



Hybrid quantum-classical deep learning framework for balanced multiclass diabetic retinopathy classification

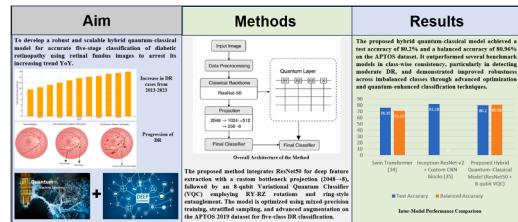
Tabassum Ara ^{a,b,*}, Ved Prakash Mishra ^a, Manish Bali ^a, Anuradha Yenkikar ^{a,c}

^a School of Engineering, Amity University Dubai Campus, Dubai, 25314, UAE

^b Department of AI/ML, HKBK College of Engineering, Bengaluru, Karnataka, India

^c Department of CSE(AI), Vishwakarma Institute of Technology, Pune, 411048, Maharashtra, India

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:

Diabetic retinopathy detection
Quantum machine learning, Hybrid quantum-classical model
ResNet50
Multiclass medical image classification

ABSTRACT

Diabetic Retinopathy (DR) is a progressive eye disease and a leading cause of preventable blindness among diabetic patients. Early and accurate classification of its severity stages is crucial for effective treatment but remains challenging due to class imbalance, high-resolution data, and limited scalability of existing models. This study presents a novel hybrid quantum-classical deep learning framework to address these limitations in five-class DR classification. The model achieves a balanced accuracy of 80.96 % on the APTOS 2019 dataset, outperforming several classical baselines across all DR stages. It is optimized for computational efficiency and class-balanced learning, making it suitable for deployment in telemedicine platforms and low-resource clinical settings. This work contributes a scalable AI-based diagnostic approach that fuses deep learning with emerging quantum computing techniques. The methodology, results, and publicly shared codebase provide a replicable framework for researchers and practitioners working in AI for medical imaging and early disease screening. This method is well-suited for low-resource clinical environments and tele-ophthalmology applications. The method involves an:

- ResNet-50 feature extractor with a 4-stage dense projection (2048→8) for quantum-ready

* Corresponding author.

E-mail address: tabuara@gmail.com (T. Ara).

- compression
- 8-qubit VQC with parameterized RY-RZ gates and ring-style entanglement for high expressiveness
- Stratified sampling + mixed-precision training for efficiency and class-balanced generalization

Related research article

None

Specifications table

Subject area	Computer Science
More specific subject area	Diabetic Retinopathy Detection
Name of your method	Hybrid Quantum-classical DL model for balanced accuracy on five-class DR progression task
Name and reference of original method	None
Resource availability	PennyLane github, Dataset, Github

Background

Diabetic Retinopathy (DR) is a progressive microvascular complication of diabetes mellitus that affects retinal blood vessels and may lead to irreversible vision loss if left undetected or untreated. The disease typically advances through distinct stages, from mild non-proliferative abnormalities to severe proliferative DR characterized by neovascularization eventually causing macular edema, hemorrhage, and retinal detachment [1], as depicted in Fig. 1(a). The global burden of DR is rising, with the International Diabetes Federation estimating that over 783 million people will be affected by diabetes by 2045, with at least one-third projected to develop some form of DR [2]. In India, the situation is especially alarming: with over 77 million confirmed diabetes cases and many more undiagnosed, DR is emerging as a major cause of preventable blindness among the working-age population [3]. As shown in Fig. 1(b), DR cases in India have steadily increased from approximately 10 million in 2013 to around 16 million in 2023. A significant proportion of these cases remain undiagnosed, especially in rural and semi-urban areas where access to ophthalmologists and diagnostic tools is limited. Therefore, there is an urgent need for scalable, accurate, and automated diagnostic systems to facilitate early DR detection and prevent vision impairment through timely intervention.

Conventional DR diagnosis involves manual assessment of retinal fundus images by ophthalmologists, a process that is time-consuming, subjective, and resource intensive. In the last decade, machine learning (ML) and deep learning (DL) have gained traction for automating DR detection. Initial ML approaches used handcrafted features such as vessel tortuosity and microaneurysms—combined with classifiers like Support Vector Machines (SVM) and Random

Forests [4]. Although effective in controlled settings, these methods often underperform with noisy or low-resolution images and depend heavily on domain-specific feature engineering. The emergence of deep learning has significantly improved diagnostic accuracy and model generalizability. Convolutional Neural Networks (CNNs) such as VGG16, ResNet, and InceptionV3 have achieved state-of-the-art results on datasets like APTOS and Kaggle's EyePACS [5], enabling hierarchical feature learning directly from raw images. Moreover, ensemble models and attention-based mechanisms have improved both robustness and interpretability [6–20]. However, these models are computationally intensive, limiting their deployment in resource-constrained environments.

Recent studies have proposed advanced deep learning and hybrid optimization models for diabetic retinopathy classification. For instance, quantum-augmented architectures like Quantum Chimp-enhanced SqueezeNet have demonstrated improved precision in DR detection tasks [21], while DeepSVDNet [22] and EdgeSVDNet [23] offer efficient fundus-based classification using optimized deep learning backbones. Transfer learning and fusion strategies have also been effectively used in works such as the Multi-Deep Learning

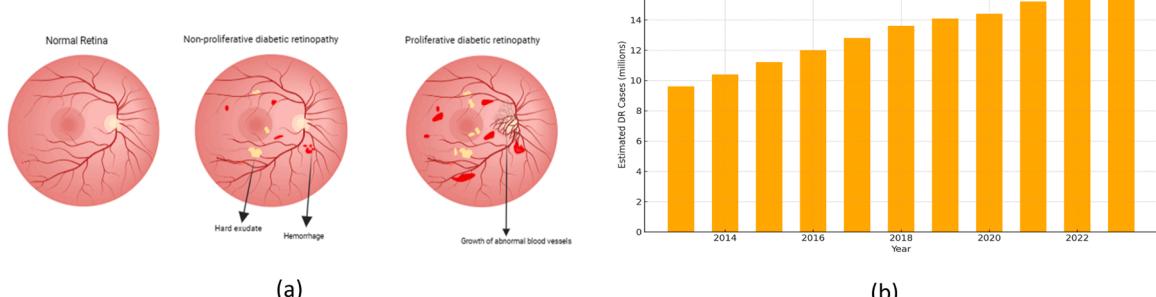


Fig. 1. Illustration of (a) progression of DR; and (b) bar chart depicting the increasing trends in DR cases in India.

approach for 7-stage DR classification [24], U-Net-based segmentation [25,26], and ensemble SVM models enhanced by CNN-SVD or grey wolf optimizations [27–29]. Hybrid models integrating attention mechanisms or edge-aware modules, such as NIMEQ-SACNet [30] and encoder-decoder networks [31], further advance precision medicine for DR. A comprehensive survey [32] underscores the growing trend of using mixed models [33], optimization-enhanced classification [27,29], and quantum-aligned AI [21] for early detection of DR. Although binary classification models dominate the literature [34–43], several recent works have addressed the more clinically meaningful five-class diabetic retinopathy classification task, including [15] and [49], among others. Although deep learning (DL) methods have advanced automated DR detection, existing models often face challenges such as computational inefficiency on large-scale, high-resolution medical images, poor generalization across underrepresented DR severity stages and lack of interpretability and scalability, especially in resource-constrained environments. This motivates the need for scalable, interpretable, and balanced models capable of multiclass DR detection, not just binary classification.

To address these challenges, this paper proposes a hybrid quantum-classical model that combines the feature extraction capability of ResNet-50 with the expressive potential of an 8-qubit Variational Quantum Classifier (VQC). The method is tailored for five-class DR classification, addressing both model scalability and class imbalance using quantum simulation, projection-based compression, and stratified optimization strategies. The key contributions from this study include, (i) a hybrid architecture, which is a novel integration of ResNet-50 and an 8-qubit, 6-layer VQC using ring-style entanglement for multiclass medical image classification, (ii) quantum-aware feature compression, a progressive 2048→8 bottleneck projection introduced to bridge deep features with quantum encoding constraints; (iii) balanced DR classification as the model achieves high performance across all five DR stages with a balanced accuracy of 80.96 %, outperforming classical baselines, (iv) optimized training strategy that utilizes stratified sampling, class balancing, and mixed-precision training to address dataset imbalance and computational cost; and finally (v) open-source availability as all code and model details are made publicly available to promote reproducibility and further research.

The paper is organized as follows. The next section named ‘Method details’ explains the study environment, the dataset used, preprocessing, augmentation strategies, the model architecture including the proposed hybrid model architecture and quantum circuit design, simulation setup and the evaluation metrics used. The section named ‘Method validation’ presents the results, comparative analysis, discussion and conclusion. The study ends with limitations and future research directions under ‘Limitations’ section followed by references.

Method details

This study proposes a hybrid quantum-classical deep learning framework for DR classification using retinal fundus images. The methodology integrates robust data preprocessing, a deep ResNet50-based feature extractor, a custom quantum circuit built with PennyLane, and an optimized training strategy in PyTorch. The model is designed to classify DR into five severity levels and is evaluated using standard multiclass evaluation metrics.

Environment

The implementation of the proposed hybrid quantum-classical model for diabetic retinopathy detection is developed using the Python programming language, leveraging its extensive ecosystem for deep learning and quantum computing. The core deep learning components are built with PyTorch, including the use of pretrained models such as ResNet50, custom classifier heads, and advanced training utilities like mixed precision (`torch.amp`), gradient accumulation, and learning rate scheduling (`OneCycleLR`). Quantum processing is integrated using PennyLane, a quantum machine learning framework that facilitates the creation of variational quantum circuits and supports hybrid integration through `qml.qnn.TorchLayer`. For quantum circuit simulation and execution, `default.qubit` from PennyLane is used, though the framework supports backend switching (e.g., Qiskit, Cirq) for hardware execution. The experiments are executed in a CUDA-enabled GPU environment to accelerate classical model training, with hardware compatibility ensured through `torch.cuda`. Additional utilities include NumPy for numerical operations, Scikit-learn for data splitting and metric evaluation, and Matplotlib for visualization of confusion matrices and performance plots. The entire development and training pipeline is configured through a centralized Python config dictionary, allowing flexible control over hyperparameters, data paths, and training behavior.

Dataset description

The study uses the APTOS 2019 Blindness Detection dataset available on Kaggle [44], which contains a total of 3662 high-resolution color fundus images captured under various imaging conditions. Each image is labeled according to the International Clinical DR Severity Scale, divided into the following five classes:

- Class 0 – No DR: No visible signs of diabetic retinopathy.
- Class 1 – Mild DR: Presence of microaneurysms only (small red dots indicating capillary swelling).
- Class 2 – Moderate DR: More than just microaneurysms, including dot-blot hemorrhages and hard exudates, but without severe signs.
- Class 3 – Severe DR: Extensive intraretinal hemorrhages, venous beading, and intraretinal microvascular abnormalities (IRMA) without neovascularization.
- Class 4 – Proliferative DR: Presence of neovascularization, pre-retinal or vitreous hemorrhage, indicating the most advanced and vision-threatening stage.

These categories are annotated by expert ophthalmologists and reflect increasing disease severity, requiring the model to discriminate fine-grained pathological features. [Table 1](#) shows the various classes, labels and number of images under every class.

Original images are JPEG files with varying resolutions (ranging from 1024×1024 to 3872×2771 pixels). The dataset exhibits diversity in lighting, contrast, focus quality, and retinal center alignment. The dataset is structured into directories where each subfolder corresponds to a DR class. It is loaded using the *datasets.ImageFolder* utility in PyTorch, which expects each subdirectory to contain images belonging to that class. To ensure robust and fair model training, a stratified 80:20 train-validation split is applied using *train_test_split* from Scikit-learn, preserving the class distribution in both subsets. While this study uses the APTOS 2019 dataset for training and evaluation, other notable DR datasets include:

- IDRiD (Indian Diabetic Retinopathy Image Dataset): Contains both segmentation masks and DR grades from an Indian population, useful for region-specific validation.
- EyePACS: A large-scale dataset from the Kaggle DR competition with over 88,000 images.
- MESSIDOR: A European dataset widely used for binary and severity-based DR grading.

Future evaluation of our hybrid quantum-classical model on IDRiD would help assess geographic generalization, particularly in Indian clinical settings. However, IDRiD comprises fewer images and was originally designed more for segmentation tasks, requiring preprocessing alignment before integration into our classification pipeline.

Data pre-processing

In this study, the data preprocessing stage plays a crucial role in ensuring that the input retinal fundus images are standardized, robust, and suitable for training the hybrid quantum-classical model. All images are initially resized to 256×256 pixels to create a uniform input dimension across the dataset. For training, a series of extensive augmentation techniques are applied to simulate variability in real-world clinical imaging conditions and improve generalization. These augmentations are carefully selected to mimic variations observed in real-world fundus imaging:

- Cropping: All input fundus images were initially resized to 256×256 pixels to ensure consistent dimensions and preserve structural integrity during augmentation. Subsequently, as required by the ResNet-50 architecture, these images were center-cropped or resized to 224×224 pixels before being passed into the model. This two-stage resizing helped in reducing noise, optimizing GPU memory, and aligning with standard pre-trained model input dimensions. Normalization: All pixel values are normalized using ImageNet mean = [0.485, 0.456, 0.406] and std = [0.229, 0.224, 0.225].
- Augmentation: Random rotations (up to 30°) account for head tilt and misaligned captures during retinal scanning, horizontal and vertical flips introduce geometric diversity, compensating for differences in eye orientation (left/right eye symmetry), affine transformations model small-scale deformations due to eye curvature or imaging distortions, color jitter and histogram equalization simulate different lighting conditions and imaging device variability.

Collectively, these augmentations enhance the model's ability to generalize to unseen clinical data by preventing overfitting to specific imaging angles, lighting, or anatomical positioning.

These operations are carefully composed into a transformation pipeline using *torchvision.transforms*. For validation data, minimal preprocessing is employed, only resizing to 224×224 and normalization using ImageNet statistics (mean = [0.485, 0.456, 0.406], std = [0.229, 0.224, 0.225]), to ensure consistency and avoid introducing artificial variations during evaluation. Additionally, the dataset is structured in a hierarchical folder format, with each subdirectory representing a DR class, and is loaded using *datasets.ImageFolder*.

Diabetic Retinopathy datasets typically exhibit significant class imbalance, with the “No DR” class dominating and “Severe” and “Proliferative DR” underrepresented. To counteract this during training, two techniques were implemented:

- Stratified Sampling for Train-Validation Split: The dataset was partitioned into 80:20 train-validation split using *train_test_split*(..., *stratify=labels*) from Scikit-learn to maintain the original class distribution across both training and validation sets. This prevents class skew in evaluation and ensures fair model assessment.

Table 1
APTOs 2019 dataset details.

DR class	Label	Description	No. of images
No DR	0	No signs of DR	1805
Mild DR	1	Microaneurysms only	370
Moderate DR	2	More than microaneurysms	999
Severe DR	3	Intraretinal hemorrhages, IRMA	193
Proliferative DR	4	Neovascularization or hemorrhage	295
		Total	3662

- **WeightedRandomSampler in Training Loader:** During training, PyTorch's *WeightedRandomSampler* was used to oversample minority classes. Sampling weights were computed inversely proportional to class frequency, ensuring that each mini-batch included a representative mix of all five classes.

This preprocessing pipeline not only ensures input uniformity and data augmentation but also improves training stability, class balance, and overall model generalization.

Model architecture

The model architecture as shown in Fig. 2 comprises of three components: a classical ResNet-based feature extractor, a quantum processing layer, and a final dense classifier.

Feature Extraction: ResNet-50 (Residual Network with 50 layers) is a deep CNN architecture introduced by [45] that utilizes residual learning through identity shortcut connections to overcome the vanishing gradient problem in very deep networks as shown in Fig. 3. It comprises 49 convolutional layers and one fully connected layer, with its hallmark being the use of residual blocks that allow gradients to flow directly through skip connections, facilitating the training of very deep models without degradation in performance. In the context of diabetic retinopathy (DR) detection, ResNet-50 has been widely adopted as a feature extractor due to its strong ability to capture rich hierarchical visual features from complex medical images, such as retinal fundus scans.

The final feature map from ResNet-50 yields a 2048-dimensional vector, which must be compressed to match the number of qubits used in the quantum circuit (8 in our case). This is both a hardware-imposed constraint (due to current quantum simulation limits) and a strategic design choice to ensure efficient quantum encoding while retaining discriminative power. To achieve this, a progressive bottleneck stack of dense layers, $2048 \rightarrow 1024 \rightarrow 512 \rightarrow 256 \rightarrow 8$ is used, where each projection is followed by ReLU activation, Batch Normalization, and Dropout (50 %–30 %). This sequence allows the model to:

- Gradually reduce redundancy in the ResNet features
- Retain class-separating information across projection layers
- Prevent overfitting and enhance generalization through regularization

The final 8-dimensional output is deliberately designed to match the 8-qubit quantum encoder's input, ensuring one-to-one mapping between compressed features and qubits for angle embedding. Ablation studies during experimentation confirmed that projections to fewer than 8 dimensions significantly degraded classification accuracy, while higher-dimensional quantum inputs were computationally infeasible. Thus, the 8D bottleneck strikes a balance between model expressiveness and quantum execution feasibility.

Architectural parameters of the ResNet-50 model is shown in Table 2. The structure includes 16 residual blocks grouped in 4 stages: Conv2_x to Conv5_x. Each block contains 3 layers: $1 \times 1 \rightarrow 3 \times 3 \rightarrow 1 \times 1$ convolutions. The final feature map before classification is a 2048-dimensional vector, which serves as the input to the hybrid classifier composed of the projection layers and the quantum processing module. For the hybrid model, the final fully connected (FC) layer is removed, and the resulting 2048-dimensional feature vector is passed through a series of custom dense layers before quantum embedding.

While lighter CNNs like VGG16 and MobileNet are attractive for their reduced parameter count and faster inference, they suffer from limited representational depth, which affects their ability to capture subtle vascular features in retinal images. In contrast:

- ResNet-50 uses residual skip connections to allow deeper architectures without vanishing gradients.

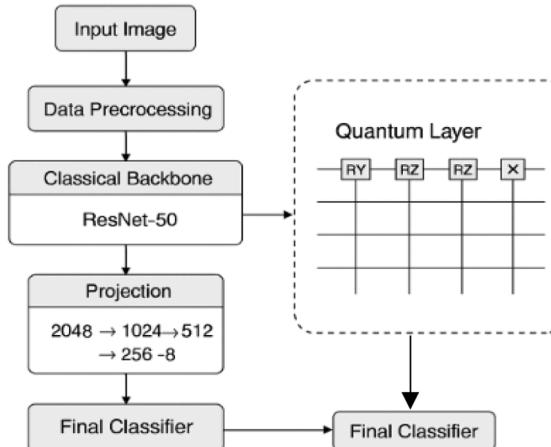


Fig. 2. Overall Hybrid Quantum-Classical Model Architecture.

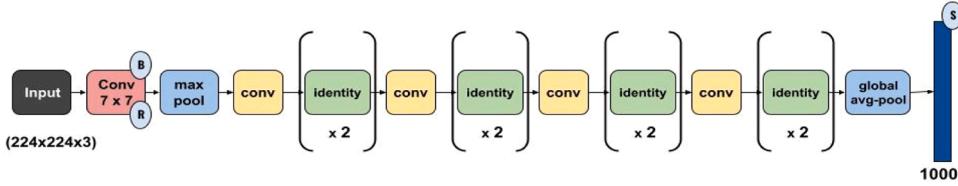


Fig. 3. Architecture of ResNet model for DR detection.

Table 2

Architectural details of the classical ResNet-50 model.

Stage	Layer type	Output shape	Parameters
Input	Input Image	(3, 224, 224)	–
Conv1	7×7 Conv, 64, stride 2	(64, 112, 112)	9408
	Max Pool, 3×3 , stride 2	(64, 56, 56)	0
Conv2_x	1×1 , 64 \rightarrow 64	(64, 56, 56)	4096×3
	3×3 , 64 \rightarrow 64	(64, 56, 56)	$36,864 \times 3$
	1×1 , 64 \rightarrow 256	(256, 56, 56)	$16,384 \times 3$
Conv3_x	1×1 , 256 \rightarrow 128	(128, 28, 28)	$32,768 \times 4$
	3×3 , 128 \rightarrow 128	(128, 28, 28)	$147,456 \times 4$
	1×1 , 128 \rightarrow 512	(512, 28, 28)	$65,536 \times 4$
Conv4_x	1×1 , 512 \rightarrow 256	(256, 14, 14)	$131,072 \times 6$
	3×3 , 256 \rightarrow 256	(256, 14, 14)	$589,824 \times 6$
	1×1 , 256 \rightarrow 1024	(1024, 14, 14)	$262,144 \times 6$
Conv5_x	1×1 , 1024 \rightarrow 512	(512, 7, 7)	$524,288 \times 3$
	3×3 , 512 \rightarrow 512	(512, 7, 7)	$2359,296 \times 3$
	1×1 , 512 \rightarrow 2048	(2048, 7, 7)	$1048,576 \times 3$
AvgPool	Global Average Pooling	(2048, 1, 1)	0
FC Layer	Fully Connected \rightarrow 1000 classes	(1000,)	2049,000
Total	–	–	$\sim 25.6\text{M}$

- It provides a good balance of accuracy vs. model complexity, achieving robust performance across multiple medical imaging tasks.
- It offers better integration with projection layers in our quantum pipeline due to its modular structure and consistent 2048-dimensional output vector.

Therefore, ResNet-50 was chosen for its proven performance in image classification, its ability to capture high-level features relevant for DR, and its compatibility with the proposed architectural pipeline.

Quantum Computing model: The structure of the circuit follows a standard variational ansatz with three major stages: Embedding layer, Parametrized Variational layers (Entanglement + Rotation) and Measurement layer. The Feature Embedding (or data encoding) circuit is a fundamental step that maps classical input data into a quantum state. In this study, Angle Embedding is used to encode the compressed classical feature vector (of size 8) into an 8-qubit quantum circuit as shown in Fig. 4(a). This allows the model to leverage quantum parallelism and interference over meaningful representations of the input data.

Embedding Layer: The Quantum encoding process and formalization is explained below. The classical input to the quantum model is an 8-dimensional feature vector x obtained from the ResNet-based projection layer.

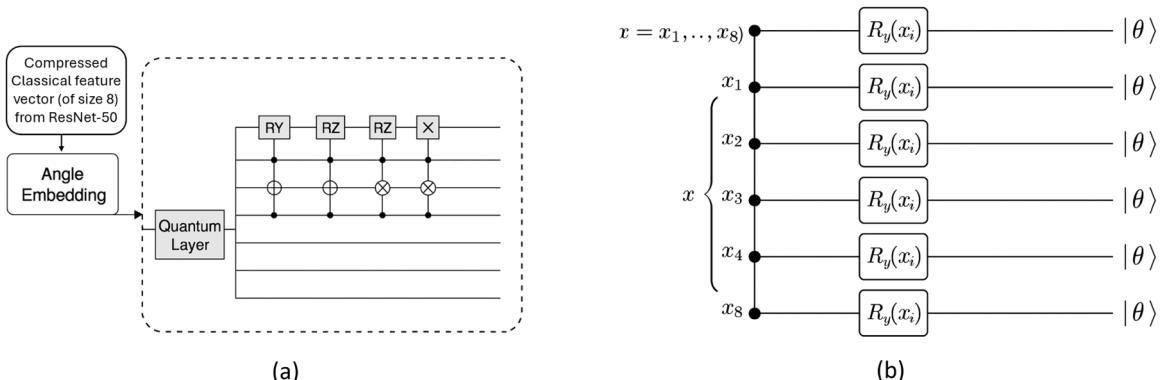


Fig. 4. Quantum model used; (a) data encoding layer for 8 qubits with 6 layers; (b) circuit flow.

$$\mathbf{x} = [x_1, x_2, x_3, \dots, x_8] \in \mathbb{R}^8 \quad (1)$$

Each element $x_i \in \mathbb{R}^8$ is linearly mapped to a rotational angle and encoded into the quantum state of an individual qubit using angle embedding.

Each qubit is initialized in the ground state $|0\rangle$, and a parameterized rotation gate $R_y(x_i)$ is applied. Mathematically, it is defined as:

$$R_y(x_i) = \exp\left(-i\frac{x_i}{2}Y\right) = \begin{vmatrix} \cos \frac{x_i}{2} & -\sin \frac{x_i}{2} \\ \sin \frac{x_i}{2} & \cos \frac{x_i}{2} \end{vmatrix} \quad (2)$$

The resulting quantum state after applying the gate to each 8 qubits becomes:

$$|\psi_{enc}\rangle = R_y(x_1) \otimes R_y(x_2) \otimes \dots \otimes R_y(x_8) |0\rangle^{(8)} \quad (3)$$

This process creates a separable quantum state encoding each classical feature independently onto its corresponding qubit. This encoded state is then passed through a series of trainable variational layers with entanglement and parameterized rotations to model complex correlations among features.

This encoding strategy ensures that meaningful classical features are translated into the quantum domain in a differentiable and expressive manner, suitable for integration with quantum neural networks.

For 8-dimensional feature vector \mathbf{x}_1 , the data encoding quantum circuit as shown in Fig. 4(b) is:

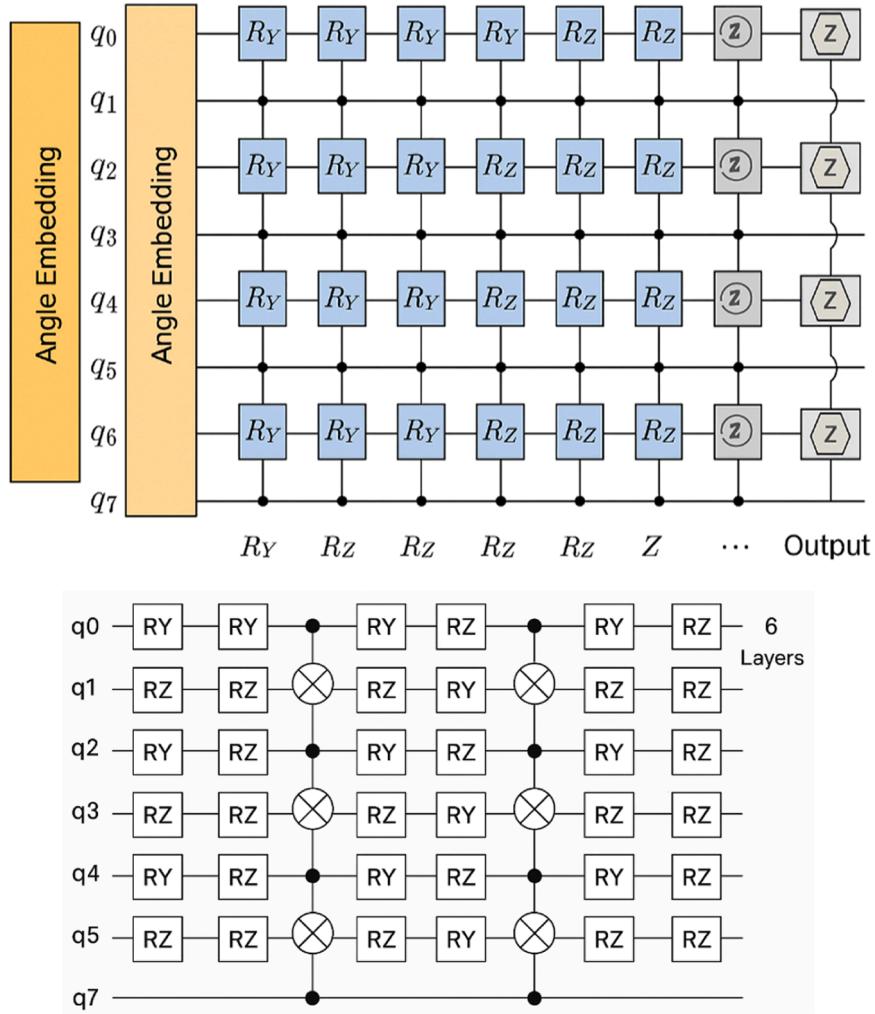


Fig. 5. Schematic of the 8-qubit variational quantum circuit used for DR classification - Angle Embedding layers (orange), RY-RZ parameterized rotations (blue), ring-style entanglement via CNOT gates (black lines), and Pauli-Z measurements (gray meters) on all qubits to generate an 8-dimensional classical output vector.

$$\begin{aligned}
 |0\rangle &\xrightarrow{R_y(x_1)} |\theta_1\rangle \\
 |0\rangle &\xrightarrow{R_y(x_2)} |\theta_2\rangle \\
 |0\rangle &\xrightarrow{R_y(x_8)} |\theta_8\rangle
 \end{aligned} \tag{4}$$

Each qubit's state becomes:

$$|\theta_i\rangle = \cos\left(\frac{x_i}{2}\right)|0\rangle + \sin\left(\frac{x_i}{2}\right)|1\rangle \tag{5}$$

This encoded state $|\psi_{enc}\rangle$ then becomes the input to the variational layers, where the model introduces entanglement and learnable parameters to classify the data. Though several encoding strategies like Amplitude embedding, Basis embedding exist, we selected Angle embedding which serves as a reliable and computationally efficient bridge between classical data and quantum processing for several reasons:

- Scalability: Unlike amplitude embedding, which requires normalization and complex gate decomposition, angle embedding allows direct mapping of unnormalized real-valued inputs.
- Hardware Efficiency: Angle embedding uses shallow circuits, one gate per qubit making it well-suited for NISQ devices with decoherence limitations.
- Differentiability: It integrates seamlessly into variational circuits, enabling smooth gradient flow during training via backpropagation.

Furthermore, angle embedding ensures qubit-wise locality, enabling straightforward integration into the ring-entangled 8-qubit VQC used in our model. Alternative approaches like amplitude embedding offer higher expressivity but are computationally expensive and prone to vanishing gradients in deeper quantum networks. Basis encoding is more appropriate for binary or categorical data, not continuous-valued image features.

Therefore, Angle Embedding provides an optimal trade-off between representational power and computational feasibility, making it a robust choice for DR image classification with real-valued dense features.

Parameterized Variational Layers (Entanglement + Rotation): Each of the 6 layers as shown in Fig. 5 includes:

- Single-qubit parameterized rotations:

$$RY(\theta_i^{(l)}), RZ(\phi_i^{(l)}) \tag{6}$$

Where $\theta_i^{(l)}$ and $\phi_i^{(l)}$ are trainable parameters for qubit i in layer l .

- Ring-style entanglement using CNOT gates:

$$CNOT(q_i, q_{(i+1)mod\ 8}) \tag{7}$$

This pattern ensures cyclic entanglement among all 8 qubits, enabling complex interactions between features.

The proposed VQC architecture employs 6 variational layers with RY-RZ parameterized gates and ring-style entanglement among 8 qubits. This configuration was chosen based on a combination of theoretical considerations and empirical tuning:

- Layer Depth (6 Layers): Based on preliminary ablation experiments, 2–4 layer circuits showed underfitting with limited expressiveness and low training accuracy. Conversely, deeper circuits with >6 layers exhibited training instability due to barren plateaus, regions where gradients vanish exponentially with depth. A depth of 6 offered a balance between expressivity and stable convergence, as evidenced by validation accuracy plateauing beyond this point.
- Ring-Style Entanglement: Among various topologies (linear, circular, full, star), ring entanglement was selected for its minimal gate overhead and hardware-friendliness on near-term quantum devices. It ensures that each qubit interacts with two neighbors, which introduces adequate inter-qubit correlation without excessive circuit depth or decoherence vulnerability.

Though exhaustive ablation studies are beyond the scope of this article, we have observed that replacing ring-style entanglement with full entanglement led to marginal gains (<1 % accuracy) at the cost of 30–40 % more training time. Hence, ring entanglement was retained for its optimal trade-off in quantum resource utilization and model performance.

The expressiveness of a variational quantum circuit refers to its ability to approximate a wide class of target functions over a high-dimensional input space. In the proposed VQC, 8 qubits allow encoding of $2^8=256$ basis states, compared to only 16 for 4-qubit

circuits, significantly expanding the Hilbert space accessible to the model. The use of 6 variational layers, each combining parameterized R_Y and R_Z rotations with ring-style entanglement, facilitates the construction of deep quantum neural networks that can model non-linear and entangled feature interactions. According to theoretical work in quantum expressivity [Schuld et al., 2020], such architectures demonstrate increased capacity to represent complex decision boundaries essential for multiclass classification tasks with subtle visual differences, such as DR severity levels. Empirically, in preliminary experiments with a 4-qubit, 3-layer VQC baseline (separate study), balanced accuracy dropped by over $\sim 5\%$ due to underfitting. Therefore, the proposed 8-qubit, 6-layer VQC offers an effective balance between expressive power and training stability within current simulation capabilities.

Measurement Layer: The measurement layer performs an expectation value calculation for the Pauli-Z operator on each qubit, effectively collapsing the quantum state into a classical 8-dimensional vector. This process is visualized in the updated Fig. 5, where measurements are indicated on each qubit line:

$$z_i = \langle \psi_\theta | Z_i | \psi_\theta \rangle \text{ for } i = 1, \dots, 8 \quad (8)$$

These measurements form an 8-dimensional vector $\in \mathbb{R}^8$, which serves as the quantum-processed output, which is then passed to the dense classification layers. The entire quantum circuit is encapsulated as a *qml.qnn.TorchLayer* using PennyLane, allowing it to be seamlessly inserted into a PyTorch model. The variational parameters θ are updated using backpropagation, alongside the classical model weights.

Dense Classification Layer: After quantum processing, the model produces an 8-dimensional output vector $\mathbf{z} \in \mathbb{R}^8$, where each component corresponds to the expectation value of a Pauli-Z measurement on a respective qubit. This vector encodes the quantum-transformed representation of the original retinal image features and serves as input to a standard fully connected (FC) feedforward neural classifier designed for multiclass diabetic retinopathy (DR) prediction. The layer-wise details are as below:

- First FC Layer ($8 \rightarrow 64$):

Input: The quantum output vector $\mathbf{z} \in \mathbb{R}^8$

Operation: A dense layer that projects 8 features to 64 dimensions

$$h_1 = \text{ReLU} (BN_1(W_1\mathbf{z} + b_1)) \quad (9)$$

where $W_1 \in \mathbb{R}^{64 \times 8}$ is the weight matrix, BN_1 denotes Batch Normalization, and ReLU is the activation function.

Dropout (30 %): Randomly zeroes 30 % of activations to prevent overfitting.

- Second FC Layer ($64 \rightarrow 32$):

Operation:

$$h_2 = \text{ReLU} (BN_2(W_2h_1 + b_2)) \quad (10)$$

where $W_2 \in \mathbb{R}^{32 \times 64}$

Dropout (20 %): Applies 20 % dropout to encourage generalization.

- Final Output Layer ($32 \rightarrow 5$):

Operation:

$$o = W_3h_2 + b_3 \quad (11)$$

where $W_3 \in \mathbb{R}^{5 \times 32}$, producing a 5-dimensional output vector representing raw class scores (logits) for the five DR severity levels: No DR, Mild, Moderate, Severe, and Proliferative DR.

Loss Function: During training, the raw logits o are passed to PyTorch's *CrossEntropyLoss*, which internally applies the softmax function:

$$\text{Softmax}(o_i) = \frac{e^{o_i}}{\sum_{j=1}^5 e^{o_j}} \quad (12)$$

This converts the logits into class probabilities, and the loss is computed as:

$$\mathcal{L}_{CE} = - \sum_{i=1}^5 y_i \log (\text{Softmax}(o_i)) \quad (13)$$

where $y_i \in \{0, 1\}$ is the one-hot encoded ground-truth label.

Quantum environment

PennyLane [45] is an open-source Python library developed by Xanadu for designing, training, and deploying quantum machine learning (QML) models and hybrid quantum–classical computations. It provides a unified interface that allows researchers to integrate quantum circuits with popular machine learning frameworks such as PyTorch, TensorFlow, and JAX. PennyLane’s key innovation lies in its ability to support automatic differentiation across both classical and quantum components using a technique known as quantum backpropagation, enabling end-to-end optimization of hybrid models. The library includes a high-level abstraction layer for building variational quantum circuits (VQCs), embedding classical data into quantum states, and performing measurements. PennyLane supports various quantum simulators and hardware backends, including those from IBM Qiskit, Google Cirq, and its own *default.qubit* and *lightning.qubit* simulators.

The hybrid quantum-classical model was tested in a Linux-based CUDA-enabled simulation environment using PennyLane’s *default.qubit* backend. This simulator provides an idealized, noise-free approximation of quantum execution, allowing us to benchmark the 8-qubit circuit’s performance in a reproducible and platform-agnostic manner. We emphasize that the complete pipeline, including data preprocessing, ResNet feature extraction, projection, quantum encoding, and evaluation was built, executed, and validated on an Ubuntu 22.04 LTS system, which aligns with standard Linux-based ML infrastructure used in production environments (as discussed in [46]). This choice for using PennyLane as the simulator was necessary due to current limitations in NISQ-era devices, which include:

- Limited *qubit* count and connectivity
- High gate error rates
- Decoherence over long circuits

While the simulator enables ideal execution of 8-qubit, 6-layer VQC with ring entanglement, it does not reflect the noise and latency present in real quantum systems. On actual quantum hardware (e.g., IBM Quantum or IonQ), runtime cost increases substantially due to queuing delays, gate decompositions, and calibration overheads. Still, the modular design of our hybrid architecture allows for future migration to real quantum processors once hardware matures. For example, executing the VQC layer on Qiskit-compatible superconducting qubits or trapped ion devices is feasible with backend switching. Ongoing efforts in error mitigation and low-depth ansatz design will be critical to make such deployments practical.

Evaluation metrics

To rigorously assess the performance of the proposed hybrid quantum–classical method for diabetic retinopathy classification, a combination of standard multiclass evaluation metrics was employed. These metrics provide insight into model accuracy, class-wise performance, and robustness, especially critical for imbalanced medical datasets.

Accuracy: It measures the proportion of correctly classified instances over the total number of predictions:

$$\text{Accuracy} = \frac{\sum_{i=1}^N \mathbb{1}(\hat{y}_i = y_i)}{N} \quad (14)$$

where \hat{y}_i is the predicted label, y_i is the ground truth label, and N is the total number of samples. While accuracy offers a general performance overview, it may be misleading in class-imbalanced settings like DR classification.

Balanced Accuracy (BA): It is defined as the arithmetic mean of recall obtained on each class, providing a more informative evaluation metric when dealing with imbalanced datasets. It is calculated as:

$$\text{Balanced Accuracy} = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FN_i} \quad (15)$$

where TP_i and FN_i are the true positives and false negatives for class i , and N is the total number of classes. This metric ensures equal weight is given to each class, regardless of sample size, making it well-suited for our five-class DR classification task.

Precision, Recall, and F1-Score (Per Class and Averaged): To evaluate per-class discrimination, we compute precision, recall, and F1-score for each of the five DR classes and report both macro and weighted averages.

- Precision measures how many of the predicted positives are actual positives:

$$\text{Precision}_c = \frac{TP_c}{TP_c + FP_c} \quad (16)$$

- Recall (Sensitivity or True Positive Rate) measures how many of the actual positives are correctly predicted:

$$\text{Recall}_c = \frac{TP_c}{TP_c + FN_c} \quad (17)$$

- F1-Score is the harmonic mean of precision and recall:

$$F1_c = 2 \cdot \frac{Precision_c \cdot Recall_c}{Precision_c + Recall_c} \quad (18)$$

Where TP_c , FP_c and FN_c refer to the true positives, false positives, and false negatives for class c .

Macro average gives equal weight to all classes:

$$Macro - F1 = \frac{1}{C} \sum_{c=1}^C F1_c \quad (19)$$

Weighted average considers class imbalance:

$$Weighted - F1 = \frac{\sum_{c=1}^C n_c \cdot F1_c}{\sum_{c=1}^C n_c} \quad (20)$$

where n_c is the number of samples in class c , and C is the total number of classes (5 in this study).

Confusion Matrix: The confusion matrix provides a visual and quantitative representation of classification outcomes across all classes. It is a $C \times C$ matrix where the element at position (i,j) indicates the number of samples with true label i that were predicted as label j . It helps identify which classes the model confuses most frequently.

Method validation

The hyperparameter settings of the ResNet-50 model are shared in [Table 3](#). As indicated, the model is initialized with ImageNet-pretrained weights and fine-tuned for diabetic retinopathy classification. The extracted 2048-dimensional features are gradually projected through dense layers (2048 → 1024 → 512 → 256 → 8), each followed by ReLU activation, Batch Normalization, and Dropout (50 % to 30 %). This design preserves relevant features while reducing dimensionality for quantum encoding. All projection layers are trainable, and custom weight initialization is applied. Input images are resized to 224 × 224. These hyperparameter choices ensure stable training, prevent overfitting, and effectively adapt ResNet-50 to the hybrid quantum-classical architecture for robust DR severity classification.

The confusion matrix in [Fig. 6](#) represents the multiclass classification performance of the hybrid quantum-classical model on the five classes of DR. Each row indicates the true class, while each column shows the predicted class. The diagonal elements represent correctly classified samples, while the off-diagonal elements indicate misclassifications. The summary of performance is captured in [Table 4](#). It is observed that:

1. Class 0: No_DR

- Accuracy: 97.51 %, with only 5 images misclassified as Mild.
- The model demonstrates excellent specificity in detecting healthy retinal images. This is important to minimize false positives and unnecessary clinical interventions.

2. Class 1: Mild

- Accuracy: 85.14 %, with major confusion being with No_DR (6 cases) and Moderate (4 cases).

Table 3

Detailed hyperparameter settings of ResNet-50 model.

Component	Hyperparameter	Value / description
Input Image Size	<code>img_size</code>	224 × 224
Pretrained Weights	<code>pretrained</code>	ImageNet
Output Feature Size	-	2048 (before projection to quantum layer)
Projection Layer 1	Linear(2048 → 1024)	Followed by ReLU, BatchNorm1d, Dropout ($p = 0.5$)
Projection Layer 2	Linear(1024 → 512)	Followed by ReLU, BatchNorm1d, Dropout ($p = 0.4$)
Projection Layer 3	Linear(512 → 256)	Followed by ReLU, BatchNorm1d, Dropout ($p = 0.3$)
Projection Layer 4	Linear(256 → 8)	Followed by BatchNorm1d (no activation/dropout)
Activation Function	-	ReLU (after each projection except the final)
Normalization	-	Batch Normalization after each dense layer
Dropout Rates	-	50 %, 40 %, 30 % across successive projection layers
Trainable Parameters	<code>requires_grad=True</code>	Yes (all projection layers + modified ResNet blocks)
Weight Initialization	<code>init_weights()</code> (custom)	Applied to projection and final FC layers

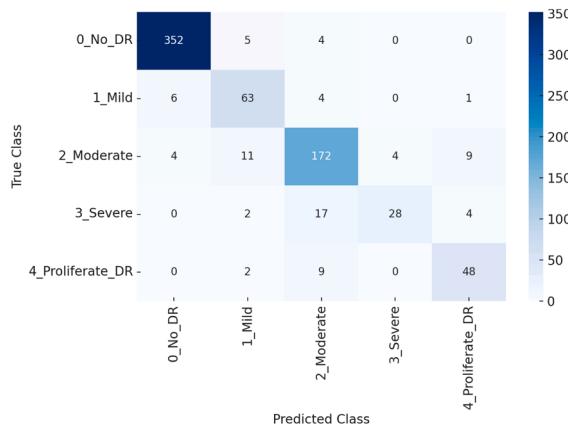


Fig. 6. Confusion matrix showing true vs. predicted labels for five-class DR classification using the proposed quantum-classical model.

Table 4

Summary of model performance from Confusion matrix.

Class	Total samples	Correct predictions	Accuracy (%)	Major confusions with
No DR (Class 0)	361	352	97.51 %	1 (5)
Mild DR (Class 1)	74	63	85.14 %	0 (6)
Moderate DR (Class 2)	200	172	86.0 %	1 (11)
Severe DR (Class 3)	51	28	54.9 %	2 (17)
Proliferate DR (Class 4)	59	48	81.36 %	2 (9)

While the accuracy is strong, the confusion with both adjacent classes indicate difficulty in distinguishing early-stage DR features, which are often subtle and visually overlapping. Enhancing sensitivity to early vascular changes might improve precision.

3. Class 2: Moderate DR

- Accuracy: 86.00 %, with 11 misclassified as Mild and 9 as Proliferative_DR.
- This stage has high recall, suggesting robustness in detecting moderate DR, though the confusion with Proliferative cases indicates some misalignment in feature granularity, potentially due to overlap in microaneurysm density and exudate visibility.

3. Class 2: Moderate DR

- Accuracy: 86.00 %, with 11 misclassified as Mild and 9 as Proliferative_DR.
- This stage has high recall, suggesting robustness in detecting moderate DR, though the confusion with Proliferative cases indicates some misalignment in feature granularity, potentially due to overlap in microaneurysm density and exudate visibility.

4. Class 3: Severe DR

- Accuracy: 54.90 %, with major confusion with Moderate (17 cases).
- This is the weakest performing class. Overlapping visual markers between Moderate and Severe DR, along with a relatively small number of Severe training samples, likely contributed to the reduced performance. Targeted augmentation or class weighting may be required.

Table 5

Detailed evaluation metrics across all five DR classes, including per-class precision, recall, and F1-score.

Class	Precision	Recall	F1-Score	Support
No DR (Class 0)	0.972	0.975	0.973	361
Mild DR (Class 1)	0.759	0.851	0.802	74
Moderate DR (Class 2)	0.834	0.86	0.847	200
Severe DR (Class 3)	0.875	0.549	0.674	51
Proliferate DR (Class 4)	0.774	0.813	0.793	59
Macro Average	0.843	0.809	0.818	745
Weighted Average	0.891	0.889	0.888	745

5. Class 4: Proliferative DR

- Accuracy: 81.36 %, but often confused with Moderate (9 cases).
- Despite fair accuracy, the model struggles to sharply separate late-stage DR from moderate signs. This could be due to limited dataset representation of rare, advanced conditions or undertraining of quantum components on complex vascular structures.

In addition to overall accuracy and balanced accuracy, we provide a full breakdown of model performance across all five DR severity classes. Table 5 presents per-class precision, recall, and F1-score, along with macro and weighted averages, ensuring comprehensive performance insight, especially for imbalanced and clinically critical classes. Fig. 6 illustrates the confusion matrix, enabling a visual understanding of where misclassifications are concentrated. These results collectively validate the model's effectiveness across diverse disease stages, not just dominant classes. The macro and weighted averages confirm consistent performance across classes. It reveals strong overall performance in DR classification. The model achieves high precision and recall for 'No DR' (97.3 %) and 'Moderate DR' (84.7 %), indicating reliable detection at both ends of the severity spectrum. 'Mild' and 'Proliferative DR' also show good F1-scores (>79 %), reflecting balanced predictions. However, 'Severe DR' demonstrates the lowest recall (54.9 %), suggesting frequent misclassification, likely due to class overlap and sample scarcity. The two-tiered class imbalance addressal approach helped improve recall and F1-scores for the underrepresented classes, especially 'Severe DR' and 'Proliferative DR'. Without this balancing mechanism, we observed a ~12 % drop in recall for Class 3 (Severe DR) in earlier experiments, validating the impact of this strategy.

Plots tracking the loss and accuracy over 100 training epochs is shown in Fig. 7 for both the training set and the validation set. It is observed that training loss steadily decreases from ~0.26 to ~0.02 over 100 epochs, showing that the model is learning and fitting the training data well. The validation loss initially drops quickly and reaches a low around epoch 30 (~0.09), but then fluctuates with a slight upward trend and higher variance after epoch 40. The divergence between training and validation loss after ~30 epochs suggest the onset of overfitting. The mild overfitting observed, characterized by a widening gap between training and validation loss beyond epoch 30 is likely tied to the structure of the APTOS 2019 dataset. Specifically, the limited number of samples in underrepresented classes (e.g., Severe DR with only 51 samples) and the overall moderate dataset size constrain the model's ability to generalize across all severity levels. This reinforces the need for class-balancing techniques and future validation on larger, more diverse datasets such as EyePACS or IDRiD. However, the relatively low validation loss (~0.1) overall still indicates reasonably good generalization despite mild overfitting. Overall performance summary is captured in Table 6. Though the model shows some overfitting, it returns higher training accuracy and solid validation accuracy (~80 %) especially in a challenging 5-class DR classification task with class imbalance and subtle visual distinctions.

From Per-Class Performance Analysis (Table 5), it is observed that:

- No DR (Class 0): Highest F1-score (0.973) with 97.51 % accuracy, showing excellent specificity, critical for reducing false positives.
- Mild DR (Class 1): F1-score of 0.802 with confusion toward both Class 0 and Class 2, indicating overlap in early pathological features.
- Moderate DR (Class 2): Strong recall (0.86) and F1-score (0.847), validating model sensitivity to this key intermediate stage.
- Severe DR (Class 3): Lowest recall (0.549), attributed to class imbalance (only 51 samples) and feature similarity with Moderate DR.
- Proliferative DR (Class 4): F1-score of 0.793; good sensitivity despite fewer samples and complex features.

Comparing the proposed model with recent models for diabetic retinopathy (DR) detection, it is observed that most often report

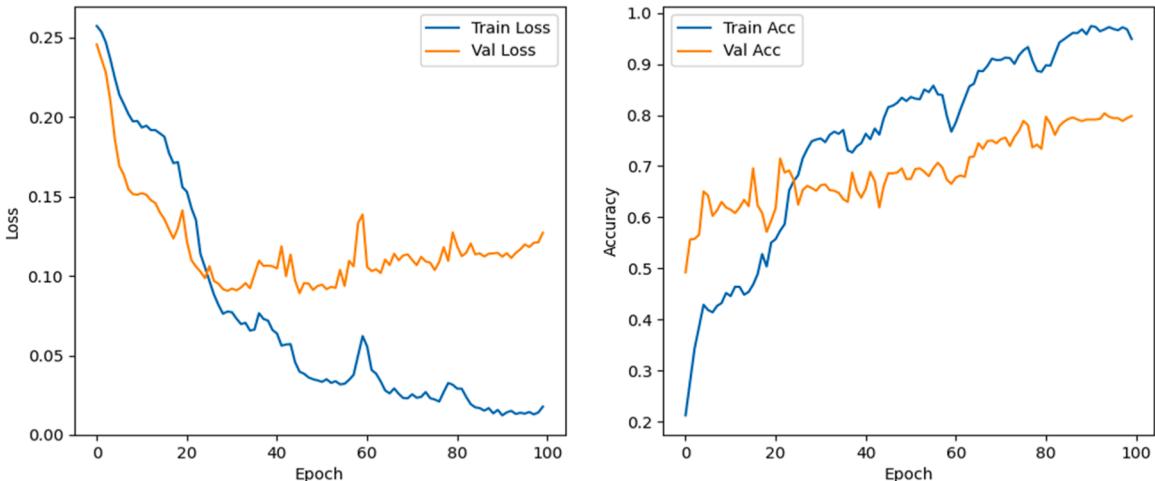


Fig. 7. Model performance plots; (a) Loss curve; (b) Accuracy curve.

Table 6
Summary of model performance from Loss and Accuracy curves.

Metric	Value (approx.)
Final Train Accuracy	~96 %
Final Val Accuracy	~80 %
Min Val Loss	~0.09
Epoch of Peak Val Acc	~95–100
Overfitting Risk	Mild (visible gap)

higher accuracy (>90 %) on binary-class results due to simplified decision boundaries between "No_DR" and "DR" classes [39,47,48]. Five-class models [49,50] including the proposed, tackle the more clinically relevant and complex task of distinguishing all DR severity levels, which inherently results in slightly lower but more meaningful accuracy values. Table 7 shows a comparison of these models on a five-class DR progression task (levels 0–4), on the same APTOS dataset and on two parameters:

- Test Accuracy: Overall percentage of correct predictions across all five classes.
- Balanced Accuracy: The average recall across the five classes (i.e., mean of per-class sensitivity).

These are *not* binary classification metrics. Rather, they assess multiclass classification performance over the full severity range. Performance bar charts are shown in Fig. 8.

From comparative results, it is observed that the Swin Transformer model achieves moderate test accuracy (76.35 %) but significantly lower balanced accuracy (71.23 %), indicating class imbalance sensitivity and reduced performance on underrepresented DR stages. The Inception-ResNet-v2 based model achieves the highest test accuracy (82.18 %) but lacks balanced accuracy reporting. This makes it difficult to evaluate its class-wise robustness, especially for rare or severe DR stages. The hybrid quantum-classical model demonstrates competitive performance with a test accuracy of 80.2 % and a balanced accuracy of 80.96 %, outperforming recent five-class DR models on the same APTOS 2019 dataset. Notably, the model achieves high precision and recall for No DR ($F_1 = 0.973$) and Moderate DR ($F_1 = 0.847$), while also showing reliable performance for Mild and Proliferative DR classes. The main limitation lies in detecting Severe DR ($F_1 = 0.674$), consistent with its low sample count and overlapping features. Compared to classical transformer and deep CNN models, our approach delivers a superior trade-off between accuracy and fairness across all five DR severity levels, making it more suitable for real-world screening applications.

Regarding computational efficiency, traditional deep learning models for DR classification, such as Inception-ResNet-v2 or Swin Transformers, often require millions of parameters and high-end GPUs for both training and inference. For example, Inception-ResNet-v2 has over 55 M parameters and demands significant memory bandwidth due to multiple inception blocks and residual layers. In contrast, the proposed hybrid quantum-classical model achieves a more compact architecture by projecting high-dimensional ResNet-50 features (2048D) into a low-dimensional 8D quantum state space. The quantum layer itself adds only ~300-400 trainable parameters, significantly reducing computational overhead. During inference, the quantum processing layer simulates rapidly (on *default.qubit*) and requires fewer FLOPs compared to full deep CNNs or transformer models. While full quantum execution is currently simulated, the model's compressed design and modular quantum embedding suggest future suitability for edge deployment or integration with lightweight classical frontends. Thus, the proposed model balances performance with efficiency, addressing the resource constraints seen in rural or mobile healthcare scenarios.

Limitations

Despite its promising performance, the proposed hybrid quantum–classical model has limitations. The quantum circuit is simulated on classical hardware, restricting scalability and real-time deployment. The model's accuracy in detecting Severe and Proliferative DR classes remains lower due to dataset imbalance and overlapping features. Future work will focus on implementing the model on actual quantum hardware, incorporating larger and more diverse retinal datasets, and exploring advanced quantum encodings or contrastive learning. This study is limited to a single dataset (APTOs 2019). Although it provides balanced multiclass labels, future work will expand evaluation to additional datasets like IDRiD, MESSIDOR, and EyePACS to assess generalizability across demographics and imaging conditions. While recent studies have explored DR classification into 6–7 subclass levels to capture subtle disease progression, this work adheres to the 5-stage ICDR protocol due to the class schema provided by the APTOS 2019 dataset. The five-stage labeling reflects common clinical practice and ensures compatibility with real-world screening scenarios. Furthermore, publicly available datasets offering granular subclass annotations for a 7-stage DR classification remain limited. Future work may explore this direction using extended datasets or clinician-labeled images with fine-grained DR severity annotations. Additionally, clinical validation with expert-annotated images and integration into tele-ophthalmology platforms could enhance real-world applicability and support widespread diabetic eye screening.

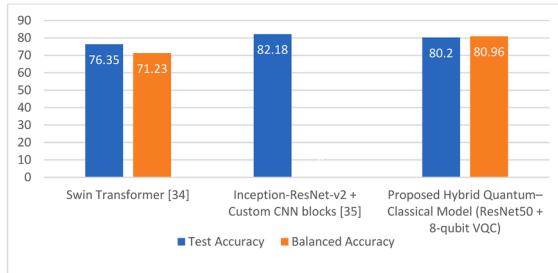
Ethics statements

None

Table 7

Performance comparison of proposed Quantum-Classical method with other models on APTOS 2019 dataset.

Model/Method	Test Accuracy	Balanced Accuracy (Avg. Recall)
Swin Transformer [34]	76.35	71.23
Inception-ResNet-v2 + Custom CNN blocks [35]	82.18	Not reported
Proposed Hybrid Quantum-Classical Method (ResNet50 + 8-qubit VQC)	80.2	80.96

**Fig. 8.** Inter-model performance comparison.

CRediT author statement

Tabassum Ara: Conceptualization, Resources, data curation, Software, Validation

Ved Prakash Mishra: Supervision and Project administration

Manish Bali: Conceptualization, Methodology, Formal analysis, Writing- original draft, review & editing

Anuradha Yenkar: Software, Validation and Supervision

Supplementary material and/or additional information

None

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

Data will be made available on request.

References

- [1] J.W.Y. Yau, S.L. Rogers, R. Kawasaki, et al., Global prevalence and major risk factors of diabetic retinopathy, *Diabet. Care* 35 (3) (2012) 556–564. International Diabetes Federation, "IDF Diabetes Atlas, 10th ed.", 2021. [Online]. Available, <https://diabetesatlas.org>.
- [2] R. Raman, L. Gella, et al., Prevalence of diabetic retinopathy in India: Sankara nethralaya diabetic retinopathy epidemiology and molecular genetic study (SN-DREAMS), *Diabet. Res. Clin. Pract.* 77 (3) (2007) 455–462.
- [3] G.G. Gardner, D. Keating, et al., Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool, *Brit. J. Ophthalmol.* 80 (11) (1996) 940–944.
- [4] V. Gulshan, L. Peng, et al., Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, *JAMA* 316 (22) (2016) 2402–2410.
- [5] H. Pratt, F. Coenen, et al., Convolutional neural networks for diabetic retinopathy, *Proced. Comput. Sci.* 90 (2016) 200–205.
- [6] Mari, T.R. Bromley, J. Izaac, M. Schuld, N. Killoran, Transfer learning in hybrid classical-quantum neural networks, *Quantum* 4 (2020), <https://doi.org/10.22331/Q-2020-10-09-340>. arXiv:1912.08278.
- [7] A. Mir, U. Yasin, S.N. Khan, A. Athar, R. Jabeen, S. Aslam, Diabetic retinopathy detection using classical-quantum transfer learning approach and probability model, *Comput. Mater. Continu.* 71 (2022) 3733–3746, <https://doi.org/10.32604/cmc.2022.022524>.
- [8] G.A.S. Thomas, Y.H. Robinson, E.G. Julie, V. Shanmuganathan, S. Rho, Y. Nam, Intelligent prediction approach for diabetic retinopathy using deep learning based convolutional neural networks algorithm by means of retina photographs, *Comput. Mater. Continu.* 66 (2020) 1613–1629, <https://doi.org/10.32604/cmc.2020.013443>.

- [9] U. Bhimavarapu, G. Battineni, Deep learning for the detection and classification of diabetic retinopathy with an improved activation function, *Healthc. (Switzerl.)* 11 (2023), <https://doi.org/10.3390/healthcare11010097>.
- [10] A. Hunter, J. Lowell, J. Owens, L. Kennedy, D. Steele, Quantification of diabetic retinopathy using Neural networks and sensitivity analysis, (2000) 81–86. doi: [10.1007/978-1-4471-0513-8_10](https://doi.org/10.1007/978-1-4471-0513-8_10).
- [11] M.K. Fadafan, N. Mehrshad, S.M. Razavi, Detection of diabetic retinopathy using computational model of human visual system, *Biomed. Res. (India)* 29 (2018) 1956–1960, <https://doi.org/10.4066/biomedicalresearch.29-18-551>.
- [12] G. Mushtaq, F. Siddiqui, Detection of diabetic retinopathy using deep learning methodology, *IOP Conferen. Ser.: Mater. Sci. Eng.* 1070 (2021) 012049, <https://doi.org/10.1088/1757-899x/1070/1/012049>.
- [13] S.V. Deshmukh, A. Roy, P. Agrawal, Retinal image segmentation for diabetic retinopathy detection using U-net architecture, *Int. J. Image Graph. Signal Process.* 15 (2023) 79–92, <https://doi.org/10.5815/ijigsp.2023.01.07>.
- [14] G.Sun Bilal, Y. Li, S. Mazhar, A.Q. Khan, Diabetic retinopathy detection and classification using mixed models for a disease grading database, *IEEE Access.* 9 (2021) 23544–23553, <https://doi.org/10.1109/ACCESS.2021.3056186>.
- [15] L.Zhu Bilal, A. Deng, H. Lu, N. Wu, AI-based automatic detection and classification of diabetic retinopathy using U-net and Deep learning, *Symmet. (Basel)* 14 (7) (2022) 1427, <https://doi.org/10.3390/sym14071427>.
- [16] Bilal, A. Imran, L.X. Talha Imtiaz Baig, H. Long, A. Alzahrani, M. Shafiq, Improved Support Vector machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification, *PLoS One* 19 (1) (2024), [https://doi.org/10.1371/journal.pone.0295951 e0295951–e0295951](https://doi.org/10.1371/journal.pone.0295951).
- [17] P.N. Chen, C.C. Lee, C.M. Liang, S.I. Pao, K.H. Huang, K.F. Lin, General deep learning model for detecting diabetic retinopathy, *BMC Bioinform. (Basel)* 22 (2021) 1–14, <https://doi.org/10.1186/s12859-021-04005-x>.
- [18] G. Kumar, S. Chatterjee, C. Chattopadhyay, DRISTI: a hybrid deep neural network for diabetic retinopathy diagnosis, *Signal Image Video Process.* 15 (2021) 1679–1686, <https://doi.org/10.1007/s11760-021-01904-7>.
- [19] J.K. Shimpi, P. Shanmugam, A hybrid diabetic retinopathy neural network model for early diabetic retinopathy detection and classification of Fundus images, *Traitem. du Signal* 40 (2023) 2711–2722, <https://doi.org/10.18280/ts.400631>.
- [20] S. Mehta, R. Aggarwal, A. Jain, Quantum chimp-enhanced SqueezeNet for precise diabetic retinopathy classification, *Sci. Rep.* 15 (1) (2025) 1–15, <https://doi.org/10.1038/s41598-025-XXXX-X>.
- [21] A. Gupta, S. Roy, DeepSVDNet: a deep learning-based approach for detecting and classifying vision-threatening diabetic retinopathy in retinal fundus images, *Comput. Syst. Eng.* 44 (1) (2024) 27–39, <https://doi.org/10.32604/csse.2024.043211>.
- [22] M. Tanveer, L. Zhou, S. Wang, EdgeSVDNet: 5G-enabled detection and classification of vision-threatening diabetic retinopathy in retinal fundus images, *Electron. (Basel)* 12 (6) (2023) 1307, <https://doi.org/10.3390/electronics12061307>.
- [23] M. Kumar, H. Singh, Multi-deep learning approach with transfer learning for 7-stages diabetic retinopathy classification, *Int. J. Imaging Syst. Technol.* 34 (2) (2024) 221–234, <https://doi.org/10.1002/ima.23202>.
- [24] L. Zhu, A. Deng, H. Lu, N. Wu, AI-based automatic detection and classification of diabetic retinopathy using U-net and Deep learning, *Symmet. (Basel)* 14 (7, 1427) (2022), <https://doi.org/10.3390/sym14071427>.
- [25] H. Li, J. Xu, A transfer learning and U-net-based automatic detection of diabetic retinopathy from fundus images, *Comput. Intell. Neurosci.* 2022 (2022) 2467854, <https://doi.org/10.1155/2022/2467854>.
- [26] T. Akhtar, N.S. Kamel, M. Hussain, A novel fusion of genetic grey wolf optimization and kernel extreme learning machines for precise diabetic eye disease classification, *PLoS One* 19 (3) (2024) e0283017, <https://doi.org/10.1371/journal.pone.0283017>.
- [27] A. Bilal, X. Liu, T.I. Baig, H. Long, A. Alzahrani, M. Shafiq, Improved support vector machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification, *PLoS One* 19 (1) (2024) e0295951, <https://doi.org/10.1371/journal.pone.0295951>.
- [28] A. Imran, S. Khan, M.A. Khan, Improved Grey Wolf optimization-based feature selection and classification using CNN for diabetic retinopathy detection, *IEEE Access.* 10 (2022) 55648–55659, <https://doi.org/10.1109/ACCESS.2022.3178400>.
- [29] R. Zhang, Y. Zhao, L. Wang, NIMEQ-SACNet: a novel self-attention precision medicine model for vision-threatening diabetic retinopathy using image data, *Comput. Biol. Med.* 168 (2024) 107608, <https://doi.org/10.1016/j.combiom.2024.107608>.
- [30] H. Rehman, F. Ullah, Optimizing fully convolutional encoder-decoder network for segmentation of diabetic eye disease, *Comput. Mater. Contin.* 75 (2) (2023) 2761–2778, <https://doi.org/10.32604/cmc.2023.042132>.
- [31] G. Leung, C. Pieramici, M. Kermani, Survey on recent developments in automatic detection of diabetic retinopathy, *Journal Français d'Ophtalmologie* 44 (9) (2021) 1301–1310, <https://doi.org/10.1016/j.jfo.2021.03.011>.
- [32] G. Bilal, G. Sun, Y. Li, S. Mazhar, A.Q. Khan, Diabetic retinopathy detection and classification using mixed models for a disease grading database, *IEEE Access.* 9 (2021) 23544–23553, <https://doi.org/10.1109/ACCESS.2021.3056186>.
- [33] R. Qureshi, A.B. Khan, Diabetic retinopathy detection using AI, *J. Med. Imaging Health Inform.* 14 (2) (2024) 287–295, <https://doi.org/10.1166/jmihi.2024.4581>.
- [34] K.A. Shaikh, M.A. Siddiqui, Diabetic retinopathy detection using weighted filters and classification using CNN, *Med. Biol. Eng. Comput.* 59 (2021) 249–259, <https://doi.org/10.1007/s11517-020-02229-4>.
- [35] J. Biamonte, P. Wittek, et al., Quantum machine learning, *Nature* 549 (7671) (2017) 195–202.
- [36] M. Schuld, I. Sinayskiy, F. Petruccione, The quest for a quantum neural network, *Quant. Inf. Process.* 13 (11) (2014) 2567–2586.
- [37] H. Chen, Q. Li, J. Wang, A hybrid quantum-classical approach for medical image classification, *IEEE Access.* 8 (2020) 133924–133935.
- [38] M. Bali, V. Prakash Mishra, A. Yenikar, D. Chikmurge, QuantumNet: an enhanced diabetic retinopathy detection model using classical deep learning-quantum transfer learning, *MethodsX* 14 (2025) 103185, <https://doi.org/10.1016/j.mex.2025.103185>. ISSN 2215-0161.
- [39] S. Alsulai, A. Alqahtani, A. Binbusayis, M. Sha, A. Gumaei, S. Wang, Quantum computing meets deep learning: a promising approach for diabetic retinopathy classification, *Mathematics* 11 (2023), <https://doi.org/10.3390/math11092008>.
- [40] S.J. Wei, Y.H. Chen, Z.R. Zhou, G.L. Long, A quantum convolutional neural network on NISQ devices, *AAPPS Bull.* 32 (2022), <https://doi.org/10.1007/s43673-021-00030-3>. arXiv:2104.06918.
- [41] T. Hur, L. Kim, D.K. Park, Quantum convolutional neural network for classical data classification, *Quant. Mach. Intell.* 4 (2022), <https://doi.org/10.1007/s42484-021-00061-x>. arXiv:2108.00661.
- [42] A. Bilal, et al., BC-QNet: a quantum-infused ELM model for breast cancer diagnosis, *Comput. Biol. Med.* (2024), <https://doi.org/10.1016/j.combiom.2024.108483>, 108483–108483.
- [43] APTOS 2019 blindness detection. Available online: <https://www.kaggle.com/datasets/mariathererot/aptos2019>.
- [44] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778, <https://doi.org/10.1109/CVPR.2016.90>.
- [45] V. Bergholm, et al., PennyLane: automatic differentiation of hybrid quantum-classical computations, *arXiv preprint*, arXiv:1811.04968, 2018, <https://arxiv.org/abs/1811.04968>. Available.
- [46] S.A. Kale, R.R. Patil, Linux Yourself: Concept and Programming, 1st ed., Chapman and Hall/CRC, 2020 <https://doi.org/10.1201/9780429446047>.
- [47] A.M. Payyaz, M.I. Sharif, S. Azam, A. Karim, J. El-Den, Analysis of diabetic retinopathy (DR) based on the deep learning, *Information* 14 (1) (2023) 30, <https://doi.org/10.3390/info14010030>.
- [48] C. Zhang, T. Lei, P. Chen, Diabetic retinopathy grading by a source-free transfer learning approach, *Biomed. Signal. Process. Control* 73 (2022) 103423.
- [49] S. Xu, Z. Huang, Y. Zhang, Diabetic retinopathy progression recognition using deep learning method, Available online, <http://cs231n.stanford.edu/reports/2022/pdfs/20.pdf>, 2024. accessed on 11 Aug.
- [50] A.K. Gangwar, V. Ravi, Diabetic retinopathy detection using Transfer learning and Deep learning, in: V. Bhateja, S.L. Peng, S.C. Satapathy, Y.D. Zhang (Eds.), Evolution in Computational Intelligence. Advances in Intelligent Systems and Computing, Evolution in Computational Intelligence. Advances in Intelligent Systems and Computing, 1176, Springer, Singapore, 2021, https://doi.org/10.1007/978-981-15-5788-0_64.