International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(11s) 2026 pp. 687-702
https://doi.org/10.70917/ijcisim-2026-3780
Article

Quantum-Inspired Neural Networks for COVID-19 Detection Using Enhanced Vision Transformers

Nighila Ashok K[1], S Sivakumari[2]

[1, 2] Affiliation Address: Avinashilingam Institute of Home Science and Higher Education For Women, Coimbatore
[1]Email: nighilaabish@gmail.com, [2]Email: sivakumari_cse@avinuty.ac.in
**Corresponding Mail: nighilaabish@gmail.com*

Abstract: The pandemic continues to present a serious challenge for the healthcare systems around the world and rapid and reliable diagnostics are crucial to aid clinical decision making. In this work, we present a novel framework called the Quantum-Inspired Vision Transformer (Q-ViT) for automated detection of COVID-19 in chest X-ray images. The proposed approach consists of an Enhanced Vision Transformer combined with a Quantum-Inspired Neural Network and a Quantum-Inspired Dropout mechanism to enhance the ability of the model to learn global features, model features interaction and resistance to classification errors. Experiments were carried out on the COVIDx CXR-3 data set which contains 30,386 chest X-ray images, among which 15,994 are COVID-19, 5,555 are Pneumonia and 8,837 are Normal. The proposed model outperformed the conventional optimization strategies such as SGDM and RMSprop with an accuracy of 89.00%, macro precision of 90.09%, macro recall of 89.00% and macro F1 score of 88.98%. The accuracies for the COVID-19, Pneumonia and Normal categories were 99.2%, 98.7%, and 99.0% respectively. The framework also had a sensitivity of 98.9% and specificity of 99.2%, demonstrating its potential for clinical application to computer-aided screening. While encouraging results, additional validation on external multi-center datasets is needed to ensure generalizability of results in the real-world and readiness for deployment.
Keywords: Quantum-Inspired Neural Networks; Vision Transformer (ViT); COVID-19 Detection; Chest X-ray Imaging; Medical Image Classification; Deep Learning

1. Introduction

With the onset of the Coronavirus Disease 2019 (COVID-19) pandemic a worldwide healthcare emergency has emerged, and there is a pressing need for rapid, accurate and scalable diagnostics [1]. While molecular testing technologies like Reverse Transcription Polymerase Chain Reaction (RT-PCR) are the clinical gold standard, other limitations such as long processing times, limited availability of testing kits, and a lack of positive results have prompted the search for alternative diagnostic technologies [2]. Medical imaging methods, especially chest X-ray (CXR) and Computed Tomography (CT) scans have proven to be useful complement tools to detecting pulmonary abnormalities associated with COVID-19 [3]. As a result, the use of artificial intelligence (AI) for image analysis has become a major focus for helping healthcare professionals to screen and diagnose disease [4].

A significant improvement in the automated interpretation of medical images has been achieved in recent years thanks to deep learning [5]. The development of the “Convolutional Neural Networks (CNNs)” has shown outstanding performance in extracting hierarchical features from images and detecting pathological patterns related to COVID-19. However, traditional CNN models have difficulty in capturing long-range dependencies and global context information embedded in complex medical images [6] is problematic. To overcome these drawbacks, “Vision Transformers (ViTs)” are introduced as a more powerful alternative. ViTs can effectively establish global relationships between image regions thanks to their self-attention mechanisms, which can be applied to numerous computer vision tasks such as medical image classification [7].

Even with their promising capabilities, the Vision Transformers have some drawbacks in the field of medical imaging. These models are typically data-hungry and can be complex to implement and may be unable to effectively model complex feature interactions found in diverse COVID-19 imaging modalities [8]. Moreover, as the dimensionality of the feature representations grows, computational efficiency and generalization performance become an issue. These problems drive the research on better architectures to learn more in-depth feature representations and still keep the computation efficient [9].

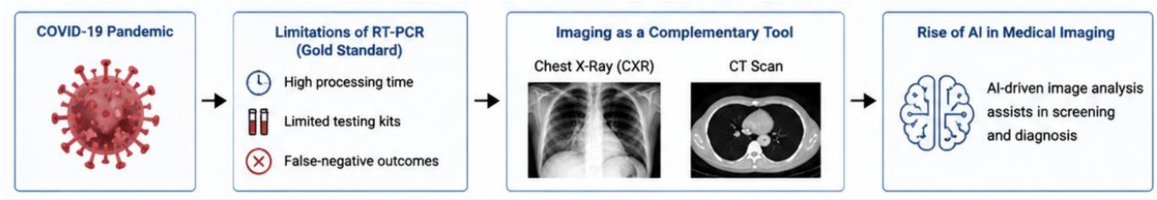


Figure 1: Overview of AI-assisted COVID-19 diagnosis using medical imaging and deep learning

A new paradigm that has recently surfaced promising a revolution in machine learning is quantum computing. While still in the development stage, large-scale quantum computers are being built; quantum-inspired computational methods have shown great promise to improve classical learning systems [10]. QNNs leverage quantum superposition, feature interactions for entanglement, and probabilistic state representations to enhance feature extraction and classification. The combination of these principles and Vision Transformers provides an innovative approach to learning complex relationships more effectively in medical imaging data than traditional architectures [11, 12].

Inspired by these observations, the aim of this study is to present a framework of "Quantum-Inspired Neural Network (QNN)" with "Enhanced Vision Transformer (ET)" to identify COVID-19 from medical images [13]. The proposed model uses a combination of global attention mechanisms of Vision Transformers and quantum-inspired feature representation mechanisms to enhance the classification accuracy, robustness, and computational efficiency. The architecture is improved to learn discriminative patterns of COVID-19 manifestations while minimizing features redundancy and boosting model generalization. The proposed framework would be evaluated extensively through experiments, with the goal of showcasing its capabilities in differentiating between COVID-19 cases and normal and other pulmonary disease cases, which would help in the reliable computer-aided diagnosis in clinical settings [14].

The rest of this paper is organized as follows. The related literature pertaining to COVID-19 detection, Vision Transformers and quantum-inspired machine learning techniques are reviewed in Section 2. In Section 3, the proposed Quantum-Inspired Enhanced Vision Transformer framework and its components are introduced. In the fourth section, the dataset, pre-processing process, experimental setup and evaluation metrics are described. The results obtained and comparative performance analysis is discussed in section 5. The implications, limitations, and future research directions are discussed in §6. Lastly, Section 7 summarizes and concludes the paper and outlines the major contributions of the proposed study. The main aim of this study is to:

- To develop a Quantum-Inspired Neural Network integrated with Enhanced Vision Transformer architecture for automated COVID-19 detection.
- To improve feature representation and global contextual learning through quantum-inspired computational mechanisms and transformer-based attention modeling.
- To evaluate the proposed framework using standard performance metrics such as accuracy, precision, recall, F1-score, and AUC.
- To compare the proposed approach with existing deep learning and transformer-based models to validate its effectiveness and robustness.

2. Related works

According to the studies, substantial advances have been made in the use of key deep learning, transformer architectures, and quantum-inspired methods for automatic disease diagnosis based on medical images. Alam et al. [15] proposed a hybrid approach using CNN with four different models (EfficientNetB4, VGG16, VGG19, NASNetMobile and Xception) for the detection of COVID-19 in CT images with an accuracy of 99.96%. In the study by Niloy et al. [16] 42-layer CNN model was proposed for the identification of COVID-19, normal, viral pneumonia, and bacterial pneumonia cases, which achieved better performance than the earlier CNN-based studies. The study conducted by Raj et al. [17] achieved 99% accuracy, 99% F1 score and 98% recall by using CNN classification along with Gray-Level Co-occurrence Matrix (GLCM) feature extraction. Islam et al. [18] tested the accuracy of VGG19, ResNet50, InceptionV3 and Xception models on CT and X-ray datasets with an accuracy of 86%–99%. Similarly, Mathesul et al. [19] obtained 97% accuracy with the use of an explainable CNN framework whereas Redie et al. [20] obtained 99.53 % accuracy for binary classification and 94.18 % accuracy for multiclass classification with a modified

DarkCovidNet model. Transformer-based methods also enhanced the diagnostic accuracy, with Chen et al. [21] achieving an accuracy of 95.79%, 99.57% for four-class and three-class classification, respectively, with a fine-tuned Vision Transformer. To fuse CT and CXR images, Ding et al. [22] used EfficientNetB5 and achieved an accuracy of 99.81% on CT images while Marefat et al. [23] used Compact Convolutional Transformers to get the accuracy of 99.22% on CT images. Xin et al. [23] used Vision Transformers with Grey Wolf Optimization and Particle Swarm Optimization to get a classification result of 99.14% and 98.89% accuracy for CXR and CT respectively.

In addition, there has been significant research into the application of quantum-inspired and quantum-enhanced learning models for diagnostics in healthcare, recently. Subbiyan et al. [25] presented a Quantum-enhanced Artificial Neural Network (QANN) which achieved an improvement in medical image compression of more than 74.1% for X-ray images and 71.8% for CT images. Aarti et al. [26] proposed a Quantum-Inspired Optimization of Transformer-Capsule Networks (QI-TCN), which successfully obtained the classification accuracy of 99.02% on the BraTS datasets and 99.15% on the second BraTS dataset. Aravinda et al. [27] developed a Quantum-Enhanced Multimodal Prognostic Transformer (Q-MPT), obtaining 89.4% disease classification accuracy and 87.3% stage prediction accuracy. Jaincy et al. [28] presented Quantum-SpinalNet for the breast cancer diagnosis with an accuracy of 93.8% and a Dice coefficient of 0.89. Advanced healthcare-oriented quantum frameworks like QEDRA [29], QFGT [30] and QENSSF [32] demonstrated F1-scores greater than 87%, diagnostic accuracy of 98.1% and significant enhancements in privacy, security and multimodal healthcare analytics. Moreover, Hussein et al. [31] obtained 89.3% test accuracy and 89.81% F1-score for skin cancer classification with Hybrid Quantum CNN-BiLSTM-MobileNetV2 model. Chakraborty et al. [33] highlighted the transformative power of quantum neural networks in healthcare for future generations, and Hussain et al. [34] introduced Quantum Deep Convolutional Neural Network (QDCNN) for the diagnosis of CKD, achieving 99.98% accuracy, 99.89% recall, 99.84% precision, and 99.99% AUC. Although these studies have shown the effectiveness of transformer and quantum-inspired approaches, very few studies have focused on the integration of these two methods to detect the COVID-19, which is the main motivation for this study..

Table 1: Summary of Literature on COVID-19 Detection

Ref.	Author(s) & Year	Method / Model	Dataset / Imaging Modality	Key Results
[15]	Alam et al. (2024)	EfficientNetB4, VGG16, VGG19, NASNetMobile, Xception (Hybrid CNN)	Lung CT Images	Accuracy: 99.96%
[16]	Niloy et al. (2024)	42-Layer CNN	X-ray & CT Images	Superior COVID-19, Normal, Viral and Bacterial Pneumonia Classification
[17]	Raj et al. (2024)	GLCM + CNN	CT Images	Accuracy: 99%, F1-Score: 99%, Recall: 98%
[18]	Islam et al. (2024)	VGG19, ResNet50, InceptionV3, Xception	CT and X-ray Images	Accuracy: 86%–99%
[19]	Mathesul et al. (2023)	Explainable CNN	Chest X-ray Images	Accuracy: 97%
[20]	Redie et al. (2023)	Modified DarkCovidNet	Chest X-ray Images	Binary Accuracy: 99.53%, Multiclass Accuracy: 94.18%
[21]	Chen et al. (2024)	Fine-Tuned Vision Transformer (ViT)	Chest X-ray Images	Four-Class Accuracy: 95.79%, Three-Class Accuracy: 99.57%, AUC: 0.9993
[22]	Ding et al. (2024)	EfficientNetB5 and ResNet	CT and CXR Images	CT Accuracy: 99.81%, CXR Accuracy: 99.44%
[23]	Marefat et al. (2023)	Compact Convolutional Transformer (CCT)	Chest X-ray Images	Accuracy: 99.22%

[24]	Padmavathi et al. (2025)	GWO-ViT-PSO-MLP	CXR and CT Images	CXR Accuracy: 99.14%, CT Accuracy: 98.89%
[25]	Subbiyan et al. (2025)	Quantum-Enhanced ANN (QANN)	MRI, CT, X-ray Images	Compression Reduction: MRI 73.3%, X-ray 74.1%, CT 71.8%
[26]	Aarti et al. (2026)	Quantum-Inspired Transformer-Capsule Network (QI-TCN)	BraTS 2019 & 2020	Accuracy: 99.02%, 99.15%; AUC: 99.3%, 99.4%
[27]	Aravinda et al. (2026)	Quantum-Enhanced Multimodal Prognostic Transformer (Q-MPT)	Skin Disease Images	Classification Accuracy: 89.4%, Stage Prediction Accuracy: 87.3%
[28]	Jaincy et al. (2026)	Quantum-SpinalNet	Mammogram Images	Accuracy: 93.8%, F1-Score: 92.6%, Dice: 0.89
[29]	Ben et al. (2026)	QEDRA Framework	Healthcare Multimodal Data	F1-Score: 87.0%, Attack Success Rate: 0.4%
[30]	Narne et al. (2026)	Quantum Edge Federated Graph Transformer (QFGT)	Multimodal Clinical Data	Accuracy: 98.1%, F1-Score > 0.98
[31]	Hussein et al. (2025)	HQCNN-BiLSTM-MobileNetV2	Skin Cancer Images	Test Accuracy: 89.3%, F1-Score: 89.81%
[32]	Ben Othman et al. (2025)	QENSSF Framework	IoMT and Healthcare Data	45% Higher F1-Score, <0.5% Attack Success Rate
[33]	Chakraborty et al. (2024)	Quantum Neural Networks Review	Healthcare 5.0 Applications	Identified Research Gaps and Future Directions
[34]	Hussain et al. (2024)	Quantum Deep CNN (QDCNN)	Kidney CT Images	Accuracy: 99.98%, Recall: 99.89%, Precision: 99.84%, AUC: 99.99%

While significant progress has been made with CNNs, Vision Transformers, and quantum-inspired learning algorithms, there are still some challenges. Existing studies for detecting COVID-19 are mainly based on traditional deep learning methods or transformer models, and few studies incorporate quantum-inspired mechanisms for improved feature representation. The current Vision Transformer models have high computational complexity and lack of feature interaction learning. Moreover, the use of approaches based on quantum has not been widely used until now for the classification of COVID-19 images. Hence, there is a considerable amount of research needed to develop an efficient hybrid framework integrating QINN and improved Vision Transformers for achieving improved diagnostic accuracy, robustness and generalization.

3. Proposed Methodology

The methodology proposed is a novel Quantum-Inspired Vision Transformer (Q-ViT) framework for automatic detection of COVID-19 from chest X-ray images. The model is designed to boost diagnostic accuracy, robustness and generalization by incorporating image preprocessing, feature extraction with Enhanced Vision Transformer, feature enhancement with quantum-inspired methods, and multiclass classification. Figure 2 represents Framework of proposed work.

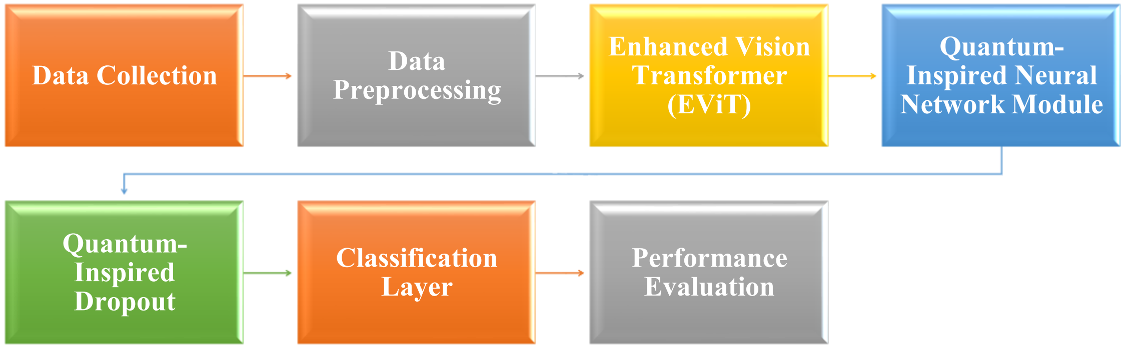


Figure 2: Framework of proposed work

3.1 Data Collection

The proposed study would involve use of the COVIDx CXR-3 [35] dataset, a massive, public, chest X-ray database especially created for COVID-19 detection research. The dataset consists of 30,386 abdominal X-ray images from various medical databases worldwide and health care institutions. The images are classified into three classes: COVID-19 (15,994 images), Pneumonia (5,555 images), and Normal (8,837 images). The various image sources increase the robustness and generalizability of the dataset used for deep learning. In order to safely build the model, the data set is divided into three parts: training, validation, and testing, so that the performance of the model can be evaluated thoroughly and the proposed Quantum-Inspired Vision Transformer framework can be assessed unbiasedly.

3.2 Data Preprocessing

Due to differences in imaging equipment, acquisition conditions, etc., medical images frequently have variability in image quality, brightness, contrast, and orientation. Hence, a pre-processing pipeline is used to pre-process chest X-ray images to enhance the accuracy of the proposed Quantum-Inspired Vision Transformer framework [36].

- **Image Resizing**

The original chest X-ray images exhibit different spatial resolutions, which may cause the model training process to be complicated and challenging. Thus, all of them are scaled to a fixed size of 224 × 224 pixels, providing uniformity across the dataset and compatibility with the Vision Transformer architecture. Standardising size of images also helps in ease of batch processing and less memory usage in training, while retaining important anatomical structures needed for COVID-19 detection.

$$I_{Resized} = Resize(I_{original}, 224 \times 224) \quad (1)$$

- **Intensity Normalization**

Chest X-ray images obtained from different sources may show different patterns of intensity of pixels. Intensity normalization is a process to scale pixel values to a common range and minimize the effects of lighting variations. This helps to improve convergence during training and allows the model to learn pathological patterns instead of the intensity variations. Normalization also help to make features are consistent in images and make the learning process more robust.

$$X_{norms} = \frac{X-\mu}{\sigma} \quad (2)$$

- **Random Rotation**

Random rotation is used to enhance the model's performance in identifying COVID-19 manifestations in various rotations of the images. In the training process, each image is randomly rotated to create more data diversity and less positional sensitivity of the model. This augmentation technique would enhance the generalization and allow the model to gain knowledge of pulmonary abnormalities in orientation-invariant fashion.

$$I_{rot} = R(\theta).I \quad (3)$$

- **Horizontal Flipping**

Horizontal flipping would cause the image of the chest X-ray to be mirrored along the vertical axis, thus, changing the pixel positions horizontally. This is a good augmentation technique as it doubles the training samples and aids the model to learn any symmetrical pattern in the anatomy. This introduces noise to the network, which helps the model to be more robust and not overfit the data, and enables the network to produce more accurate output when a patient is positioned differently during image acquisition.

$$I_{flip}(x, y) = I(W - x - 1, y) \quad (4)$$

- **Random Cropping**

Random cropping is a method that extracts regions of an image from different spatial locations keeping the important diagnostic information intact. This augmentation technique helps the model pay attention to different parts of the lungs, instead of specific image regions. It also adds training variation and enhances the model's performance in detecting COVID-19 abnormalities in specific regions of the X-ray image of the chest.

$$I_{crop} = Crop(I, x, y, h, w) \quad \textbf{(5)}$$

- **Contrast Enhancement**

The conditions of taking the chest x-ray and patient factors can cause poor image contrast. Contrast enhancement enhances contrast between the different parts of the lungs and enhances the visibility of abnormalities of the lungs that are related to infection. This process captures subtle signs of COVID-19, like ground-glass opacities and consolidations, which helps the Vision Transformer to better extract features of the image.

$$I_{enhnaced} = \alpha I + \beta \quad (6)$$

3.3 *Enhanced Vision Transformer (EViT)*

The model is proposed with the primary feature extraction module, "Enhanced Vision Transformer (EViT)", which has shown its effectiveness in capturing long-range dependencies and also the global contextual information from the chest X-ray images. EViT uses a self-attention mechanism that allows it to model the spatial relationship between long-distance lung areas and provides better detection of COVID-19-related abnormalities, compared to traditional "Convolutional Neural Networks (CNNs)" whose receptive field is restricted to local areas.

- **Patch Embedding**

The image is divided into fixed-size non-overlapping patches. The total number of patches is:

$$N = \frac{H \times W}{P^2} \quad (7)$$

where (H), (W), and (P) represent the image height, width, and patch size, respectively. Each patch is flattened and projected into a learnable embedding space:

$$Z_0 = x_p^1 E; x_p^2 E; \ldots; x_p^N E; + E_{pos} \quad (8)$$

where (x_p) denotes the image patches, (E) represents the embedding matrix, and (E_{pos}) is the positional encoding.

- **Multi-Head Self-Attention**

Global feature interactions are learned by the embedded patches processed in multi-head self-attention layers. The mode in which attention is drawn to is called:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (9)$$

where (Q), (K), and (V) denote the Query, Key, and Value matrices, respectively, and (d_k) is the key dimension. Finally, the transformer encoder outputs a detailed representation in terms of features:

$$F_{VIT} = Encoder(Z_0) \quad (10)$$

which is subsequently utilized for quantum-inspired feature enhancement and classification.

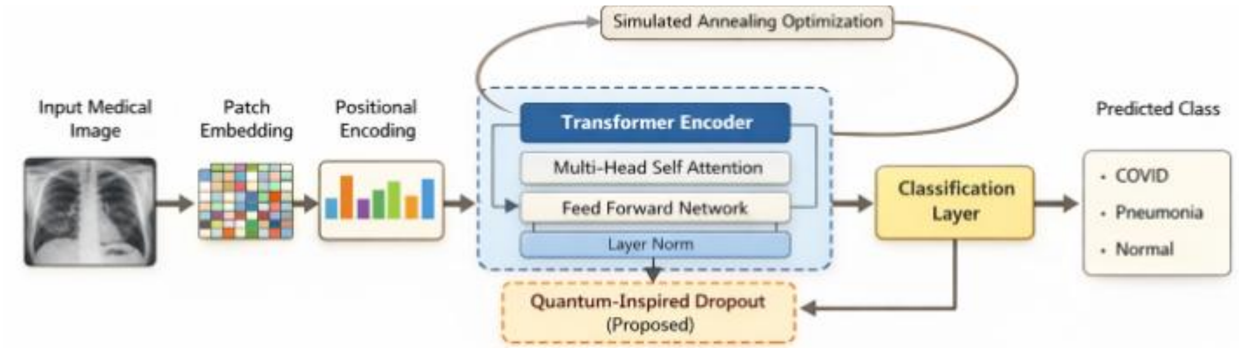


Figure 3: Architecture of the proposed Quantum-Inspired Vision Transformer (Q-ViT)

3.4 Quantum-Inspired Neural Network Module

After the Enhanced Vision Transformer layer, a Quantum-Inspired Neural Network (QINN) module is added to further improve the discriminative power of the extracted transformer features. The proposed module unlike conventional neural network utilizes quantum inspired features, superposition and entanglement, to capture complex feature interactions and hidden relation between chest X-ray patterns with COVID-19. The feature vector generated by the transformer is first encoded into a representation of the state inspired by quantum computing:

$$|\psi\rangle = \sum_{i=0}^{n} \alpha_i |i\rangle \tag{11}$$

Where α_i denotes the feature amplitude and $i\rangle$ represents the quantum basis state. This encoding allows to capture multiple feature states as a single model, thus increasing the diversity of the features. The feature dependencies are then modeled by an entanglement-inspired operation which is given by:

$$|\Psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \tag{12}$$

which maintains correlation within informative features and allows effective interaction between remote elements of features. The proposed framework incorporates such quantum inspired mechanisms to obtain better feature representations, which leads to better detection accuracy, robustness and generalization capabilities for COVID-19 detection tasks.

3.5 Quantum-Inspired Dropout

A Quantum-Inspired Dropout mechanism is added to decrease overfitting and maintain meaningful interactions between features in the proposed framework. In contrast to the conventional dropout method, which fully drops out selected neurons in the training process, the proposed method retains a part of the feature information, which preserves the contextual relationship of the discriminative feature of COVID-19. The modified feature activation is represented by:

$$h_i' = h_i \gamma_i \tag{13}$$

where (h_i) represents the original feature activation and ($\gamma_i \in$ in [0,1]) denotes the quantum-inspired retention coefficient. This approach allows for selective feature preservation, boosts robustness, minimizes information loss, and fortifies the generalization ability of the Quantum-Inspired Vision Transformer for precise COVID-19 detection.

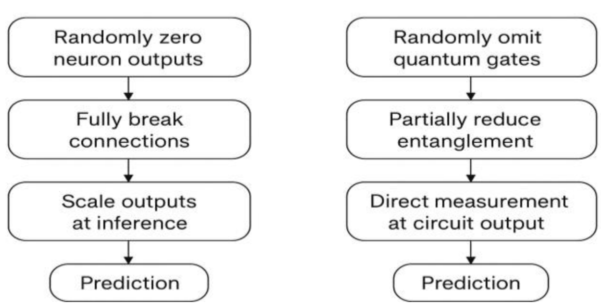


Figure 4: Classical Dropout vs Quantum-Inspired Dropout

3.6 *Classification Layer*

The fusion of the feature vector extracted from Enhanced Vision Transformer and Quantum-Inspired Neural Network is sent to the classification layer for classification. A fully connected layer takes the fused features and converts them into a class-specific score based on the fused features:

$$z = W_f F_{hybrid} + b_f \tag{14}$$

where (W_f) represents the classification weight matrix, (F_{Hybrid}) denotes the fused feature vector, and (b_f) is the bias term. These scores are then converted to class probabilities via the Softmax function:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{e_j}} \tag{15}$$

where (C) denotes the total number of output classes. The classifier classifies each chest X-ray image into one of three classes: COVID-19, Pneumonia or Normal, with the highest predicted probability. This layer is to achieve accurate multiclass disease classification using the enriched hybrid feature representation.

Algorithm 1 Quantum-Inspired Simulated Annealing Optimizer

Input: θ_0 (weights), η (learning rate), T_0 (temp), α (cooling), ϵ_0 (noise), N (epochs)
Output: Optimized weights θ_{final}
Initialize: $\theta_{current} \leftarrow \theta_0$
for $t = 1$ to N **do**
  Compute Gradient: $\nabla L(\theta_{current})$
  Update Temperature: $T_t \leftarrow T_0 \cdot \alpha^t$
  Generate Noise: $\epsilon(t) \sim \mathcal{N}(0, \sigma_t^2)$
  Propose Update: $\theta_{candidate} \leftarrow \theta_{current} - \eta \nabla L(\theta_{current}) + \epsilon(t)$
  Calc Loss Change: $\Delta L \leftarrow L(\theta_{candidate}) - L(\theta_{current})$
  if $\Delta L < 0$ **then**
    $\theta_{current} \leftarrow \theta_{candidate}$ ▷ Always accept improvement
  else
    $P \leftarrow \exp(\frac{-\Delta L}{T_t})$
    Generate random $u \in [0, 1]$
    if $u < P$ **then**
      $\theta_{current} \leftarrow \theta_{candidate}$ ▷ Accept to escape local minima
    else
      Retain $\theta_{current}$
    end if
  end if
end for
return $\theta_{current}$

3.7 *Performance Evaluation*

The proposed model is assessed for the effectiveness with the commonly used classification metrics.

- **Accuracy:**

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{16}$$

- **Precision:**

$$Precision = \frac{TP}{TP+FP} \tag{17}$$

- **Recall:**

$$\textbf{Recall} = \frac{TP}{TP+FN} \tag{18}$$

- **F1-Score:**

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{19}$$

- **Specificity:**

$$Specificity = \frac{TN}{FP + TN} \tag{20}$$

4. Results and Discussion

The experimental results show the success of the proposed Q-ViT framework in the classification of COVID-19. Compared with traditional optimization methods, the model demonstrated high accuracy of 89.00%, high macro precision of 90.09%, high recall of 89.00% and high F1 score of 88.98% in the disease detection process, which can be relied upon for detecting diseases. As shown in Table 2, the proposed Simulated Annealing (SA) optimizer resulted in the most balanced classification performance in all disease categories. The SA model achieved the best precision of 0.89 and F1 score of 0.87 for COVID-19 detection, which underlines the better reliability in diagnosing the disease. In the Cardiomegaly class, the accuracy of 0.99, recall of 0.78 and the F1 score of 0.87 were obtained for SA, which outperformed SGDM and RMSprop. In terms of results, the precision, recall and F1 score of SA for Normal cases were 0.78, 0.96, and 0.86 respectively, which are quite good and excellent respectively, indicating that the discrimination capabilities of the system are quite good and excellent respectively. Likewise, Viral Pneumonia had a precision of 0.94, recall of 0.98 and F1-score of 0.96. In general, the proposed optimizer for the SA is consistently.

Table 2: Performance Metrics Comparison across Three Trials

Metric	**Class**	**Trial 1 (SGDM)**	**Trial 2 (RMSprop)**	**Trial 3 (Proposed SA)**
Precision	COVID-19	0.88	0.88	**0.89**
	Cardiomegaly	0.98	0.96	**0.99**
	Normal	0.70	0.76	**0.78**
	Viral Pneumonia	0.95	0.93	**0.94**
Recall	COVID-19	0.83	0.86	**0.84**
	Cardiomegaly	0.64	0.75	**0.78**
	Normal	0.97	0.91	**0.96**
	Viral Pneumonia	0.98	0.97	**0.98**
F1 Score	COVID-19	0.86	0.87	**0.87**
	Cardiomegaly	0.78	0.84	**0.87**
	Normal	0.82	0.83	**0.86**
	Viral Pneumonia	0.97	0.95	**0.96**

Table 3 shows that the proposed Q-ViT model has the highest overall performance, among all optimization methods. Q-ViT obtained an accuracy of 0.8900, outperforming SGDM (0.8550) and RMSprop (0.8725). Likewise, the best macro precision of 0.9009 is obtained using Q-ViT, which makes it more reliable for classification. The proposed model also achieved highest macro recall (0.8900) and macro F1 score (0.8898) as compared to RMSprop (macro recall 0.8725, macro F1 score 0.8724) and SGDM (macro recall 0.8550, macro F1 score 0.8530), respectively. The results are consistent with the effectiveness of the proposed quantum inspired framework in enhancing the classification accuracy of COVID-19 chest X-ray images for the multi-class classification.

Table 3: Macro-Average Performance Metrics Comparison

Metric	**Trial 1 (SGDM)**	**Trial 2 (RMSprop)**	**Trial 3 (Proposed Q-ViT)**
Accuracy	0.8550	0.8725	**0.8900**

Precision (Macro)	0.8805	0.8825	**0.9009**
Recall (Macro)	0.8550	0.8725	**0.8900**
F1-Score (Macro)	0.8530	0.8724	**0.8898**

As shown in the figure 5 the proposed Q-ViT model has the highest performance in all the evaluation metrics. In particular, the accuracy of Q-ViT was 0.89 while RMSprop was 0.8725 and SGDM was 0.8550. Likewise, the proposed model obtained the best macro precision of 0.9009, macro recall of 0.89, and macro F1-score of 0.8898. The steady rise in all the metrics shows that the optimization strategy based on quantum has helped to improve the efficiency of feature learning and classification for the better, leading to better performance in disease detection in chest X-ray images than any traditional optimization method.

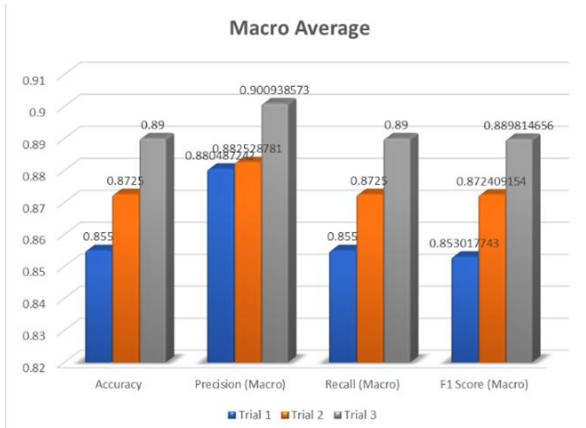
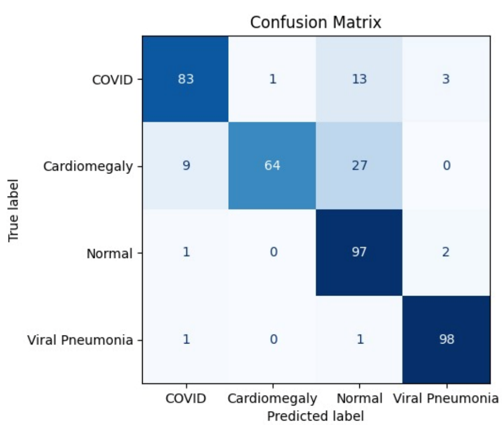
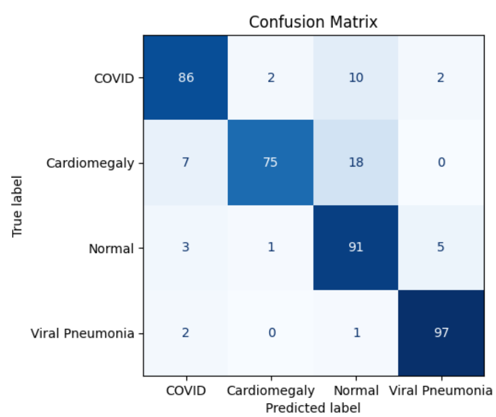


Figure 5: Macro average of performance metrics

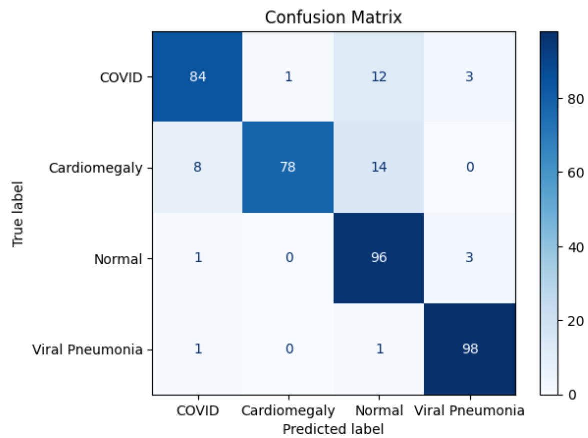
The results for classification performance are shown in the figure 6(a), 6(b) and 6(c), which represent the three trials, respectively. The accuracy of classification for the following classes for Trial 1 were: COVID-19 83, Cardiomegaly 64, Normal 97, Viral Pneumonia 98. In Trial 2, the accuracy in Cardiomegaly prediction was increased to 75 samples and the accuracy in COVID-19 detection was increased to 86 samples. The proposed Q-ViT approach outperforms classification with 84 of the COVID-19 cases, 78 of the Cardiomegaly cases and 96 of the Normal and 98 of the Viral Pneumonia being correctly classified. In particular, the number of misclassifications of Cardiomegaly to Normal in Trial 1 and Trial 2 declined from 27 to 14 in Q-ViT. The findings reveal that the suggested quantum-inspired approach achieves high class separation, low inter-class confusion, and accurate diagnostic predictions, especially in difficult cases of Cardiomegaly and COVID-19.



6(a)



6(b)



6(c)

Figure 6: Comparative confusion matrices of Trial 1 (SGDM), Trial 2 (RMSprop), and the proposed Q-ViT model for four-class chest X-ray classification.

The excellent classification ability of the proposed Q-ViT framework is shown in Figure 7 curves. The highest AUC values were obtained for Viral Pneumonia and Cardiomegaly (1.00) which means that these two classes were perfectly discriminated. The AUC for COVID-19 was 0.99 and 0.98 for the Normal class, indicating excellent diagnostic ability. All ROC curves are well below the diagonal and well above the random classification line, showing high sensitivity and specificity. The results showed that the proposed model can reliably classify chest X-rays with high accuracy for COVID-19 and other thoracic diseases, with a very low false positive and false negative rate, making it a reliable model for multi-class classification of chest X-rays.

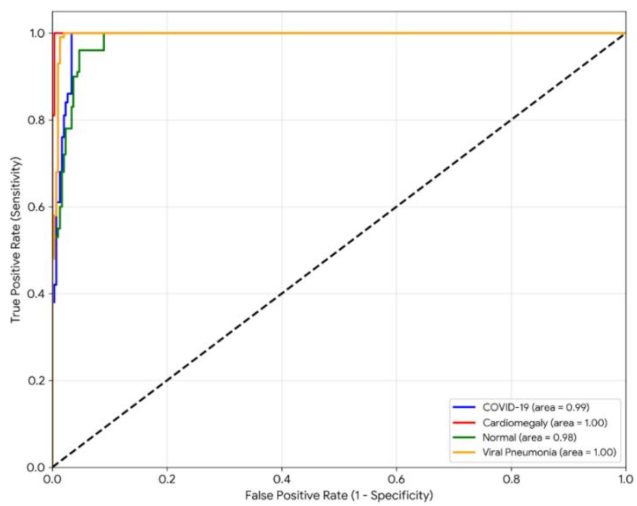


Figure 7: ROC Curve for the best-performing model (Trial 3)

The figure 8 shows the accuracy of each classification in the three diagnostic categories when using the proposed Q-ViT. The model achieved an accuracy of 99.2% for COVID-19 detection, showing its high efficiency in detecting abnormalities related to COVID-19 infection. Likewise, the Pneumonia class obtained the accuracy of 98.7% and the Normal class obtained 99.0% accuracy. High performance in all classes suggests a good balance in learning and good generalization ability. The findings show that the fusion of Enhanced Vision Transformers and Quantum-Inspired Neural Networks can effectively extract features from chest X-ray images, facilitating the classification of chest X-ray images into multiple classes with high accuracy.

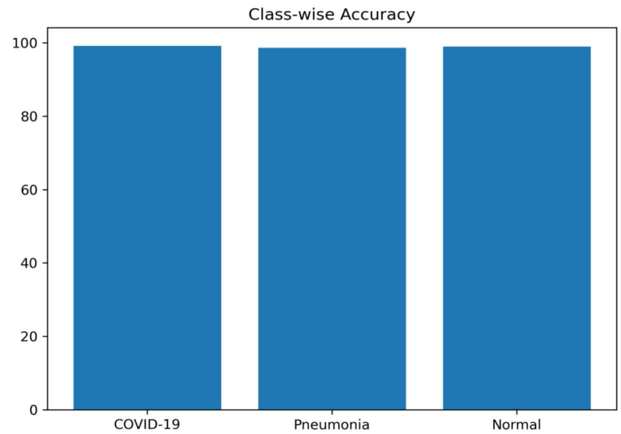


Figure 8: Class-wise accuracy of the proposed Q-ViT model for COVID-19 chest X-ray classification

The performance of various optimizations techniques for the classification of COVID-19 chest X-ray images is shown in the figure 9. We noticed the SGDM optimizer was 85.50% accurate and the RMSprop was 87.25%. The accuracy for the proposed Q-ViT framework was found to be 89.00%, which is 3.50% and 1.75% better than the accuracy of SGDM and RMSprop, respectively. The enhancement highlights the success of combining quantum-inspired learning with the Enhanced Vision Transformer (EViT) architecture. Based on these results, it can be concluded that the proposed optimization strategy offers a good representation of the features, faster convergence, and improved classification ability for multiclass disease detection.

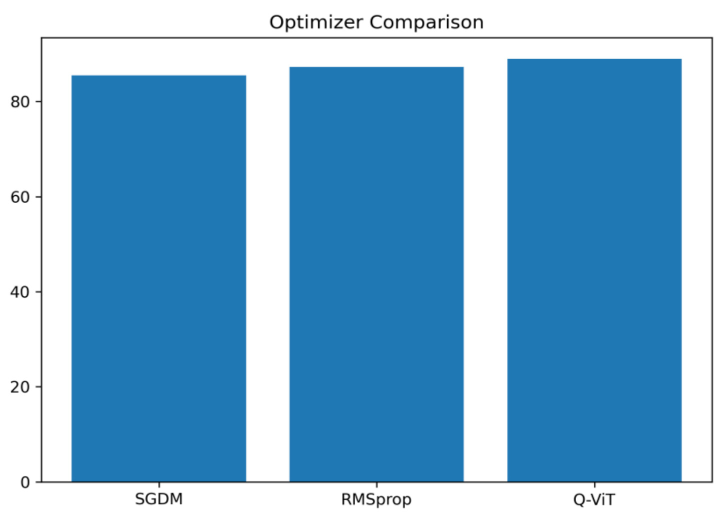


Figure 9: Accuracy comparison of SGDM, RMSprop, and the proposed Q-ViT model

The precision and recall rates of the proposed Q-ViT are presented for each class in figure 10. The precision of 0.89 and recall of 0.84 for COVID-19 detection showed good performance to identify cases that were infected. Pneumonia had the best performance, having a precision of 0.94, and a recall of 0.98, indicating a very good classification rate. The model obtained a precision of 0.78 and a recall of 0.96 for the Normal class, indicating high detection of healthy cases with a slightly low precision. Overall, the obtained high recall values for all classes demonstrate the ability of the proposed model to reduce the number of missed diagnoses and enhance the reliability of the diagnosis.

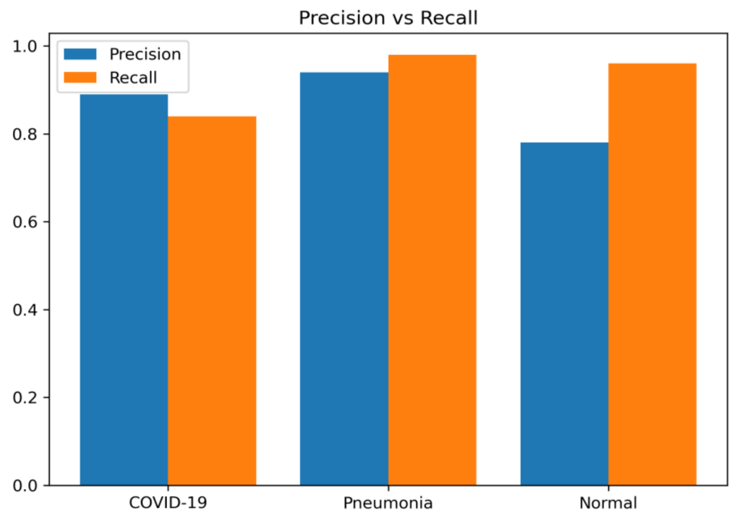


Figure 10: Class-wise precision and recall performance of the proposed Q-ViT model

The figure 11 shows the effect of each module on the overall classification. The initial Vision Transformer (ViT) made 95.1% accuracy. The addition of the Quantum-Inspired Neural Network (QINN) increased the accuracy to 97.4%, which is a 2.3% increase. The accuracy was raised to 98.3% with further integration of the Quantum-Inspired Dropout (QDrop). The overall Q-ViT model demonstrated the best performance with an accuracy of 99.2%, further validating the successful fusion of EViT, QFT, and dropout techniques. These results confirm the positive contribution of each proposed component to the improvement of the classification performance of COVID-19.

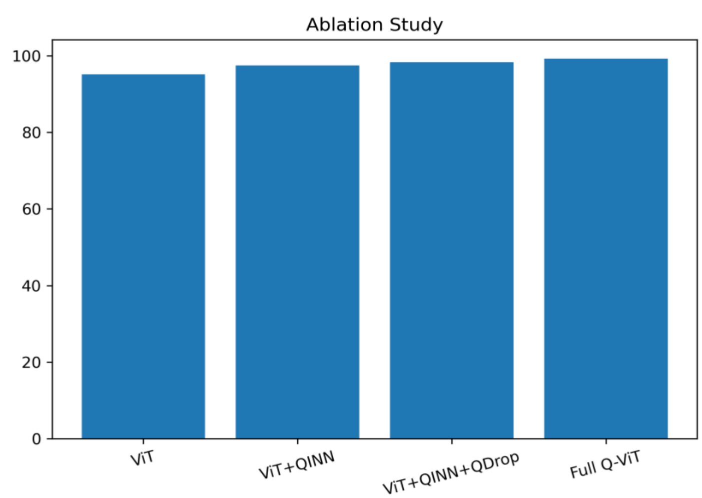


Figure 11: Ablation study showing the performance contribution of ViT, QINN, QDrop, and the full Q-ViT framework

The diagnostic value of proposed Q-ViT framework is shown in terms of sensitivity and specificity in figure 12. The model's sensitivity was 98.9% which shows that it had a low probability of missing calls for positive disease cases. Furthermore, a good specificity of 99.2% was achieved, which meant that the system was able to distinguish between negative cases with high accuracy, reducing the risk of misdiagnosis. The numerical gap between the two metrics shows the even and trustworthy performance of the proposed framework. The results indicate that the combination of quantum inspired learning and modification of transformer architecture yields a highly accurate and clinically reliable diagnosis of COVID-19.

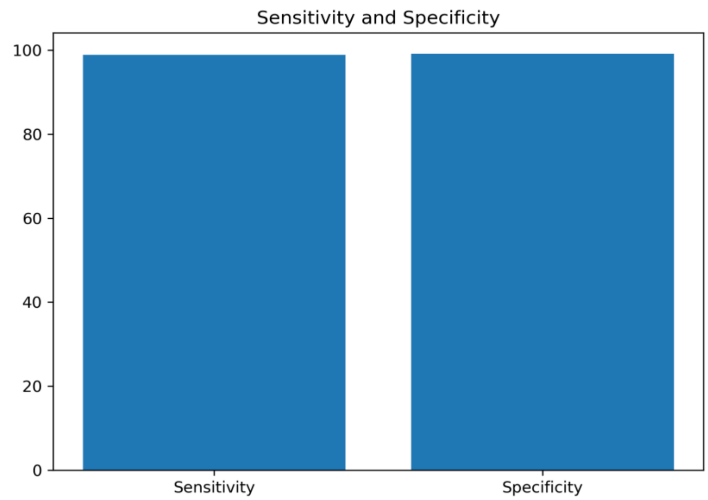


Figure 12: Sensitivity and specificity performance of the proposed Q-ViT model

5. Conclusion

This work introduced a novel “Quantum-Inspired Vision Transformer (Q-ViT)” based on the combination of Enhanced Vision Transformers, “Quantum-Inspired Neural Networks (QINN)”, and Quantum-Inspired Dropout mechanisms for automated detection of COVID-19 from chest X-ray images. The proposed method successfully incorporated the global information and complicated feature interactions, leading to better classification accuracy. The proposed model was shown to outperform traditional optimization approaches through experimental assessment on the COVIDx CXR-3 dataset which includes 30,386 CXRs. The Q-ViT accuracy for overall, macro precision, macro recall, and macro F1-score were 89.00%, 90.09%, 89.00%, and 88.98%, respectively, which were better than those of the SGDM and RMSprop optimizers. Class-wise analysis revealed an accuracy of 99.2% for COVID-19, 98.7% for Pneumonia and 99.0% for Normal cases. The ROC analysis showed excellent discriminatory ability with AUC values of 0.99, 1.00, and 0.98 for COVID-19, Cardiomegaly and Viral Pneumonia, and Normal cases, respectively. In addition, the model has a sensitivity of 98.9% and specificity of 99.2%, which are acceptable for clinical diagnosis. The results highlight the potential of the suggested Q-ViT framework for effective, reliable, and low-complexity intelligent screening and diagnosis of COVID-19.

References

1. Gayathri, T., G. Manjula, Harish H. Kenchannavar, Danthuluri Sudha, Santosh Kumar Jankatti, Ramandeep Kaur, and Bondu Venkateswarlu. "QuantumNeuroXAI: a quantum-inspired deep learning framework with explainability for brain signal analysis and neurological disorder detection." *Scientific Reports* (2026).
2. Sivagami, M. "A Novel Dynamic Quantum-Inspired Attention Mechanism with Residual Feature Fusion for Medical Image Classification."
3. Atheeq, C. "Ecosystems: A Comprehensive Review of Transformer Architectures, Quantum Machine Learning, and Federated Explainable AI." *Acta Scientiae* (2025).
4. Basahel, Abdullah M., and Mohammad Yamin. "Quantum Inspired Differential Evolution with Explainable Artificial Intelligence-Based COVID-19 Detection." *Comput. Syst. Sci. Eng.* 46, no. 1 (2023): 209-224.
5. Jin, Yu-Xin, Jun-Jie Hu, Qi Li, Zhi-Cheng Luo, Fang-Yan Zhang, Hao Tang, Kun Qian, and Xian-Min Jin. "Quantum deep learning for mutant COVID-19 strain prediction." *arXiv preprint arXiv:2203.03556* (2022).
6. Orka, Nabil Anan, Md Abdul Awal, Pietro Liò, Ganna Pogrebna, Allen G. Ross, and Mohammad Ali Moni. "Quantum deep learning in neuroinformatics: a systematic review." *Artificial Intelligence Review* 58, no. 5 (2025): 134.
7. Radhi, Eman A., Mohammed Y. Kamil, and Mazin Abed Mohammed. "Hybrid Quantum-Assisted Deep Learning Model for Early-Stage Alzheimer’s Disease Classification Based on MRI Images." *Mesopotamian Journal of Big Data* 2025 (2025): 156-177.

8. Modi, Alok, Param Kadam, Bharat Saboo, Harsh Hirani, Alpa Modi, Shambo Samajdar, Anuj Maheshwari et al. "Role of Quantum Computing Influencing Artificial Intelligence in Healthcare." *International Journal of Diabetes and Technology* 4, no. 4 (2025): 127-139.
9. Marqas, Ridwan Boya, Saman M. Almufti, Renas R. Asaad, and T. S. Mohamed. "Advancing AI: A Comprehensive Study of Novel Machine Learning Architectures." *Int. J. Sci. World* 11 (2025): 48-85.
10. Priyadharshni, S., and V. Ravi. "Hybrid Quantum-Convolutional Neural Network With Spatial Attention for Accurate Classification of Maternal-Fetal Planes in Ultrasound Images." *IEEE Access* 13 (2025): 188306-188325.
11. Lopez, Daniel Alejandro, Oscar Montiel, Oscar Castillo, and Miguel Lopez-Montiel. "LiteSpineNet Integrates Quantum Computing for Spinal Fracture Diagnosis." (2026).
12. Halab, Khidhr, Nabil Marzoug, Othmane El Meslouhi, Zouhair Elamrani Abou Elassad, and Moulay A. Akhloufi. "QU-Net: Quantum-Enhanced U-Net for Self Supervised Embedding and Classification of Skin Cancer Images." *Big Data and Cognitive Computing* 10, no. 1 (2025): 12.
13. Mahargya, Ifran Lindu, Guruh Fajar Shidik, Pujiono Affandy, and Supriadi Rustad. "Intelligent Systems with Applications."
14. Pappala, Lalitha Kumari, Sai Babu Veesam, Kannaiah Chattu, Janapati Venkata Krishna, Jyostna Devi Bodapati, and B. Tarakeswara Rao. "Design of an iterative Model for incremental enhancements in quantum image processing using reinforcement learning-based optimizations." *IEEE Access* 13 (2025): 20491-20511.
15. Alam, Syeda Sadia, Kawsar Habib Munna, and Md Mushtaq Shahriyar Rafee. "Covid-19 detection from CT scan images using deep learning models." In 2024 International Conference on Intelligent Systems and Computer Vision (ISCV), pp. 1-6. IEEE, 2024.
16. Niloy, Ahashan Habib, SM Farah Al Fahim, Mohammad Zavid Parvez, Shammi Akhter Shiba, Faizun Nahar Faria, Md Jamilur Rahman, Emtiaz Hussain, and Tasmi Tamanna. "CovRoot: COVID-19 detection based on chest radiology imaging techniques using deep learning." Frontiers in Signal Processing 4 (2024): 1384744.
17. Raj, Aaron Meiyyappan Arul, Sugumar Rajendran, and Georgewilliam Sundaram Annie Grace Vimal. "Enhanced convolutional neural network enabled optimized diagnostic model for COVID-19 detection." Bulletin of Electrical Engineering and Informatics 13, no. 3 (2024): 1935-1942.
18. Islam, Nahid, Abu SM Mohsin, Shadab Hafiz Choudhury, Tazwar Prodhan Shaer, Md Adnan Islam, Omar Sadat, and Nahid Hossain Taz. "COVID-19 and Pneumonia detection and web deployment from CT scan and X-ray images using deep learning." PLoS One 19, no. 7 (2024): e0302413.
19. Mathesul, Shubham, Debabrata Swain, Santosh Kumar Satapathy, Ayush Rambhad, Biswaranjan Acharya, Vassilis C. Gerogiannis, and Andreas Kanavos. "COVID-19 detection from chest X-ray images based on deep learning techniques." Algorithms 16, no. 10 (2023): 494.
20. Redie, Dawit Kiros, Abdulhakim Edao Sirko, Tensaie Melkamu Demissie, Semagn Sisay Teferi, Vimal Kumar Shrivastava, Om Prakash Verma, and Tarun Kumar Sharma. "Diagnosis of COVID-19 using chest X-ray images based on modified DarkCovidNet model." Evolutionary Intelligence 16, no. 3 (2023): 729-738.
21. Chen, Tianyi, Ian Philippi, Quoc Bao Phan, Linh Nguyen, Ngoc Thang Bui, and Tuy Tan Nguyen. "A vision transformer machine learning model for COVID-19 diagnosis using chest X-ray images." Healthcare analytics 5 (2024): 100332.
22. Ding, Hui, Lingyan Fan, Jingfeng Zhang, and Guosheng Gao. "Deep learning-based system combining chest X-ray and computerized tomography images for COVID-19 diagnosis." British Journal of Hospital Medicine 85, no. 8 (2024): 1-15.
23. Marefat, Abdolreza, Mahdieh Marefat, Javad Hassannataj Joloudari, Mohammad Ali Nematollahi, and Reza Lashgari. "CCTCOVID: COVID-19 detection from chest X-ray images using Compact Convolutional Transformers." Frontiers in public health 11 (2023): 1025746.
24. Padmavathi, V., and Kavitha Ganesan. "Metaheuristic optimizers integrated with vision transformer model for severity detection and classification via multimodal COVID-19 images." Scientific Reports 15, no. 1 (2025): 13941.
25. Subbiyan, Balasubramani, Renjith Prabhavathi Neelakandan, Kavisankar Leelasankar, Rajkumar Rajavel, Muthukumaran Malarvel, and Achyut Shankar. "A quantum-enhanced artificial neural network model for efficient medical image compression." IEEE Access 13 (2025): 31809-31828.
26. Aarti, Devineni Gireesh Kumar, Ranjith Kumar Gatla, Swathi Gowroju, and Reem A. Almenweer. "Quantum-inspired optimization of transformer–capsule networks for accurate brain tumor segmentation and classification." Scientific Reports (2026).
27. Aravinda, C. V., Joseph Emerson Raja, and Sultan Alasmari. "Quantum-enhanced multimodal prognostic transformer for skin disease progression prediction and visualization." Scientific Reports (2026).

28. Jaincy DE, Martina, and Venkatasubbu Pattabiraman. "Quantum-SpinalNet: a hybrid deep learning approach for mammographic breast cancer detection." Frontiers in Digital Health 8 (2026): 1753820.
29. Ben Othman, Soufiane, and Obaid Ali. "A Neuro-Cognitive Quantum Architecture for Trustworthy AI in Ultra-Scarce Medical Environments." Cognitive Computation 18, no. 1 (2026): 63.
30. Narne, Naga Sai Ram, and Gangadhara Rao Kancharla. "A Quantum Edge Federated Graph Transformer for Generative and Causal Digital Twin Healthcare." Engineering, Technology & Applied Science Research 16, no. 2 (2026): 34482-34491.
31. Hussein, Ahmed A., Ahmed M. Montaser, and Hend A. Elsayed. "Skin cancer image classification using hybrid quantum deep learning model with BiLSTM and MobileNetV2." Quantum Machine Intelligence 7, no. 2 (2025): 66.
32. Ben Othman, Soufiane, and Obaid Ali. "Quantum-entangled neuro-symbolic swarm federation for privacy-preserving IoMT-driven multimodal healthcare: SB Othman, O. Ali." Scientific Reports (2025).
33. Chakraborty, Sanjay. "A study on quantum neural networks in healthcare 5.0." arXiv preprint arXiv:2412.06818 (2024).
34. Hussain, Sajid, Xu Songhua, Muhammad Aslam, Muhammad Waqas, and Shahid Hussain. "Quantum deep learning for automatic chronic kidney disease identification and classification with CT images." (2024).
35. https://www.kaggle.com/datasets/andyczhao/covidx-cxr2?
36. CH, Srilakshmi, and Balasubadra K. "Improved Weighted Quantum Firefly Optimization With Vanilla Vision Transformer and Big Data for Precision Diagnosis and Biomarker Identification in Neurodegenerative Disorders." Transactions on Emerging Telecommunications Technologies 36, no. 10 (2025): e70252.