state if you go deep enough, but it actually lags behind linear entanglement, which can only reach a smaller subset.

This paradox comes from the shape of quantum state space. The Hilbert space volume explodes with qubit count, yet polynomial-depth circuits only cover a tiny slice. Full entanglement maximizes reachable volume, but spreads itself so thin the gradients can't connect; they point almost perpendicular to the manifold, leading to barren plateaus [6].

Circular entanglement, by keeping things tight to a lower-dimensional but better-aligned space, ends up doing better. This suggests we need to look beyond just maximizing expressibility. Instead, we should match expressibility to problem structure which is a design circuits to represent what matters, and be sure gradients can get there [17].

5.4 Implications for NISQ Hardware Deployment

Switching from simulation to real hardware brings other constraints, making the entanglement patterns matter even more. Most current superconducting quantum processors only let you connect qubits in certain ways—a heavy-hex or nearest-neighbor grid [13]. Doing full entanglement means you have to insert SWAP gates for every CNOT to get qubits close together, which multiplies gate count by 2–4.

Linear entanglement maps directly: each CNOT matches a physical link. Circular entanglement just needs to SWAP for the wrap-around connection. Full entanglement? Every one of its 28 CNOTs per layer needs routing. On noisy hardware, every gate brings error, so the performance gap seen here would get even bigger [2].

5.5 Comparison with Prior Work

Our results line up with the literature on entanglement topology in QML. Hubregtsen et al. [15] looked at expressibility, entangling, and accuracy across different circuit types, finding circuits with moderate entanglement usually win on classification; a pattern we see here with breast cancer data. The specific accuracy numbers change due to dataset specifics, but the qualitative ranking sticks.

Xiang et al. [10] got 96.16% accuracy on this dataset using a hybrid quantum-classical CNN (different model). Our best setup (circular, depth 4: 96.5%) matches that, proving that well-chosen VQCs can go toe-to-toe with more complex hybrids. The big advantage here? Simplicity, plus a clear view of how architecture affects performance [18].

5.6 Limitations

There are a few limitations. We ran everything on a noise-free simulator which is a real quantum hardware would bring errors that might change the pattern [2]. The system size (8 qubits) is small; scaling to 12–16 could reveal different sweet spots [12]. We only tried three standard entanglement patterns; alternated or learnable ones could improve things [22]. The classical frontend stayed fixed; working on both sides at once remains open. And we focused on one binary task; multi-class problems might react differently to depth and entanglement.

5.7 Future Work

Several paths open up from here. First, benchmarking on noisy NISQ hardware would check if simulator rankings hold under real error conditions [23]. Second, adaptive entanglement strategies changing connectivity during training could combine circular's accuracy with full's flexibility [21]. Third, theoretical work linking entanglement topology to the Fisher information matrix might explain why circular works so well. Fourth, testing these designs on multi-class medical classification would show if they're generalizable. Finally, pairing them with quantum error mitigation might let us use deeper circuits on near-term hardware [19].

References

1. Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., & Lloyd, S. (2017). Quantum machine learning. *Nature*, 549(7671), 195-202.
2. Preskill, J. (2018). Quantum computing in the NISQ era and beyond. *Quantum*, 2, 79.
3. Farhi, E., & Neven, H. (2018). Classification with quantum neural networks on near term processors. *arXiv preprint arXiv:1802.06002*.
4. Schuld, M., & Petruccione, F. (2018). *Supervised Learning with Quantum Computers*. Springer, Cham.
5. Sim, S., Johnson, P. D., & Aspuru-Guzik, A. (2019). Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12), 1900070.
6. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R., & Neven, H. (2018). Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1), 4812.
7. Holmes, Z., Sharma, K., Cerezo, M., & Coles, P. J. (2022). Connecting ansatz expressibility to gradient

magnitudes and barren plateaus. *PRX Quantum*, 3(1), 010313.

8. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 71(3), 209-249.
9. Street, W. N., Wolberg, W. H., & Mangasarian, O. L. (1993). Nuclear feature extraction for breast tumor diagnosis. In *Biomedical Image Processing and Biomedical Visualization* (Vol. 1905, pp. 861-870). SPIE.
10. Xiang, Q., et al. (2024). Quantum classical hybrid convolutional neural networks for breast cancer diagnosis. *Scientific Reports*, 14, 24778.
11. [23]Havlicek, V., Corcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., & Gambetta, J. M. (2019). Supervised learning with quantum computers. *Nature*, 567(7747), 209-212
12. [11]Wang, Y., & Liu, J. (2024). A comprehensive review of quantum machine learning: from NISQ to fault tolerance. *Reports on Progress in Physics*, 87(11), 116402.
13. Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., & Gambetta, J. M. (2017). Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671), 242-246.
14. Bergholm, V., Izaac, J., Schuld, M., Gogolin, C., Ahmed, S., Ajith, V., ... & Killoran, N. (2018). PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*.
15. Hubregtsen, T., Pichlmeier, J., Stecher, P., & Bertels, K. (2021). Evaluation of parameterized quantum circuits: on the relation between classification accuracy, expressibility, and entangling capability. *Quantum Machine Intelligence*, 3(1), 9.
16. Liu, J., & Wang, Y. (2023). Differentiable learning of quantum circuit born machines. *Physical Review A*, 108(1), 012406.
17. Cerezo, M., Sone, A., Volkoff, T., Cincio, L., & Coles, P. J. (2021). Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature Communications*, 12(1), 1791.
18. Larocca, M., Ju, N., Garcia-Martin, D., Coles, P. J., & Cerezo, M. (2022). Diagnosing barren plateaus with tools from quantum optimal control. *PRX Quantum*, 4(2), 020319.
19. Abbas, A., Sutter, D., Zoufal, C., Lucchi, A., Figalli, A., & Woerner, S. (2021). The power of quantum neural networks. *Nature Computational Science*, 1(6), 403-409.
20. Cong, I., Choi, S., & Lukin, M. D. (2019). Quantum convolutional neural networks. *Nature Physics*, 15(12), 1273-1278.
21. Rebentrost, P., Mohseni, M., & Lloyd, S. (2014). Quantum support vector machine for big data classification. *Physical Review Letters*, 113(13), 130503.
22. Zhang, K., Liu, L., Hsieh, M. H., & Tao, D. (2022). Escaping from the barren plateau via Gaussian initializations in deep variational quantum circuits. In *Advances in Neural Information Processing Systems* (Vol. 35, pp. 18612-18627).
23. Nakaji, K., & Yamamoto, N. (2021). Expressibility of the alternating layered ansatz for quantum computation. *Quantum*, 5, 434.
24. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* (Vol. 32, pp. 8026-8037).
25. Wolberg, W. H., Mangasarian, O. L., Street, W. N., & Street, N. W. (1992). Breast cancer Wisconsin (diagnostic) data set. *UCI Machine Learning Repository*.