Quantum Machine Intelligence (2025) 7:66
https://doi.org/10.1007/s42484-025-00288-y

RESEARCH ARTICLE

Skin cancer image classification using hybrid quantum deep learning model with BiLSTM and MobileNetV2

Ahmed A. Hussein[1] · **Ahmed M. Montaser**[1] · **Hend A. Elsayed**[2]

Received: 16 September 2024 / Accepted: 21 May 2025

Abstract
Skin cancer image classification is known to be extremely complex due to the subtle visual differences between benign and malignant lesions. In this study, we propose a novel hybrid model that leverages the hierarchical feature extraction capabilities of the hybrid quantum convolutional neural network (HQCNN), the temporal dynamics captured by the bidirectional long short-term memory neural networks (BiLSTM) model, and the efficient feature extraction capabilities of MobileNetV2. We evaluated the proposed model on a clinically relevant skin cancer dataset, using images resized to 32×32 and 128×128 pixels to investigate the impact of resolution on classification performance. The HQCNN model augmented with BiLSTM and MobileNetV2 achieved a training accuracy of 97.7% and a test accuracy of 89.3% on 128×128-pixel color images, along with an F1 score of 89.81% and a recall of 94.33% for malignant cases, confirming clinical reliability and strong sensitivity in cancer detection. These results demonstrate robust feature extraction, improved contextual learning, and strong generalization for complex medical image classification tasks.

Keywords Image classification · Hybrid quantum convolutional neural network (HQCNN) · Skin cancer image classification · BiLSTM · MobileNetV2

1 Introduction

Skin cancer can occur due to ultraviolet (UV) rays, classified as UVA, UVB, and UVC. UVC rays do not reach the skin's surface since they are primarily absorbed by the atmosphere. UVA and UVB radiation have the potential to raise the risk of skin cancer. UVB radiation has been linked to the direct cause of melanoma skin cancer. Malignant melanoma, the term for skin cancer, is caused by abnormal growth of skin cells and pigmentation (Tuncer et al. 2024a; Panelos and Massi 2009).

Dermatologists are advised to advise surgical removal of the damaged skin to avoid future disease development if an abnormality in melanoma cells is detected in the early stages (Simoes et al. 2015; Femiano et al. 2008). Given the much-increased chance of a favorable outcome, it is critical to accurately diagnose and confirm malignant melanoma in its early stages. Regretfully, simple skin cancer detection is not made easier by traditional diagnostic techniques (Cheng et al. 2018). One of the deadliest forms of cancer is melanoma, which is included in the group of deadly cancers (Naeem et al. 2020). Machine learning-based diagnostic systems and automatic diagnosis have been widely used in many industries, including medicine (Erten et al. 2023; Kirik et al. 2023). Experts' workloads have been greatly lowered by these apps. In addition, decision support systems are being created to provide a preliminary diagnosis in places without professionals (Baig et al. 2016). The integration of artificial intelligence (AI) represents an opportunity to revolutionize disease detection and significantly enhance diagnostic accuracy, simplify the diagnostic process, and mitigate the risk of human error, which is especially valuable when analyzing data collected from monitoring devices. This underscores the transformative role of AI in medical diagnosis, particularly in complex disease detection scenarios such as skin cancer (Hassan et al. 2024). Even in areas without access to a specialist doctor, machine learning (ML) helps in the early detection of BC. The medical imaging community is becoming increasingly interested in using ML

✉ Ahmed A. Hussein
ahmed.ayman@techedu.sohag.edu.eg

[1] Electrical Technology Department, Faculty of Technology & Education, Sohag University, Sohag, Egypt

[2] Electrical Engineering Department, Faculty of Engineering, Damanhour University, Damanhour, Egypt

Published online: 13 June 2025

and deep learning (DL) to increase the accuracy of cancer screening (Saber et al. 2025).

Quantum machine learning (QML) is a new topic of study that combines quantum processing and traditional machine learning (ML). The primary objective of QML is to leverage quantum phenomena to enhance the performance of conventional machine learning methods. Many studies are being conducted on the quantum equivalent of classical neural networks, known as quantum neural networks (QNNs), spurred by the impressive performance of NNs in the classical setting. Despite a great deal of interest in the topic, there is currently no concrete evidence in the literature that QNNs perform better than their classical counterparts (Kashif et al. 2022). These days, thanks to their strong computational capabilities, quantum neural networks (QNNs) have had some success classifying images. However, the training procedure gets more difficult, and the accuracy tends to decline as the number of layers in quantum neural networks rises, leading to poor network stability. Therefore, a hybrid classical convolutional quantum neural network model is proposed (Wang et al. 2024a). The operational architecture of quantum computers is fundamentally different from that of classical computers. Quantum computers use quantum bits (qubits), as opposed to binary digits (bits: 0 or 1) used by traditional computers (Wang et al. 2022).

Due to its distinct structure, which consists of convolutional layers and pooling layers for image classification, quantum convolutional neural network (QCNN) has drawn a lot of interest as a potential method. Because of its structure, structured information may be extracted with fewer parameters, enabling local pattern identification in tasks involving image classification (Houssein et al. 2022). This led to the development of QCNN, which researchers have explored and used extensively in the field of image categorization (Cong et al. 2019).

Numerous research papers have suggested a range of hybrid quantum–classical convolutional neural networks (HQCNN) designed for multi-classification applications. It has been established that they are more robust than conventional QNNs since they can handle noisy data more effectively and by using fewer coherent quantum bits (Liang et al. 2021; Doan et al. 2022). Due to the potential benefits of incorporating quantum principles to improve several aspects of machine learning algorithms and computations, the hybrid quantum–classical neural network (HQCNN) architecture has recently gained a lot of interest (Xu, et al. 2024). Hybrid quantum–classical neural networks (HQCNN), a hybrid framework, has been proposed for image categorization more recently. Within this paradigm, the classification is carried out by the classical counterpart, while the quantum component is represented by a parameterized circuit intended to extract critical picture features. HQCNN has proven to be more effective than its conventional version in a variety of image classification tasks, demonstrating improved generalizability and robustness in a wide range of applications (LeCun et al. 1989).

A popularly recognized cutting-edge model that is highly renowned for maintaining contextual interactions is the recurrent neural network, or RNN. Text and video sequences have been processed by the computer vision community using RNNs. Recognizing its effectiveness, we designed our challenge to play to the strengths of RNNs: analyzing patch sequences from the same region and finally classifying the input sequence into one of four categories. We used Bayesian least squares to classify entire image regions. One well-known RNN variant for modeling text and video sequences is called BiLSTM. Because of their bidirectional construction, they have been frequently employed for activity recognition in movies and have proven to be effective in modeling contextual information for the future (Tripathi et al. 2021). Recurrent neural networks (RNNs) have a subtype called bidirectional long short-term memory neural networks (BiLSTM), which analyzes layers in sequential input repeatedly and records long-term dependencies that aid in context understanding (Bharal and Krishna 2021). Long-term relationships between signals can be captured more effectively by LSTM, a unique kind of recurrent neural network (RNN), than by conventional RNNs (Cui et al. 2020). The single backward propagation mechanism in LSTM, on the other hand, is limited to accessing prior information included in the sequence data and is unable to gather information from a subsequent context. It has been suggested in recent years that the BiLSTM can access both forward and backward features in a specific amount of time (Gupta et al. 2022). BiLSTM can incorporate contextual information, improve feature extraction from the original sequence, and increase the accuracy of model output—especially for sequence-based tasks—by combining a forward LSTM and a backward LSTM (Xiaoyan et al. 2022).

Descriptiveness and discriminative power of features extracted are crucial for achieving strong classification performance in image classification problems (Hussein et al. 2024). Designed primarily for mobile and embedded vision applications, MobileNet is a lightweight CNN architecture (Howard, et al. 2017). A batch normalization and a pointwise convolution layer come after each block in the architecture, which consists of many depthwise separable convolutional layers organized into blocks. These blocks are designated to extract multi-level features and operate as the network's backbone. To further reduce spatial dimensions and produce feature vectors, which are then input into fully connected layers for classification, MobileNet also includes a global average pooling layer. MobileNet focuses on efficiency and compactness, along with its careful consideration of model size, computational complexity, and accuracy, making it an excellent choice

for contexts with limited resources. In many domains, including computer vision, pre-trained MobileNet models are frequently employed as feature extractors in transfer learning applications including picture classification, object recognition, and semantic segmentation (Reka et al. 2024).

Recent advances in deep learning have significantly improved the accuracy of skin cancer classification. Traditional convolutional neural networks (CNNs) have demonstrated powerful feature extraction capabilities, while more advanced architectures, such as vision transformers (ViTs) and hybrid models, have further improved classification performance. For example, Ozdemir and Pacal (Ozdemir and Pacal 2025a) presented a deep learning model that leverages ConvNeXtV2 blocks and self-attention mechanisms, achieving high accuracy on the ISIC 2019 dataset. Their results highlight the effectiveness of deep learning in medical image analysis. However, purely classical deep learning models remain computationally expensive and rely on large datasets. Building on these advances, Pacal et al. (Pacal et al. 2024) proposed an improved Swin Transformer model that integrates hybrid variable window-based multi-head self-attention (HSW-MSA) and SwiGLU-based MLP, achieving 89% accuracy on the ISIC 2019 dataset. This model demonstrated the power of self-attention mechanisms in improving lesion classification. These developments further underscore the potential of hybrid architectures combining CNNs and transformers to improve diagnostic accuracy. Despite significant advances in deep learning for skin cancer classification, current models still face challenges in balancing classification accuracy, computational efficiency, and robustness of feature extraction. While CNN-based architectures excel at learning local features, transformer-based models enhance global contextual understanding. However, purely classical deep learning models remain computationally expensive and heavily rely on large-scale datasets. To address the limitations of individual architecture, we propose an integrated model combining HQCNN, BiLSTM, and MobileNetV2. HQCNN enables quantum-enhanced local feature extraction, capturing spatial patterns that may be subtle in skin lesion images. BiLSTM introduces a temporal and contextual dimension, which helps in modeling complex, distributed features across the lesion area. MobileNetV2, known for its efficiency and strong representational power in resource-constrained environments, serves as a high-performance feature extractor. The combination ensures improved robustness, generalizability, and computational efficiency—key qualities for real-time clinical applications. Our approach introduces a new quantum-optimized feature convolution layer, which improves the learning ability of deep networks while maintaining computational efficiency. The practical contributions of this work are as follows:

- This study pioneers the application of hybrid quantum–classical neural networks (HQCNN) to detect skin cancer, combining the strengths of deep learning and quantum computing.
- Enhanced feature learning using BiLSTM and MobileNetV2. The proposed model captures both spatial and sequential dependencies, ensuring superior lesion classification compared to traditional CNN-based models.
- Computational efficiency unlike purely classical models that require extensive computational resources. Our approach improves feature learning using quantum-inspired transformations, paving the way for scalable and efficient AI-based medical diagnosis.
- Real-world medical impact. This work contributes to automated skin cancer detection powered by AI, providing a highly accurate and effective tool for early diagnosis that may assist dermatologists in making clinical decisions in real time.

2 Related works

Recent advances in deep learning have significantly improved the accuracy of skin cancer classification. Traditional CNN-based models such as VGG16, ResNet, EfficientNet, and MobileNet have been widely used due to their powerful feature extraction capabilities. For example, Ozdemir and Pacal proposed a deep learning model incorporating ConvNeXtV2 blocks and self-attention mechanisms, achieving high accuracy on the ISIC 2019 dataset (Ozdemir and Pacal 2025a). Lesion classification has improved with transformer-based topologies. In tasks involving medical imaging, the Swin Transformer and DeiT have shown excellent generalization. Using the ISIC 2019 dataset, Pacal et al. demonstrated a hybrid Swin Transformer model that combines SwiGLU-based MLP with hybrid variable window-based multi-head self-attention (HSW-MSA), with 89% accuracy (Pacal et al. 2024). A promising method that combines the ideas of deep learning and quantum computing is hybrid quantum–classical neural networks, or HQCNN. A hybrid quantum–classical CNN model created by Wang et al. showed increases in classification efficiency while lowering parameter overhead. However, the present limitations of quantum hardware continue to hinder practical application (Wang et al. 2024b). The cutting-edge models for automatic skin cancer diagnosis that we examined are described below.

A skin cancer classification method based on YOLOv5 and hyperspectral imaging was developed by Huang et al. They achieved an accuracy of 79.20% by using a total of

1854 images for three classes (squamous cell carcinoma, basal cell carcinoma, and seborrheic keratosis) (Huang et al. 2023). Dandu et al. concentrated on using transfer learning to categorize skin cancer. They took advantage of the 10,000-image collection from the 2018 International Skin Imaging Collaboration. The accuracy value in the method they described was 90.96% (Dandu et al. 2023). Keerthana et al. created a hybrid convolutional neural network approach for the diagnosis of skin cancer. Utilizing the 1279 photos from the ISBI 2016 dataset, they were able to achieve an accuracy of 86.23% (Keerthana et al. 2023). He and colleagues used a seven-point checklist dataset and CNN to predict the occurrence of skin cancer. The average accuracy rate achieved by their suggested method was 76.80% (He et al. 2023). Gupta et al. classified skin cancer images using deep learning techniques, especially CNNs, and obtained an accuracy of 83.20% (Gupta et al. 2021). Nivyashree and Pramila detected malignant and benign skin lesions using novel densely connected convolutional and compared the accuracy with convolutional neural network and obtained an accuracy of 80.80% (Nivyashree et al. 2023). Angelina and Ulfitria classified skin cancer using VGG16 deep learning network and achieved 83.75% accuracy (Angelina and Ulfitria 2024). Kartikeya Agarwal et al. classified skin cancer images using convolutional neural network. The highest accuracy of the model achieved was more than 86.65% (Agarwal and Singh 2022).

Pacal et al. proposed a CNN-ViT-based hybrid model for early skin cancer diagnosis that integrates convolutional and transformer-based architectures within the MetaFormer backbone. Their approach introduces a self-focusing mechanism to enhance focus on the lesion region and suppress background noise. Evaluated on the ISIC 2019 and HAM10000 datasets, the model achieved classification accuracy of 92.54% and 95.01%, respectively, outperforming several state-of-the-art CNN and ViT baselines while maintaining a lightweight design of 35.01 million parameters (Pacal et al. 2025). Ozdemir and Pacal proposed a hybrid deep learning model that integrates ConvNeXtV2 blocks and focal self-attention mechanisms to enhance skin cancer detection. Their approach, optimized for mobile and real-time applications, combines local and global feature extraction to tackle class imbalance and variability in the ISIC 2019 dataset. The model achieved an accuracy of 93.60%, precision of 91.69%, recall of 90.05%, and F1 score of 90.73%, outperforming several state-of-the-art CNNs and vision transformer models across eight skin cancer classes (Ozdemir and Pacal 2025b). Ince et al. introduced a novel deep learning framework that integrates ConvNeXtV2 and Global Response Normalization (GRN) within a U-Net architecture to enhance the segmentation of cerebral vascular occlusions in diffusion-weighted imaging (DWI). This hybrid model leverages the advanced feature extraction capabilities of ConvNeXtV2 and the normalization benefits of GRN to improve the accuracy and robustness of lesion segmentation (Ince et al. 2025a). Ince et al. conducted a comparative study on U-Net, U-Net + +, and Attention U-Net architectures for brain stroke segmentation using diffusion-weighted imaging from the ISLES 2022 dataset. Their results demonstrated that Attention U-Net achieved the highest segmentation accuracy, with an IoU of 0.8223 and a dice similarity coefficient of 0.9021, outperforming its counterparts in most metrics. This study emphasizes the effectiveness of attention mechanisms in improving lesion boundary delineation, especially in complex or low-contrast medical imaging tasks (İnce et al. 2025). Ozdemir et al. proposed a lightweight hybrid deep learning model combining CNNs and vision transformers using InceptionNeXt blocks with integrated grid and block attention mechanisms for lung cancer detection. Trained on the IQ-OTH/NCCD and chest CT datasets, respectively. Its design enables efficient multi-scale feature extraction and spatial-context awareness while maintaining a low parameter count (18.1 M), demonstrating both computational efficiency and superior diagnostic performance in real-time clinical scenarios (Ozdemir et al. 2025). Alnowaiser et al. proposed a hybrid model that combines pre-trained CNNs (DenseNet-121 and VGG-16), BiLSTM for temporal feature learning, and optimized SVM classification using the grey wolf algorithm. Their method demonstrated excellent performance for breast cancer classification, achieving accuracy up to 99.86% on the MIAS dataset, highlighting the value of hybrid deep learning architectures in medical image analysis (Alnowaiser et al. 2024). Elbedwehy et al. proposed a hybrid model that integrates pre-trained CNNs (AlexNet and ConvNeXt) with a custom Adam optimizer, leveraging feature concatenation to enhance spatial representation. Their method achieved an outstanding classification accuracy of 99.85% on a CT kidney disease dataset, underscoring the effectiveness of combining classical and modern architectures with optimization techniques in medical image classification (Elbedwehy et al. 2024).

These studies exemplify the critical role of medical imaging experts in optimizing data preprocessing, segmentation, and interpretation workflows, thereby ensuring clinically relevant deployment of deep learning systems. They also highlight the need for collaborative efforts between AI researchers and image processing technicians to ensure the development of robust, interpretable, and practically deployable models for cancer diagnostics. These findings underscore not only the technical advancements in deep learning but also the essential role of medical image processing technicians in clinical deployment. Their contributions span image acquisition, expert annotation, noise reduction, normalization, and real-time data interpretation, which are vital for model training and validation. As deep learning

models become increasingly complex, collaboration with medical image professionals ensures accurate, clinically relevant results, while also facilitating trust and interpretability in AI-assisted diagnostics. While modern deep learning techniques, such as CNNs, transformers, and hybrid quantum–classical models, have demonstrated strong performance in medical imaging, balancing accuracy, efficiency, and generalization remains a challenge. CNN-based models extract spatial features efficiently but often struggle with long-range dependencies. Transformer-based architectures improve understanding of global context but require significant computational resources, making them less feasible for real-time medical applications. Hybrid quantum–classical neural networks (HQCNNs) offer quantum-enhanced feature transformations, but their real-world applicability remains limited by hardware limitations and computational overhead.

This study proposes a hybrid classical-quantum deep learning model that integrates MobileNetV2 and BiLSTM with HQCNN to address these limitations. The BiLSTM component improves sequential feature extraction, enhancing the model's ability to capture contextual dependencies in lesion images. MobileNetV2 acts as an efficient feature extractor, reducing computational cost while preserving important image details. Meanwhile, the HQCNN component applies quantum-inspired transformations in simulations, potentially improving feature representations. Our approach delivers improved lesion classification accuracy by combining quantum-inspired and classical feature representations. Our model enhances feature learning and stronger generalization, as the inclusion of BiLSTM allows the model to capture contextual dependencies, improving robustness across diverse medical images. The computational cost trade-off is that while MobileNetV2 reduces the number of parameters, the quantum layer introduces additional simulation overhead. The feasibility of real-world implementation remains an open challenge for future work. By leveraging quantum-inspired feature transformations while preserving the strengths of classical deep learning, this study contributes to the development of hybrid AI techniques for skin cancer classification. However, the computational feasibility of quantum deployment in the real world remains an area for future investigation.

3 The work environment

The following hardware requirements are part of the computer environment where the experiments were conducted using CPU resources in the Google Colab environment:

- Python version: 3.10.12Python version: 3.10.12
- Compiler: GCC 11.4.0
- Processor (CPU): Intel(R) Xeon(R) CPU @ 2.20 GHz, single core
- Memory (RAM): 12.67 GB
- Operating System: Linux-6.1.85 + -x86_64-with-glibc2.35

4 Problem formulation

Skin cancer detection relies heavily on visual assessment, which is prone to human error. Traditional methods have often been challenged by subtle differences in visual features between malignant and benign conditions. Skin cancer detection is a significant research area in medical imaging, where the ability to accurately classify cancerous and non-cancerous lesions from images can significantly impact patient outcomes. In this study, we present a novel model that combines three state-of-the-art deep learning techniques: the hybrid quantum convolutional neural network (HQCNN), the bidirectional long short-term memory (BiLSTM) network, and MobileNetV2. MobileNetV2 is used here as an effective feature extractor due to its high ability to analyze data quickly and efficiently while minimizing the number of required parameters, making it ideal for handling large and complex datasets such as medical images. This innovative combination of the three models represents a novel and underexplored addition to the existing literature and represents an important step toward improving the performance of deep learning models in medical applications. To evaluate the effectiveness of the proposed model, experiments were conducted on a skin cancer dataset, and the results demonstrated the high classification accuracy of the proposed model, highlighting its tremendous potential in increasing diagnostic accuracy.

5 Proposed methodology

Figure1 depicts the proposed methodology's overall description.

5.1 Data preparation

In this study, we used the skin cancer dataset, which can be downloaded from Kaggle and is divided into two classes, benign and malignant (Fanconi 2019). To achieve balance, we normalized the images in each of the two classes each pair of images belongs. Each dataset contains 80% of the data used for training and 20% for testing the model. The training set is the initial set of images, which includes 2637 images, of which 1440 are benign and 1197 are malignant. There are 660 images in the second batch, known as the test set, of which 300 are malignant and 360 are benign. Before

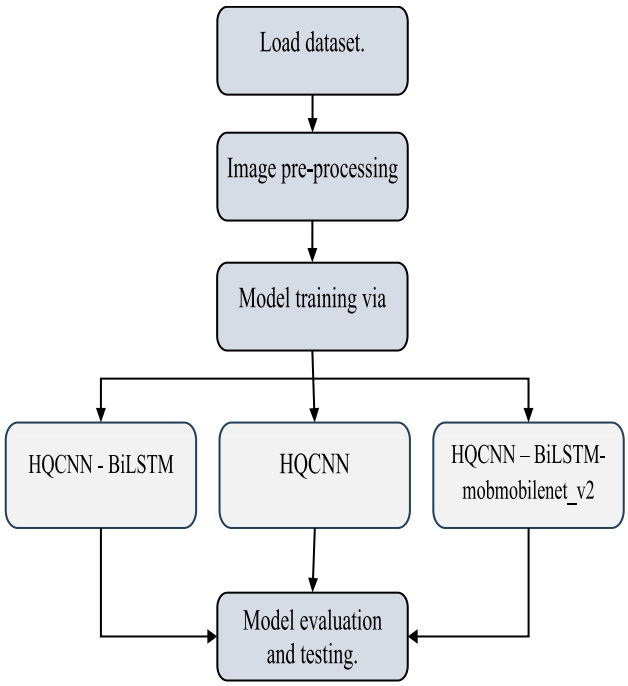

Fig. 1 Diagrammatic illustration of the suggested project

feeding the image data to the neural network, to address class imbalance, we applied random sampling to the training set, ensuring that each class contained a maximum of 1100 images. This prevented bias toward the majority class while maintaining sufficient data for model training. The test set remained unchanged. Before feeding images into the neural network, all datasets were preprocessed and normalized to ensure consistency in model training.

5.2 Image preprocessing

The preprocessing procedures of the dataset used to train and assess the hybrid quantum convolutional neural network are shown in Fig. 2 and comprise the following:

- Image resizing: To evaluate the impact of input resolution on classification performance, all images were resized to 32×32 and 128×128 pixels. The 32×32 resolution was chosen for consistency with previous quantum neural network studies, where low-dimensional inputs are necessary due to the limitations of current quantum circuits. This reduction helps reduce computational costs and simulation time during training. However, we recognize that downscaling can result in the loss of clinically relevant lesion features. To evaluate this, we compared the same benign and malignant lesion images visually at 128×128 and 32×32 resolutions, with the lower-resolution images upscaled for interpretability. As shown in Fig. 3, although fine details are attenuated at 32×32 resolution, key lesion features—such as shape, asymmetry, and underlying pigmentation—remain discernible. This demonstrates that meaningful classification is still possible even at lower resolutions, particularly in quantum-augmented models.

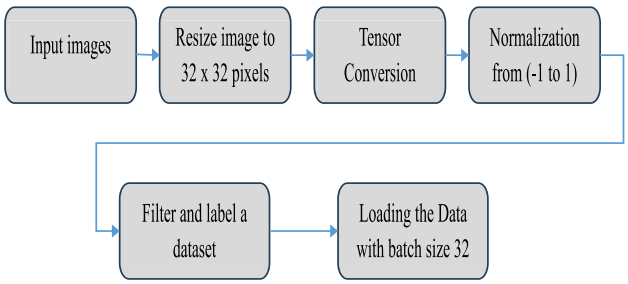

Fig. 2 Image preprocessing scheme

- Convert images to PyTorch tensors. They are multidimensional arrays used in PyTorch to train models.
- Normalization of image data where pixel values are scaled to the range $(-1, 1)$ is necessary and useful for achieving stability and learning more effectively in the neural network training process.
- Data augmentation (planned for future work): While data augmentation was considered to enhance model generalization, it was not implemented in the current implementation. Future experiments will explore augmentation techniques such as:
  - Random rotation ($\pm 20°$) to simulate different lesion orientations
  - Horizontal and vertical flipping to account for differences in clinical imaging views
  - Brightness adjustments to make the model stable to lighting conditions
- Filter and label the dataset and split it into training and test sets.
- Load the training and test dataset into the data loader with a batch size of 32. This step allows the model to process images in small batches instead of all at once (Fig. 2).

5.3 Building a hybrid quantum convolutional neural network (HQCNN)

The proposed network architecture integrates classical convolutional neural networks with quantum computing elements designed for image classification. The architecture takes advantage of the strengths of both convolutional and quantum neural networks to enhance the generalization ability of the model while maintaining computational efficiency. After data preparation and image preprocessing, the data

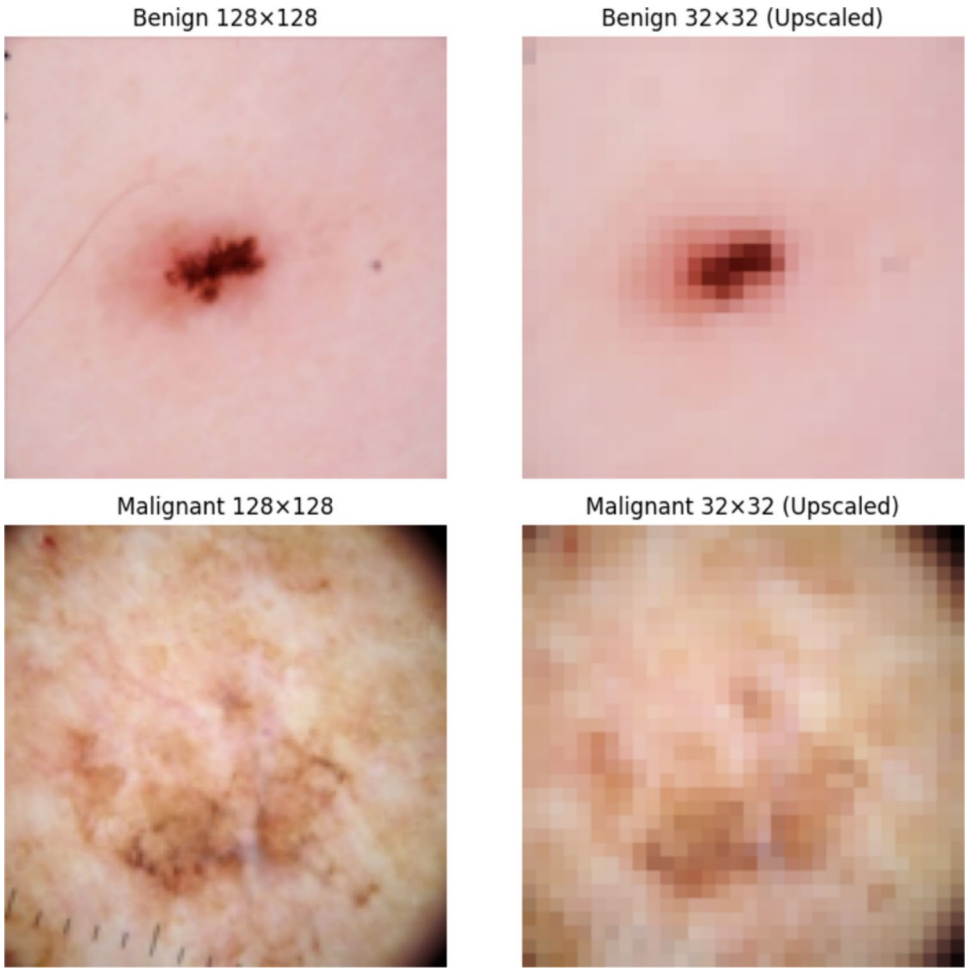

Fig. 3 Resolution comparison for benign and malignant lesions

is loaded into the network, which starts with two convolutional layers, each using eight filters with a kernel size of 3×3, followed by max pooling layers. These layers extract hierarchical features from the input images, reducing spatial dimensions while retaining essential information. These layers follow a flattening process, which converts 2D feature maps to a 1D vector. This vector is passed through a fully connected layer (dense) that is reduced to 32 dimensions, followed by layer normalization and ReLU activation. A dropout layer with a dropout rate of 0.3 is applied to prevent overfitting by randomly zeroing out some elements.

A notable component of this model is the integration of a hybrid quantum layer, which incorporates parameterized quantum circuits (PQC) into the classical architecture. In this layer, classical features are encoded into a quantum state via quantum gates. A Hadamard gate is first applied to initialize each qubit in a state of superposition, enabling the exploration of multiple feature paths simultaneously. Then, RY gates—whose rotation angles are derived from classical features—introduce trainable quantum transformations. The quantum state is measured in the computational basis, and the expectation values are extracted and returned as quantum-enhanced features. This process allows the model to capture entangled, non-linear relationships that may be difficult to learn with classical layers alone.

Gradient computation within the hybrid quantum layer is performed using a finite-difference method by shifting the RY gate's parameter ($\theta \pm \pi/2$) and measuring the corresponding change in output. This ensures smooth backpropagation and enables efficient training of the hybrid quantum–classical model.

The network was trained using the Qiskit Aer simulator with a single qubit for 30 epochs using the Adam optimizer, learning rate = 0.0001, weight decay = 0.0001, and batch size = 32. A learning rate scheduler (ReduceLROnPlateau) was applied to dynamically adjust the learning rate if the validation loss did not improve for three consecutive epochs. Additionally, early stopping was used, stopping training after 20 consecutive epochs without an improvement in the validation loss. This approach prevented overfitting, where continued training after 30 epochs degraded performance on the validation set. By stopping early, we ensured better generalization while also reducing computational overhead. Figure 4 illustrates the

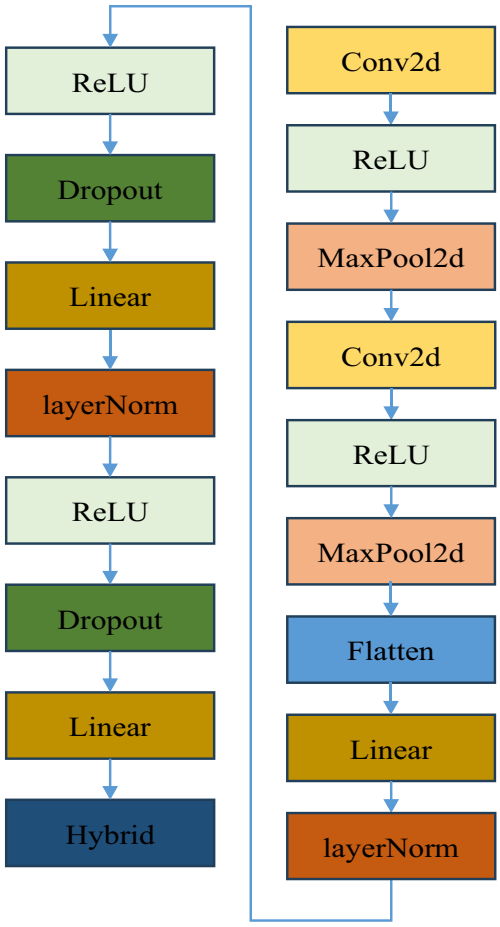

Fig. 4 Hybrid classical convolutional-quantum neural network architecture (HQCNN)

layers used to build the network. The detailed layer-by-layer structure of the proposed HQCNN model is shown in Table 1.

5.4 Quantum circuit engineering and hardware considerations

The hybrid quantum layer in our model is implemented using a two-layer, deep, single-qubit quantum circuit. Each layer applies a Hadamard (H) gate to initialize the qubit into a superposition state, followed by a parameterized RY (θ) spin gate. The input angle θ is derived from the output of the previous fully connected classical layer. After the final layer, a measurement is performed based on the Z computation, and the resulting expectation value is fed back into the classical network for further processing. Figure 5 illustrates the quantum circuit used in the hybrid layer of the model.

The gradients for this quantum layer are calculated using the finite-difference parameter shift method, as shown in Eq. 1.

$$\frac{\partial\langle O\rangle}{\partial\theta} \approx \frac{\langle O\left(\theta+\frac{\pi}{2}\right)\rangle - \langle O\left(\theta-\frac{\pi}{2}\right)\rangle}{2} \tag{1}$$

This allows the quantum layer to integrate seamlessly into the classical backpropagation process. All quantum simulations were performed using a Qiskit Aer simulator to avoid hardware noise and resource constraints. Practical implementation on real quantum devices faces several challenges:

- Decoherence and gate betrayal, which can lead to errors during the implementation of multi-layer circuits
- Limited qubit connectivity and coherence time, limiting the possible circuit depth
- Shot noise due to limited measurement samples

To address these limitations, future versions of this form may include the following:

- Noise-aware training using noisy simulators or real equipment
- Improved shallow quantum circuits to reduce error propagation
- Error mitigation techniques such as zero-noise extrapolation

5.5 Building a hybrid quantum convolutional neural network (HQCNN) with BiLSTM

This model was created by combining convolutional neural network (CNN) layers, a bidirectional long short-term memory (BiLSTM) layer—an advanced form of recurrent neural network (RNN) architecture—and quantum neural network (QNN) layers. Incorporating BiLSTM into HQCNN enhances feature learning by capturing long-range dependencies between extracted feature maps. While CNN layers excel at extracting local features, BiLSTM operations feature bidirectional sequences, preserving contextual relationships that might otherwise be lost. This approach is particularly valuable in medical imaging, where lesion textures and spatial patterns exhibit complex dependencies. Experimental results confirm that incorporating BiLSTM significantly improved classification accuracy, demonstrating its potential for enhancing feature representations for skin cancer detection. The following sections explain each component of this network and its function:

- The structure of this network starts with two convolutional layers Conv1 and Conv2 with a 3 ×3 filter each followed by a ReLU activation function and maximum pooling where these layers extract features from the input data.

Table 1 Detailed layer-by-layer architecture of the proposed hybrid quantum convolutional neural network (HQCNN)

Layer name	Type	Output shape	Parameters	Description
Input	RGB Image	(3, 128, 128)	—	Input image resized and normalized
Conv1	Conv2D (3 × 3)	(8, 128, 128)	224	8 filters, kernel size 3 × 3, padding = 1
ReLU1	Activation	(8, 128, 128)	—	Non-linear activation
Pool1	MaxPooling (2 × 2)	(8, 64, 64)	—	Reduces spatial dimensions by half
Conv2	Conv2D (3 × 3)	(8, 64, 64)	584	Second convolution layer
ReLU2	Activation	(8, 64, 64)	—	Non-linear activation
Pool2	MaxPooling (2 × 2)	(8, 32, 32)	—	Further reduces spatial resolution
Flatten	Reshape	(8192,)	—	Converts feature maps to 1D vector
FC1	Linear	(32,)	262,176	Fully connected layer
LN1	LayerNorm	(32,)	64	Stabilizes training
ReLU3	Activation	(32,)	—	Non-linear activation
Dropout1	Dropout ($p = 0.3$)	(32,)	—	Regularization to prevent overfitting
FC2	Linear	(32,)	1056	Second dense layer
LN2	LayerNorm	(32,)	64	Layer normalization
ReLU4	Activation	(32,)	—	ReLU activation
Dropout2	Dropout ($p = 0.3$)	(32,)	—	Prevents overfitting
FC3	Linear	(1,)	33	Prepares input to quantum circuit
Hybrid	Quantum Layer (1Q)	(1,)	Qiskit Circuit	Uses parameterized RY gate; gradient via finite-difference
Output	Softmax	(2,)	Derived from hybrid	Converts output to class probabilities

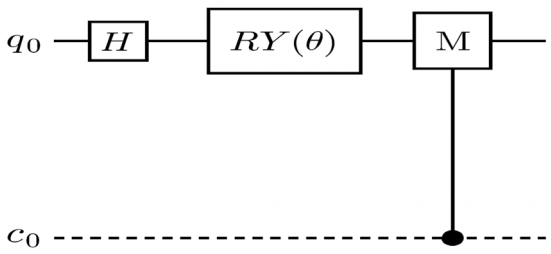

Fig. 5 Single-qubit variable quantum circuit

- Flattening is where the convolutional layer output is flattened to a 1D vector in this step, preparing it for the fully connected layers.
- Fully connected layers Fc1 and Fc2 are two fully connected layers. The first layer reduces the feature dimensions to 32 followed by layer normalization and ReLU activation. The second layer processes the data further followed by layer normalization and ReLU activation.
- Long short-term memory (LSTM) layer processes the output from the first fully connected layer, which is particularly useful for capturing temporal dependencies even though the input data is spatial. The LSTM output is then passed through the second fully connected layer.
- Hybrid quantum layer. This layer uses a 1-qubit quantum circuit to represent the quantum portion of the model. This layer offers quantum computing, which may have benefits including quicker convergence and better data expression for specific kinds of information.
- After the fully connected layers, a dropout layer with a rate of 0.3 is implemented to minimize overfitting by randomly setting certain neuron outputs to zero during training.

The model was trained using the Adam optimizer, chosen for its adaptive learning ability, with a learning rate of 0.0001, dynamically controlled by the ReduceLROnPlateau scheduler. The initial learning rate was set to 0.001 and adjusted based on the validation loss to prevent stagnation. Adam has been widely recognized for its fast convergence and adaptive parameter updates, making it particularly suitable for deep learning-based medical image classification. Recent studies have shown that Adam consistently outperforms traditional optimizers such as SGD and RMSprop, particularly when dealing with complex feature distributions and non-convex loss surfaces (Krushelnytskyi 2023). In the context of hybrid classical-quantum neural networks, the choice of optimizer is even more critical due to the unique gradient noise introduced by quantum circuits. Experimental research has shown that Adam provides more stable and efficient gradient updates than SGD or RMSprop, particularly in the optimization of variable quantum circuits, where gradient fluctuations can hinder convergence (Zaman et al. 2024; Arthur and Date 2022). Alternative optimizers,

such as momentum-driven SGD, were tested but resulted in slower convergence and lower accuracy. AdamW was also considered due to its improved handling of weight decay, but no significant advantage was observed in our experimental setup.

The class cross-entropy (CCE) loss function was chosen to train our hybrid classical-quantum model due to its effectiveness in multi-class classification tasks. Since our dataset was stratified to maintain class balance, CCE ensures stable gradient updates across both benign and malignant classes. Recent literature confirms that CCE performs optimally when class balance is maintained, as demonstrated in medical image classification studies (Scholz et al. 2024; Yeung et al. 2022). The focal loss is most useful on highly imbalanced datasets, whereas the Dice and Tversky losses are primarily designed for segmentation tasks, making them less suitable for this classification problem. Therefore, CCE remains the most computationally efficient and theoretically justified choice for our model.

The dropout rate was set to 0.3, chosen empirically after evaluating values between 0.2 and 0.5 to achieve a balance between regularization and generalization. Low dropout (0.2) resulted in overfitting, while high dropout (0.5) resulted in underfitting. The batch size was set to 32, chosen after testing values of 16, 32, and 64. A batch size of 32 offered the best trade-off between convergence speed and stability. Smaller batch sizes resulted in unstable gradients, while larger batch sizes increased memory consumption without significant performance gains. The ReLU activation function was selected after comparing it with LeakyReLU and ELU. While all activation functions were tested, ReLU showed the best convergence speed and training stability, with no significant performance improvements observed using alternate activations. The model was trained for 30 epochs, chosen based on early convergence observations. Extending training beyond 30 epochs did not yield further performance improvements. The Aer simulator from the Qiskit library was used as the quantum circuit simulator to ensure a balance between computational feasibility and quantum layer integration. These settings were carefully optimized through experimental testing and validation, ensuring strong training stability and strong generalization performance. Figure 6 illustrates the layers used to build the model. Table 2 shows the detailed structure of the proposed HQCNN with BiLSTM model layer by layer.

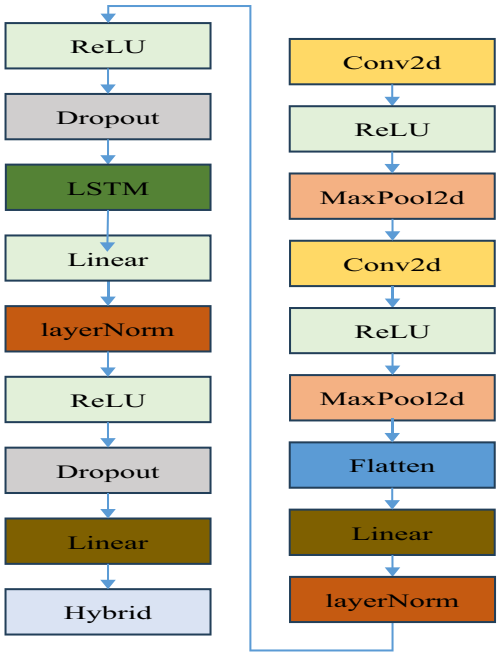

Fig. 6 Hybrid classical convolutional-quantum neural network architecture (HQCNN) with BiLSTM

Table 2 Detailed layer-by-layer architecture of the proposed hybrid quantum convolutional neural network (HQCNN) with BiLSTM

Layer name	Type	Output shape	Parameters
Input	RGB Image (128 × 128)	(3, 128, 128)	—
Conv1	Conv2D (3 → 8)	(8, 128, 128)	224
ReLU1	Activation	(8, 128, 128)	—
Pool1	MaxPool2D (2 × 2)	(8, 64, 64)	—
Conv2	Conv2D (8 → 8)	(8, 64, 64)	584
ReLU2	Activation	(8, 64, 64)	—
Pool2	MaxPool2D (2 × 2)	(8, 32, 32)	—
Flatten	Flatten	(8192)	—
FC1	Linear (8192 → 32)	(32)	262,176
LayerNorm1	LayerNorm	(32)	64
Dropout1	Dropout ($p = 0.3$)	(32)	—
LSTM	BiLSTM (32 → 32)	(32)	8320
FC2	Linear (32 → 32)	(32)	1056
LayerNorm2	LayerNorm	(32)	64
Dropout2	Dropout ($p = 0.3$)	(32)	—
FC3	Linear (32 → 1)	(1)	33
Hybrid	1-Qubit Qiskit Layer	(1,)	Qiskit Circuit
Output	Softmax	(2)	Derived from hybrid layer

5.6 Building a hybrid quantum convolutional neural network (HQCNN) with BiLSTM and MobileNetV2

The pre-trained MobileNetV2 is used as a feature extractor at the beginning of the model; only the feature extraction portion of MobileNetV2 is utilized, except for the final classification layers, when the MobileNetV2 architecture harvests high-level features from the input photos. The model's last convolutional layers produced these features, which are vectors of 1280 dimensions. The mathematical

relationship governing the feature extraction process in MobileNetV2 is expressed using deep convolutions and point convolutions. Equations 2 and 3 define deep convolutions and point convolutions, respectively, which are the two main components of the model's feature extraction mechanism.

$$Z_{\text{depthwise}} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n,c} W_{m,n,c} \tag{2}$$

$$Z_{\text{pointwise}} = \sum_{c=0}^{C-1} Z_{\text{depthwise},i,j,c} W_{\text{pointwise},c,k} \tag{3}$$

Where $Z_{\text{depthwise}}$ represents the output of depthwise convolution applied to each channel separately, $W_{\text{pointwise}}$ is the pointwise convolution kernel, and C is the number of channels.

- The 1280-dimensional feature vector is flattened in the following phase, which is called flattening.
- Fully connected layers in which the vector that has been flattened are passed through a sequence of fully connected layers, the first of which shrinks the feature size to 32 dimensions before applying LayerNorm to stabilize the learning process. It is expressed by the mathematical relationship in Eq. 4.

$$fc = \text{ReLU}\left(W_{fc} X_{flat} + b_{fc}\right) \tag{4}$$

- LSTM layer where the output is reshaped and passed through the LSTM network. The recursive nature of the LSTM layer is governed by Eqs. 5, 6, 7, 8, 9, 10, 11 which specify gate activations, cell state updates, and hidden state transitions. These equations enable the model to preserve important temporal dependencies in the extracted feature representations.

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{5}$$

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{6}$$

$$\widetilde{C}_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{7}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \tag{8}$$

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{9}$$

$$h_t = o_t \odot \tanh\left(C_t\right) \tag{10}$$

For bidirectional processing, the final output of BiLSTM is:

$$H_t = \left[h_t^{\text{forward}}; h_t^{\text{backward}}\right] \tag{11}$$

- Dropout layers are included at a rate of 30% after each fully connected layer to prevent overfitting.
- Quantum circuit where it is applied after the final fully connected layer and acts as a classical-quantum interface allowing the gradient to be calculated from during the quantum circuit as shown in Eq. 12.

$$\left|\psi_{\text{output}}\right\rangle = U_\theta \left|\psi_{\text{input}}\right\rangle \tag{12}$$

where U_θ is a parameterized quantum unitary transformation. The expectation value of a quantum measurement on the computational basis is computed as follows, in accordance with Eq. 13.

$$\langle Z \rangle = \langle \psi\_\{\varepsilon\{output\}\} | Z | \psi\_\{\varepsilon\{output\}\} \rangle \tag{13}$$

By leveraging quantum feature encoding, this step enhances the representational power of the network while maintaining computational efficiency.

- Output layer where after the quantum circuit the output is concatenated with its probability to form a two-dimensional output where this output is interpreted as the class probabilities for the binary classification task, and this structure is designed to be flexible which allows the model to learn complex patterns in image data. Figure 7 shows the layers used to build the model. Table 3 shows the detailed layer-by-layer architecture of the HQCNN model with BiLSTM and MobileNetV2.

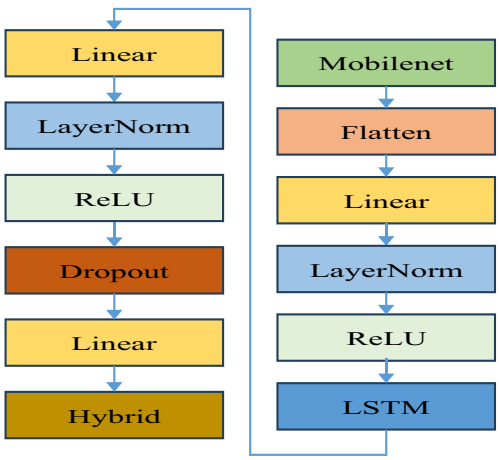

Fig. 7 Hybrid quantum convolutional neural network (HQCNN) architecture using BiLSTM and MobileNetV2 microstructure

Table 3 Detailed layer-by-layer architecture of the proposed hybrid quantum convolutional neural network HQCNN with BiLSTM and MobileNetV2

Layer name	Type	Output shape	Parameters
Input	RGB Image	(3, 128, 128)	—
MobileNetV2	Pre-trained feature extractor	(1280, 4, 4)	~2.2M
Flatten	Flatten	(20480)	—
FC1	Linear → LayerNorm →ReLU	(8)	163,848
LSTM	LSTM (8→16)	(16)	1,632
FC2	Linear → LayerNorm →ReLU	(16)	272
Dropout1	Dropout (p=0.3)	(16)	—
FC3	Linear	(1)	17
Hybrid	1-Qubit Quantum Circuit	(1,)	Qiskit Circuit
Output	Softmax-like (2-class vector)	(2)	Derived from hybrid

5.6.1 Mathematical formulation of the proposed HQCNN + BiLSTM + MobileNetV2 architecture

To provide a clearer understanding of how the HQCNN, BiLSTM, and MobileNetV2 components interact within the proposed hybrid classical-quantum model, this section presents a mathematically supported step-by-step formulation of the decision function. This formalization highlights how the feature representations are gradually transformed—from classical to quantum domains—ultimately leading to the final classification output. The flow is organized as follows:

- Features are extracted from the input image via MobileNetV2 as shown in Eq. 14.

$$z_1 = f_{\text{Mobile}}(x) \in R^{1280\times4\times4} \tag{14}$$

- The output feature maps are flattened to a one-dimensional vector as shown in Eq. 15.

$$z_2 = \text{Flatten}(z_1) \in R^{20480} \tag{15}$$

- The dimensions are reduced to 8 via a dense layer followed by LayerNorm and ReLU activation as shown in Eq. 16.

$$z_3 = \text{ReLU}\big(\text{LayerNorm}\big(W_1 z_2 + b_1\big)\big) \in R^8 \tag{16}$$

- The bidirectional LSTM processes the vector to capture the temporal relationships as shown in Eq. 17.

$$z_4 = f_{\text{BiLSTM}}(z_3) = \overrightarrow{h_T} \oplus \overleftarrow{h_1} \in R^{16} \tag{17}$$

- Another fully connected layer enhances the features after BiLSTM as shown in Eq. 18.

$$z_5 = \text{ReLU}\big(\text{LayerNorm}\big(W_2 z_4 + b_2\big)\big) \in R^{16} \tag{18}$$

- The 16-dimensional output is mapped to a quantum rotation angle θ as shown in Eq. 19.

$$\theta = W_3 z_5 + b_3 \in R \tag{19}$$

- The single-qubit Hadamard quantum circuit is applied, and the $R(\theta)$ gate parameter is used and then measures the expectation value in the Z-basis as shown in Eq. 20.

$$q = <0|U^{\dagger}(\theta)ZU(\theta)|0> \; where \; U(\theta) = RY(\theta)\cdot H \tag{20}$$

- The single output q∈ [− 1,1] is mapped to a two-class probability vector as in Eq. 21.

$$\hat{y} = \text{Softmax}\big(\big[q, 1-q\big]\big) \in R^2 \tag{21}$$

- The entire structure can be summarized in a nested function as in Eq. 22.

$$\hat{y} = \text{Softmax}\big(\mathcal{Q}\big(f_{\text{FC2}}\big(f_{\text{BiLSTM}}\big(f_{\text{FC1}}\big(\text{Flatten}\big(f_{\text{Mobile}}(x)\big)\big)\big)\big)\big)\big) \tag{22}$$

6 Results and discussion

6.1 Model results with skin cancer dataset

6.1.1 Skin cancer dataset of size 32 × 32

The performance of HQCNN, HQCNN with BiLSTM, HQCNN with BiLSTM, and MobileNetV2 models in image classification was the primary focus of this work. As shown in Table 4, all models underwent 30 epochs of training using a batch size of 32 and an image size of 32 × 32 pixels. Additional experiments demonstrated that the model achieved stable accuracy and convergence in the loss within this range, with minimal improvement observed after 30 epochs. Furthermore, extending the training to 300 epochs resulted in overfitting,

Table 4 Model results after 30 epochs for image size 32 × 32 skin cancer dataset

Learning rate	Test accuracy, %	Training accuracy, %	Time (s)	Batch size	Model
0.0001	79.39	82.37	2660	32	HQCNN
0.0001	78.9	81.04	2616	32	HQCNN + BILTSM
0.0001	86.67	88.78	2981	32	HQCNN + BILTSM + MobileNetV2

where the validation loss increased despite the continuous improvements in training accuracy. Therefore, 30 epochs were chosen to balance model performance and computational efficiency and prevent overfitting, as the results showed the following:

- The HQCNN model performs well, with a training accuracy of 82.37%, a testing accuracy of 79.39%, and a training time of 2660 s.
- The training accuracy dropped to 81.04% and the testing accuracy to 78.9% when BiLSTM was added to the HQCNN model. On the other hand, the training time dropped to 2616 s, suggesting that the computational efficiency was increased by including BiLSTM.
- With a training time of 2981 s and training and testing accuracies of 88.78% and 86.67% respectively, the proposed new HQCNN model with BiLSTM and MobileNetV2 has been hailed as a breakthrough in the field. This has demonstrated the effectiveness and robustness of the model and made it an effective tool for early detection and diagnosis of skin cancer, which is an important addition to the field of medical image analysis.

Additionally, since the quantum circuits were simulated rather than implemented on real quantum hardware, this introduced additional computational overhead. The hybrid quantum–classical model was trained in the Google Colab environment using a single-core Intel Xeon CPU at 2.20 GHz with 12.67 GB of RAM, and the total training time for 30 epochs was 2660 s (~ 44 min). While this is manageable for research, further improvements—such as reducing circuit depth or testing on real quantum hardware—will be explored in future work to enhance its computational feasibility for clinical applications.

To evaluate the effectiveness of the proposed classical-quantum neural network HQCNN with BiLSTM and MobileNetV2, we evaluated its classification performance on the melanoma dataset, comparing it with the HQCNN and HQCNN + BiLSTM models. Table 5 presents the detailed classification results, including precision, recall, F1 score, and accuracy for benign and malignant cases.

The findings unequivocally demonstrate that while adding MobileNetV2 enhances performance, adding BiLSTM enhances feature extraction. In particular:

- HQCNN + BiLSTM + MobileNetV2 achieved the highest accuracy (87%), outperforming HQCNN (79%) and HQCNN + BiLSTM (79%).
- Precision-recall trade-off: HQCNN has a high precision (90%) for malicious cases, but struggles with recall (70%), meaning it misses a significant number of malicious cases. The proposed model (HQCNN + BiLSTM + MobileNetV2) corrects this by achieving 90% recall for

Table 5 Model performance metrics after 30 epochs for skin cancer dataset

Model	Class	Precision (%)	Recall (%)	F1 score (%)	Accuracy (%)
HQCNN	Benign (0)	72	90	80	79
-	Malignant (1)	90	70	79	-
Macro Avg	-	81	80	79	79
Weighted Avg	-	82	79	79	79
HQCNN + BiLSTM	Benign (0)	90	69	78	79
	Malignant (1)	71	91	80	-
Macro Avg	-	81	80	79	79
Weighted Avg	-	82	79	79	79
HQCNN + BiLSTM + MobileNetV2	Benign (0)	91	84	87	87
-	Malignant (1)	82	90	86	-
Macro Avg	-	87	87	87	87
Weighted Avg	-	87	87	87	87

malicious cases while maintaining 82% precision, ensuring better sensitivity for early detection.
- Balanced classification: The overall average F1 score of 87% confirms that the model maintains high classification performance in both benign and malignant cases.

These improvements validate the integration of quantum-enhanced feature extraction with deep learning models, demonstrating that hybrid classical-quantum architectures offer significant advantages over purely classical methods. The accuracy, loss, and confusion matrices of the hybrid quantum convolutional neural network model, a hybrid quantum convolutional neural network with BiLSTM, a hybrid quantum convolutional neural network with BiLSTM, and MobileNetV2 after 30 epochs are shown in Figs. 8, 9, 10, 11, 12 and 13.

To assess the effect of image resolution on model performance, we initially evaluated the

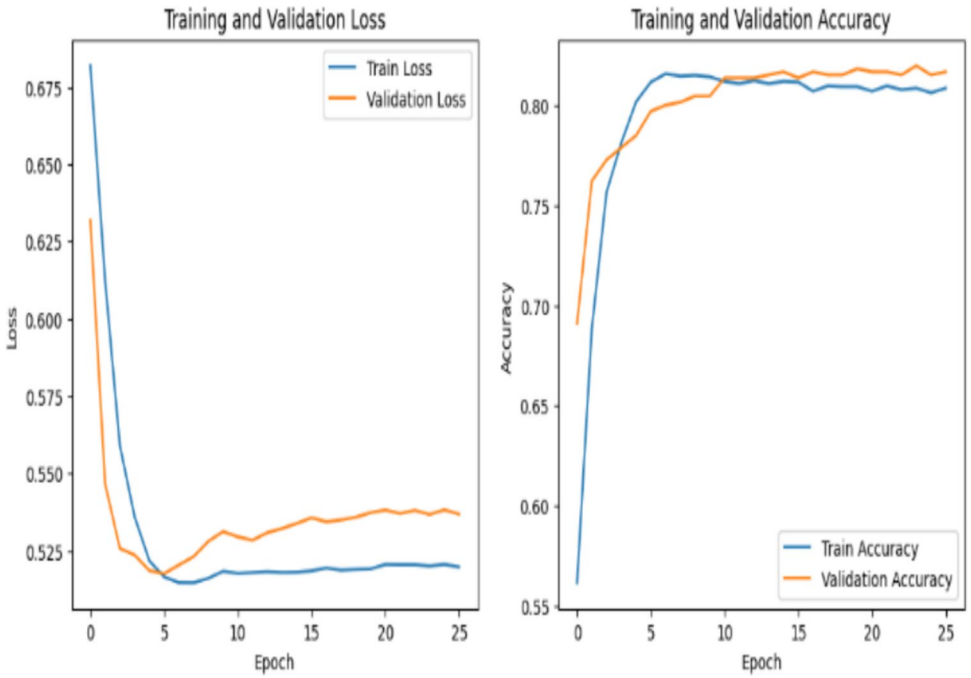

Fig. 8 Accuracy of training and testing data and loss function for the HQCNN model over 30 epochs

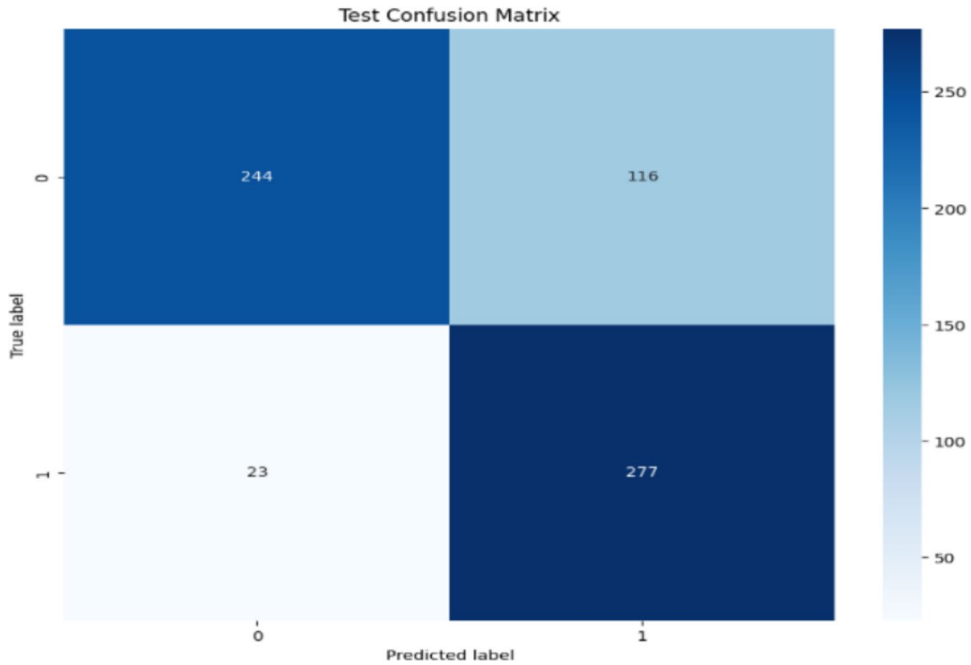

Fig. 9 Confusion matrix of HQCNN model

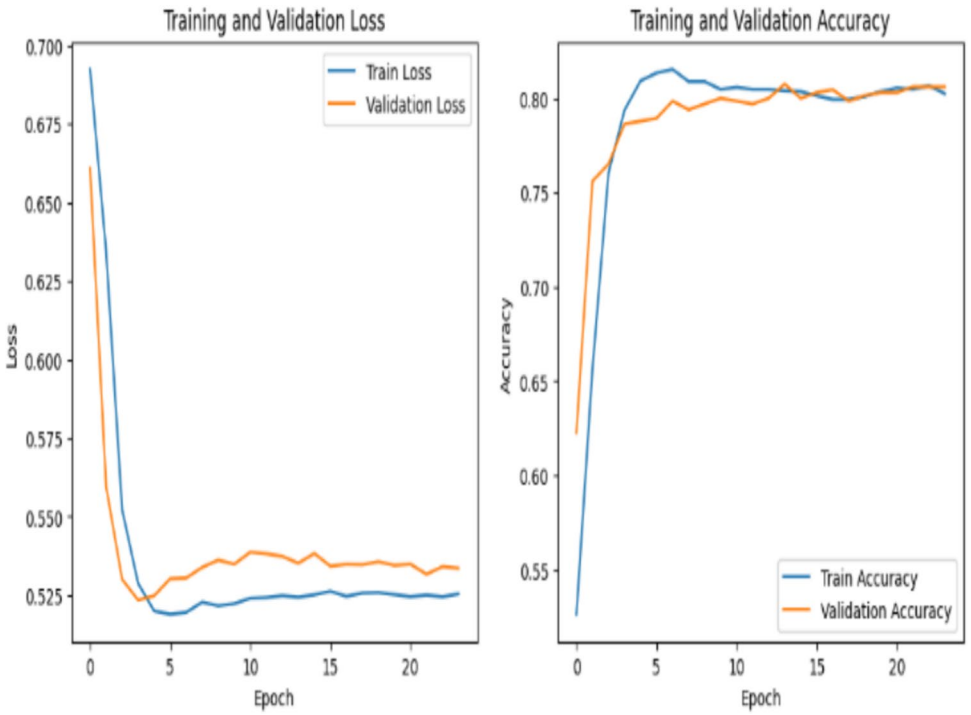

Fig. 10 Loss function and accuracy of training and testing data for HQCNN with BiLSTM model for 30 epochs

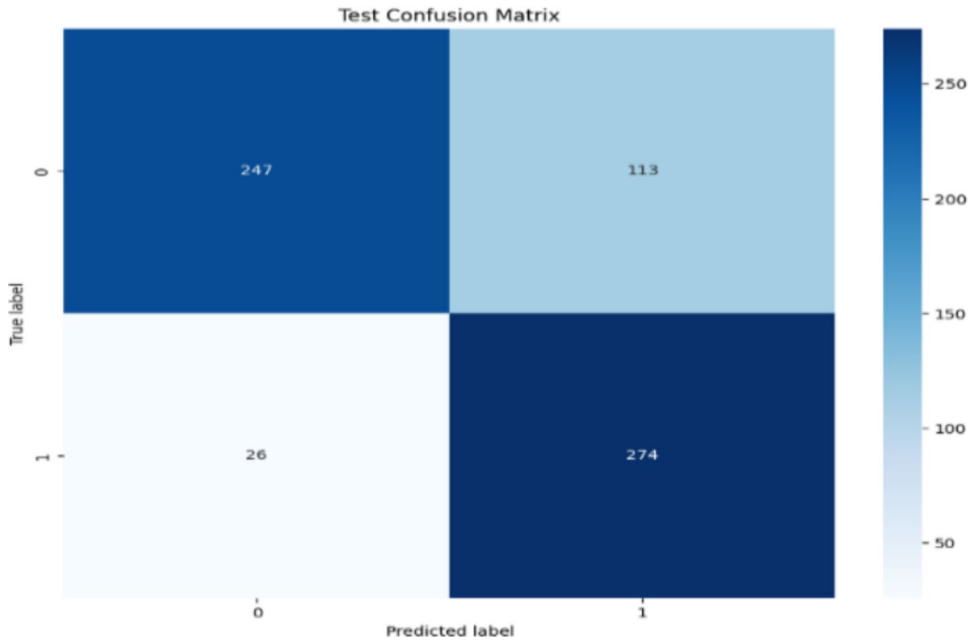

Fig. 11 Confusion matrix of HQCNN with BiLSTM model

HQCNN+BiLSTM+MobileNetV2 model using 32 × 32-pixel images, establishing a baseline. Expanding this analysis, we tested a 64 × 64-pixel resolution, which led to improved training accuracy (0.9586) while maintaining balanced test performance (0.8717), indicating that a moderate increase in resolution enhances model learning without significantly affecting generalization.

Additionally, we examined the effect of data augmentation (random horizontal flips and rotations) at 64 × 64 resolution. While augmentation slightly increased the training accuracy (96.73%), it decreased the test accuracy (84.67%), suggesting that data augmentation may have led to overfitting in this configuration. These results highlight the trade-off between accuracy, augmentation, and generalization, emphasizing the need for careful selection of preprocessing techniques. A comparative summary is provided in Table 6, with the corresponding performance curves plotted in Figure 14.

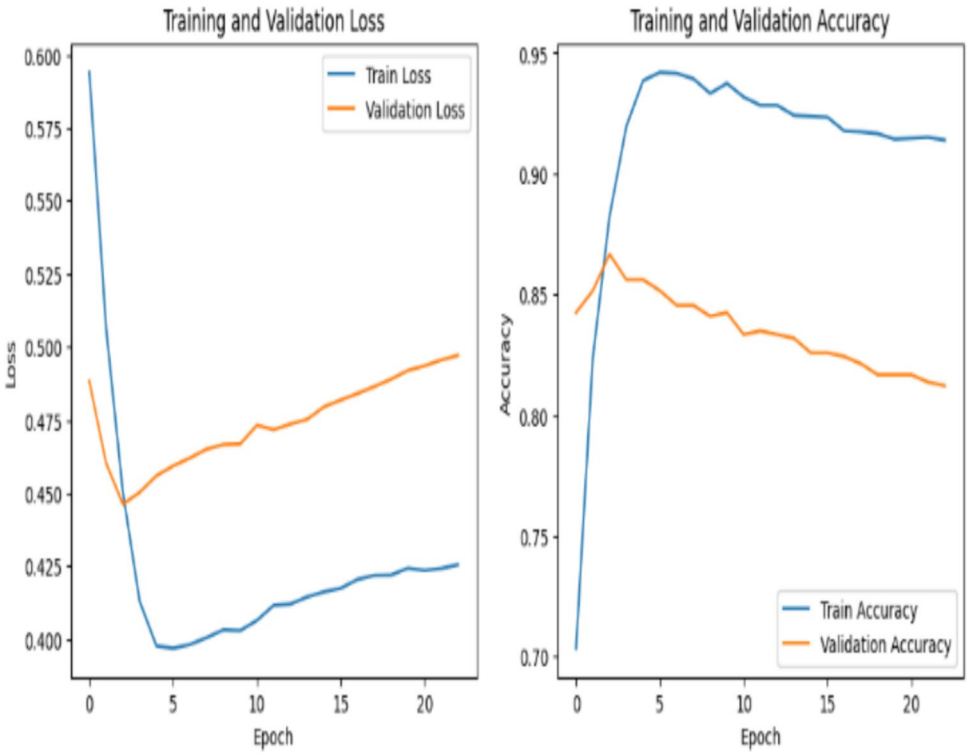

Fig. 12 Accuracy and loss function of test and training data for HQCNN using BiLSTM and the MobileNetV2 model across 30 epochs

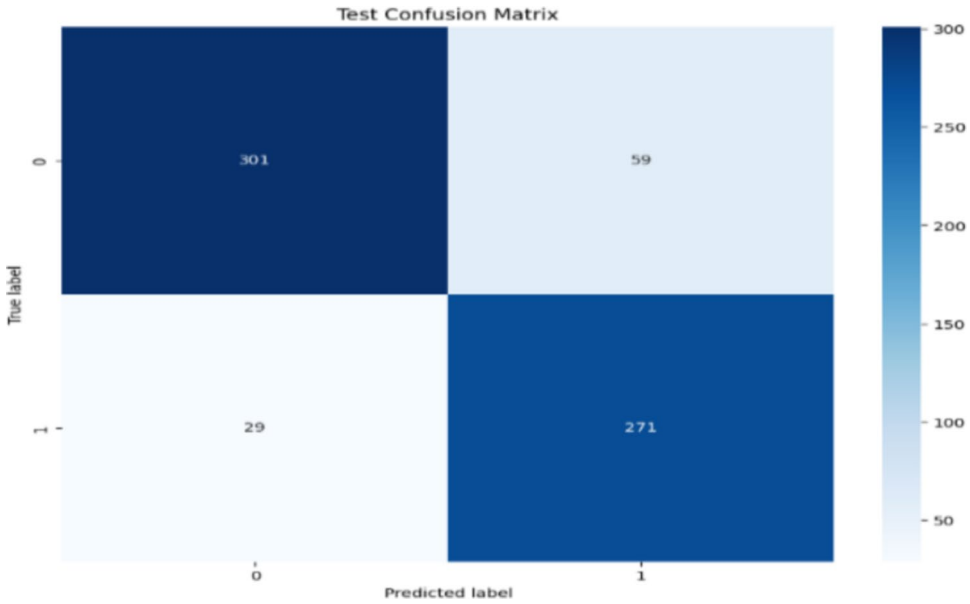

Fig. 13 Confusion matrix of HQCNN with BiLSTM and MobileNetV2 model

Table 6 Comparing the performance of the 64 × 64 model with and without data augmentation

HQCNN + BILTSM + MobileNetV2	Training accuracy, %	Test accuracy, %
Without augmentation	95.8	87.17
With augmentation	96.7	84.6

6.1.2 Skin cancer dataset of size 128 × 128

The experiment was also run using the same dataset but with a modified image size of 128 × 128 pixels. Training was done for 30 epochs using a batch size of 32 and a learning rate of 0.0001, as indicated in Table 7. The HQCNN with BiLSTM

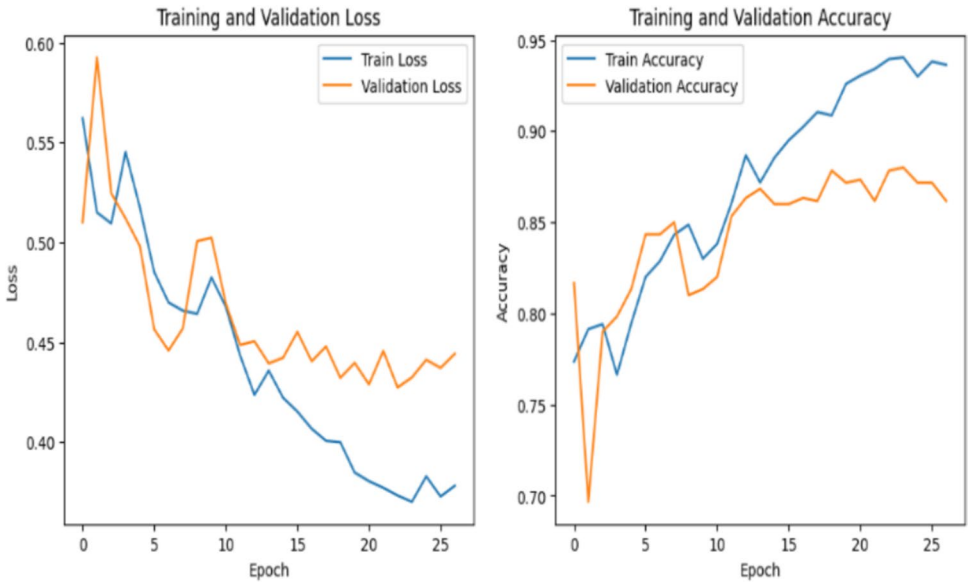

Fig. 14 Training and testing accuracy-loss curves for the HQCNN + BiLSTM + MobileNetV2 model at 64 × 64

Table 7 Model results after 30 epochs for image size 128 × 128 skin cancer dataset

Learning rate	Test accuracy, %	Training accuracy, %	Time (s)	Batch size	Model
0.0001	84.17	90.68	2865	32	HQCNN
0.0001	83.83	90.18	3194	32	HQCNN + BILTSM
0.0001	89.3	97.7	4538	32	HQCNN + BILTSM + MobileNetV2

and MobileNetV2 model outperformed the other models by a wide margin, achieving an accuracy of 97.7% on the training set and 89.3% on the test set with 128-pixel pictures, according to the results. All models' performance improved steadily when the image size was increased from 32 to 128 pixels.

Furthermore, we evaluated the real-time feasibility of the HQCNN + BiLSTM + MobileNetV2 model by measuring its inference time on a CPU-based system. The model achieved an average processing time of 0.0332 s per image at 128 × 128 resolution, demonstrating its potential for near-real-time applications. However, further improvements—such as GPU acceleration, model quantization, or deployment on optimized hardware—could enhance its efficiency in practical clinical settings.

To further validate the practicality of the proposed architecture, we conducted a detailed analysis of the computational cost associated with incorporating the quantum layer into the model. To this end, we compared the HQCNN + BiLSTM + MobileNetV2 model with and without the quantum-enhanced layer using classical simulations. The quantum-enhanced version achieved a test accuracy of 89.3%, with 2.39 million parameters and 104.32 MMac FLOPs, requiring 4538 s for training and an average of 1.0624 s per batch for inference on the CPU. In contrast, the classical counterpart achieved an accuracy of 87.0% with 2.24 million parameters and 106.55 MMac FLOPs, requiring 3713.91 s for training and 0.5816 s per batch for inference. These results demonstrate that the quantum-enhanced model yields a measurable performance boost, albeit with a moderate increase in overhead. The slight reduction in FLOPs in the hybrid model is attributed to the integrated design of the variable quantum circuit, which replaces some classical processing components. Although current limitations arise from simulating quantum operations on classical hardware, future deployment on quantum processors could alleviate these limitations. Thus, the quantum layer provides a forward-compatible and practically applicable improvement over classical architectures in the field of medical image classification.

Figures 15, 16, 17, 18, 19 and 20 present the accuracy, loss, and confusion matrices for HQCNN, HQCNN + BiLSTM, and HQCNN + BiLSTM + MobileNetV2 after 30 epochs. To validate the statistical significance of performance differences among these models, we conducted a one-way ANOVA test. The analysis yielded an F-statistic of 0.3578 with a p-value of 0.7132, indicating that the observed variations in performance are not statistically significant ($p > 0.05$). While modest differences exist, these results suggest that the improvements are not statistically meaningful. Nevertheless, practical considerations such as computational efficiency and model complexity may still influence the choice of architecture depending on specific application requirements.

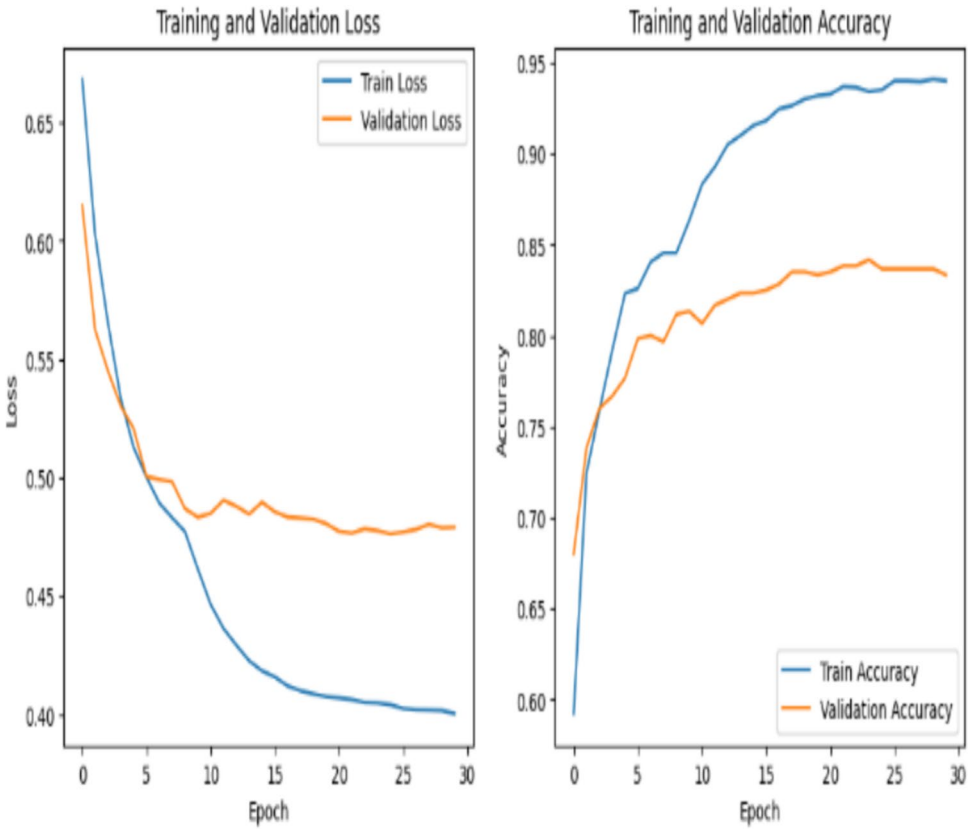

Fig. 15 Accuracy of training and testing data and loss function for the HQCNN model over 30 epochs

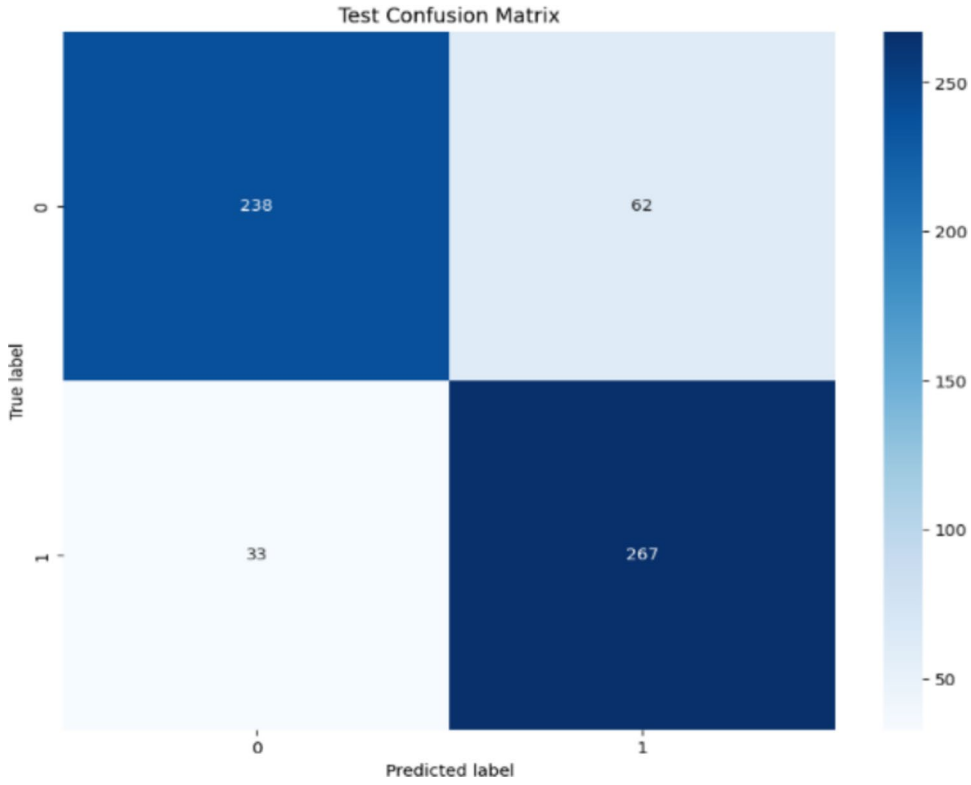

Fig. 16 Confusion matrix of HQCNN model

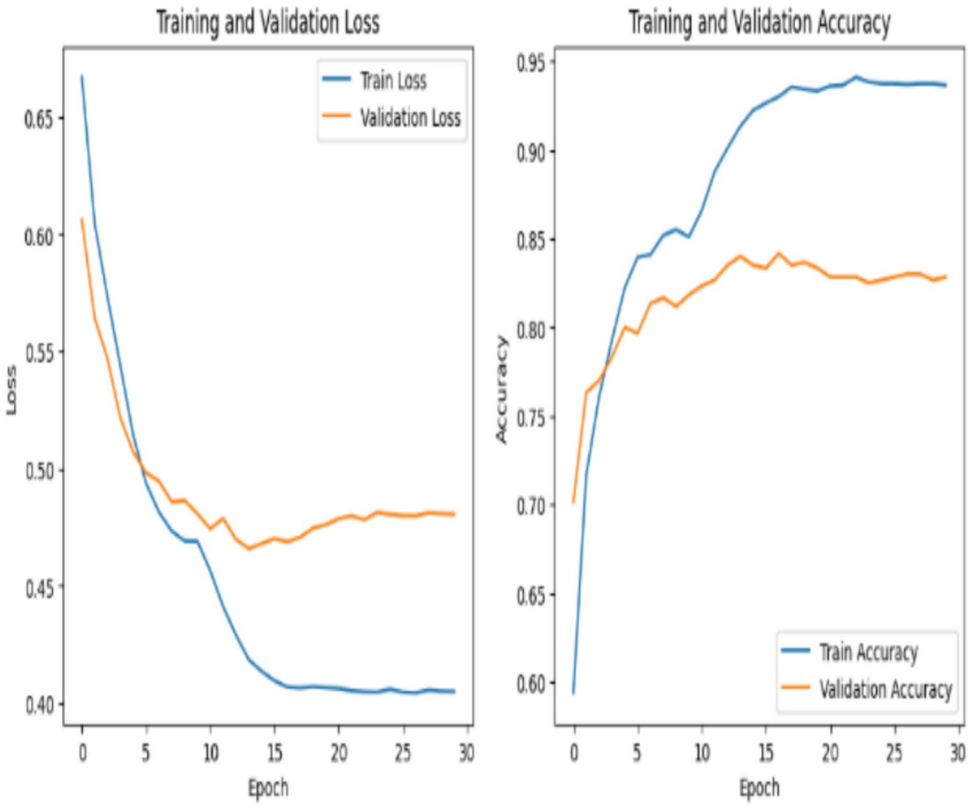

Fig. 17 Accuracy and loss function of the training and testing data for the HQCNN using the BiLSTM model across 30 epochs

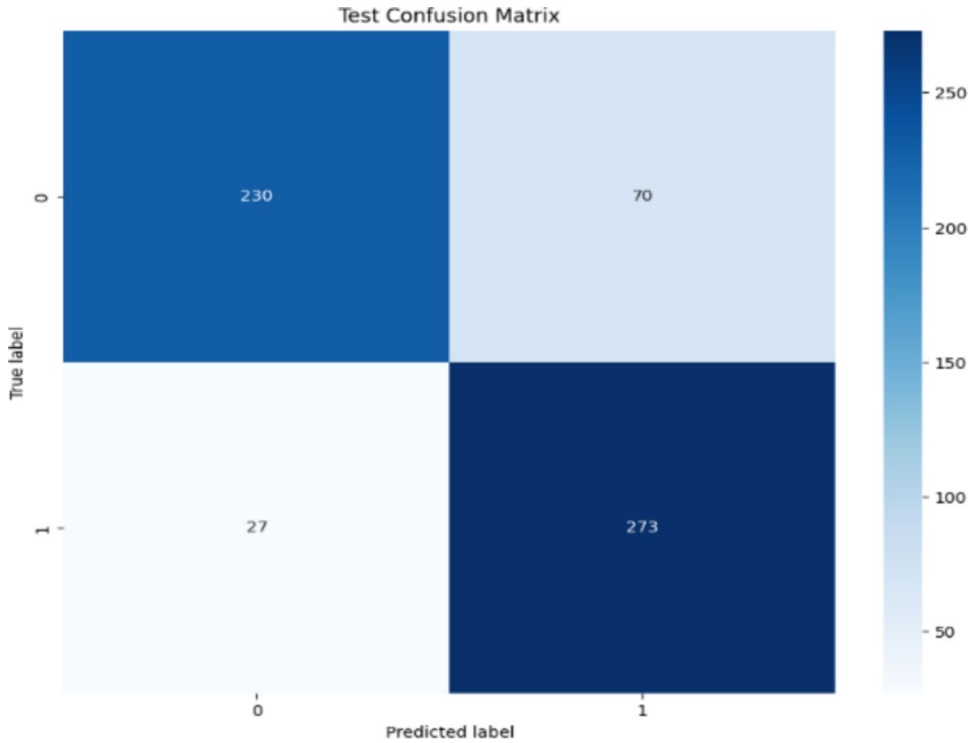

Fig. 18 Confusion matrix of HQCNN with BiLSTM model

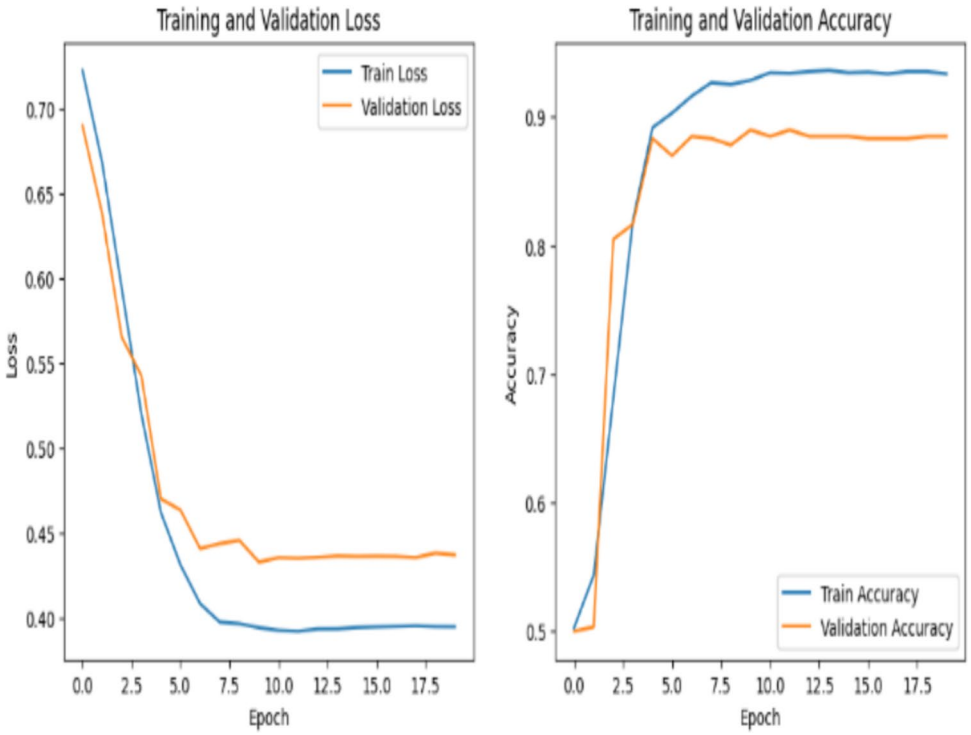

Fig. 19 Accuracy and loss function of test and training data for HQCNN using BiLSTM and the MobileNetV2 model across 30 epochs

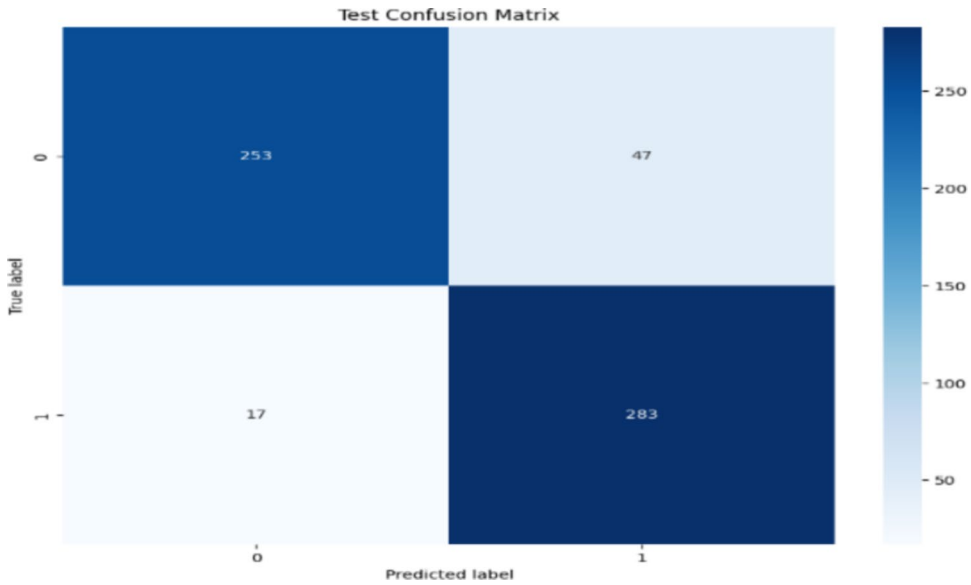

Fig. 20 Confusion matrix of HQCNN with BiLSTM and MobileNetV2 model

Table 8 Performance metrics with confidence intervals

Metric	Score	95% confidence interval
AUROC	0.9025	[0.8742–0.9272]
Balanced accuracy	0.8850	[0.8575–0.9088]
PR AUC	0.8561	[0.8136–0.8938]

To evaluate the robustness of the proposed HQCNN + BiLSTM + MobileNetV2 model beyond benchmark accuracy, we calculated additional class-balanced metrics relevant to medical image classification. These include the area under the ROC curve (AUROC), balanced accuracy, and the area under the precision-recall curve (PR AUC). We

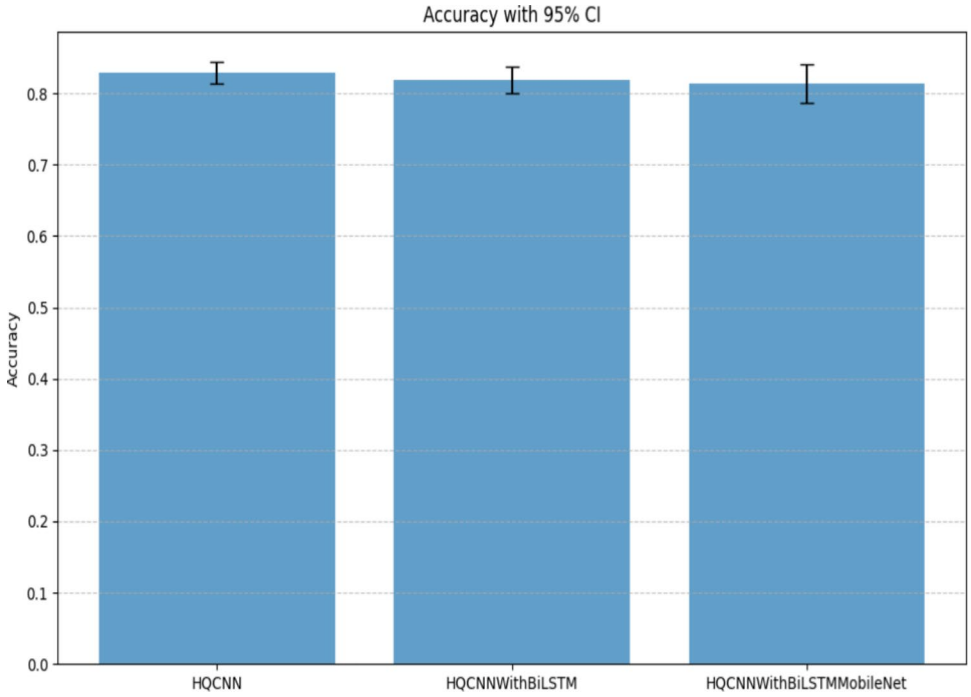

Fig. 21 Accuracy with 95% confidence intervals

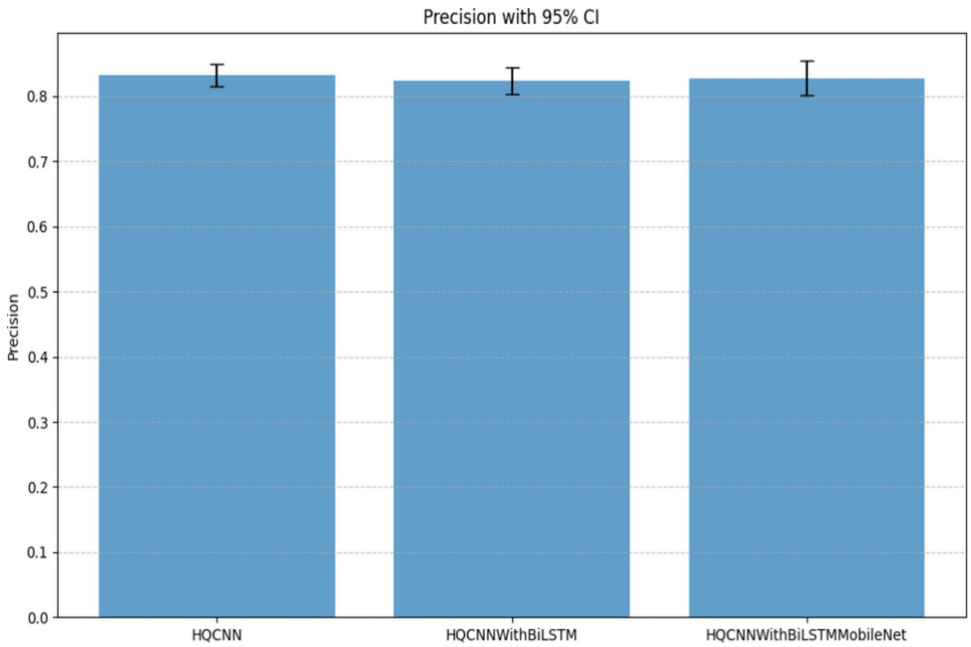

Fig. 22 Precision with 95% confidence intervals

also estimated 95% confidence intervals using a nonparametric bootstrap approach with 1000 resamplings. These metrics provide a deeper understanding of the model's discriminative ability, especially in the presence of potential class imbalance. All results support the reliability of the proposed method in distinguishing between benign and malignant skin lesions. A summary of the results is provided in Table 8.

To assess the variability in model performance, we report accuracy, precision, recall, and F1 score along with their 95% confidence intervals. Figures 21, 22, 23 and 24 illustrate these metrics, demonstrating that the observed differences between HQCNN, HQCNN + BiLSTM, and HQCNN + BiLSTM + MobileNetV2 are statistically stable. The tight confidence intervals indicate consistent performance across different runs, strengthening the reliability of our findings.

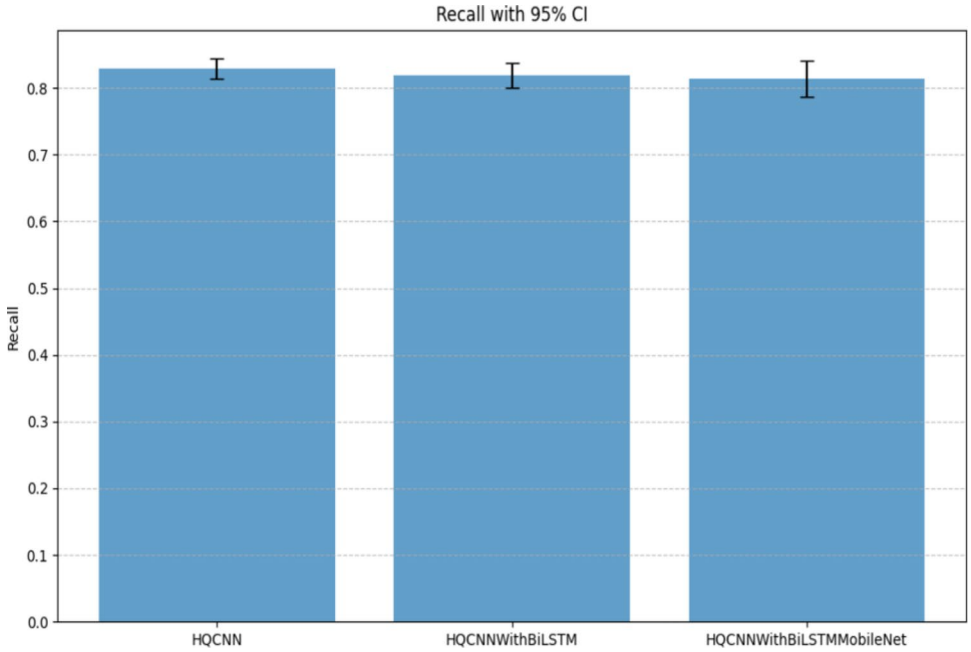

Fig. 23 Recall with 95% confidence intervals

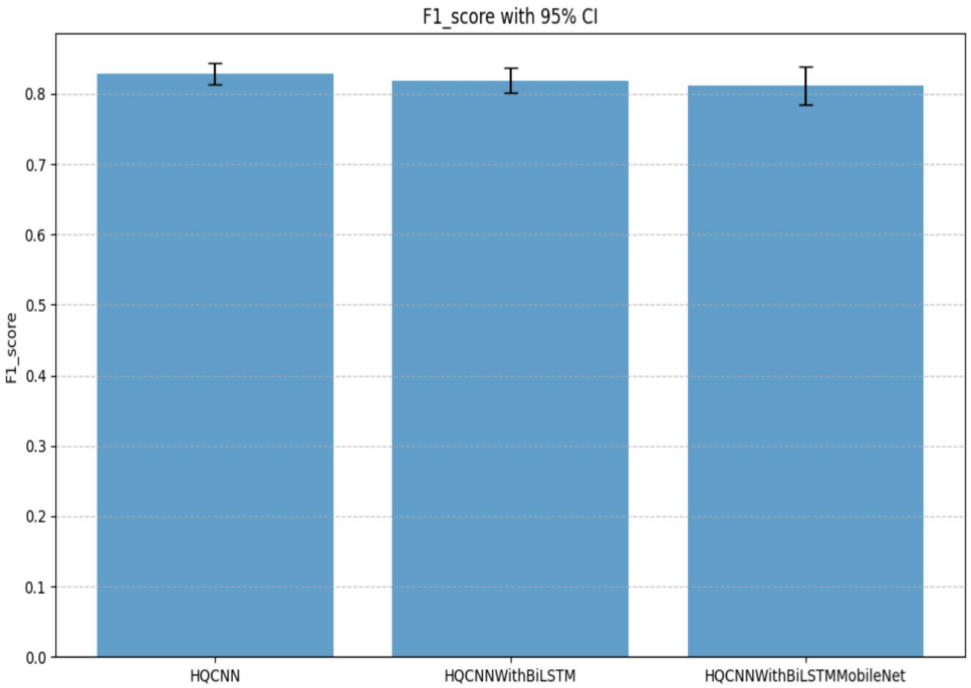

Fig. 24 F1 score with 95% confidence intervals

6.2 Excision study and model robustness

To comprehensively evaluate the contribution of each component in our proposed hybrid quantum–classical neural network (HQCNN), we conducted an ablation study. This study isolates the influence of BiLSTM and MobileNetV2 and combines them with HQCNN to analyze their individual and collective effects on classification performance. We evaluated five key performance metrics: accuracy, precision, recall, specificity, and F1 score. The following model configurations were analyzed:

- HQCNN (baseline): a hybrid quantum–classical convolutional neural network without BiLSTM or MobileNetV2
- HQCNN + BiLSTM: combining HQCNN with BiLSTM to capture sequential dependencies in lesion patterns
- HQCNN + MobileNetV2: an HQCNN with MobileNetV2 as a feature extractor to improve the representation of spatial features

Table 9 Study of the proposed ablation method after 30 epochs

Model	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1 score (%)
HQCNN	84.17	81.15	89	79.33	84.92
HQCNN + BiLSTM	83.83	79.65	91	76.67	85.03
HQCNN + MobileNetV2	87.83	85.42	91.33	84.33	88.28
HQCNN + BiLSTM + MobileNetV2	89.33	85.76	94.33	84.33	89.81

- HQCNN + BiLSTM + MobileNetV2: a complete model that combines the three components for maximum classification efficiency

All models were trained under identical conditions (batch size = 32, learning rate = 0.0001, 30 epochs, and image size 128 × 128 pixels) to ensure fair comparison. The experimental results are summarized in Table 9.

The HQCNN model alone achieves an accuracy of 84.17%, demonstrating its ability to extract meaningful features. However, it lacks advanced feature representation and temporal pattern recognition capabilities, limiting its ability to effectively distinguish complex lesion structures. The BiLSTM effect improves recall to 91%, meaning the model correctly identifies more malignant cases. However, specificity drops to 76.67%, resulting in more false positives. This trade-off highlights the ability of BiLSTM to enhance sensitivity, but it does not significantly improve overall accuracy without additional feature extraction mechanisms. The MobileNetV2 effect significantly improves accuracy, reaching 87.83%, demonstrating its effectiveness in improving feature extraction. Precision increases to 85.42%, reducing false positives and making predictions more reliable. Recall also improves slightly to 91.33%, confirming that MobileNetV2 enhances the ability to extract detailed lesion features. The full model (HQCNN + BiLSTM + MobileNetV2) achieves the best performance, with an accuracy of 89.33%, a recall of 94.33%, and an F1 score of 89.81%. The increased recall ensures that nearly all malignant cases are detected, reducing the risk of misclassification. Specificity remains at 84.33%, while maintaining a balance between sensitivity and false positives, ensuring the model is reliable and accurate.

This ablation study strongly justifies the need to integrate BiLSTM and MobileNetV2 within HQCNN. BiLSTM boosts recall, improving the model's ability to detect malignant cases, while MobileNetV2 significantly improves feature extraction, enhancing classification performance. The full model offers the best trade-off between sensitivity and specificity, providing a highly accurate and clinically applicable classification system. These results directly support the necessity of our proposed architecture and demonstrate its superiority over alternative configurations.

Table 10 Comparison of the accuracy of the suggested and current approaches

Method	Accuracy (%)
VGG16 (Tuncer et al. 2024b)	83.7
CNN (Maurya et al. 2024)	87
CNN-based two-phase evolutionary framework (Ghosh et al. 2023)	83
EfficientNetB3 (Salian and Sawarkar 2022)	87
ResNetv50 (He et al. 2016)	85
DenseNet169 (Huang et al. 2017)	87
MobileNetv3-Large (Howard, et al. 2019)	87
Mixer-B16 (Tolstikhin et al. 2021)	88
MobileViT-Small (Mehta and Rastegari 2021)	89
RepViT-m1 (Chen et al. 2025)	87
NextViT-Base (Li, et al. 2022)	87.5
ConvMixer (Trockman and Kolter 2022)	87.6
PoolFormer-M36 (Yu, et al. 2022)	88
Swin-Tiny (Dagnaw et al. 2024)	87.7
Swin-Small (Dagnaw et al. 2024)	87.7
Swin-Base (Dagnaw et al. 2024)	87.8
ViT-Base (Dagnaw et al. 2024)	88.6
EfficientNetV2-S (Dagnaw et al. 2024)	88.6
SwinV2-Small-Window16 (Pacal et al.2024)	85.92
ViT-ResNet50 (Himel et al. 2024)	80.9
Proposed model	89.3

6.3 Comparing the accuracy of one model to another

To evaluate the effectiveness of the proposed HQCNN + BiLSTM + MobileNetV2 model, we conducted a comprehensive comparison with a wide range of state-of-the-art deep learning architectures, including traditional CNN-based models, transformer-based architectures, hybrid frameworks, and quantum-inspired methods. Table 10 summarizes the classification accuracy results for these models.

The proposed model achieves the highest accuracy of 89.3%, outperforming established CNN architectures such as VGG16 (83.7%), ResNet50 (85%), and EfficientNetB3 (87%). It also demonstrates competitive or superior performance compared to state-of-the-art transformer-based and hybrid models, including NextViT-Base (87.5%),

ConvMixer (87.6%), PoolFormer-M36 (88%), Swin Transformer variants (87.7–87.8%), and even ViT-Base (88.6%) and EfficientNetV2-S (88.6%) models. These results confirm the robustness of our architecture, particularly in capturing domain-specific features relevant to medical imaging.

While the proposed model demonstrates first-class classification performance, it is not immune to occasional misclassifications. Misclassifications typically occur in visually ambiguous situations, such as benign lesions simulating malignant lesions under varying lighting conditions or due to poor image quality. These factors are common challenges in dermatological imaging and suggest opportunities for future improvement through advanced preprocessing, lesion localization, or multimodal inputs. In terms of computational trade-offs, our model balances accuracy and architectural efficiency, integrating MobileNetV2 (a lightweight CNN backbone), BiLSTM (temporal context modeling), and a quantum simulation layer. Although quantum circuit simulation introduces overhead, the resulting performance gains—particularly in AUROC and PR AUC—justify this cost. Transformer-based models such as Swin, ConvNeXt, and ViT are known for their strong performance on large-scale datasets, but they often require significantly more computational resources. Overall, the proposed hybrid quantum–classical approach offers a compelling trade-off between accuracy and complexity in the context of skin cancer image classification. It outperforms or rivals many state-of-the-art SOTA models, demonstrating the potential of integrating quantum-inspired layers into practical medical AI systems. Future work will explore architectural improvements inspired by these baselines, with a focus on improving generalization, real-time efficiency, and clinical reliability.

Figure 25 illustrates the classification accuracy of various deep learning models. While early CNN-based models such as VGG16 and traditional architectures such as ResNet50 and DenseNet169 achieve moderate performance, newer models—including MobileViT-Small, PoolFormer-M36, and ConvMixer—demonstrate improved generalization. However, several Vision Transformer variants, such as Swin-Tiny, ViT-Base, and SwinV2-Small, despite their architectural sophistication, do not surpass the accuracy of our proposed model. The HQCNN + BiLSTM + MobileNetV2 architecture achieves the highest classification accuracy (89.3%), highlighting its superior feature extraction and classification capabilities. This confirms the effectiveness of combining hybrid quantum convolution, recurrent learning, and efficient feature representation. It is important to note that the reported performance values for some of the benchmark models were obtained from the published literature under similar skin lesion classification tasks. In future

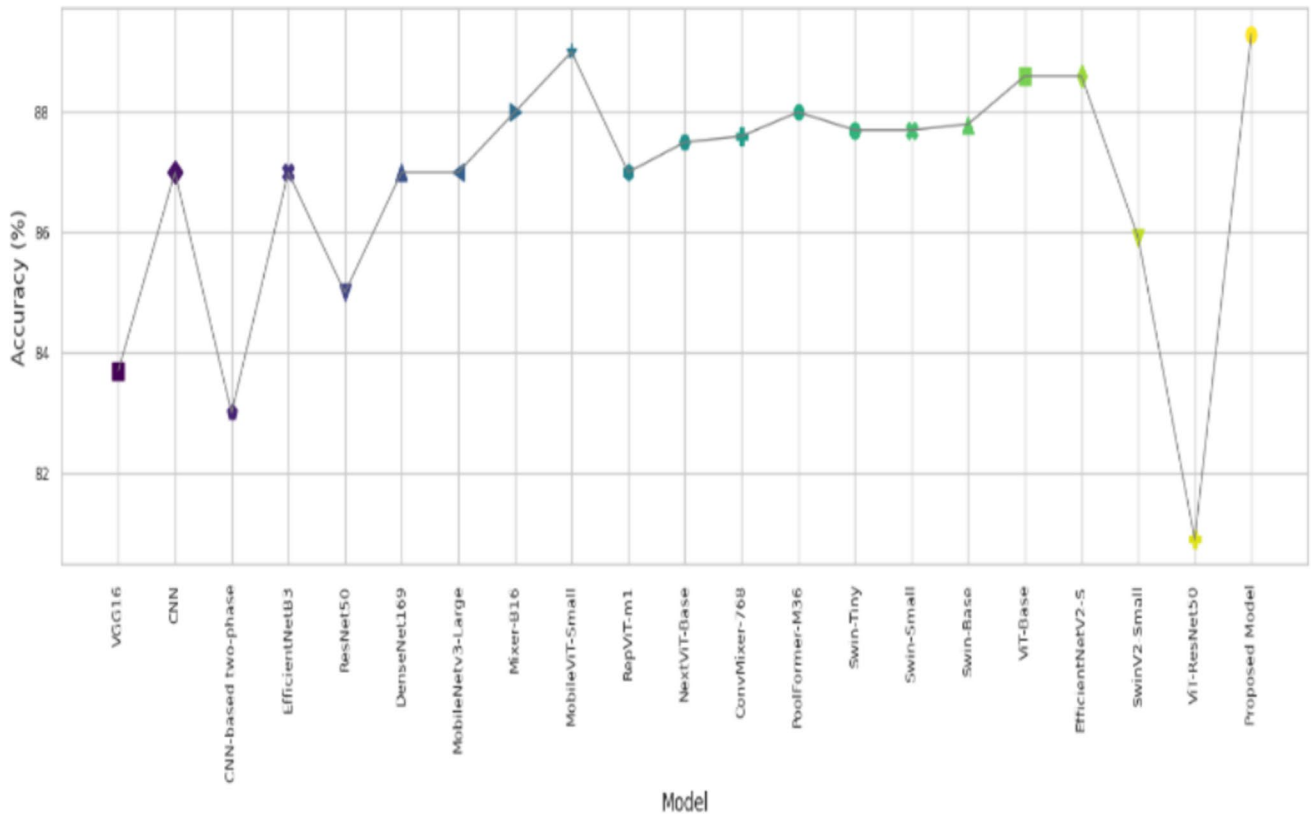

Fig. 25 Comparison of classification model accuracies

work, we aim to explore the integration of lightweight self-attention mechanisms into hybrid quantum models to further enhance performance and efficiency in medical image classification tasks.

6.4 Ethical considerations and clinical implications

There are ethical advantages and disadvantages to integrating AI into medical imaging, especially when it comes to dataset bias and the possibility of incorrect categorization and clinical distribution.

- Dataset bias and generalizability: AI models for skin cancer detection rely on large datasets, yet these datasets may not fully represent diverse skin types, lesion variations, or imaging conditions. Dataset biases can reduce the generalizability of the model, impacting diagnostic performance for underrepresented patient populations. Future work should focus on curating more diverse datasets to mitigate this problem.
- AI misclassification and clinical risks: While our model demonstrates high accuracy, no AI system is infallible. False negatives pose a significant risk, potentially delaying critical treatment. Similarly, false positives can lead to unnecessary biopsies and patient anxiety. To address this, AI-based diagnosis should function as a decision support tool rather than a standalone diagnostic system, ensuring that dermatologists verify the model's predictions.
- Regulatory and ethical compliance: AI deployment in healthcare must align with ethical and regulatory standards, with an emphasis on patient safety, explainability, and fairness. Future studies should explore explainable AI (XAI) techniques to enhance physician trust and model transparency. Additionally, adversarial robustness testing should be considered to assess the reliability of AI under real-world conditions. Our study ensures compliance with these ethics.

6.5 Clinical applications and publication considerations

The proposed hybrid classical-quantum model (HQCNN + BiLSTM + MobileNetV2) has potential applications in automated skin cancer classification and clinical decision support. However, several challenges must be addressed before real-world deployment, including model interpretability, external validation, computational feasibility, and regulatory compliance.

1. Potential clinical applications
   - AI-assisted dermatology: The model can serve as a decision support system for dermatologists, helping analyze images of skin lesions and prioritize high-risk cases for further evaluation.
   - Telemedicine and Telediagnosis: Due to its computational efficiency with MobileNetV2, the model has the potential to be deployed in cloud-based platforms or mobile applications to facilitate remote screening and diagnosis in resource-poor areas.
   - Integration with clinical imaging devices: The model can be further optimized for integration with dermatoscopes, enhancing real-time assessment of skin lesions in dermatology clinics.
2. Clinical publication considerations
   - The model was trained on the Kaggle skin cancer dataset, but further validation on multi-institutional datasets is necessary before clinical deployment.
   - Testing on skin images from real-world hospital settings will be an essential next step.
   - While quantum simulation introduces additional computational overhead, we have optimized our model using MobileNetV2, making it suitable for deployment on low-resource devices.
   - Future research will explore real quantum devices (e.g., IBM Quantum and Rigetti) to evaluate computational efficiency.
   - To be clinically approved, the model must adhere to regulatory standards for medical AI such as the Food and Drug Administration (FDA) and CE (Conformité Européenne) certification, ensuring patient safety, transparency, and fairness.

6.6 Limitations and future work

This study presents a hybrid classical-quantum model for skin cancer classification, but several challenges must be addressed before real-world deployment.

- While this study uses quantum circuit simulations, we acknowledge the importance of evaluating the performance of real quantum devices to assess the feasibility of practical deployment.
- The computational overhead of quantum simulation increases training time, which may hinder real-time deployment.
- Quantum noise and decoherence. Unlike simulations, real quantum processors introduce significant noise, which can degrade classification performance.
- Scalability issues: The number of available qubits remains a limiting factor, limiting the depth and complexity of quantum-enhanced layers.

- The dataset used lacked patient- or lesion-level metadata, preventing strict patient-level separation between training and test sets. This raises the potential for information leakage if multiple images from the same patient were inadvertently split across subsets. Future research should leverage datasets that contain annotated lesion or patient identifiers to allow for more robust clinical validation.
- The dataset may not fully capture the diverse demographics of patients, leading to potential bias.
- Variation in skin color, lesion appearance, and imaging conditions may affect classification performance.
- While the hybrid quantum–classical model leverages MobileNetV2 for efficiency, the quantum layers introduce additional computational overhead, making deployment on low-power medical devices challenging.
- Although MobileNetV2 is known for its lightweight design and efficiency, it may not be able to handle subtle lesion features, especially in small or imbalanced medical image datasets. This can limit its ability to detect complex patterns critical for accurate diagnosis. In our model, we mitigate this by combining BiLSTM and HQCNN units to enrich the feature space and capture more complex structures.
- Future research will focus on the following:
- Reducing the depth of quantum circuits to improve computational efficiency.
- Exploring hybrid classical-quantum architectures that reduce reliance on simulations.
- Deploy lightweight versions of the model for real-time clinical applications.
- Future research will focus on enhancing model interpretability (e.g., Grad-CAM, SHAP) to ensure compliance with clinical guidelines.
- Future work will focus on external validation using multi-institutional datasets to improve model generalization.
- To address model robustness, we demonstrate that our use of MobileNetV2 for feature extraction, dropout regularization (0.3), and learning rate scheduling enhances generalization, while future work will evaluate adversarial robustness (FGSM and PGD) and noise sensitivity (Gaussian and salt-and-pepper noise) to further enhance reliability.
- Explore more expressive backbones—such as attention-based CNNs or transformer models—to further enhance lesion sensitivity, especially in underrepresented cases.

6.7 Ethical considerations and clinical integration

While the proposed HQCNN-BiLSTM-MobileNetV2 model shows promising results in skin cancer image classification, it is essential to address several ethical and practical aspects. The dataset used in this study may not fully represent the broad diversity of global populations in terms of skin tones, lesion types, and imaging conditions, which could affect the model's generalizability and fairness. Furthermore, like all AI-based systems, the model carries a risk of misclassification, especially in borderline or ambiguous cases. Such errors could lead to inappropriate clinical decisions if used autonomously. Therefore, the proposed model is intended strictly as a clinician-assistive tool, designed to support dermatologists in the diagnostic process by providing a second opinion—not to replace expert judgment. Future work will explore model robustness across diverse populations and real-world clinical settings to ensure safe and equitable deployment.

7 Conclusions

This research integrates MobileNetV2 as a feature extractor into HQCNN with BiLSTM architecture, providing an advanced technique for image classification in medical diagnosis. The proposed model demonstrates significant improvements over existing methods. Its true effectiveness is evident in complex medical challenges, such as skin cancer image classification. Compared to HQCNN and the baseline HQCNN with BiLSTM, which achieved test accuracies of 84.17% and 83.83%, respectively, the proposed model achieves a significant improvement, reaching an accuracy of 89.3% at an image resolution of 128×128 pixels. These results highlight the potential of integrating deep learning and quantum-inspired methods for medical imaging. However, it is important to recognize the current limitations of quantum computing in practical applications. This study relies on quantum simulators rather than real quantum devices, due to limited coherence and scalability, which hinders large-scale applications. Furthermore, while our hybrid approach demonstrates improvements in accuracy, the computational efficiency of quantum neural networks requires further research. Future research should focus on testing hybrid quantum–classical models on real quantum devices to evaluate their feasibility. Furthermore, improving quantum circuit designs and exploring alternative techniques for extracting quantum-enhanced features could enhance the efficiency and robustness of these models. Despite these limitations, this study lays the foundation for the development of hybrid quantum deep learning in medical imaging, providing a promising direction for future research.

Author Contributions A.H. conceptualized the study, developed the hybrid quantum deep learning model, and performed the primary analysis. A.M. contributed to the implementation of the quantum neural network and data preprocessing. H.E. assisted in conducting the experiments and interpreting the results. A.H. and A.M. wrote the main manuscript text, and H.E. prepared Figs. 1-5 and Tables 1-3. All authors reviewed and approved the final version of the manuscript.

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB). This work is partially supported by Sohag University and Damanhour University.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication All authors consent to the publication of this work.

Competing interests The authors declare no competing interests.

References

Agarwal K, Singh T (2022) Classification of skin cancer images using convolutional neural networks;*arXiv preprint* arXiv:2202.00678. Accessed 12 Jan 2025

Alnowaiser K, Saber A, Hassan E, Awad WA (2024) An optimized model based on adaptive convolutional neural network and grey wolf algorithm for breast cancer diagnosis. PLoS ONE 19(8):e0304868

Angelina C, Ulfitria RU (2024) Classification of skin cancer using ResNet and VGG deep learning network. Proceedings of the 11th International Applied Business and Engineering Conference, ABEC 2023, September 21st, 2023, Bengkalis, Riau, Indonesia

Arthur D, Date P (2022) Hybrid quantum-classical neural networks," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, IEEE, pp. 49–55

Baig MM, Hosseini HG, Lindén M (2016) Machine learning-based clinical decision support system for early diagnosis from real-time physiological data.Conference: TENCON 2016 - 2016 IEEE Region 10 Conference IEEE 2016:2943–2946

Bharal R, Krishna OV (2021) Social media sentiment analysis using CNN-BiLSTM. International Journal of Science and Research (IJSR) 10:656–661

Chen K, Lin J, You B, Luo H (2025) Fully automatic surface defect detection of CFRP using computer vision and an augmented YOLOv8 model. J Perform Constr Facil 39(3):04025006

Cheng L, Lopez-Beltran A, Massari F, MacLennan GT, Montironi R (2018) Molecular testing for BRAF mutations to inform melanoma treatment decisions: a move toward precision medicine. Mod Pathol 31(1):24–38

Cong I, Choi S, Lukin MD (2019) Quantum convolutional neural networks. Nat Phys 15(12):1273–1278

Cui Z, Ke R, Pu Z, Wang Y (2020) Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values. Transp Res Part C Emerg Technol 118:102674

Dagnaw G. H, M. El Mouhtadi, and M. Mustapha (2024) Skin cancer classification using vision transformers and explainable artificial intelligence. J Med Artif Intell. 7:14-14

Dandu R, M. V. Murthy, and Y. B. R. Kumar (2023) Transfer learning for segmentation with hybrid classification to detect melanoma skin cancer. Heliyon. 9:e15416

Doan A.-D, M. Sasdelli, D. Suter, and T.-J. Chin (2022) A hybrid quantum-classical algorithm for robust fitting. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 417–427.

Elbedwehy S, Hassan E, Saber A, Elmonier R (2024) Integrating neural networks with advanced optimization techniques for accurate kidney disease diagnosis. Sci Rep 14(1):21740

Erten M et al (2023) Automated urine cell image classification model using chaotic mixer deep feature extraction. J Digit Imaging 36(4):1675–1686

Fanconi C, (2019) Skin cancer: malignant vs. benign-processed skin cancer pictures of the ISIC archive.

Femiano F, Lanza A, Buonaiuto C, Gombos F, Di Spirito F, Cirillo N (2008) Oral malignant melanoma: a review of the literature. J Oral Pathol Med 37(7):383–388

Ghosh A, Jana ND, Das S, Mallipeddi R (2023) Two-phase evolutionary convolutional neural network architecture search for medical image classification. IEEE Access 11:115280–115305

Gupta B, Prakasam P, Velmurugan T (2022) Integrated BERT embeddings, BiLSTM-BiGRU and 1-D CNN model for binary sentiment classification analysis of movie reviews. Multimed Tools Appl 81(23):33067–33086

Gupta S, A. Panwar, and K. Mishra (2021) Skin disease classification using dermoscopy images through deep feature learning models and machine learning classifiers," in *IEEE EUROCON 2021–19th International Conference on Smart Technologies*, IEEE. 170–174.

Hassan E, Saber A, Elbedwehy S (2024) Knowledge distillation model for acute lymphoblastic leukemia detection: exploring the impact of nesterov-accelerated adaptive moment estimation optimizer. Biomed Signal Process Control 94:106246

He K, X. Zhang, S. Ren, and J. Sun. (2016) Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 770–778.

He X, Wang Y, Zhao S, Chen X (2023) Co-attention fusion network for multimodal skin cancer diagnosis. Pattern Recognit 133:108990

Himel GMS, Islam MM, Al-Aff KA, Karim SI, Sikder MKU (2024) Skin cancer segmentation and classification using vision transformer for automatic analysis in dermatoscopy-based noninvasive digital system. Int J Biomed Imaging 2024(1):3022192

Houssein EH, Abohashima Z, Elhoseny M, Mohamed WM (2022) Hybrid quantum-classical convolutional neural network model for COVID-19 prediction using chest X-ray images. J Comput des Eng 9(2):343–363

Howard A. et al., (2019) Searching for mobilenetv3. Conference: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). 1314–1324.

Howard A.G, et al., (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications," *arXiv preprint* arXiv.org/abs/1704.04861. Accessed 20 Dec 2024

Huang G, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger (2017) Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition.* 4700–4708.

Huang H-Y, Hsiao Y-P, Mukundan A, Tsao Y-M, Chang W-Y, Wang H-C (2023) Classification of skin cancer using novel hyperspectral imaging engineering via YOLOv5. J Clin Med 12(3):1134

Hussein AA, Montaser AM, Elsayed HA (2024) Classification of spine images using hybrid quantum neural network classifier. Quantum 16:17

İnce S, Kunduracioglu I, Bayram B, Pacal I (2025a) U-Net-based models for precise brain stroke segmentation. Chaos Theory Appl 7:50–60. https://doi.org/10.51537/chaos.1605529

Ince S, Kunduracioglu I, Algarni A, Bayram B, Pacal I (2025) Deep learning for cerebral vascular occlusion segmentation: a novel ConvNeXtV2 and GRN-integrated U-Net framework for diffusion-weighted imaging. Neuroscience 574:42–53. https://doi.org/10.1016/j.neuroscience.2025.04.010

Kashif M, Al-Kuwari S (2022) Demonstrating quantum advantage in hybrid quantum neural networks for model capacity. IEEE international conference on rebooting computing (ICRC). IEEE 2022:36–44

Keerthana D, Venugopal V, Nath MK, Mishra M (2023) Hybrid convolutional neural networks with SVM classifier for classification of skin cancer. Biomedical Engineering Advances 5:100069. https://doi.org/10.1016/j.bea.2022.100069

Kirik S et al (2023) FGPat18: Feynman graph pattern-based language detection model using EEG signals. Biomed Signal Process Control 85:104927

Krushelnytskyi I (2023) Comparing the performance of different optimizers in convolutional neural networks for leukocyte classification. InterConf 33(155):431–448. https://doi.org/10.51582/interconf.19-20.05.2023.038

LeCun Y et al (1989) Backpropagation applied to handwritten zip code recognition. Neural Comput 1(4):541–551

Liang Y, Peng W, Zheng Z-J, Silvén O, Zhao G (2021) A hybrid quantum–classical neural network with deep residual learning. Neural Netw 143:133–147

Li J, et al., (2022) Next-vit: next generation vision transformer for efficient deployment in realistic industrial scenarios.*arXiv preprint* arXiv:2207.05501. Accessed 10 Jan 2025

Maurya R, A. K. Bais, T. Gopalakrishnan, M. K. Dutta, N. N. Pandey, and S. M. YV (2024) Skin lesion classification using deep feature fusion and selection using XGBoost classifier," in *2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, IEEE. 1–5.

Mehta S, M. Rastegari (2021) Mobilevit: light-weight, general-purpose, and mobile-friendly vision transformer," *arXiv preprint* arXiv.org/abs/2110.02178. Accessed 14 Jan 2025

Naeem A, Farooq MS, Khelifi A, Abid A (2020) Malignant melanoma classification using deep learning: datasets, performance measurements, challenges and opportunities. IEEE Access 8:110575–110597

Nivyashree T, Pramila PV (2023) Detection of malignant and benign skin lesions using the influence of activation function and accuracy analysis in densely connected convolutional network compared over convolutional neural network. Intelligent Computing and Control for Engineering and Business Systems (ICCEBS). IEEE 2023:1–6

Ozdemir B, Aslan E, Pacal I (2025) Attention enhanced InceptionNeXt-based hybrid deep learning model for lung cancer detection. IEEE Access 13:27050–27069. https://doi.org/10.1109/ACCESS.2025.3539122

Ozdemir B, Pacal I (2025a) A robust deep learning framework for multiclass skin cancer classification. Sci Rep 15(1):4938

Ozdemir B, Pacal I (2025b) An innovative deep learning framework for skin cancer detection employing ConvNeXtV2 and focal self-attention mechanisms. Results in Engineering 25:103692

Pacal I, Ozdemir B, Zeynalov J, Gasimov H, Pacal N (2025) A novel CNN-ViT-based deep learning model for early skin cancer diagnosis. Biomed Signal Process Control 104:107627

Pacal I, M. Alaftekin, and F. D. Zengul (2024) Enhancing skin cancer diagnosis using Swin transformer with hybrid shifted window-based multi-head self-attention and SwiGLU-based MLP. J Imaging Inform Med. 1–19.

Panelos J, Mass D (2009) Emerging role of Notch signaling in epidermal differentiation and skin cancer. Cancer Biol Ther 8(21):1986–1993. https://doi.org/10.4161/cbt.8.21.9921

Reka S.S, H. L. Karthikeyan, A. J. Shakil, P. Venugopal, and M. Muniraj (2024) Exploring quantum machine learning for enhanced skin lesion classification: a comparative study of implementation methods. IEEE Access 1-1

Saber A, Elbedwehy S, Awad WA, Hassan E (2025) An optimized ensemble model based on meta-heuristic algorithms for effective detection and classification of breast tumors. Neural Comput Appl 37(6):4881–4894

Salian S.R, S. D. Sawarkar (2022) Melanoma skin lesion classification using improved efficientnetb3. Jordanian J Comput Inf Technol 8:45-48

Scholz D, A. C. Erdur, J. A. Buchner, J. C. Peeken, D. Rueckert, and B. Wiestler (2024) Imbalance-aware loss functions improve medical image classification. Med Imaging Deep Learn. 1341–1356

Simoes MCF, Sousa JJS, Pais AACC (2015) Skin cancer and new treatment perspectives: a review. Cancer Lett 357(1):8–42

Tolstikhin IO et al (2021) Mlp-mixer: an all-mlp architecture for vision. Adv Neural Inf Process Syst 34:24261–24272

Tripathi S, Singh SK, Lee HK (2021) An end-to-end breast tumour classification model using context-based patch modelling – a BiLSTM approach for image classification. Comput Med Imaging Graph 87:101838. https://doi.org/10.1016/j.compmedimag.2020.101838

Trockman A, J. Z. Kolter (2022) Patches are all you need?,*arXiv preprint* arXiv:2201.09792. Accessed 22 Dec 2024

Tuncer T, P. D. Barua, I. Tuncer, S. Dogan, and U. R. Acharya (2024a) A lightweight deep convolutional neural network model for skin cancer image classification. Appl Soft Comput 162:111794 https://doi.org/10.1016/j.asoc.2024.111794

Tuncer T, Barua PD, Tuncer I, Dogan S, Acharya UR (2024b) A lightweight deep convolutional neural network model for skin cancer image classification. Appl Soft Comput 162:111794

Wang Y, Wang Y, Chen C, Jiang R, Huang W (2022) Development of variational quantum deep neural networks for image recognition. Neurocomputing 501:566–582>

Wang A, J. Hu, S. Zhang, and L. Li (2024a) Shallow hybrid quantum-classical convolutional neural network model for image classification. Quantum Inf Process 23:1 https://doi.org/10.1007/s11128-023-04217-5

Wang A, Hu J, Zhang S, Li L (2024b) Shallow hybrid quantum-classical convolutional neural network model for image classification. Quantum Inf Process 23(1):17

Xiaoyan L, Raga RC, Xuemei S (2022) GloVe-CNN-BiLSTM model for sentiment analysis on text reviews. J Sens 2022(1):7212366

Xu Z, et al., (2024) Parallel proportional fusion of spiking quantum neural network for optimizing image classification," *arXiv preprint* arXiv.org/abs/2404.01359. Accessed 9 Jan 2025

Yu W, et al., (2022) Metaformer is actually what you need for vision. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 10819–10829.

Yeung M, Sala E, Schönlieb C-B, Rundo L (2022) Unified focal loss: generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. Comput Med Imaging Graph 95:102026. https://doi.org/10.1016/j.compmedimag.2021.102026

Zaman K, T. Ahmed, M. A. Hanif, A. Marchisio, and M. Shafique (2024) A comparative analysis of hybrid-quantum classical neural networks," *arXiv preprint* arXiv.org/abs/2402.10540. Accessed 7 Jan 2025