SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I11P101

Volume 13 Issue 11, 1-24, November 2026
© 2026 Seventh Sense Research Group®

Original Article

Q-SWINER: A Novel Hybrid Quantum-Classical Enhanced Swin Transformer Model for Efficient Lung Cancer Classification

S. Radhika[1], G. Sharada[2]

[1]*Department of Computer Science and Engineering, Malla Reddy University, Hyderabad, India.*
[2]*Department of Computer Science and Engineering, CMR College of Engineering and Technology, Hyderabad, India.*

[1]*Corresponding Author: radhikasefur@gmail.com*

Received: 18 July 2026 Revised: 17 October 2026 Accepted: 22 October 2026 Published: 27 November 2026

Abstract *- Lung Cancer (LC) is a major cause of cancer deaths around the world, and effective diagnostic methods facilitate early prediction and analysis, and hence improve the prognosis of patients with LC. Machine Learning (ML) and Deep Learning (DL) frameworks are widely used for early LC detection but require significant computational resources for large, complex datasets, limiting diagnostic capabilities. To address this limitation, this article presents the Quantum-Enhanced Swin Transformer V2 Model (Q-SWINER), which improves segmentation and classification results. The overall structure consists of four parts, namely image data collection, data pre-processing, quantum segmentation and feature extraction, and quantum classification networks. We tested the model on two heterogeneous datasets (CT and PET/CT) using accuracy, precision, recall, specificity, and F1-score, and compared it with the classical Swin Transformer and existing quantum attention models. The proposed hybrid quantum model achieves 99.2%, 98.9%, 98.7%, 99%, and 99.0% accuracy, precision, recall, specificity, and F1-score, respectively, for lung cancer image classification, as shown in the simulation results on Dataset-1. In Q-SWINER, we replace the vector-parameterized components of the Swin Transformer with fixed-structure quantum circuits, yielding a new class of quantum Swin Transformer models.*

Keywords *- Lung Cancer, Swin Transformer V2, Deep learning, Quantum Machine Learning, Quantum Self-Attention, Medical image segmentation.*

1. Introduction

Lung cancer is an abnormal growth of lung tissue cells, especially from alveolar cells [1-5]. It is a common disease and has high mortality rates all around the world, and it continues to be a health issue that needs to be addressed as soon as possible and in an innovative way. Breast, colorectal, and prostate cancer are the three next most common cancers and are responsible for fewer deaths than Lung Cancer [6]. Factors that affect their occurrence include passive smoking, environmental pollution, occupational risk factors, and genetic susceptibility [7].

Socio-Economic inequalities, inadequate use of health care services, and insufficient early detection programs lead to late detection in many patients, when treatment is less effective [8-10]. Lung cancer comprises two distinct groups of tumors, Non-Small Cell Lung Cancer (NSCLC) and Small-Cell Lung Cancer (SCLC), as shown in Figures 1(a) and 1(b). Fine-grained classification of LC is crucial for prognosis, as it helps distinguish malignant from benign lesions. It is also an important non-invasive diagnostic tool for identifying the LC stage.

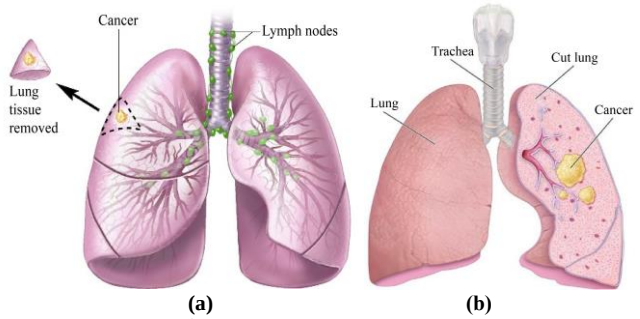


Fig. 1 Types of Lung Cancer: (a) NSCLC, and (b) SCLC.

Early, effective detection and treatment of Lung Cancer (LC) are key to improving patient outcomes. The National Lung Screening Trial (NLST) has shown that low-dose helical CT screening is very effective at reducing death rates among those at high risk. A problem with LC screening is, however, its tendency to yield False-Positive (FP) results, resulting in

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

unnecessary medical interventions, higher health care costs, and psychological distress.

In response, CAD systems have been introduced to overcome these limitations, helping increase early-stage cancer detection while also attempting to decrease FP rates in diagnosis [11, 12]. Meanwhile, the field has seen advances such as liquid biopsies, a minimally invasive approach that analyzes genes and biomarkers in blood samples to diagnose LC early in the disease course [13-16].

More accurate and efficient diagnostics are still needed. Recent innovations in Artificial Intelligence (AI) and ML seem to be the transformative solution. In the field of medical image processing, highly relevant results can be obtained using Artificial Intelligence, especially DL approaches, such as Convolutional Neural Networks (CNNs). CNNs can extract important features from imaging data, enhancing the accuracy and reliability of cancer diagnosis. In recent years, Transformer Networks have gained prominence for their effectiveness in sophisticated feature extraction, deep learning frameworks, and high performance in LC categorization using histopathology and CT scan images [17].

Vision Transformers and Swin Transformers play a significant role in various image processing applications. However, these algorithms are expensive to train and operate [18]. Specifically, they require significant computational resources and memory capacity as their magnitude increases. Past work has approximated self-attention to reduce its quadratic memory usage and computation [19, 20]. For more complex data, resource-efficient architectures are increasingly needed to achieve high performance. Recently, researchers have explored quantum ML methods to design complex transformer learning models.

CNN-, Vision Transformer-, and Swin Transformer-based methods for lung cancer image analysis show significant improvement; however, several challenges remain. CNN-based models can extract local features well, but may not capture long-range dependencies in complex lung images. Transformer-based models such as Swin Transformers can capture broader context, but their attention mechanisms and complexity can impose high computational and memory demands for high-resolution, multimodal medical images. While several quantum machine learning works have focused on quantum classifiers, quantum kernels, and hybrid quantum-classical models for cancer applications, their use in combining hierarchical Swin Transformer architectures for lung nodule segmentation and classification remains underexplored. Moreover, current research tends to focus on either classification or individual quantum-enhanced components, rather than an integrated quantum-enhanced attention mechanism within a segmentation and classification framework. Thus, a hybrid architecture is needed to integrate Swin Transformers' hierarchical representations with quantum attention and classification to achieve more precise segmentation and classification in heterogeneous lung cancer datasets.

Based on the above limitations, the research gap can be summarized as follows: Existing lung cancer deep learning methods effectively represent the images and classify lung cancer, and existing quantum machine learning studies have proven the feasibility of quantum-enhanced learning, but there is limited work to integrate quantum-enhanced multi-head self-attention with a hierarchical Swin Transformer architecture to simultaneously segment and classify lung cancer in heterogeneous lung imaging datasets. This gap will encourage the development of the proposed Q-SWINER framework.

1.1. Problem Structure

This research focuses on a new, efficient, and accurate framework for segmenting and classifying lung cancer in heterogeneous medical images, a major challenge. While both Conventional Deep Learning and Transformer-Based Models have shown great potential, their growing complexity and attention requirements can make it difficult to handle high-resolution lung images and larger datasets. Although window-based attention reduces the computational cost of global self-attention, the classic attention mechanism remains computationally expensive for more complex feature representations.

Conversely, current quantum machine learning solutions for lung cancer diagnosis have mainly focused on learning quantum classifiers and enhancing feature processing using quantum computing, rather than applying quantum attention mechanisms in hierarchical Transformer-based lung segmentation models.

This creates a methodological gap in solving detailed lesion segmentation, multi-scale feature extraction, and lung cancer classification within a single hybrid quantum-classical framework. Based on this, the research problem is: How to design a Q-SWINER architecture that integrates quantum-enhanced multi-head self-attention, a Swin Transformer-based encoder-decoder structure, and quantum classification layers, and whether this integration can improve segmentation and classification performance on heterogeneous lung cancer image datasets.

1.2. Research Contribution

Accordingly, the study presents a novel hybrid quantum-classical framework called Q-SWINER, based on the Swin Transformer, for lung cancer image segmentation and classification, addressing the above-mentioned research gap. The framework combines multi-head self-attention with the Swin Transformer, adds a U-shaped encoder-decoder structure for lesion segmentation, and uses quantum variational classification for the final prediction. This study has made the following contributions:

1. The hierarchical window-based Swin Transformer is enhanced by introducing a Quantum-Enhanced Multi-Head Self-Attention (Q-MHSA) mechanism to achieve quantum-based feature interaction and attention representation.
2. A U-shaped Q-SWINER (U-shaped quantum attention network) encoder-decoder framework is designed for lung lesion segmentation, and multi-scale features and skip connections are used to maintain the spatial information during the encoding and decoding process.
3. The extracted representations are processed using a quantum variational classification layer in the framework, which is used to make the ultimate classification of lung cancer.

Next, we test the proposed framework on heterogeneous lung imaging datasets and in a quantum-circuit simulation environment, benchmarking segmentation and classification performance against the classical Swin Transformer and existing methods using multiple evaluation metrics. Together, these works fill this gap by integrating quantum-enhanced attention, hierarchical Swin Transformer feature learning, U-shaped lesion segmentation, and quantum variational classification into a single lung cancer image analysis system.

Q-SWINER's novelty lies in combining quantum-enhanced multi-head self-attention with the hierarchical Swin Transformer structure to segment and classify lung cancer images. Unlike previous quantum lung cancer diagnosis works, which mostly use quantum classifiers like VQC, QSVC, or QSVM for classification [21-24, 30], our framework applies quantum processing directly to the Swin Transformer's attention mechanism via the proposed Quantum Multi-Head Self-Attention (Q-MHSA).

Moreover, Q-SWINER has a U-shaped encoder-decoder network for multi-scale lung lesion segmentation, followed by a quantum variational classification network for final prediction. Thus, the proposed framework combines quantum-enhanced attention, hierarchical Transformer feature extraction, U-shaped segmentation, and quantum classification in a single framework. The framework is further validated on a heterogeneous collection of CT and PET/CT lung imaging datasets, enabling the proposed approach to be tested across varying imaging characteristics rather than limited to a single dataset.

1.3. Summary of the Paper

The rest of the manuscript is structured as follows: Section 2 reviews studies by various authors. Section 3 describes the datasets, followed by the pre-processing, the Q-SWINER architecture, and the quantum classification layers. Section 4 presents the experimental implementation, results, and comparative evaluation. Section 5 summarizes the study and proposes future directions to improve the proposed quantum-based deep learning framework.

2. Related Works

In healthcare, Munshi et al. (2024) [21] used Quantum Machine Learning (QML) models to predict lung cancer, namely the Variational Quantum Classifier (VQC) and Pegasos Quantum Support Vector Machine Classifier (QSVC). Their method used quantum approaches to increase detection efficiency: Pegasos QSVC achieved the highest accuracy of 85%. The study showed reduced computational resources, which was an advantage, but the quantum implementation's accuracy was lower than that of classical methods, indicating room for improvement.

To improve the efficiency of LC identification, Martis et al. (2024) [22] proposed a hybrid method that combines DL with quantum computing. They used pre-trained algorithms for quantum- enhanced feature extraction for categorization. The results showed strong performance: overall Accuracy (92.12%), sensitivity (94%), specificity (90%), and F1-score (93%). The system achieved improved performance, but implementation would be difficult because it requires maintaining both classical and quantum elements.

Dutta et al. (2024) [23] proposed a novel QKSVM with Quantum Approximate Optimization Algorithm (QAOA) embedding for LC classification. They benchmarked various multi-quantum kernel versions, such as Angle embedding and IQP embedding, with a QKSVM and classical SVM. After PCA reduced the feature vector to 4 dimensions, the QAOA-based QKSVM achieved 94.64% test accuracy and an AUC of 82.33%. One limitation was the small number of feature dimensions, which may not capture the full complexity of all cancer types.

Ghobadi and Afsaneh (2024) [24] proposed a hybrid quantum machine learning framework to classify ten cancer types, such as Lung Adenocarcinoma (LUAD), using RNA-seq data. They integrated classical feature selection and Fisher scores with another QSVM for classification. The LUAD classification results also showed high performance, with QSVM achieving 99% accuracy, 95% precision, 99% recall, and a 97% F1-score. In clinical applications, hardware constraints limit the use of quantum simulators and could impede scalability for broader use in biomarker discovery.

Crasta et al. (2024) [25] developed a novel DL system for LC detection, using a 3D-VNet architecture for nodule segmentation and a 3D-ResNet for nodule classification. This approach achieved strong performance, with a 99.34% Dice Similarity Coefficient (DSC), 99.2% classification accuracy, and a 0.4% false positive rate. While the model shows potential, its complexity could make it difficult to implement in resource-limited settings, such as healthcare.

Kumar et al. (2024) [26] proposed a framework for LC prediction using multiple DL methods and Transfer Learning (TL)-the fusion model identified squamous cells with 100%

accuracy and the other categories with 90% accuracy. While the model shows excellent performance, concerns remain about possible overfitting, particularly when applied to unobserved or varied datasets.

In cancer research, Mukku et al. (2023) [27] examined using QML with an adaptation of the traditional K-means clustering algorithm, known as the Q-means algorithm, for medical image segmentation.

The study also contrasted Q-means performance with traditional machine learning methods, emphasizing quantum computing's value in detecting subtle patterns for cancer diagnosis. However, the research was mostly theoretical, and the proposed systems' advantages were not widely validated.

Wahab Sait et al. (2023) propose a computationally efficient DL-based method for identifying lung cancer using PET/CT images [28]. The method combined DenseNet-121, deep autoencoders, and MobileNet V3-Small to achieve the highest accuracy (98.6%) and Cohen's Kappa (95.8%) while reducing computational demands.

Shah et al. (2023) [29] developed a DL ensemble 2D CNN approach for lung nodule identification in the LUNA 16 dataset. They used three types of CNNs with different architectures in an ensemble, achieving 95% combined accuracy. This ensemble approach improved detection robustness by leveraging each model's strengths. The study also points out potential implementation issues, such as the computational overhead of ensemble techniques.

Table 1. Overview of existing research

S. No	Author & Year	Technology	Results Obtained	Advantages	Limitations
1	Munshi et al. (2024)	Variational Quantum Classifier (VQC) and Pegasos Quantum Support Vector Machine Classifier	85% classification accuracy	Reduced computational resources	Lower accuracy compared to classical methods
2	Martis et al. (2024)	A hybrid framework combining DL with quantum computing	Accuracy: 92.12%, Sensitivity: 94%, Specificity: 90%, F1-score: 93%	Impressive overall performance	High Computation
3	Dutta et al. (2024)	QKSVM with QAOA embedding	94.64% test accuracy, 82.33% AUC	High test accuracy	Limited to minimal feature dimensions
4	Ghobadi & Afsaneh (2024)	Quantum Support Vector Machine (QSVM)	99% accuracy, 95% precision, 99% recall, 97% F1-score (LUAD classification)	Effective biomarker discovery	Limited scalability concern
5	Crasta et al. (2024)	3D-VNet for segmentation and 3D-ResNet for classification	99.34% Dice Similarity Coefficient	Low false positives (0.4%)	High model complexity
6	Kumar et al. (2024)	Ensemble of ResNet-50, EfficientNet-B3, and ResNet-101 with transfer learning	100% accuracy for Squamous Cells	Innovative fusion model	Potential overfitting
7	Mukku et al. (2023)	Q-means algorithm for medical image segmentation	Improved cancer diagnosis	Potential for improved speed and pattern recognition	Lacks empirical validation
8	Wahab Sait (2023)	DenseNet-121, Deep Autoencoders, and MobileNet V3	98.6% accuracy, Cohen's Kappa of 95.8	Efficient computation	Limited generalizability
9	Shah et al. (2023)	Deep Ensemble 2D CNN approach	95% combined accuracy	Improved detection robustness	Computational overhead
10	Jain et al. (2020)	Hybrid classical-quantum approach with QCrush data representation	95.24% accuracy in cancer subtype classification	Successful identification of genetic markers	Dependence on specialized quantum hardware

Jain et al. (2020) [30] proposed a hybrid classical-quantum ML approach for classifying Non-Small Cell Lung Cancer (NSCLC) and its subtypes. Their approach used feature selection via univariate statistics and XGBoost, along with a proprietary data representation for a D-Wave quantum annealing machine called QCrush. The hybrid solution achieved 95.24% accuracy and identified key genetic markers. However, as a general method, this reliance on specific quantum hardware may hamper widespread use. Table 1 presents state-of-the-art previous work in this area, including key methods, data, and results

As shown in Table 1, existing studies show that classical deep learning and quantum machine learning can effectively analyze lung cancer, but their methods differ from the proposed Q-SWINER framework. Munshi et al. [21] used VQC and Pegasos QSVC for lung cancer classification, and Martis et al. [22] studied a hybrid quantum architecture for lung cancer detection. Dutta et al. [23] applied QKSVM with QAOA embedding, and Ghobadi and Afsaneh [24] used QSVM for cancer classification. These approaches illustrate the applicability of quantum learning for classification. Q-SWINER, by contrast, introduces quantum processing into the multi-head self-attention mechanism of a hierarchical Swin Transformer and integrates it into a U-shaped segmentation system. Crasta et al. [25] introduced the 3D-VNet and 3D-ResNet frameworks, which tackled both segmentation and classification, but did not include a quantum-enhanced attention mechanism. Similarly, transfer-learning and ensemble CNN models [26, 29] rely on classic feature extraction and classification methods. In this regard, the proposed work is unique in integrating a quantum-enhanced attention mechanism, hierarchical Transformer representation, lesion segmentation, and quantum variational classification into a unified quantum framework and testing it on heterogeneous lung imaging datasets.

Unlike most quantum circuits that serve as standalone classifiers or quantum feature-processing modules, Q-SWINER embeds quantum operations within the attention mechanism of a hierarchical Swin Transformer V2 framework. The proposed Q-MHSA is a quantum attention mechanism that integrates ZFeatureMap-based quantum encoding with CNOT and ISWAP entanglement in the window-based attention pipeline. Therefore, it integrates quantum attention processing with Swin V2, rather than using quantum circuits as a classification unit.

3. System Overview

Figure 2 shows the proposed hybrid quantum-classical Q-SWINER network. It consists of four steps: (1) heterogeneous data collection, (2) image pre-processing, (3) a modified Swin Transformer combined with a quantum multi-head attention model, and (4) performance assessment. These are explained in the following subsections, one module for each.

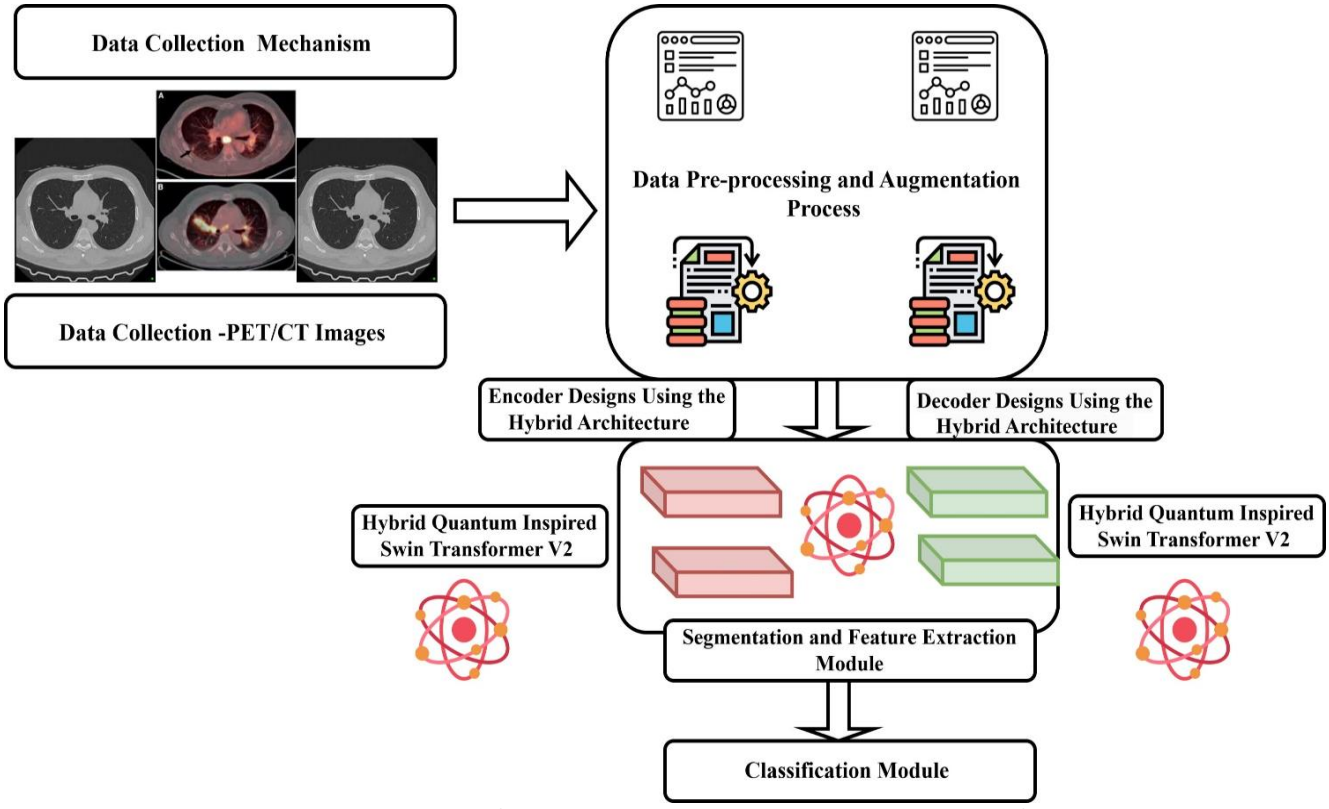
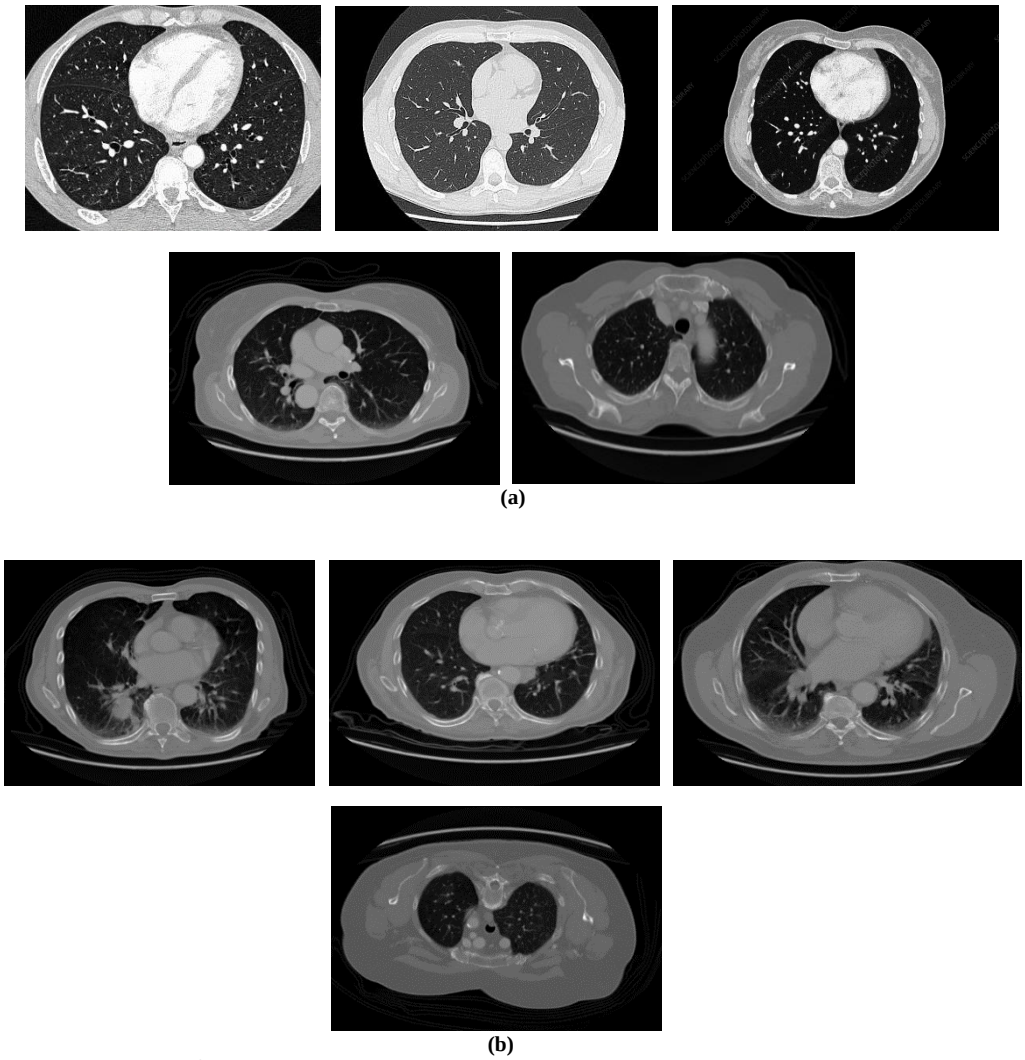


Fig. 2 Overall architecture for Q-SWINER-Proposed Quantum-Inspired Swin Transformer

3.1. Materials and Methods

In this study, two heterogeneous datasets are used, as shown in Figure 2: The Lung CT images are from the LIDC-IDRI collection [31] available from The Cancer Imaging Archive (TCIA) [33], and the PET/CT images are from the Lung-PET-CT-Dx collection [32] also available from TCIA [33]. In the first scenario (Dataset-1), we use 1,018 CT scans from the LIDC-IDRI collection. We then test the Lung-PET-CT-Dx collection (Dataset-2), with 31,562 images collected from patients in healthcare centers in China, to assess the generalisability of the proposed LC recognition method. The Computer Center and Cancer Institute at Harbin Medical University, Harbin, China, acquired the PET/CT data and authorized their use. The datasets do not contain any sensitive patient information. The dataset consists of CT and PET-CT DICOM images from lung cancer patients, along with annotations in Extensible Markup Language (XML) format, with target locations marked as bounding boxes. The pathological criteria/tissue diagnosis determined the patient's classification. Five academic thoracic radiologists with expertise in LC detection annotated tumor locations in the CT and PET images. Figures 3(a) and 3(b) shows sample images from the two datasets.



(a)

(b)

Fig. 3 Sample image datasets deployed for evaluating the recommended model: (a) Dataset-1 (LIDC-IDRI), and (b) Dataset-2 (LUNG-PET-CT-DX).

3.2. Data Preprocessing Technique

Preprocessing is done in several stages. First, we convert the DICOM volumes to 2D axial slices: we extract slices from each of the 1,018 LIDC-IDRI CT studies in Dataset-1 and from the Lung-PET-CT-DX PET/CT studies in Dataset-2. Each slice is labeled with the patient-level diagnosis and radiologist annotations (benign/malignant). We then crop out the unnecessary background and resize each slice to a uniform 512×512 pixels to enable comparison between the two sets. We use Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance contrast and Gaussian blur to reduce noise, improving image clarity and detail. We place all slices from a patient into one subset, and no two slices share a subset, before any augmentation (split ratio [70:15:15]; number of

patients and slices per subset: [to be inserted from the re-run]). The LIDC-IDRI data are extremely class-imbalanced, with the proportion of patients in the minority class (malignant) being increased from [73,589] to [146,716] in the training data set by using a Generative Adversarial Network (GAN) applied to the training set only, resulting in a balanced training set of [318,116] samples (all counts to be re-computed after patient-disjoint split). No image or its corresponding slice appears in both the validation and test sets, and these sets contain only real images, avoiding data leakage.

3.3. Segmentation Model using Swin Transformer-V2 Model

This section describes the Swin Transformer model and presents the proposed Swin Transformer model for segmenting lesion areas in lung cancer.

3.3.1. Swin Transformer Model

Figure 4 shows the Swin Transformer architecture. The Swin Transformer block consists of Layer Normalization (LN), Window-based Multi-Head Self-Attention (W-MHSA), and a Multi-Layer Perceptron (MLP). These components are stacked sequentially to form the complete Swin Transformer architecture. In Equations (1) to (4), y^{l-1}, $\hat{y}^{l}$, y^{l} and $\hat{y}^{l+1}$ denote the input, intermediate and output feature maps of successive Swin Transformer blocks. W-MHSA and SW-MHSA denote window-based and shifted-window multi-head self-attention, respectively. The feature maps are refined by the attention and feed-forward layers, while the residual (skip) connections retain the contextual information.

$$\hat{y}^{l} = W - MHSA(LN(y^{l-1})) + y^{l-1}, \quad (1)$$

$$y^{l} = MLP(LN(\hat{y}^{l})) + \hat{y}^{l}, \quad (2)$$

$$\hat{y}^{l+1} = SW - MHSA(LN(y^{l})) + y^{l}, \quad (3)$$

$$y^{l+1} = MLP\left(LN(\hat{y}^{l+1})\right) + \hat{y}^{l+1}, \quad (4)$$

In the proposed framework, we use an improved version of the original Swin Transformer described above, called Swin Transformer V2 (ST-v2). Therefore, ST-v2 is used for extracting the region of the lesion in the lung images, as it has the property of fast convergence and can process high-resolution images.

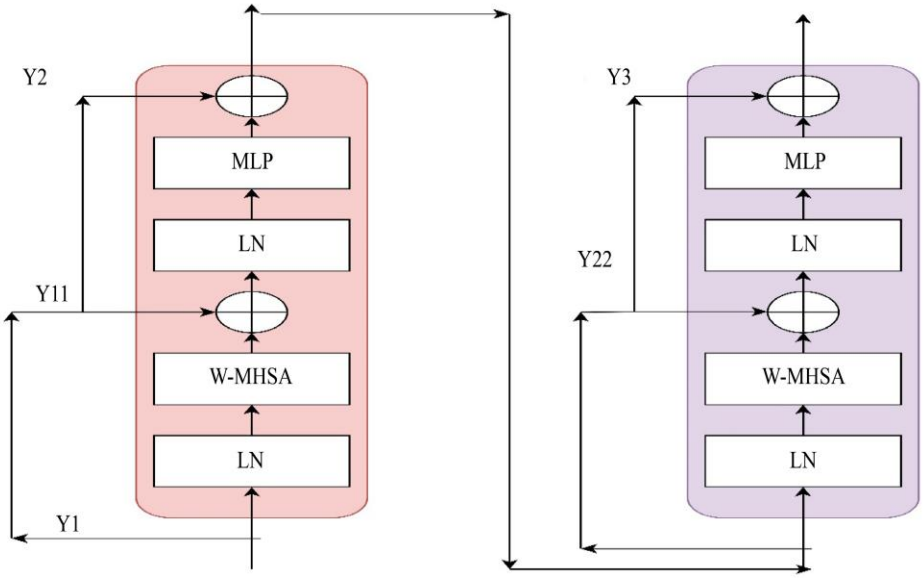


Fig. 4 Structure of the Swin Transformer

Self-attention is calculated in each window. In this mechanism, each window is flattened into a 2D token matrix, and linear projections derive the Query (Q), Key (K), and Value (V) matrices.

Each matrix is divided into a number of equal attention heads; the linear projections vary across heads, and the attention computations run in parallel. All heads produce outputs that are merged together and fed to a last linear projection for obtaining the attention output. ST-v2 uses scaled cosine attention for its computation of attention, as illustrated in Equation (5):

$$Attention(Q,K,V) = SoftMax(\frac{cos(Q,K)}{\tau} + B)V, \quad (5)$$

In Equation (5), cos(·,·) denotes cosine similarity, τ is a learnable scalar, and B is the relative-position bias. Moreover, the Log-Spaced Continuous Position Bias (Log-CPB) is integrated into ST-v2 to easily switch between window sizes, enabling flattening for interpolation during execution of algorithms learned from low-resolution data and applied to high-resolution applications, and preventing massive data transfer loss at the switch.

3.3.2. Quantum Enriched Modified Swin Transformer Model

Figure 5 shows the Quantum-based Window Multi-Head Self-Attention (W-QMHSA). As shown in Figure 5, quantum circuits are embedded within the W-MHSA, while a quantum machine learning model is embedded in the MLP layers of Q-SWINER. The following steps are followed to transform the MHSA to the Q-MHSA:

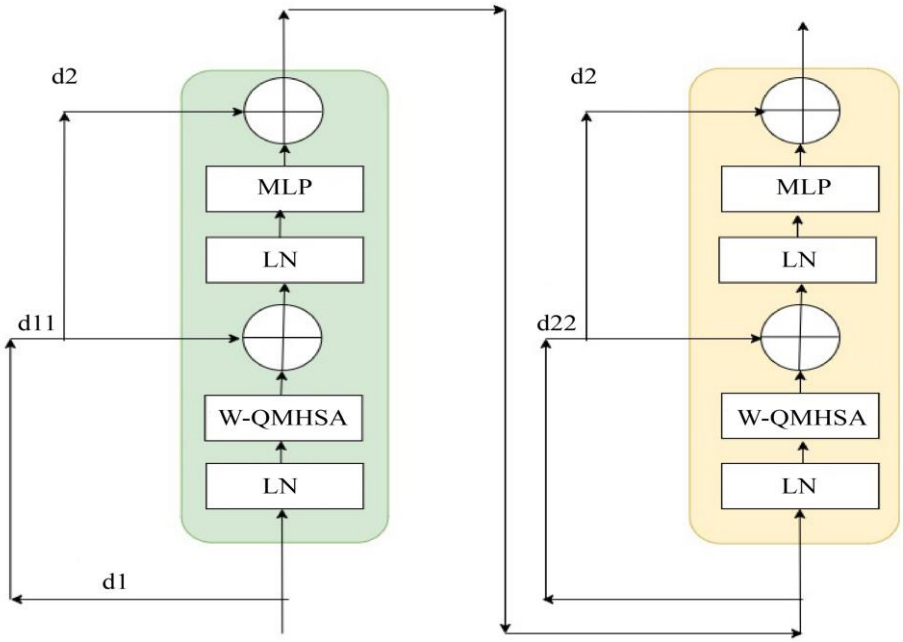


Fig. 5 Q-SWINER: An architectural overview

The corresponding Swin Transformer blocks with Q-MHSA are expressed in Equations (6) to (9), where d denotes the feature maps of the quantum-enhanced blocks (notation analogous to Equations (1) to (4)):

$$\hat{d}^{l} = W - QMHSA(LN(d^{l-1})) + d^{l-1}, \quad (6)$$

$$d^{l} = MLP\left(LN\left(\hat{d}^{l}\right)\right) + \hat{d}^{l}, \quad (7)$$

$$\hat{d}^{l+1} = SW - QMHSA\left(LN(d^{l})\right) + d^{l}, \quad (8)$$

$$d^{l+1} = MLP\left(LN\left(\hat{d}^{l+1}\right)\right) + \hat{d}^{l+1}, \quad (9)$$

The proposed Q-SWINER architecture builds on the Swin Transformer V2 by incorporating quantum operations into the traditional window-based multi-head self-attention mechanism. In addition to classical feature transformations within the attention module, the Swin V2 window-based processing generates feature representations that the ZFeatureMap module encodes into quantum states. These encoded quantum features are then passed through parameterized quantum circuits with CNOT and ISWAP entanglement operations. By measuring the resulting quantum representations, the mechanism injects them into the attention computation, forming the proposed Quantum Multi-Head Self-Attention (Q-MHSA).

Encoding Technique

In Q-SWINER, the ZFeatureMap therefore acts as the quantum encoding interface between the classical Swin V2 window-based feature representation and the subsequent quantum attention operations. Three quantum embeddings transform the classical input data into three quantum states |xs,q⟩, |xs,k⟩, and |xs,v⟩.

The ZFeatureMap implements this mapping: a quantum circuit that maps classical input data to quantum data, enabling computation on quantum hardware.

In Figure 6, the ZFeatureMap is realized by a Hadamard (H) gate followed by a parameterized phase (P) gate (a rotation about the Z axis) on each qubit, encoding the classical data.

This is accomplished using data-mapping functions, and repetitions accommodate multi-head attention networks. The mathematical encoding technique adopted is given as follows:

Define the ZFeatureMap operator

$$\hat{Z}(\text{xi}) \quad \text{as}|\psi_i\rangle = \hat{Z}(x_i)|0\rangle|0\rangle^{\otimes d_h} = \bigotimes_{j=1}^{d_h} P(2x_{ij})H|0\rangle, \quad (10)$$

Here, $\hat{H}$ represents the Hadamard gate and $\hat{P}(\cdot)$ denotes the phase gate, i.e., a rotation about the Z axis by the given angle, as depicted in Figure 6.

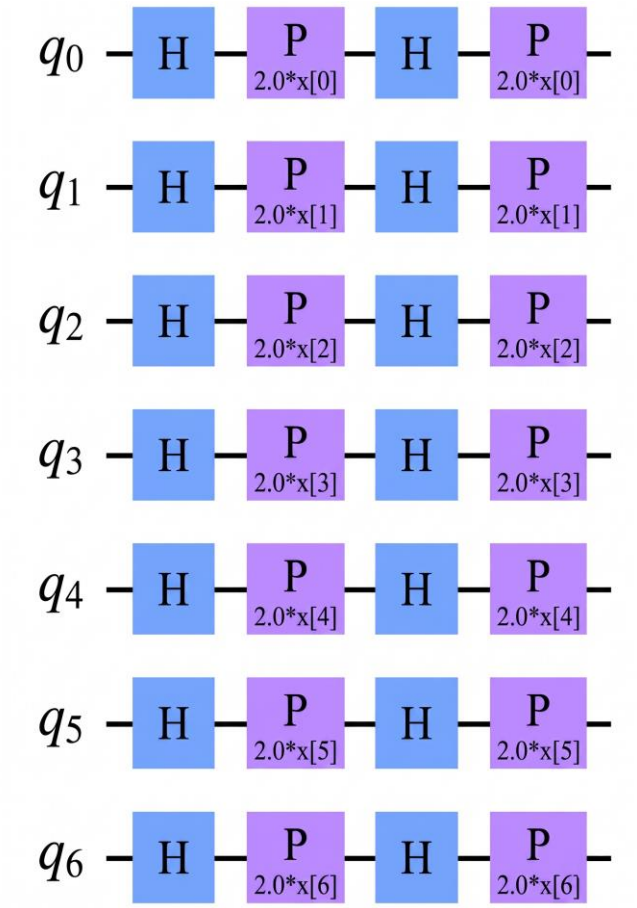


Fig. 6 Z-Feature maps for encoding classical data into quantum data

After encoding into quantum data, the key circuit is applied using the key operator $U_K(\theta_K)$ for every xi and deriving the column vector K, which is mathematically expressed as

$$K_i = \langle\psi_i|U_K\dagger(\theta_K)Z_iU_{K(\theta_K)}|\psi_i\rangle, 1 \le i \le d_t, \quad (11)$$

Where Zˆ i is a spin measurement of the ith qubit in the z direction. In the preceding stage, a query circuit is applied using the query operator $U_Q(\theta_Q)$ for each xi and acquires the column vector Q

$$Q_i = \left\langle\psi_i\left|U_Q\dagger(\theta_Q)Z_iU_{Q(\theta_Q)}\right|\psi_i\right\rangle, 1 \le i \le d_t, \quad (12)$$

Hence, the attention score matrix is obtained from the key and query vectors using Equation (13):

$$A_{ij} = -(Q_i - K_j)^2, 1 \le i,j \le d_t, \quad (13)$$

Finally, the value circuit is applied using the Value operator Vˆ(θV) to every image row and computes every q-bit individually to derive the value matrix, which is illustrated in Equation (14).

$$V_{ij} = \langle\psi_i|U_V\dagger(\theta_V)Z_jU_{V(\theta_V)}|\psi_i\rangle, 1 \le i \le d_t, 1 \le j \le d_h, \quad (14)$$

Entanglement Methods

After quantum encoding, the resulting quantum states are processed using CNOT and ISWAP entanglement operations. These operations establish interactions among the encoded quantum features and are incorporated into the Q-MHSA module to enhance feature dependency modeling within the Swin V2 attention framework.

We add two entanglement phases to the proposed Q-MHSA network to better capture long–range dependencies for each qubit and head, improving classification. As shown in Figure 7, SWAP quantum circuits can calculate the self-attention weight coefficient between queries and keys, which can be used to evaluate the adjacency between the two qubits.

After that, the qubits pass through the ISWAP test, and the quantum positional embedded bits are paired using CNOT to capture longer-range and local feature dependencies.

The entangling layers can also be repeated between trainable single-qubit rotations (RZ, RY) for each block, as shown in Equation (15).

Figure 8 depicts the entangled bits after the entanglement process. As shown in Figure 9, Q-SWINER is a quantum multi-head attention network with improved processing capability.

$$|\psi(\theta)\rangle = U(\theta)|\psi\rangle, \; U(\theta) = \prod_{l=1}^{L} U_l(\theta), \quad (15)$$

The quantum operations are applied to the feature representations obtained within the Swin V2 attention windows.

Thus, the proposed Q-MHSA retains Swin Transformer V2's localized, window-based processing while introducing quantum feature encoding and entanglement into the attention computation.

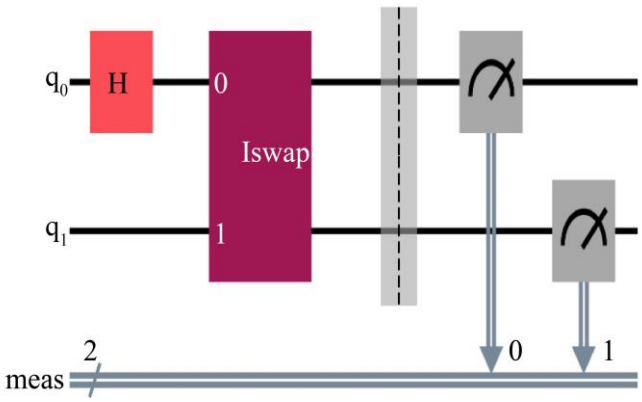


Fig. 7 Swap circuits used for entanglement

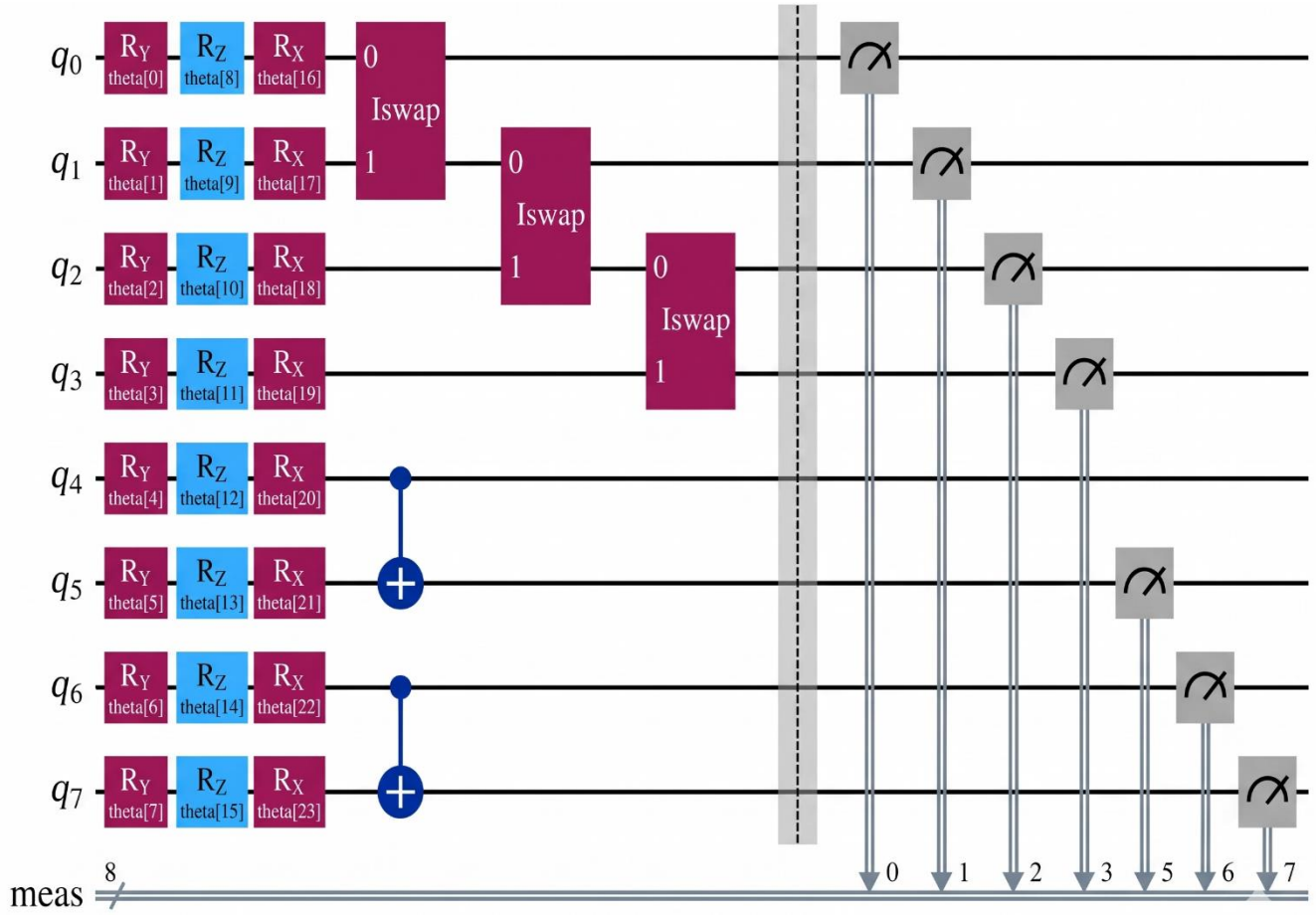


Fig. 8 Entanglement bits using swap and CNOT gates

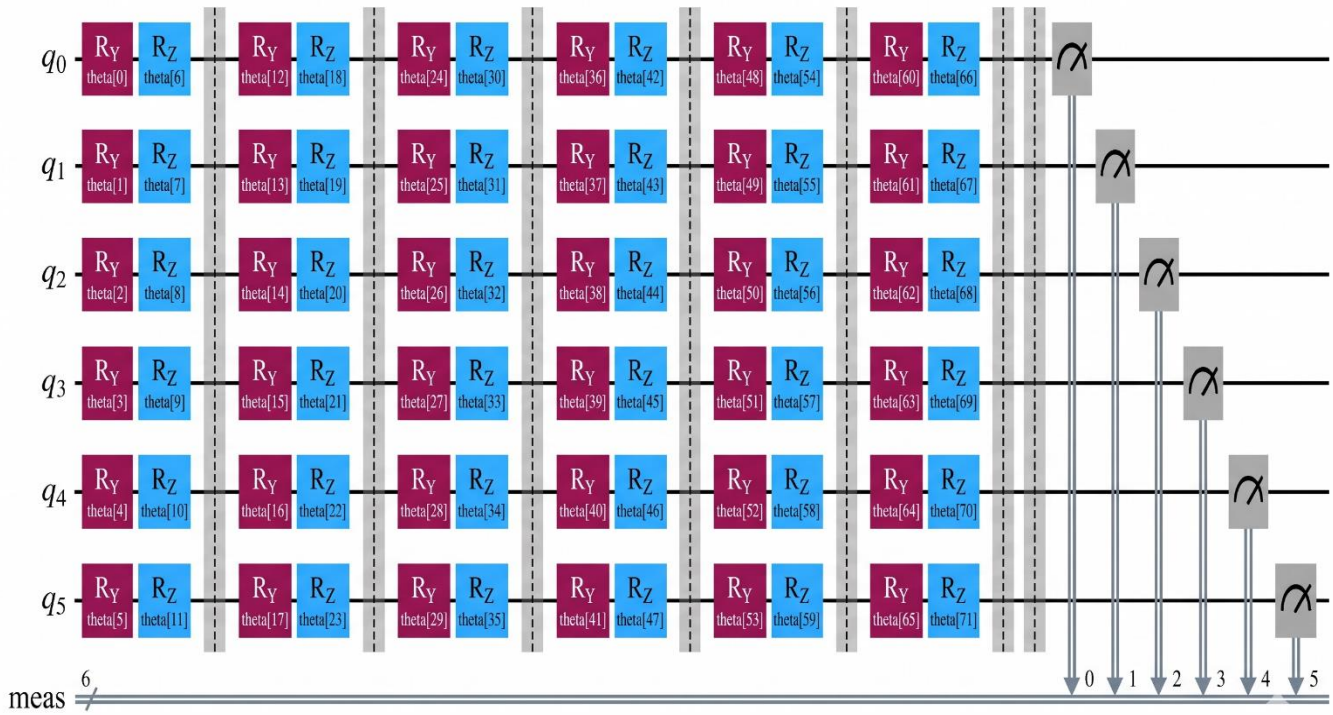


Fig. 9 Q-SWINER-Quantum-based multi-headed attention networks

Measurement Phase

After each head, a Pauli-Z operator is applied to the qubits, and classical information is obtained from the quantum inputs.

Several measurement schemes can be used (e.g., measuring in different bases or multiple qubits) to extract a wide variety of outputs from the quantum processing step. The measurements yield a real number, which is used as an "attention weight" or "enhancement from the quantum state," as in Equation (16).

$$M = \langle\psi|Z|\psi\rangle, \tag{16}$$

3.3.3. Segmentation Model-based Q-SWINER Networks

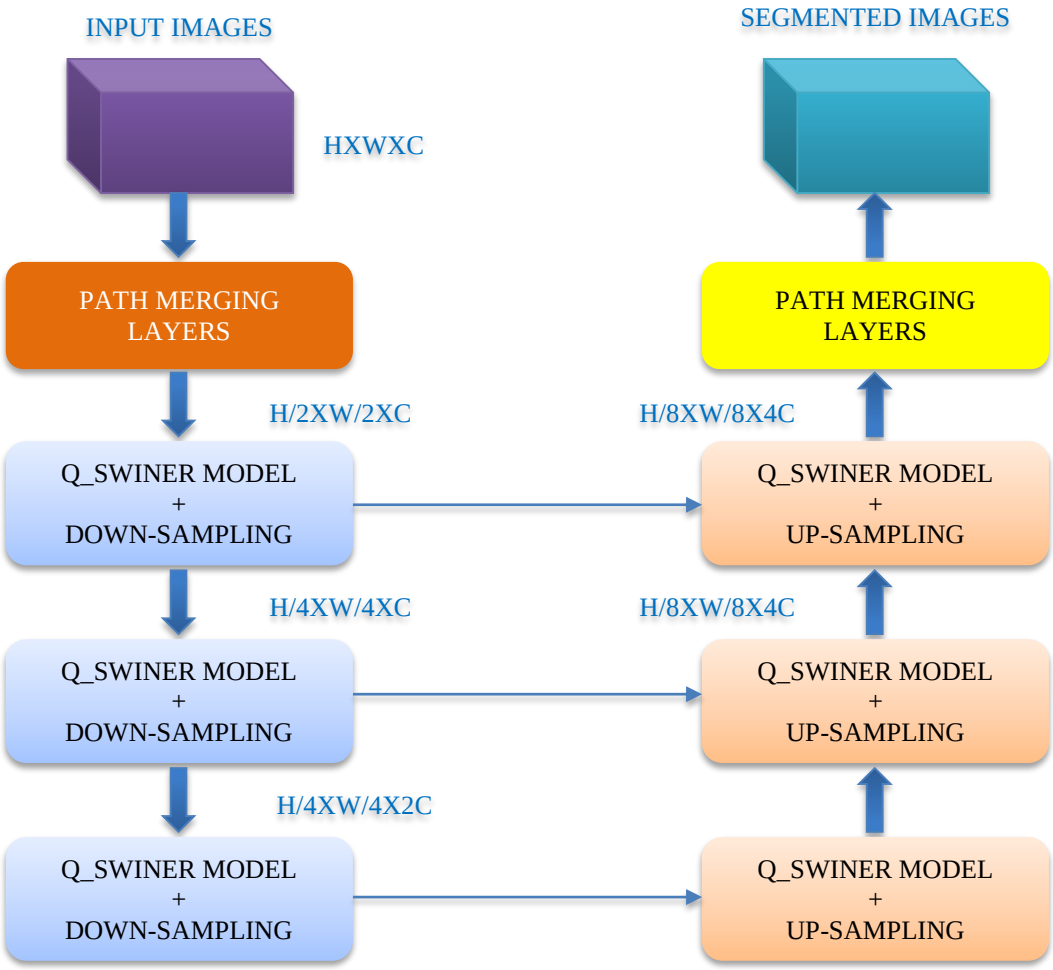


Fig. 10 U-Shaped Q-SWINER model deployed for segmentation process

Figure 10 depicts the segmentation model (Q-SWINER as a backbone), comprising an encoder, decoder, and patch merging. The encoder-decoder follows a U-shaped design adapted to lung images.

Encoder-Decoder Design

The encoder component adopts a Q-SWINER-based network with a downsampling convolutional structure and aims to progressively capture the local structures of the image to be compressed. In the first phase, patch merging layers convert the images into different patches (32x32). The extracted patches are then fed into the proposed quantum encoder, which consists of three stages that reduce the number of features as the network becomes deeper and more hierarchical. In the first three stages, features are concatenated after being input into the networks, reduced in feature resolution, and increased in dimension after the proposed networks and the down-sampling network.

The proposed encoder outputs a feature map with the same size as the left side of the structure, denoted as Q(e), and a concatenation operation denoted as ‖ is performed between them. As a result, the relationship between each of the encoder blocks is as follows:

$$Q_{Oe}^{i} = F_{Q(E_i)} \parallel F_{Q(d_i)}, \quad (17)$$

For the Proposed Model

$$Q_{Oe} = \|_{i=1}^{3}\left[F_{Q(E_i)} \parallel F_{Q(d_i)}\right], \quad (18)$$

Here, F_Q(E) denotes the function generated by each encoder phase. F_Q(dp) is the output from each downsampling layer.

In Equations (17) to (20), the symbol ‖ denotes feature concatenation along the channel dimension, and the index i runs over the three encoder/decoder stages.

The decoder primarily comprises three phases. Unlike previous U-Net architectures, the proposed decoder consists of a Q-SWINER block, an upsampling operation, and a skip connection. Specifically, the decoder will feed (decode) the information given by (encoded) the encoder. At each decoder stage, the feature maps are upsampled by 2 and fused with the encoded skip-connection features before being fed to the proposed Swin Transformer block. Incorporating the Q-SWINER block into the decoder has the following benefits:

a. Allows the decoder to effectively exploit the encoder's extracted features during the up-sampling process.
b. Improves decoding performance by capturing extended contextual relationships.

After these three phases, the resulting features have a resolution of L/4 × H/4. The decoder then combines these output features as PML layers to produce the final mask predictions through skip connections. These connections combine multi-scale features from the encoder with upsampled features from the decoder. The mathematical expression of each decoder stage that follows is:

$$D_{Qe}^{i} = O_{Qe}^{i} \parallel F_{Q(u_i)}, \quad (19)$$

For the recommended framework, the decoder phase is structured as outlined below.

$$O_Qe = \|_{i=1}^{3}[O_Qe\hat{}(i) \parallel F_Q(u_i)], \quad (20)$$

To improve feature extraction, we employ a Q-MHSA block after the segmentation model.

In the proposed approach, we pass the output of the Q-MHSA layers to the classification networks, where we replace classical MLP layers with Variational Quantum Circuits (VQC) that can work with the rich features provided by the Q-MHSA model.

As mentioned in Ref. [14, 34-39], VQC has three steps; it uses parameterized quantum gates, which can be represented as a vector $\vec{\theta}$, where each theta_i rotates around a certain axis using the corresponding unitary operator in a two-dimensional state space.

We omit the first step, as only the entangled quantum bits described in Equation (21) are active in the VQC. As a result, we use this equation as input to QMLP layers. Once the quantum bits are acquired, an ansatz A (analogous to a traditional neural network, an ordered set of gates which transform the prepared quantum state $|\varphi(\vec{x})\rangle$ by means of parameters $\vec{\theta}$:

$$|\psi(x,\theta)\rangle = A(\theta)|\varphi(x)\rangle, \quad (21)$$

Equation (21) is applied to the features from the Q-MHSA module. This technique uses quantum bits to improve classification. Ultimately, an observable $\hat{O}$ is measured on the output state $|\psi(x,\theta)\rangle$, where one of the eigenvalues of $\hat{O}$ is recorded as the prediction, similar to Equation (21):

$$\langle O\rangle = \langle\psi(x,\theta)|\hat{O}|\psi(x,\theta)\rangle, \quad (22)$$

Finally, we use the measured observable to optimize the parameters and achieve optimal performance with a classical optimization algorithm. The error function computed by utilizing the binary cross-entropy function is computationally expressed as:

$$L = -\frac{1}{K}\sum_{i=1}^{K}[y_i\, log(y'_i) + (1-y_i)\, log(1-y'_i)] + \eta\|\theta\|^2, \quad (23)$$

Where K is the number of training samples, y(i) and y′(i) are the true label and the predicted probability of the i-th sample, η is the regularisation coefficient, and ‖θ‖ is the Euclidean norm of the trainable-parameter vector. Algorithm 1 presents the proposed framework's operational procedure.

3.4. Quantum Classification Layers

Step	**Algorithm – Q-SWINER: Hybrid Quantum–Classical Model for Lung Image Classification**
Input:	Pre-processed Lung CT Images
Output:	Classification of Lung Images (Benign/Malignant)
1	Initialize the process and load the preprocessed lung image dataset.
2	Construct Quantum Multi-Head Self-Attention (Q-MHSA) using quantum circuits for parallel feature interaction and entanglement-based attention mapping.
3	Build the Q-SWINER hybrid model by integrating quantum attention blocks with the classical Swin Transformer backbone.

4	Perform segmentation to isolate the lung region using the trained Q-SWINER model.
5	Extract multi-scale spatial–spectral features using Q-MHSA-enhanced layers.
6	Forward the extracted features to the Quantum Variational Classification Network (Q-VCN) for final class prediction.
7	Evaluate performance using Precision, Accuracy, F1-Score, Recall, and AUC.
8	End process.

4. Result and Discussion

4.1. Implementation Details

We implemented the proposed framework in Python using TensorFlow, Keras, Scikit-learn, NumPy, and Pandas.

We simulated the quantum circuits in TensorCircuit and with the IBM Qiskit simulator.

The model was trained for 50 epochs using the Adam optimizer with a learning rate of 2×10^{-4}, a batch size of 48, and the cross-entropy loss function. Training was done on a workstation with an AMD Ryzen CPU, an NVIDIA Tesla GPU (model and memory to be specified), a 2 TB Hard Disk Drive (HDD), and 16 GB of RAM. Table 2 lists the hyperparameters used to create the proposed framework.

Block 2: quantum simulation details: Because simulating multi-qubit circuits directly on the full 512×512 images on a classical workstation is impossible, we simulate the quantum circuits on window-level token embeddings rather than raw pixels. The ZFeatureMap maps a d_h-dimensional feature vector into [number of qubits per head] qubits, followed by [L] depth (CNOT and ISWAP entanglement and RZ–RY–RZ rotations), and then the classification head maps a [number of qubits] dimensional feature vector into [number of qubits] qubits with an ansatz of depth [L_c]. The circuits were simulated in [state-vector mode with exact expectation values / with N shots], and gradients were calculated using [the parameter-shift rule / back-propagation through the simulator], and no hardware noise model was used [confirm]. This workstation [confirm] provides the following reported training time for the above configuration [Dataset-1].

Table 2. Hyperparameters utilized for training the recommended approach

Sl.No	Hyperparameters	Explanation
1	Learning rate	0.0002
2	No of Epochs	50
3	Batch Size	48
4	No. of Hidden Layers	120

4.2. Evaluation Metrics-Segmentation Model

To analyze the segmentation efficiency of the proposed hybrid quantum approach, we use evaluation measures such as DICE, IoU, PSNR, precision, and SSIM. Table 3 presents the calculated metrics in mathematical form. To improve efficacy, we compare the segmentation results with classical Swin Transformer approaches.

Table 3. Mathematical formulations for evaluating the segmentation metrics

Sl.no	Performance Measures	Representation
1	DICE (DSC)	$2\|S \cap T\| / (\|S\| + \|T\|)$
2	SSIM	$[(2\mu S \cdot \mu T + C1)(2\sigma ST + C2)] / [(\mu S^2 + \mu T^2 + C1)(\sigma S^2 + \sigma T^2 + C2)]$
3	IoU	$\|S \cap T\| / \|S \cup T\|$
4	Peak Signal-to-Noise Ratio (PSNR)	$10\ log10(MAXI^2 / MSE)$

S is the ground-truth mask, and T is the model's predicted mask. DSC and IoU range from 0–1, with 0 indicating no overlap and 1 indicating perfect segmentation.

PSNR (MAX is the Maximum pixel value, MSE is the Mean Squared Error) and SSIM (μ, σ, and σST are the mean, variance, and covariance of the two images, and C1 and C2 are stabilizing constants) measure image fidelity. In all four measures, a higher score indicates a better prediction and, thus, better segmentation.

4.3. Segmentation Results

4.3.1. Segmentation Experimental Outcomes (Dataset-1)

Figures 11- 14 show the segmentation efficiency of the proposed hybrid approach compared with the Swin Transformer (as a baseline).

The segmentation performance is computed for increasing numbers of epochs. Figure 11 compares the PSNR of the recommended hybrid approach with that of the classical Swin Transformer.

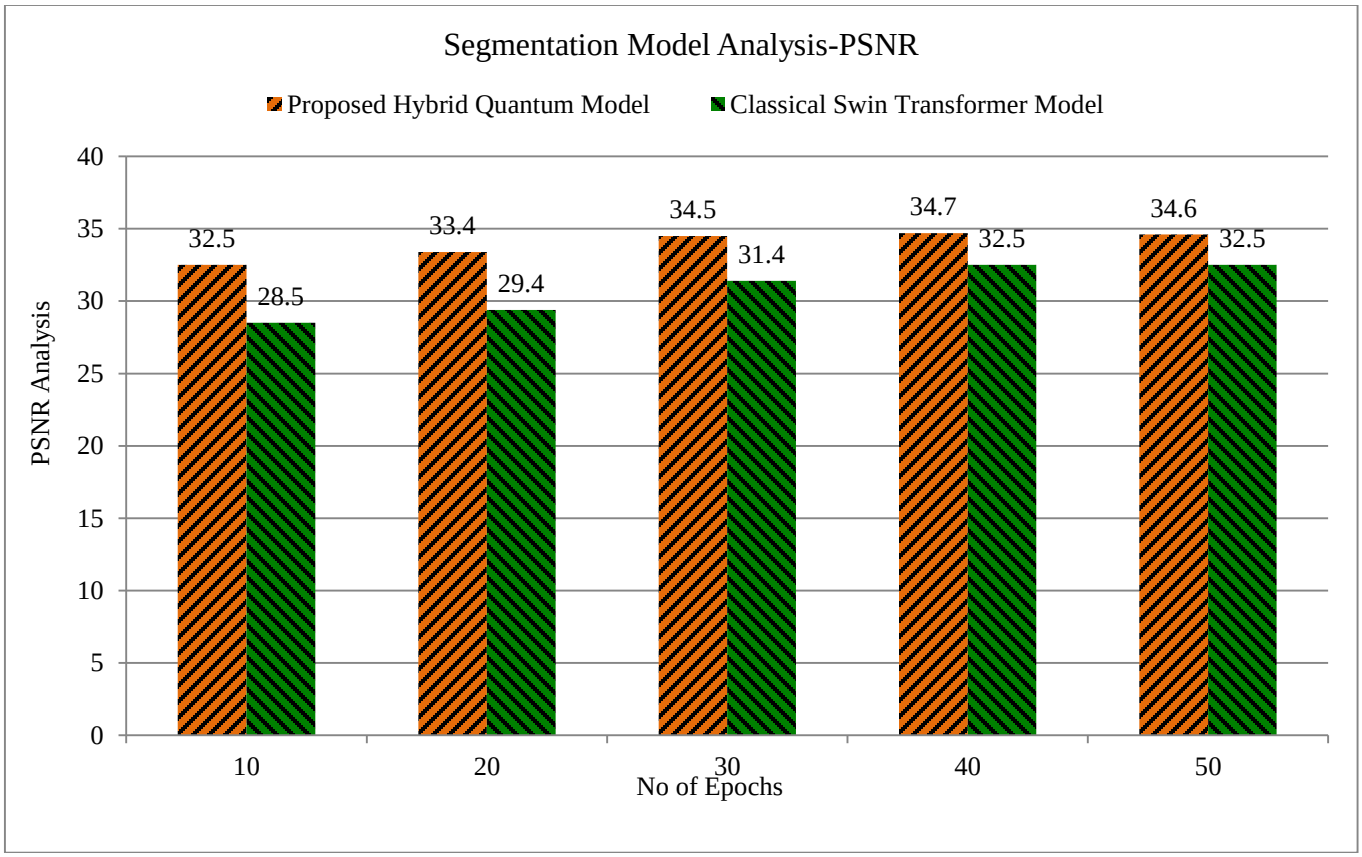


Fig. 11 PSNR of the recommended hybrid approach and the classical Swin Transformer model as epochs change

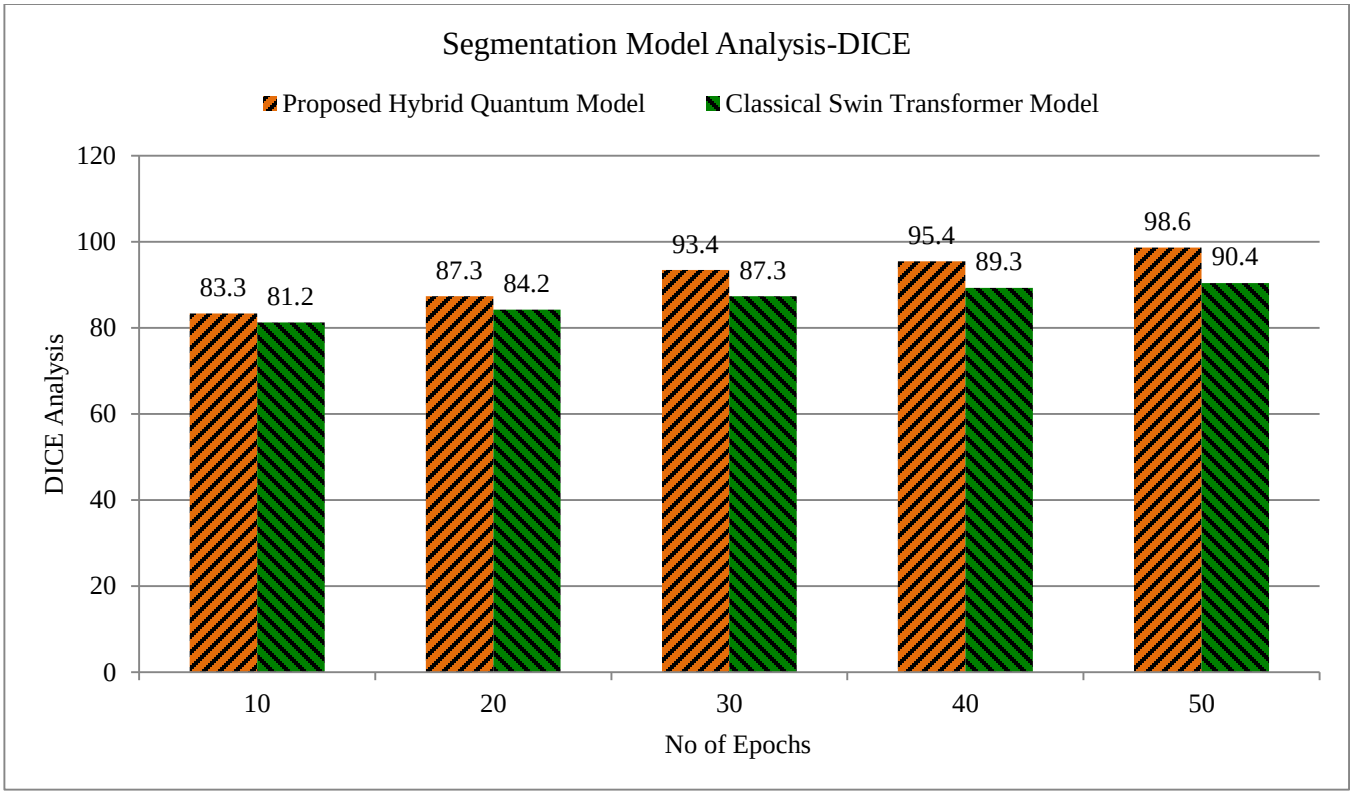


Fig. 12 DICE of the recommended hybrid model and the classical Swin Transformer model as epochs change

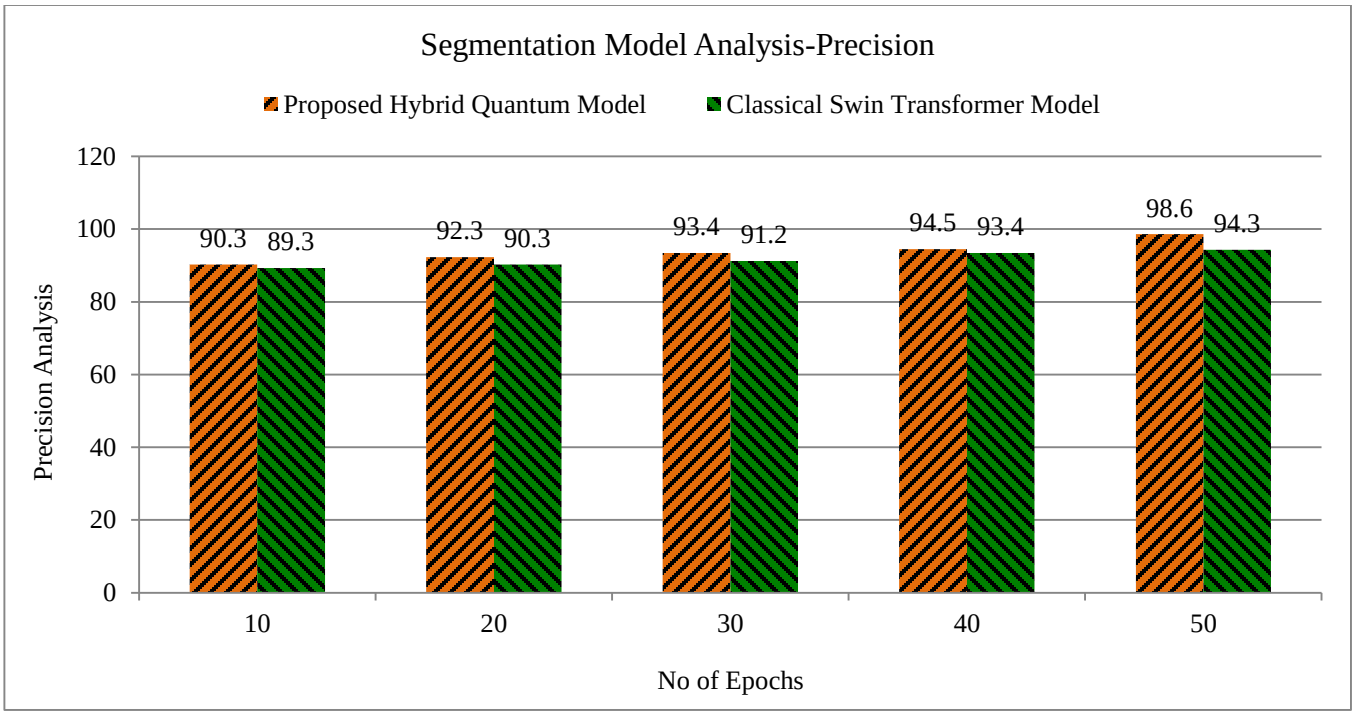


Fig. 13 Segmentation precision performance of the recommended hybrid approach and the classical Swin Transformer model as epochs change

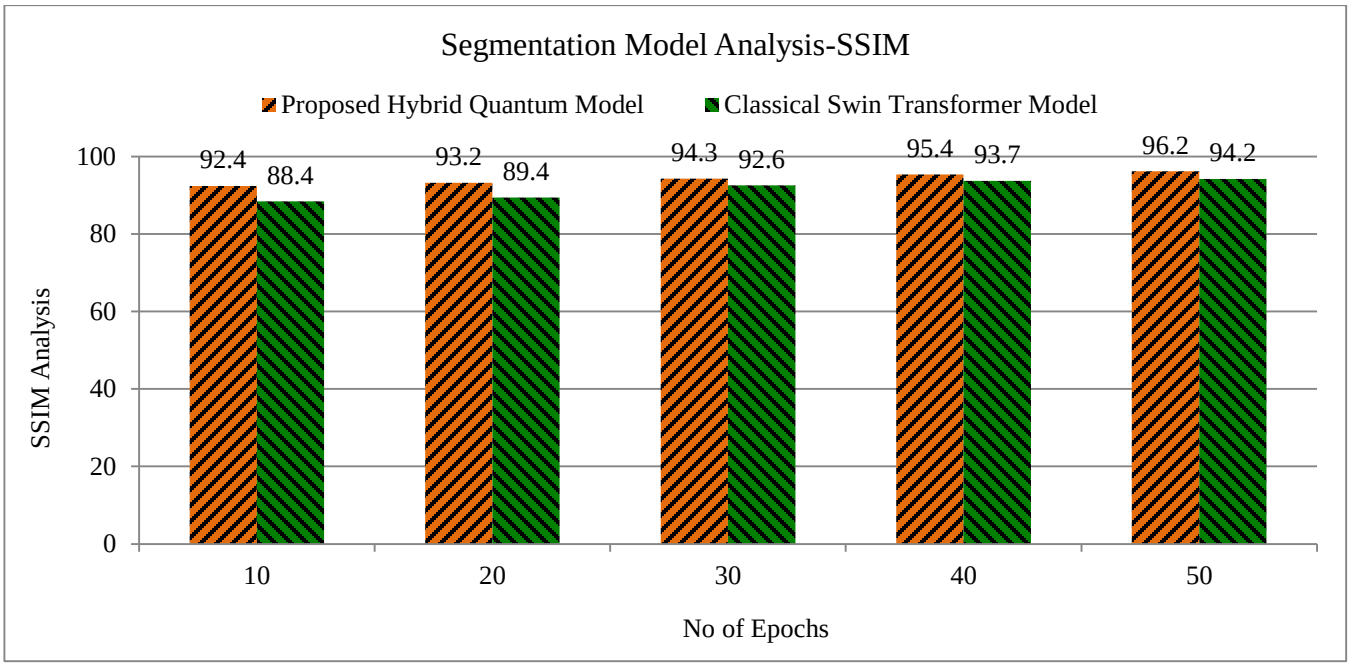


Fig. 14 Segmentation SSIM performance of the recommended hybrid approach and classical Swin Transformer model with changes in the number of epochs

The proposed model achieves a PSNR of 32.5 dB at 10 epochs and 34.6 dB at 50 epochs. Meanwhile, the classical Swin Transformer model achieves PSNRs of 28.5 and 32.5 at 10 and 50 epochs, respectively. With more epochs, the recommended approach achieves a DICE efficiency of 98.3, while the classical Swin Transformer achieves 95.3, as shown in Figure 12. Similarly, Figures 13 and 14 present the segmentation precision and SSIM performance of the proposed model and the classical Swin Transformer, respectively. As shown in Figures 11-14, it is clear that the cancer nodules were segmented from the lung cancer images using lung datasets, which shows that the recommended hybrid framework provided better results compared to the classical Swin Transformer model.

4.3.2. Segmentation Experimental Outcomes (Datasets-2)

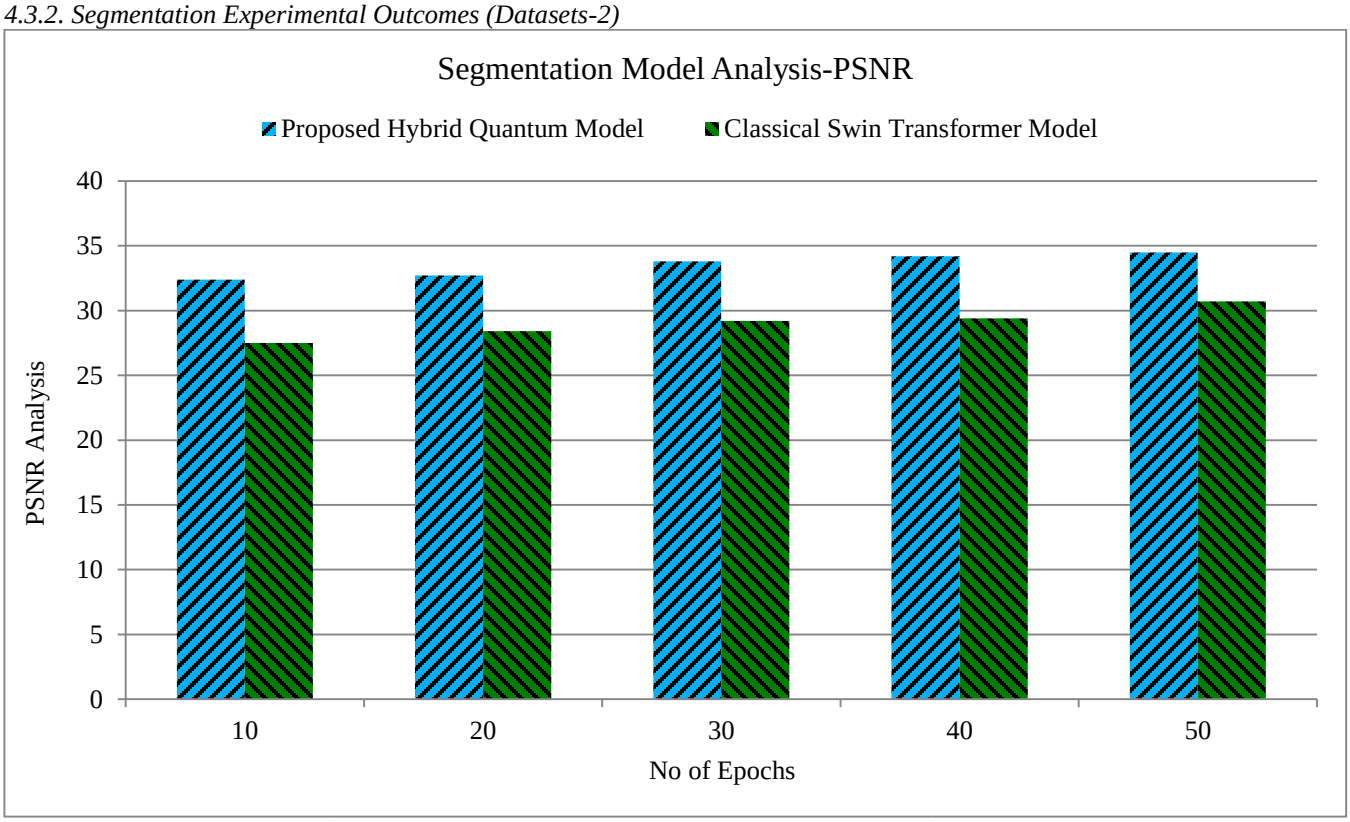


Fig. 15 PSNR of the recommended hybrid approach and the classical Swin Transformer model across epochs

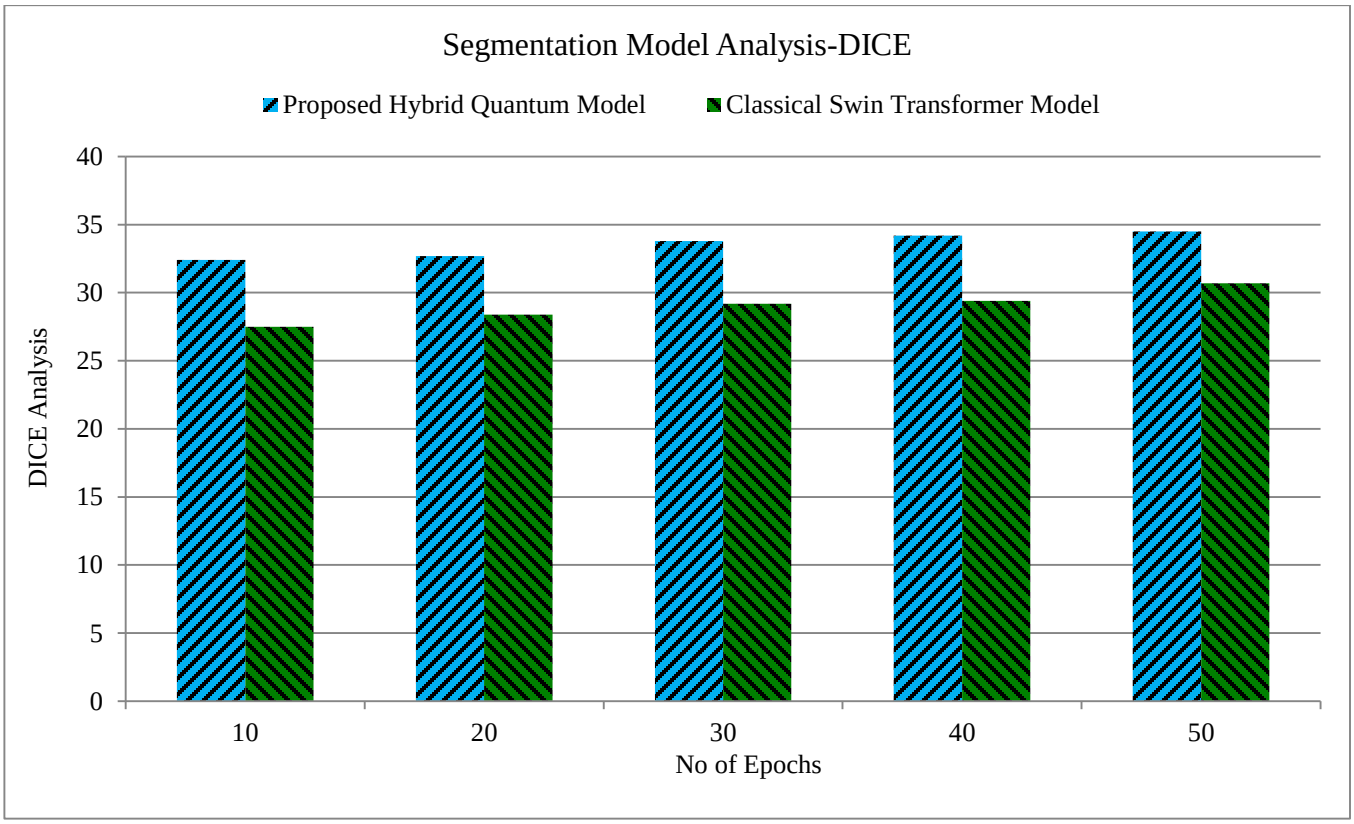


Fig. 16 DICE of the recommended hybrid approach and the classical Swin Transformer model as epochs change

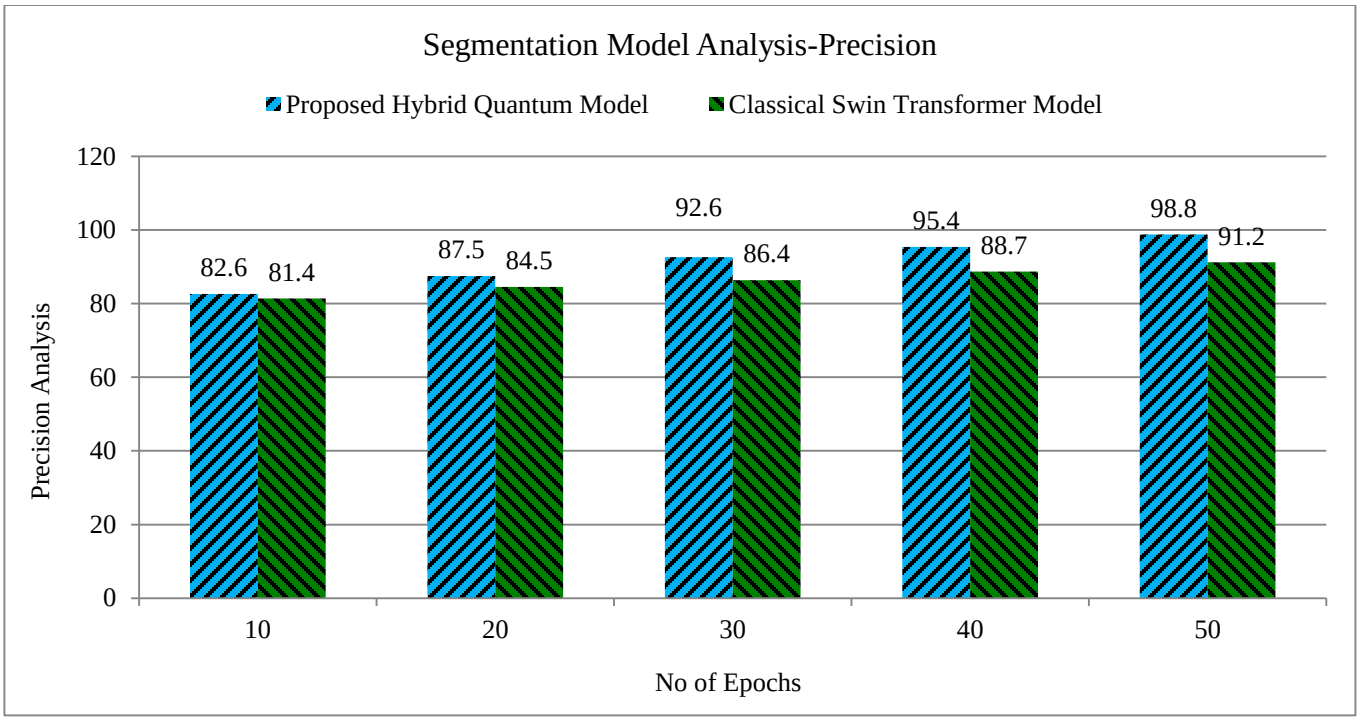


Fig. 17 Segmentation precision performance of the recommended hybrid approach and the classical Swin Transformer model as epochs change

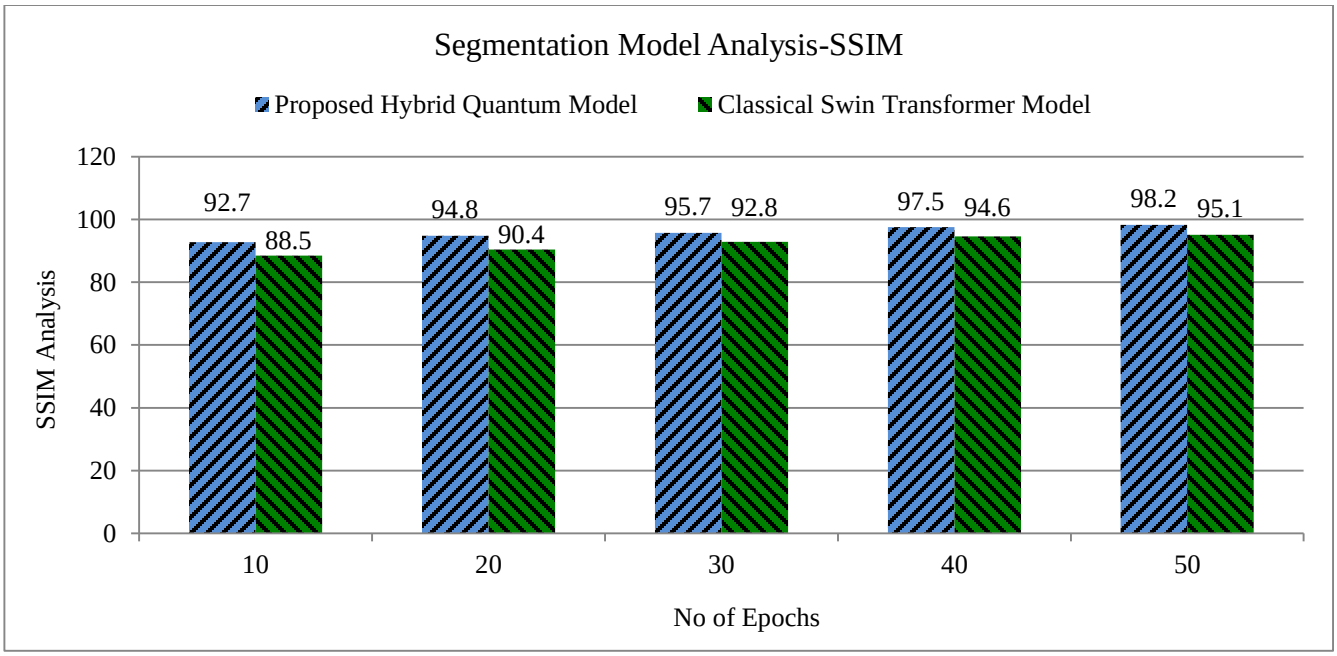


Fig. 18 SSIM performance of the recommended hybrid approach and the classical Swin Transformer model as the number of epochs changes

From Dataset 2, the proposed model reaches a PSNR of 32.4 at 10 epochs and 34.5 at 50 epochs. However, the classical Swin Transformer model yields PSNR values of 28.5 and 32.5 at 10 and 50 epochs, respectively (Figure 15). As shown in Figure 16, increasing the maximum number of epochs improves the DICE performance of the proposed approach and the classical Swin Transformer to 98.3 and 95.3, respectively. The segmentation precision and SSIM of the proposed model and the classical Swin Transformer are similar, as shown in Figures 17 and 18, respectively. Figures 15-18 show that the proposed hybrid model outperforms the classical Swin Transformer on lung cancer images when segmenting nodules. The proposed model also performed well on the larger dataset, while the classical transformer's segmentation performance decreased slightly.

4.4. Evaluation Metrics

We evaluate and directly compare precision, accuracy, recall, specificity, and F1-score with State-of-the-Art (SoTA) quantum self-attention models to establish the proposed model's superiority. To show that the proposed framework improves performance without increasing computational overhead, we compute all evaluation measures. Table 4 presents the expressions used to calculate the performance measures.

Table 4. Evaluation metrics used for experimenting with the proposed model

Performance Metrics	Mathematical Expression
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Recall	$\frac{TP}{TP + FN}$
Specificity	$\frac{TN}{TN + FP}$
Precision	$\frac{TP}{TP + FP}$
F1-Score	$2 \times \frac{Precision \times Recall}{Precision + Recall}$

4.5. Evaluation Outcomes

4.5.1. Experimental Findings using Dataset-1

Figures 19-23 show the effectiveness of the recommended model and the classical Swin Transformer in detecting lung cancer using Dataset-1.

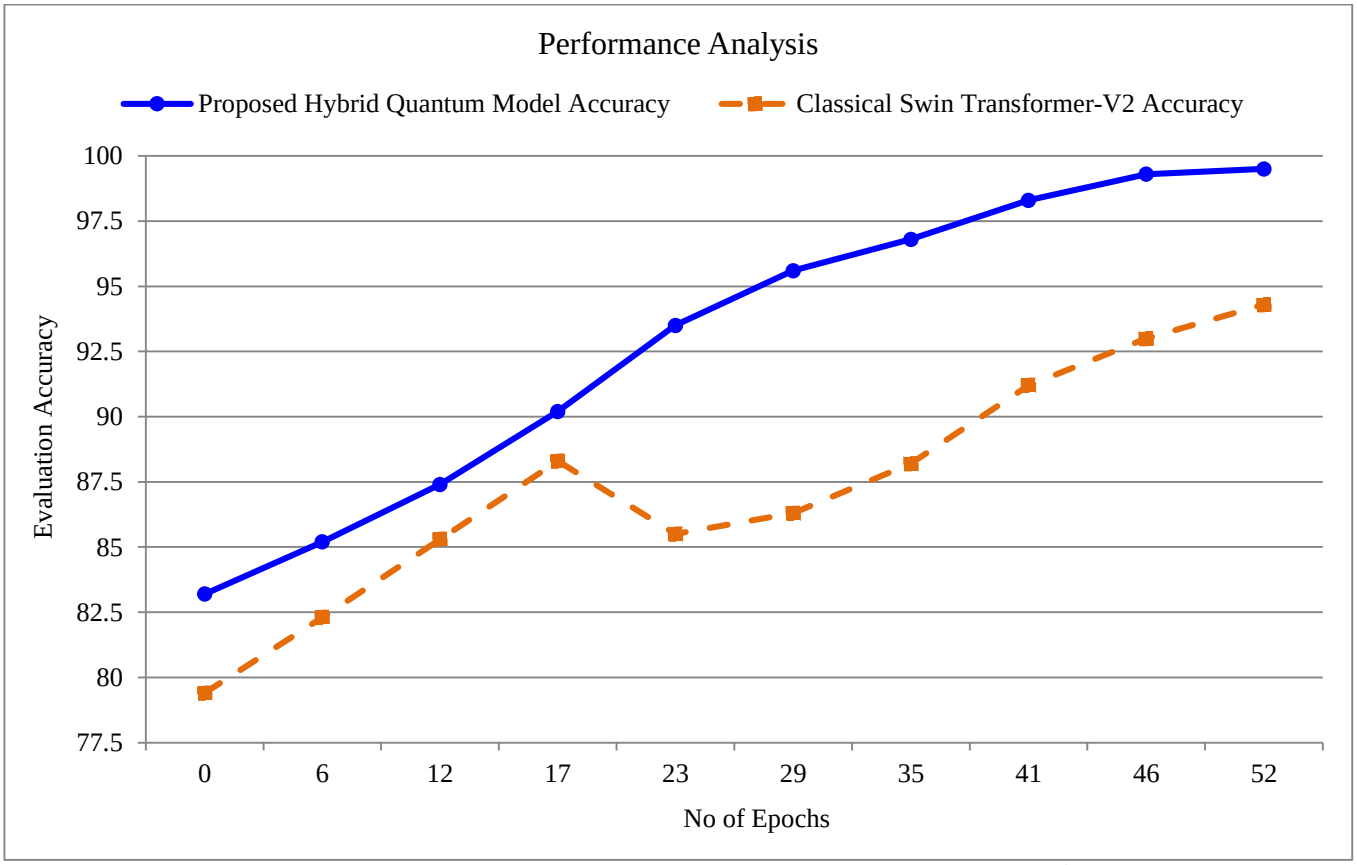


Fig. 19 Accuracy between the recommended quantum hybrid approach and classical transformers

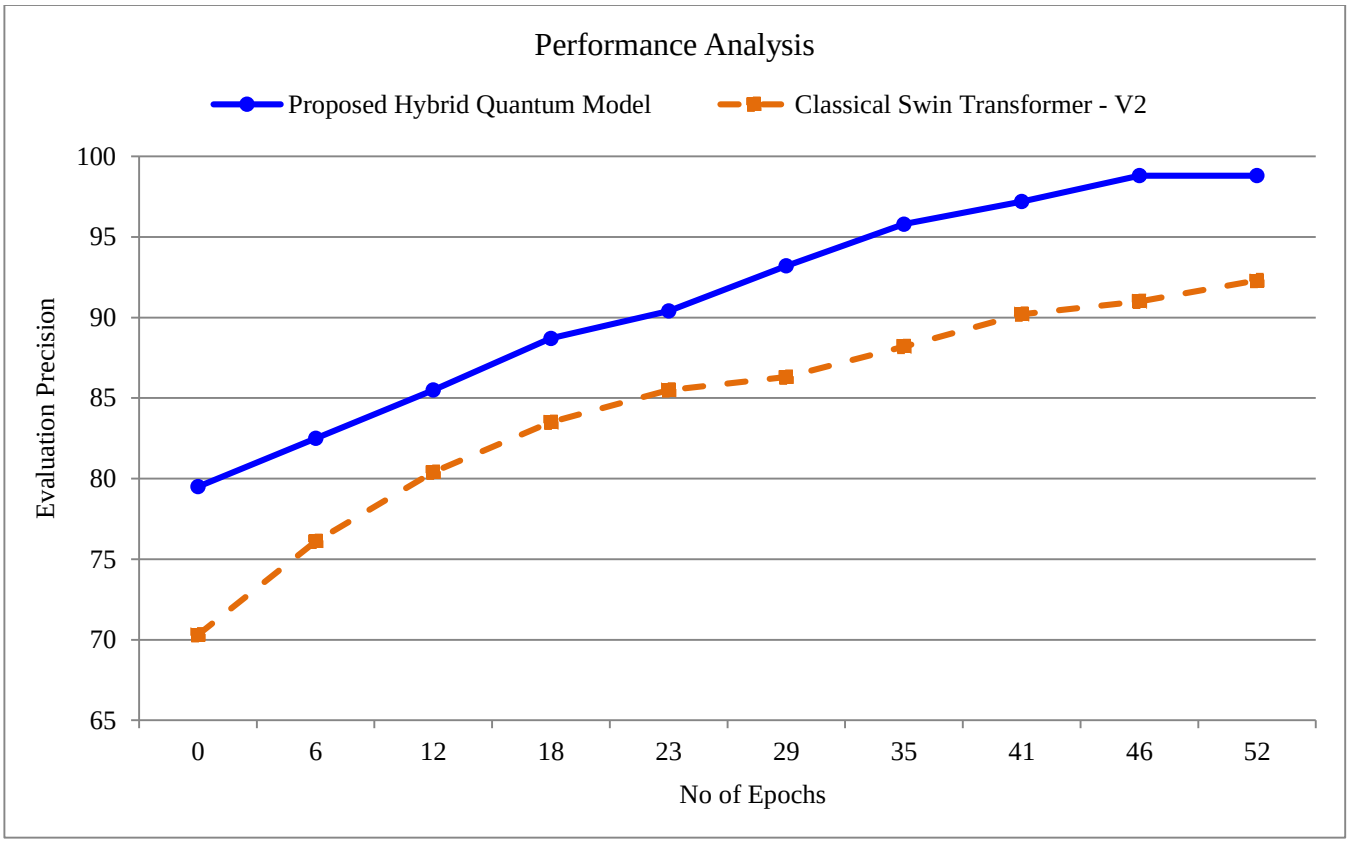


Fig. 20 Precision between the proposed quantum hybrid model and classical transformers

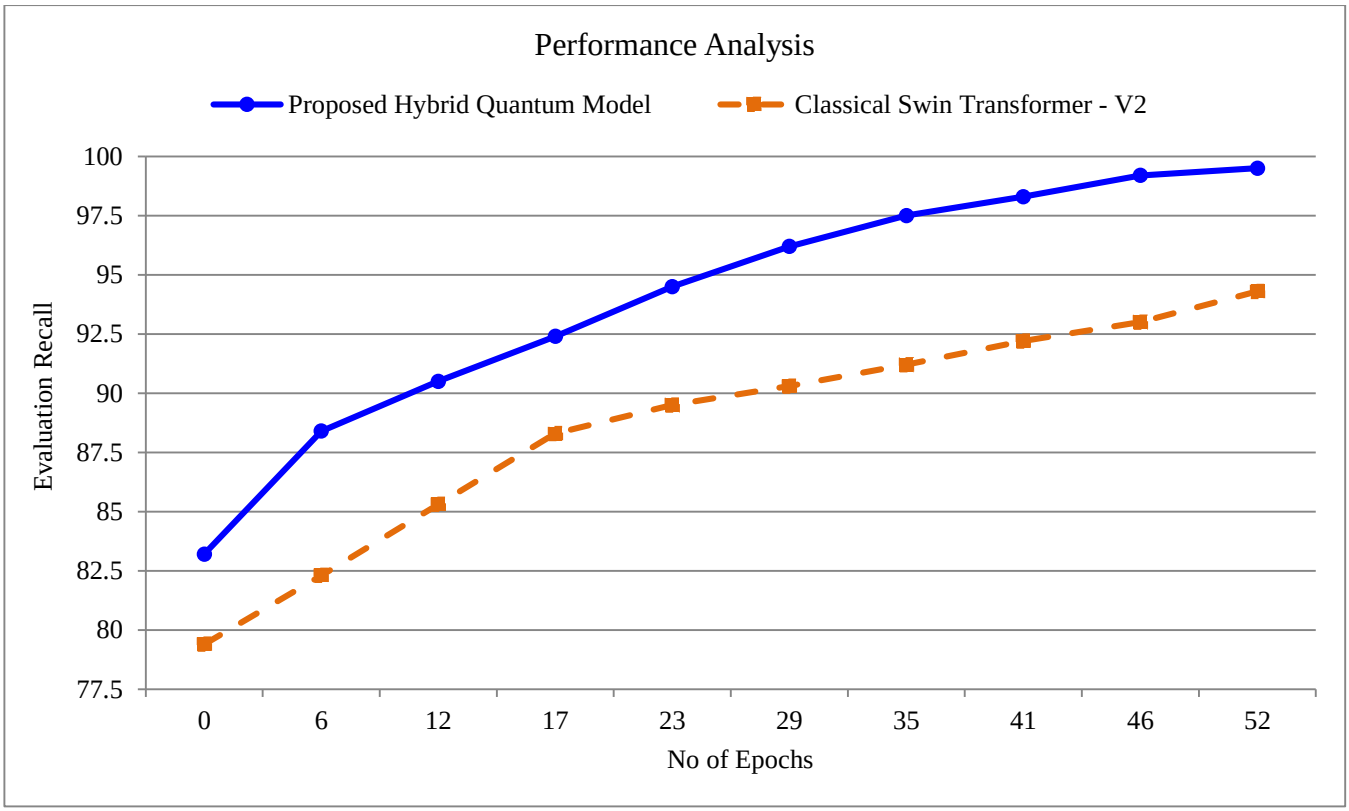


Fig. 21 Recall between the recommended quantum hybrid approach and classical transformers

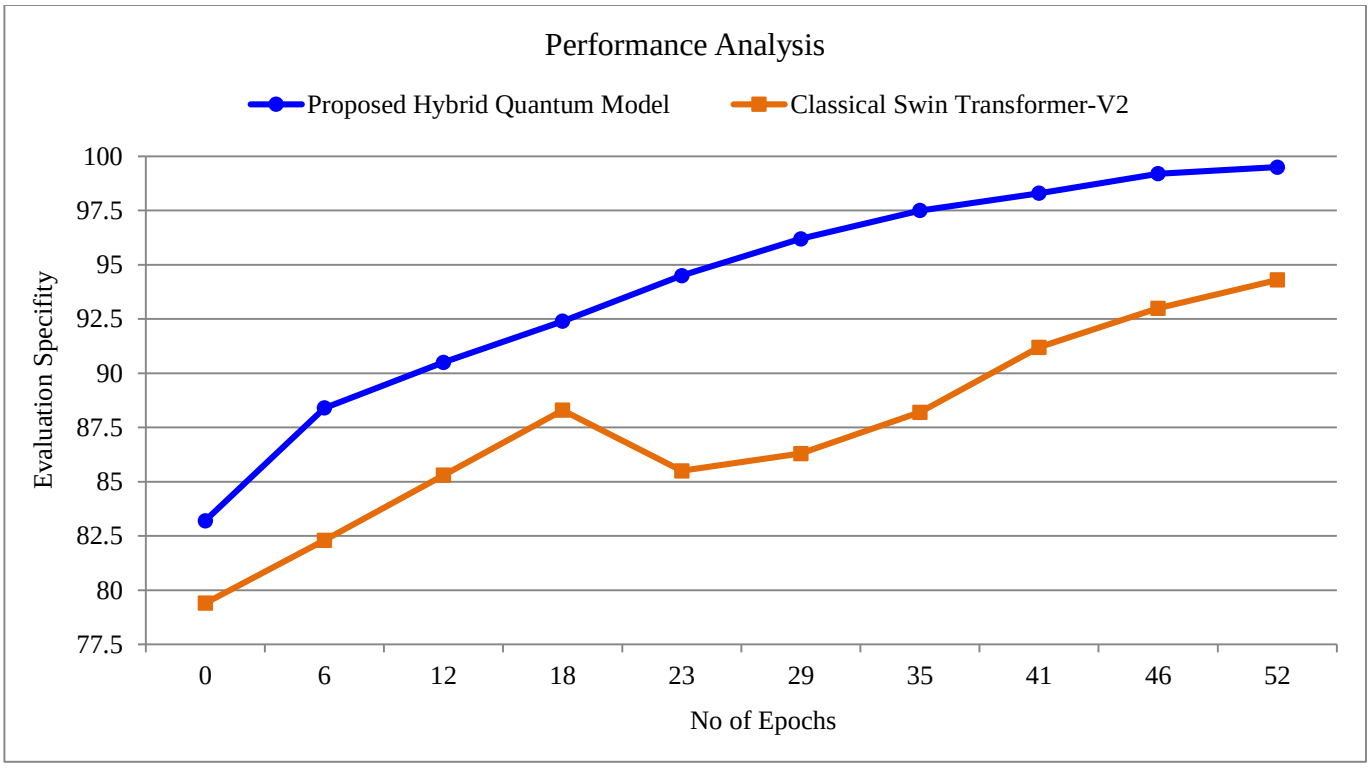


Fig. 22 Specificity performance between the proposed quantum hybrid model and classical transformers

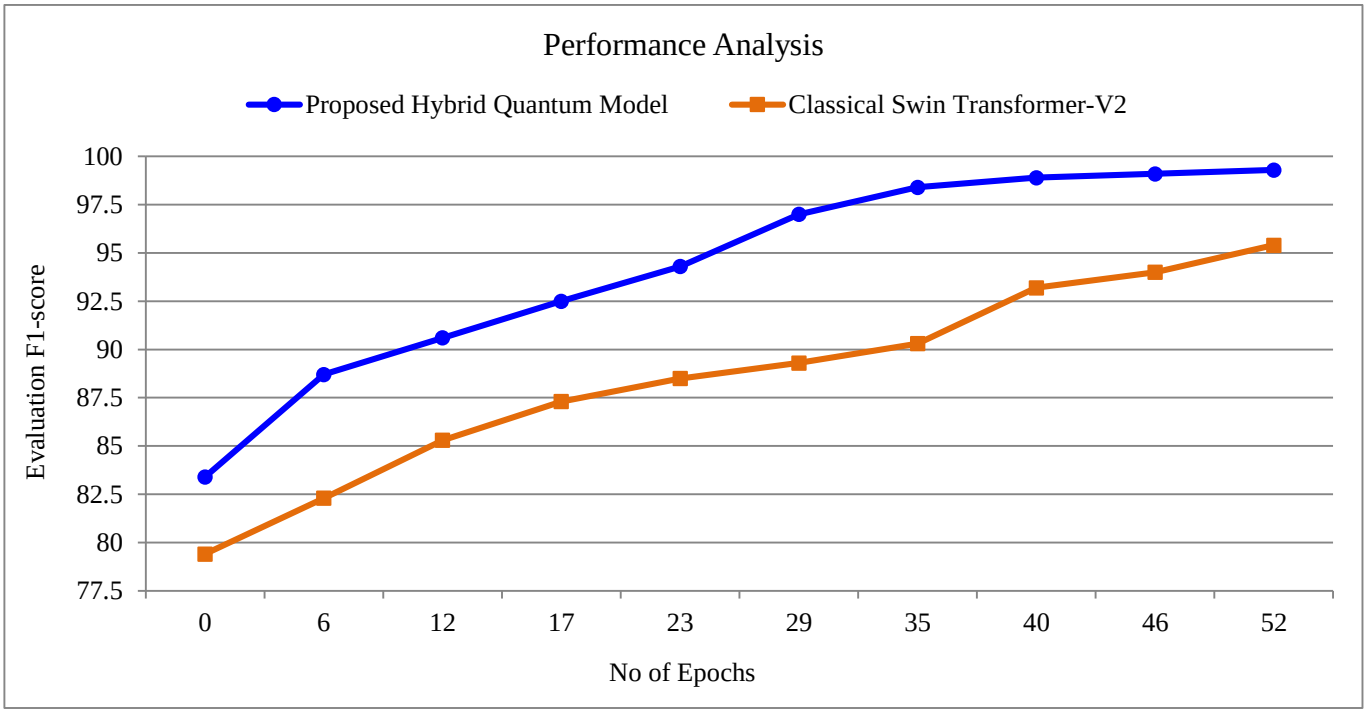


Fig. 23 F1-Score performance between the recommended quantum hybrid approach and classical transformers

The hybrid quantum MHSA improves the Q-SWINER accuracy to 99.2%, precision of 98.9%, recall of 98.7%, specificity of 99%, and the F1 score to 99.0% for Dataset-1, as shown in Figures 19-23. These results indicate that quantum-enhanced attention scales well with high-dimensional, heterogeneous image data. On the large dataset, the classical Swin Transformer achieved 95.3% accuracy, 95% precision, 94.5% recall, 94.5% specificity, and an F1-score of 95% for lung cancer detection.

4.5.2. Result Outcomes using Datasets-2

The progressive classification performance results of the proposed Q-SWINER and classical Swin Transformer models are presented in Tables 5–9, respectively, for 10, 20, 30, 40, and 50 epochs of training. The tables provide examples of the learning progression of both models as training continues on Dataset-2.

Table 5. Performance of the different models in identifying the Lung Cancers @Epochs =10

Algorithms	Performance measures (%)				
	Accuracy	Precision	Recall	Specificity	F1-score
Q-SWINER	84.0	83.2	83.0	84.2	84.3
Classical SWIN	80.2	80	79.3	80	80.2

Table 6. Performance of the different models in identifying the Lung Cancers @Epochs =20

Algorithms	Performance measures (%)				
	Accuracy	Precision	Recall	Specificity	F1-score
Q-SWINER	86.3	86.0	85.9	85.6	86.0
Classical SWIN	82.4	82.0	81.7	80.5	81.7

Table 7. Performance of the different models in identifying the Lung Cancers @Epochs =30

Algorithms	Performance measures (%)				
	Accuracy	Precision	Recall	Specificity	F1-score
Q-SWINER	91.5	90.5	90.3	91.0	90.5
Classical SWIN	85.2	84.3	83.5	82.8	82.8

Table 8. Performance of the different models in identifying the Lung Cancers @Epochs =40

Algorithms	Performance measures (%)				
	Accuracy	Precision	Recall	Specificity	F1-score
Q-SWINER	95.4	95.2	95.0	94.6	95.0
Classical SWIN	88.4	87.5	86.4	85.4	85.6

Table 9. Performance of the different models in detecting the Lung Cancers @Epochs =50

Algorithms	Performance measures (%)				
	Accuracy	Precision	Recall	Specificity	F1-score
Q-SWINER	98.3	97.5	97.3	97.1	97.2
Classical SWIN	90.5	89.2	88.9	87.4	89.0

The results in Tables 5-9 show that both models improve as the number of training epochs increases. As training progresses, accuracy, precision, recall, specificity, and F1-score steadily improve in the Q-SWINER framework, suggesting improved feature learning and classification performance. Comparison with other models in Tables 5-9 shows that the proposed hybrid model's lung cancer detection capability does not degrade when applied to Dataset-2. It achieved the best accuracy of 98.3%, precision of 97.5%, recall of 97.3%, and an F1-score of 97.2% in identifying cancerous nodules in lung cancer images. In contrast, the Swin Transformer achieved an accuracy of 90.5%, precision of 89.2%, recall of 88.9%, and an F1-score of 89.0%, and the larger dataset size posed a performance challenge. The experiments demonstrate that the Q-SWINER framework is applicable to various kinds of data and produces stable, effective results in lung cancer detection.

4.6. Comparative Analysis by Different Works

Existing algorithms based on quantum principles are considered for evaluating the proposed framework. Tables 10 and 11 show the results of the different approaches for classifying lung cancer on Dataset-1 and Dataset-2, respectively. The proposed model results in Table 11 match the epoch-50 results for Dataset-2 in Table 9. The baseline entries are those with [re-run] and should be replaced by the results from the re-run of the corresponding model on the same protocol, using Dataset-2.

Table 10. Performance assessment of the distinct algorithms in identifying the Lung Cancers utilizing dataset-1

Algorithms	Performance Measures				
	Accuracy	Precision	Recall	Training Time(Hrs)	F1-score
QMSAN-NN-Np	90%	90%	90%	15	90%
QMSAN-AA-NP	91.2%	90%	90%	18	90%
QMSAN-CB-NP	90.4%	90%	89.5%	18	89.9%
QSANN	91.3%	91%	90%	25	90%

DISCOCT	84.3%	83.5%	83%	40	83.5%
QSAN-ANA-NP	92.4%	92.0%	91.9%	18	91.5%
Proposed Model	99.2%	98.9%	98.7%	18	99.0%

Table 11. Performance analysis of the varied approaches in detecting Lung Cancer utilizing dataset-2

Algorithms	Performance Measures				
	Accuracy	Precision	Recall	Training Time(Hrs)	F1-score
QMSAN-NN-Np[35]	88.4%	87.5%	86.4%	15	86.9%
QMSAN-AA-NP[35]	89.4%	89.2%	89.0%	18	89%
QMSAN-CB-NP	88.4%	88.2%	88.0%	18	88.3%
QSANN	91.2%	90.4%	89.3%	25	89.2%
DISCOCT	80.4%	79.4%	79.2%	40	78.3%
QSAN-ANA-NP	92.4	92	91.9	18	91.5
Proposed Model	98.3%	97.5%	97.3%	18	97.2%

As the number of data points increased, Q-SWINER's performance remained consistent, while the other models' performance degraded. Thus, Q-MHSA (parameterized CNOT and ISWAP) combined with the ZFeatureMap can detect lung cancer across various datasets. Furthermore, a quantum neural network combined with Q-MHSA outperformed the other proposed models in the Swin network.

4.7. Discussion

Q-SWINER's better performance than the classical Swin Transformer and the compared quantum attention models can be attributed to the following design factors. First, the ZFeatureMap maps the window-level features to a high-dimensional Hilbert space, and the CNOT and ISWAP entanglement layers model pairwise dependencies among the window-level features in each head, giving a richer query-key similarity measure than the classical dot-product attention of the baseline Swin Transformer; gains can be seen at each point of training in Tables 5-9. Second, the quantum attention is incorporated into the hierarchical window-based structure of Swin Transformer V2, which preserves the local lesion detail and multi-scale context via shifted windows and scaled cosine attention and Log-CPB, while the quantum attention models (QMSAN, QSANN and QSAN variants) were originally designed for sequence-based data and lack the capability of preserving hierarchical spatial feature extraction. Third, we use a U-shaped encoder-decoder with skip connections to isolate the lesioned region before classification, mitigating the impact of background tissues on the classifier and, as shown in Figures 11–18, resulting in higher DICE and PSNR values. Finally, CLAHE contrast enhancement and GAN-based class balancing mitigate class imbalance.

Quantum models [21-24, 30] report 85% to 99% accuracy on tabular data, gene expression, or low-dimensional features, while classical deep learning models [25, 26, 28, 29] report 90% to 99.2% accuracy. Q-SWINER not only achieves similar or better accuracy on image data but also performs lesion segmentation and classification in one model. However, differences in datasets, input modalities, and evaluation protocols limit direct comparison.

View the gains with caution. The quantum circuits were simulated classically, and the advantage of the quantum components over the equivalent classical model should be verified by ablation studies (classical attention versus Q-MHSA on the same backbone with and without the entanglement layers and reported as mean ± std and compared via statistical tests) and repeated runs reported as mean ± std with statistical tests.

5. Conclusion and Future Enhancement

This paper presents a novel quantum Swin Transformer as a hybrid model that integrates a quantum multi-head self-attention mechanism with a quantum neural network to further improve the classification and segmentation of lung cancer images. The novel framework Q-SWINER consists of four parts: a data collection unit, data pre-processing, a segmentation model, and classification layers. We develop a U-shaped segmentation model based on Q-SWINER to improve segmentation performance on lung cancer images. Furthermore, we integrate a quantum neural network with variational quantum circuits to classify lung cancer images. We conducted a comprehensive evaluation on two datasets using multiple performance metrics, comparing the model's efficiency with the classical Swin Transformer and other State-of-The-Art quantum attention methods. The simulation results show an accuracy of 99.2%, a precision of 98.9%, and an F1 score of 99.0% for the Q-SWINER model on Dataset-1. Q-SWINER also achieved the best lung cancer detection performance across various datasets compared with other models.The Q-SWINER model is a hybrid quantum-classical prototype, tested via classical simulation of the quantum circuits. Further benefits in computational speed and performance may be realized when implemented on quantum hardware, and a full implementation for multiple

heterogeneous datasets is future work. This enhancement would also expand the model's potential applications in real-world healthcare environments and further improve its scalability and efficiency.

Conflict of Interest

The authors assert no conflict of interest

Funding

The authors declared that no funding was received for this Research and Publication

Data Availability

The datasets analyzed during this study are publicly available from The Cancer Imaging Archive [31-33].

Ethical Approval

This article does not involve studies on human individuals or animals by the author.

Author Contribution

S. Radhika and Dr.G. Sharada contributed equally to this work.

References

[1] J. Abhijith et al., "Quantum Algorithm Implementations for Beginners," *ACM Transactions on Quantum Computing*, vol. 3, no. 4, pp. 1-92, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] IBM Quantum Platform, quantum.cloud. [Online]. Available: https://quantum.cloud.ibm.com/

[3] Cirq, An Open Source Framework for Programming Quantum Computers, Google Quantum AI, 2024. [Online]. Available: https://quantumai.google/cirq

[4] Amazon Braket: Accelerate Quantum Computing Research, AWS Amazon. [Online]. Available: https://aws.amazon.com/braket/

[5] QuTiP: Quantum Toolbox in Python, qutip, 2024. [Online]. Available: https://qutip.org/

[6] Mingyu Chen et al., "QCIR: Pattern Matching based Universal Quantum Circuit Rewriting Framework," *Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design*, *Association for Computing Machinery*, New York, NY, USA, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Raban Iten et al., "Exact and Practical Pattern Matching for Quantum Circuit Optimization," *ACM Transactions on Quantum Computing*, vol. 3, no. 1, pp. 1-41, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Md Mazder Rahman, Gerhard W. Dueck, and Joseph D. Horton, "An Algorithm for Quantum Template Matching," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 11, no. 3, pp. 1-20, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[9] Nabila Abdessaied, Mathias Soeken, and Rolf Drechsler, "Quantum Circuit Optimization by Hadamard Gate Reduction," *Reversible Computation*, *Lecture Notes in Computer Science*, vol. 8507, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[10] Hui Jiang et al., "A Pattern Matching-based Framework for Quantum Circuit Rewriting," *Journal of Computer Science and Technology*, vol. 39, pp. 1312-1327, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Chengkai Zhu, and Zhenyu Huang, "Optimizing the Depth of Quantum Implementations of Linear Layers," *Information Security and Cryptology*, vol. 13837, pp. 129-147, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Jonas Nüßlein et al., "Reducing QAOA Circuit Depth by Factoring out Semi-Symmetries," *arXiv Preprint*, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Thomas Fösel et al., "Quantum Circuit Optimization with Deep Reinforcement Learning," *arXiv Preprint*, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Mateusz Ostaszewski et al., "Reinforcement Learning for Optimization of Variational Quantum Circuit Architectures," *Advances in Neural Information Processing Systems*, vol. 34, pp. 18182-18194, 2021. [Google Scholar] [Publisher Link]

[15] Lorenzo Moro et al., "Quantum Compiling by Deep Reinforcement Learning," *Communications Physics*, pp. 1-8, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Tahereh Salehi et al., "An Optimizing Method for Performance and Resource Utilization in Quantum Machine Learning Circuits," *Scientific Reports*, vol. 12, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Samuel A. Stein et al., "QuGAN: A Quantum State Fidelity-based Generative Adversarial Network," *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Broomfield, CO, USA, pp. 71-78, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Anna M. Krol et al., "Efficient Decomposition of Unitary Matrices in Quantum Circuit Compilers," *Applied Sciences*, vol. 12, no. 2, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Hartmut Führ, and Ziemowit Rzeszotnik, "A Note on Factoring Unitary Matrices," *Linear Algebra and its Applications*, vol. 547, pp. 32-44, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[20] Jessica Pointing et al., "Quanto: Optimizing Quantum Circuits with Automatic Generation of Circuit Identities," *Quantum Science and Technology*, vol. 9, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Manushi Munshi et al., "Quantum Machine Learning-based Lung Cancer Prediction Framework for Healthcare 4.0," *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*, Mysore, India, pp. 1-6, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Jason Elroy Martis et al., "Novel Hybrid Quantum Architecture-based Lung Cancer Detection using Chest Radiograph and Computerized Tomography Images," *Bioengineering*, vol. 11, no. 8, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Shivanya Shomir Dutta et al., "Quantum Kernel-based Support Vector Machine with Quantum Approximate Optimization Algorithm Embedding for Improved Lung Cancer Prediction," *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, pp. 1-6, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Mohadeseh Zarei Ghobadi, and Elaheh Afsaneh, "Potential of Quantum Machine Learning for Solving the Real-World Problem of Cancer Classification," *Discover Applied Sciences*, vol. 6, pp. 1-12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Lavina Jean Crasta, Rupal Neema, and Alwyn Roshan Pais, "A Novel Deep Learning Architecture for Lung Cancer Detection and Diagnosis from Computed Tomography Image Analysis," *Healthcare Analytics*, vol. 5, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Vinod Kumar et al., "Unified Deep Learning Models for Enhanced Lung Cancer Prediction with ResNet-50–101 and EfficientNet-B3 using DICOM Images," *BMC Medical Imaging*, vol. 24, pp. 1-21, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[27] Lalasa Mukku, and Jyothi Thomas, "Exploring the Potential of Quantum Machine Learning in Cancer Research," *Annals of Medical and Health Sciences Research*, vol. 13, pp. 690-694, 2023. [Google Scholar] [Publisher Link]

[28] Abdul Rahaman Wahab Sait, "Lung Cancer Detection Model using Deep Learning Technique," *Applied Sciences*, vol. 13, no. 22, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] Asghar Ali Shah et al., "Deep Learning Ensemble 2D CNN Approach towards the Detection of Lung Cancer," *Scientific Reports*, vol. 13, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[30] Siddhant Jain et al., "Quantum and Classical Machine Learning for the Classification of Non-Small-Cell Lung Cancer Patients," *SN Applied Sciences*, vol. 2, pp. 1-10, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[31] Samuel G. Armato III et al., "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans," *Medical physics*, vol. 38, no. 2, pp. 915-931, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[32] Li P et al., "A Large-Scale CT and PET/CT Dataset for Lung Cancer Diagnosis (Lung-PET-CT-Dx)," *The Cancer Imaging Archive*, 2020. [CrossRef] [Publisher Link]

[33] Kenneth Clark et al., "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *Journal of Digital Imaging*, vol. 26, pp. 1045-1057, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[34] Danyal Maheshwari, Daniel Sierra-Sosa, and Begonya Garcia-Zapirain, "Variational Quantum Classifier for Binary Classification: Real vs Synthetic Dataset," *IEEE Access*, vol. 10, pp. 3705-3715, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[35] Fu Chen et al., "Quantum Mixed-State Self-Attention Network," *Neural Networks*, vol. 185, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[36] Ren-Xin Zhao, Jinjing Shi, and Xuelong Li, "Qksan: A Quantum Kernel Self-Attention Network," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 10184-10195, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[37] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer, "Effect of Data Encoding on the Expressive Power of Variational Quantum-Machine-Learning Models," *Physical Review A*, vol. 103, no. 3, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[38] Jin Zheng, Qing Gao, and Zibo Miao, "Design of a Quantum Self-Attention Neural Network on Quantum Circuits," *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Honolulu, Oahu, HI, USA, pp. 1058-1063, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[39] Guangxi Li, Xuanqiang Zhao, and Xin Wang, "Quantum Self-Attention Neural Networks for Text Classification," *Science China Information Sciences*, vol. 67, 2024. [CrossRef] [Google Scholar] [Publisher Link]