ACCSCIENCE PUBLISHING

Artificial Intelligence in Health

ORIGINAL RESEARCH ARTICLE

Quantum-infused deep learning for dual-task thyroid ultrasound diagnostics

Smitirekha Behuria[1], Sujata Swain[1], Anjan Bandyopadhyay[1], Mingqiang Wang[2], and Saurav Mallik[3,4*]

[1]School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, Bhubaneswar, Odisha, India

[2]Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, California, United States of America

[3]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States of America

[4]Department of Pharmacology & Toxicology, University of Arizona, Tucson, Arizona, United States of America

***Corresponding author:**
Saurav Mallik
(smallik@arizona.edu)

Citation: Behuria S, Swain S, Bandyopadhyay A, Wang M, Mallik S. Quantum-infused deep learning for dual-task thyroid ultrasound diagnostics. *Artif Intell Health*. doi: 10.36922/AIH026050007

Received: January 30, 2026

Revised: March 9, 2026

Accepted: March 12, 2026

Published online: July 16, 2026

Abstract

Accurate classification of thyroid nodules is essential for early intervention and improved clinical outcomes; however, its diagnostic performance remains limited by inter-observer variability, limited data availability, and class imbalance in ultrasound imaging. Although existing deep learning models have demonstrated strong performance, they often struggle to capture both global semantic context and fine-grained structural details under limited data. This study introduces a hybrid quantum–classical deep learning framework that leverages quantum variational filtering in tandem with transfer learning via ViT-B16 and EfficientNet-B3 to enhance feature representation. A multi-stage preprocessing pipeline was designed to improve input diversity and class balance. The model jointly performs binary tumor classification and Thyroid Imaging Reporting and Data System (TI-RADS) level prediction, achieving 94.2% accuracy for benign-versus-malignant tumor classification and 91.8% macro-accuracy for TI-RADS prediction. The proposed approach demonstrates promising robustness and generalization in limited-data environments, highlighting the potential of integrating quantum encoding with classical deep learning architectures for medical image diagnostics.

Keywords: Thyroid nodule classification; TI-RADS prediction; Quantum variational filtering; Feature fusion; Patch-wise feature learning

1. Introduction

The accurate classification of thyroid nodules has emerged as a crucial component in contemporary medical diagnostics, owing to the thyroid gland's central role in regulating numerous physiological processes through the secretion of essential hormones. Located in the anterior neck, just below the larynx and anterior to the trachea, the thyroid gland significantly influences metabolic rate, thermoregulation, cardiovascular function, neural activity, and cognitive processes. It also contributes to the overall development and functional integrity of various organ systems.[1] Structural anomalies, such as nodular growths, can impair these regulatory functions, and the incidence of thyroid nodules continues to rise, particularly among women and older adults. While the majority of nodules are benign and asymptomatic, a subset poses a risk of malignancy, necessitating timely and accurate diagnostic interventions to inform clinical management and reduce the risk of progression.

To standardize and enhance the objectivity of thyroid nodule evaluation, the Thyroid Imaging Reporting and Data System (TI-RADS) was introduced. This framework provides a stratified risk assessment based on sonographic features, including nodule composition, echogenicity, margin definition, calcification patterns, and shape. By quantifying these visual markers, TI-RADS helps clinicians make evidence-based decisions regarding biopsy or monitoring. Despite its clinical value, the interpretation of ultrasound images remains somewhat subjective, with inter- and intra-observer variability that can affect diagnostic outcomes.[2,3]

In response to these challenges, a variety of computational methods have been explored. Classical machine learning algorithms—such as support vector machines, naive Bayes, logistic regression, and decision trees—have been applied to thyroid ultrasound image analysis. While these models offer simplicity and interpretability, they heavily rely on handcrafted features and often fail to capture the high-dimensional, non-linear patterns embedded in complex medical images. Their limited ability to generalize across imaging conditions and patient populations further restricts their clinical utility.[4]

The development of deep learning has greatly advanced medical image classification. Convolutional neural networks (CNNs), including prominent architectures such as VGG16, GoogLeNet, and ResNet, have demonstrated remarkable capabilities for learning hierarchical feature representations directly from raw image data, thereby improving classification accuracy.[5] These models have shown promise in various diagnostic tasks, including tumor detection, lesion segmentation, and anomaly classification. Nevertheless, they are not without limitations. Deep networks often require extensive training data to avoid overfitting and to generalize effectively. In medical imaging, where data can be scarce, imbalanced, or heterogeneous, this becomes a critical bottleneck. Moreover, most existing CNN architectures are computationally demanding, which can hinder their deployment in real-time clinical settings, particularly in resource-limited environments.

To address these concerns, recent studies have increasingly explored hybrid deep learning approaches that combine complementary model architectures, leveraging the strengths of multiple models to potentially achieve high effectiveness and performance. These methods aim to integrate the complementary benefits of different CNN architectures—such as the feature reuse of densely connected convolutional network (DenseNet) and the residual connections of ResNet—to capture both fine-grained local structures and global contextual information. When paired with robust preprocessing pipelines, data augmentation strategies, and techniques to address class imbalance (e.g., synthetic minority oversampling technique [SMOTE]), such hybrid models hold significant potential to deliver clinically reliable and generalizable diagnostic tools. In this study, we developed a hybrid quantum–classical deep learning framework for thyroid nodule classification and TI-RADS level prediction using ultrasound images. The proposed model combines handcrafted quantum-filtered features with semantic embeddings extracted from pretrained vision transformer–based, 16 × 16 patch size (ViT-B16) and efficient network–B3 variant (EfficientNet-B3) backbones. This approach addresses the challenges posed by high-dimensional ultrasound image data by extracting both handcrafted quantum-enhanced features and deep semantic representations through transfer learning. By integrating a quantum variational circuit[6,7] with pretrained ViT-B16 and EfficientNet-B3 models, and employing adaptive data augmentation along with SMOTE-based oversampling, the framework improves model generalization and classification accuracy—particularly in class-imbalanced and data-constrained scenarios.

1.1. Motivation

Limited data and complex patterns often challenge medical image analysis.

(i) Limited data and class imbalance: Thyroid ultrasound datasets often suffer from small sample sizes and uneven class distributions, making it difficult for traditional deep learning models to generalize well.
(ii) High-dimensional and noisy image data: Ultrasound images contain complex textures and noise artifacts, requiring advanced feature extraction techniques to capture meaningful diagnostic patterns.
(iii) Lack of interpretability and robustness: Existing methods often fail to integrate both local structural details and global semantic context, leading to reduced reliability in clinical decision support.
(iv) Emerging potential of quantum machine learning: Quantum computing offers new opportunities for encoding high-dimensional data with entangled representations, potentially enhancing classification performance in data-constrained environments.

1.2. Contributions

A hybrid quantum–classical framework enables robust, multi-task thyroid ultrasound classification.

(i) Hybrid quantum–classical framework: A novel architecture is proposed that integrates quantum variational filtering. It combines this with deep transfer-learning backbones (ViT- B16 and EfficientNet-B3) to ensure robust thyroid nodule and TI-RADS classification.
(ii) Multi-stage preprocessing pipeline: A comprehensive preprocessing pipeline was developed, including grayscale normalization, adaptive data augmentation, handcrafted multi-channel feature extraction, and SMOTE-based oversampling to enhance feature diversity and balance class distribution.
(iii) Quantum feature encoding: Patch-wise quantum filtering was implemented using 4-qubit variational circuits to capture localized structural information through entangled quantum representations.
(iv) Dual-task prediction: A unified model was developed to simultaneously perform binary tumor classification (benign versus malignant) and multi-class TI-RADS level prediction with improved generalization in data-limited settings.

2. Literature review

Thyroid tumor classification has been widely studied using various machine learning and deep learning approaches. Traditional methods relied on expert interpretation of fine needle aspiration cytology and ultrasound images, introducing subjectivity and variability in diagnosis. The advancement of artificial intelligence has enabled the automation of medical image analysis, significantly improving diagnostic accuracy and efficiency. In this section, we review existing deep learning models for thyroid tumor classification and explore emerging quantum computing approaches in medical imaging.

Accurate and early detection of thyroid nodules is clinically important and has attracted significant research interest. Various machine learning methods were initially investigated to solve this problem.[8] Previous studies on thyroid disease diagnosis can be categorized into machine learning, deep learning, and quantum machine learning approaches. In machine learning, Khan *et al.*[9] used a neural network with mutual information–based feature selection and SMOTE to handle class imbalance. Deep learning approaches have been widely explored, including CNN architectures by Srivastava *et al.*[10]; transformer- and attention-based models by Sun *et al.*[11], Wang *et al.*[12], and Zhang *et al.*[13]; and hybrid and advanced models proposed in various studies.[14–22] Reviews by Jiang *et al.*[23] and Shen *et al.*[24]summarize these developments. In addition, quantum machine learning approaches have emerged, where Maheshwari *et al.*[25] and Sajeev *et al.*[26] proposed hybrid quantum–classical models for classification tasks.

Table 1 summarizes key studies in thyroid nodule classification, highlighting the datasets used, methodological approaches, major findings, and limitations of existing methods. Several deep learning and hybrid models have been proposed for thyroid nodule classification. However, most approaches rely solely on classical architectures.

To further highlight the differences between existing approaches and the proposed hybrid framework, a comparative analysis is provided in Table 2. Classic models in medical imaging are often constrained by limited labeled data and the intensive computational cost of training. This is especially difficult in detecting subtle features, such as fine-texture changes in thyroid nodules, which are very important for proper classification. To solve such problems, our research proposes a quantum–classical hybrid framework that uses quantum filter transformations to obtain more sophisticated, fine-grained features. This novel method aims to overcome the intrinsic limitations of traditional methods and offers a promising direction for improving thyroid nodule characterization in ultrasound imaging.

Despite the progress in deep learning and quantum computing for medical imaging, several challenges remain. Conventional deep learning networks typically require substantial computational resources and a vast amount of labeled data to achieve optimal performance. On the other hand, fully quantum-based models are still in their early stages, with current quantum hardware being limited in scale and performance.

Hybrid quantum–classical models offer a practical solution by combining the advantages of quantum computing with the efficiency of classical deep learning frameworks. However, integrating quantum circuits into deep learning pipelines requires careful design and optimization to ensure compatibility and stability. This study addresses these challenges by developing a quantum-enhanced deep learning model that leverages quantum feature transformations to improve thyroid tumor classification.

3. Proposed methodology

3.1. Dataset

The dataset preparation process followed a structured pipeline to ensure data quality and fair evaluation. Initially, the Digital Database Thyroid Image (DDTI) dataset

Table 1. Summary of related work in thyroid nodule classification

Author	Year	Dataset used	Methodology	Findings	Limitations
Sun *et al.*[11]	2023	Thyroid ultrasound dataset	Vision transformer with contrastive learning for TI-RADS classification	Demonstrated strong feature learning capability using transformer architecture	Requires large training datasets
Aboudi *et al.*[15]	2023	447 ultrasound images	Bilinear convolutional neural network architecture	Achieved a higher classification accuracy than conventional CNN models	Limited dataset size and computational complexity
Maheshwari *et al.*[25]	2023	Healthcare datasets	Hybrid quantum machine learning models, including quantum SVM and MLP	Demonstrated potential benefits of quantum-enhanced learning	Limited by current quantum hardware constraints
Gomes *et al.*[27]	2020	Thyroid ultrasound dataset	Morphological and geometric features combined with traditional ML classifiers	Improved classification accuracy and enabled automated decision support	Reliance on handcrafted features and limited generalization
Liu *et al.*[28]	2017	Thyroid ultrasound images	Hybrid feature extraction using CNN features with SIFT and HOG descriptors	Demonstrated improved classification performance using hybrid representations	High computational complexity and small dataset size
Kwon *et al.*[29]	2020	Clinical thyroid ultrasound dataset	Deep CNN model based onVGG16 with data augmentation and cross-validation	Improved diagnostic consistency compared to manual evaluation	Requires large labeled datasets and significant computational resources
Zhu *et al.*[30]	2017	Thyroid ultrasound dataset	Region-of-interest extraction with transfer learning using ResNet	Improved classification accuracy using pretrained deep networks	Performance dependent on ROI extraction quality
Hang *et al.*[31]	2021	Thyroid ultrasound images	Fusion of conventional image features with deep features using Res-GAN	Enhanced feature representation and improved classification accuracy	GAN training in stability and high computational cost
Proposed work	2026	DDTI dataset (358 cleaned images)	Hybrid quantum–classical architecture integrating quantum filtering with ViT-B16 and EfficientNet-B3	Achieves improved tumor classification and TI-RADS prediction accuracy	Requires validation on larger multi-institutional datasets

Abbreviations: CNN: Convolutional neural network; DDTI: Digital Database Thyroid Image; EfficientNet-B3: Efficient network–B3 variant; GAN: Generative adversarial network; HOG: Histogram of oriented gradients; ML: Machine learning; MLP: Multilayer perceptron; Res-GAN: Residual generative adversarial network; ResNet: Residual network; ROI: Region of interest; SIFT: Scale-invariant feature transform; SVM: Support vector machine; TI-RADS: Thyroid Imaging Reporting and Data System; ViT-B16: Vision transformer–based, 16 × 16 patch size; VGG16: Visual geometry group 16-layer network.

contained 480 ultrasound images.[32] During the cleaning stage, 122 images were excluded due to incomplete or missing Extensible Markup Language (XML) metadata, including missing TI-RADS annotations or inconsistent labeling fields. After this cleaning step, 358 images remained and were used for subsequent analysis. The cleaned dataset was then divided into training (70%), validation (15%), and test (15%) sets using stratified sampling to preserve the class distribution across splits. To address class imbalance, data augmentation and SMOTE-based oversampling were applied exclusively to the training set during preprocessing. The validation and test sets were kept unchanged to ensure unbiased model evaluation. Table 3 summarizes the dataset preparation process and the distribution of images across the training, validation, and test sets.

These 358 thyroid ultrasound images were categorized into benign and malignant classes as shown in Figure 1.

Moreover, all 358 thyroid ultrasound images in the dataset have been independently assessed by specialists in medical imaging using the standardized TI-RADS. TI-RADS is a point-based stratification system that evaluates five key sonographic characteristics of thyroid nodules: composition, echogenicity, shape, margin, and echogenic foci. Suspicious features are assigned higher scores, and the cumulative sum determines the final

Table 2. Comparison of the proposed approach with existing methods for thyroid ultrasound analysis

Study	Year	Model type	Dataset	Handcrafted features	Transfer learning	Augmentation	Feature fusion	Quantum	Task
Gomes *et al.*[27]	2020	ML classifiers	Ultrasound	✓	X	X	X	X	Nodule classification
Liu *et al.*[28]	2017	CNN + SIFT/HOG	Ultrasound	✓	X	✓	✓	X	Nodule classification
Kwon *et al.*[29] _bookmark51	2020	VGG16 CNN	Ultrasound	X	✓	✓	X	X	Malignancy detection
Zhu *et al.*[30]	2017	ResNet transfer learning	Ultrasound	X	✓	✓	X	X	Nodule classification
Hang *et al.*[31]	2021	Res-GAN + CNN	Ultrasound	✓	✓	✓	✓	X	Nodule classification
Aboudi *et al.*[15]	2023	Bilinear CNN	Ultrasound	X	✓	✓	X	X	Nodule classification
Sun *et al.*[11]	2023	Vision transformer	Ultrasound	X	✓	✓	X	X	TI-RADS prediction
Maheshwari *et al.*[25]	2023	Quantum ML models	Healthcare datasets	X	X	X	X	✓	Medical classification
Proposed work	2026	Hybrid quantum deep model	DDTI	✓	✓	✓	✓	✓	Tumor classification and TI-RADS prediction

Notes: ✓: present; X: absent.
Abbreviations: CNN: Convolutional neural network; DDTI: Digital Database Thyroid Image; HOG: Histogram of oriented gradients; ML: Machine learning; Res-GAN: Residual generative adversarial network; ResNet: Residual network; SIFT: Scale-invariant feature transform; TI-RADS: Thyroid Imaging Reporting and Data System; VGG16: Visual geometry group 16-layer network.

TI-RADS classification. The scoring spectrum ranges from TI-RADS 2 (benign) to TI-RADS 5 (high suspicion of malignancy). As illustrated in Figure 2, these categories help to assess and manage clinical risk. The distribution of the statistical dataset is illustrated in Figure 3.

3.2. Preprocessing

To ensure compatibility with the proposed quantum-enhanced hybrid architecture and to improve the model's robustness and generalization capacity, a structured multi-stage preprocessing framework was implemented. This pipeline included data cleaning, grayscale normalization, data augmentation, handcrafted multi-channel feature extraction, quantum filtering, transfer learning preparation, and synthetic oversampling using SMOTE.

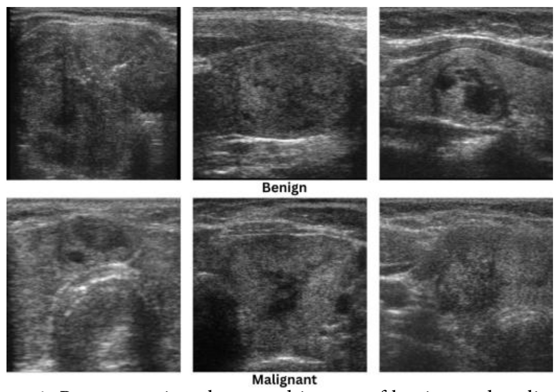

Figure 1. Representative ultrasound images of benign and malignant thyroid nodules

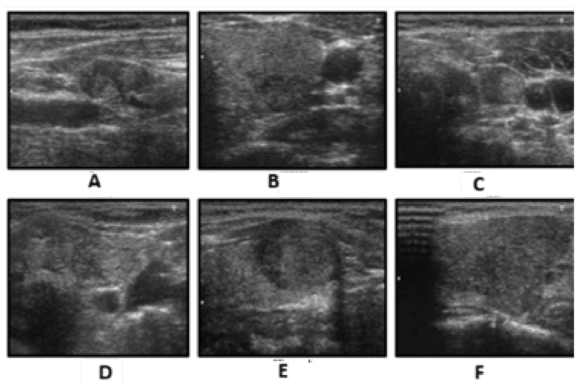

Figure 2. Representative thyroid ultrasound images across the Thyroid Imaging Reporting and Data System (TI-RADS) categories. (A) TI-RADS 2: Benign nodules with no suspicious features. (B) TI-RADS 3: Probably benign nodules with low malignancy risk. (C) TI-RADS 4a: Low suspicion for malignancy. (D) TI-RADS 4b: Intermediate suspicion for malignancy. (E) TI-RADS 4c: Moderate-to-high suspicion of malignancy. (F) TI-RADS 5: Highly suspicious nodules with a high risk of malignancy.

Table 3. Dataset preparation summary

Stage	Number of images	Description
Original DDTI dataset	480	Ultrasound images with corresponding XML metadata collected from 399 patients
Removed during cleaning	122	Images excluded due to missing TI-RADS labels or incomplete XML annotations (composition, echogenicity, margin, or classification fields)
Final usable dataset	358	Images retained after the data cleaning process
Training set (70%)	249	Used for model training, augmentation, and SMOTE-based oversampling
Validation set (15%)	49	Used for hyperparameter tuning and model selection
Test set (15%)	60	Used for final performance evaluation

Abbreviations: DDTI: Digital Database Thyroid Image; SMOTE: Synthetic minority over-sampling technique; TI-RADS: Thyroid Imaging Reporting and Data System; XML: Extensible Markup Language.

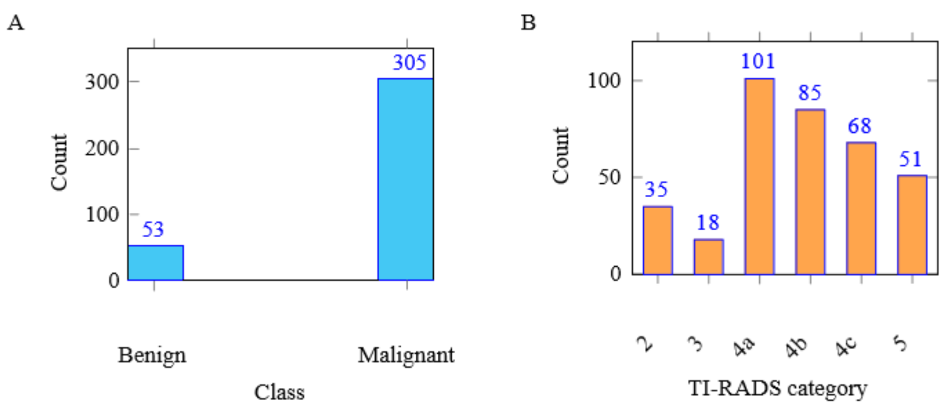

Figure 3. Class distribution in the dataset: (A) Benign versus malignant based on expert-confirmed tumor types, and (B) Thyroid Imaging Reporting and Data System (TI-RADS) score breakdown from radiologist annotations

3.2.1. Data cleaning

The original ultrasound dataset included a few image cases that were missing some crucial XML metadata, such as TI-RADS scores, nodule composition, echogenicity, margin characteristics, and calcification information. All such missing cases, along with their corresponding images, were excluded from the dataset to ensure completeness and reliability. The cleaned dataset was divided into training (70%), validation (15%), and test (15%) sets using stratified sampling based on tumor labels. Class balancing was achieved only in the training dataset through augmentation and SMOTE, while the validation and test sets retained the original distribution to avoid evaluation bias.

3.2.2. Normalization

The ultrasound images in the training dataset were also resized to a consistent resolution of 128 × 128 pixels to preserve spatial uniformity throughout the processing pipeline. Grayscale normalization was achieved using *z*-score standardization, in which each image was normalized by subtracting its mean intensity and dividing by its standard deviation.[33] Subsequently, min–max normalization scaled the pixel values to the (0, 1) range to improve numerical stability and convergence during training of the quantum and deep learning models.

3.2.3. Data augmentation

To improve generalization and prevent overfitting due to the small dataset size, a set of augmentation procedures was applied only to the training data. Augmentations included: (i) random rotation, (ii) horizontal flipping, (iii) brightness adjustment in the range (0.8, 1.2), (iv) elastic deformation, (v) additive Gaussian noise, and (vi) random zoom.

These augmentations were applied adaptively, with greater intensity for underrepresented classes (e.g., TI-RADS 2 and 3), to improve class balance and increase the effective training sample size.

3.2.4. Four-channel feature extraction

Each augmented grayscale image was transformed into a 4-channel tensor to enrich structural and textural features:

(i) Channel 1: Z-score normalized intensity map[34]
(ii) Channel 2: Vertical edge map (Sobel-X)
(iii) Channel 3: Horizontal edge map (Sobel-Y)
(iv) Channel 4: Texture descriptor using local binary patterns (LBP)

All channels were independently normalized to the (0, 1) range and smoothed with a Gaussian filter (σ = 0.5) to reduce noise artifacts before further processing.

3.2.5. Quantum filtering

To capture localized structural patterns from the ultrasound images, a quantum feature extraction module was introduced. This module operates on the four-channel feature representation produced during preprocessing, transforming spatial image patches into quantum feature embeddings using a parameterized quantum circuit.

(a) Patch formation across channels: After preprocessing, each ultrasound image is represented as a four-channel tensor. The input ultrasound image is converted into a multi-channel representation consisting of grayscale intensity, Sobel-X, Sobel-Y, and LBP features. The resulting tensor representation is expressed as **Equation 1**:

$$X \in R^{128\times128\times4} \tag{1}$$

where the four channels correspond to the extracted feature maps used for subsequent processing.

Spatial patching is performed identically across all channels. The image is divided into non- overlapping 2 × 2 spatial patches. For an image resolution of 128 × 128, the number of spatial patches is calculated as shown in **Equation 2**:

$$\frac{128}{2}\times\frac{128}{2}=4096 \tag{2}$$

For each spatial patch location, the corresponding pixel values from the four channels are collected and concatenated to form a four-dimensional feature vector:

$$x=[x_1, x_2, x_3, x_4] \tag{3}$$

where x_1, x_2, x_3, and x_4 represent features derived from the intensity, Sobel-X, Sobel-Y, and LBP channels, respectively. Thus, the four channels are concatenated, and each spatial patch is represented by a single four-dimensional vector as defined in **Equation 3**:

(b) Quantum encoding: The classical input features are encoded into the quantum circuit using parameterized rotation gates, as defined in **Equation 4**:

$$U_{encode}\left(x\right)=\frac{4}{i=1}R_Y\left(x_i\right)R_Z\left(x_i\right) \tag{4}$$

where $R_Y(\cdot)$ and $R_Z(\cdot)$ denote single-qubit rotation gates applied to the i-th qubit. This operation maps classical patch features into a quantum-state representation.

(c) Variational quantum layer: A variational layer with trainable parameters is then applied to the encoded quantum state, as defined in **Equation 5**:

$$U_{var}\left(\theta\right)=\frac{4}{i=1}R_Y\left(\theta_i\right)R_Z\left(\theta_i\right) \tag{5}$$

where $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ represents trainable parameters optimized during the hybrid quantum–classical training process.

(d) Entanglement operation: Interactions between encoded features are modeled using entanglement gates. An entanglement layer is then applied to introduce correlations between qubits using controlled-NOT gates, as defined in **Equation 6**:

$$U_{\mathrm{entangle}}=CNOT_{1,2}\cdot CNOT_{2,3}\cdot CNOT_{3,4} \tag{6}$$

The final quantum state after the encoding, variational, and entanglement operations is expressed as shown in **Equation 7**:

$$|\psi_{\mathrm{out}}\rangle=U_{\mathrm{entangle}}\cdot U_{\mathrm{var}}(\theta)\cdot U_{\mathrm{encode}}(x)|0000\rangle \tag{7}$$

The circuit consists of two rotation layers and one entanglement layer, resulting in a circuit depth of three.

(e) Measurement and feature extraction: The quantum feature values are obtained by measuring the expectation values of the Pauli-Z operators on each qubit, as defined in **Equation 8**:

$$f_i = \langle \psi_{out} | Z_i | \psi_{out} \rangle, \quad i = 1, 2, 3, 4 \tag{8}$$

This measurement produces a four-dimensional quantum feature vector for each spatial patch. Given 4,096 spatial patches, the quantum module generates 4,096 patch-level feature vectors.

(f) Feature aggregation: To obtain a compact representation of the entire image, patch-level features are aggregated via global average pooling across all patches, yielding the final quantum feature vector. The quantum circuit produces a feature vector q with dimensionality d_q, as defined in **Equation 9**:

$$q \in \mathrm{R}^{d_q}, \quad d_q = 4 \tag{9}$$

This aggregated quantum representation is subsequently fused with semantic features extracted from the pretrained ViT-B16 and EfficientNet-B3 networks in the feature fusion stage.

(g) Implementation details: The quantum circuits were implemented using the Qiskit framework and executed on the Qiskit Aer statevector simulator. Each circuit evaluation used 1,024 measurement shots to estimate expectation values. No hardware noise model was applied. During training, the variational parameters were optimized jointly with the classical deep learning components using the Adam optimizer, and gradients were computed using the parameter-shift rule.

3.2.6. Transfer learning

To complement the quantum-filtered handcrafted features, the same images were resized to 224 × 224 and passed through pretrained ViT-B16 and EfficientNet-B3 backbones. These models, pretrained on ImageNet, were fine-tuned end-to-end during training. The high-level semantic features extracted from both architectures were later fused with the handcrafted quantum features, thereby enabling encoding of both global context and local structure.

3.2.7. Synthetic oversampling and data augmentation

To address class imbalance in the dataset, SMOTE[35] was employed together with data augmentation. Both strategies were applied carefully to prevent information leakage during model evaluation.

(a) Data augmentation: Data augmentation was applied to the training images before feature extraction to increase the training data diversity and reduce overfitting. The augmentation operations included random rotation, horizontal flipping, brightness adjustment within the range (0.8, 1.2), elastic deformation, Gaussian noise addition, and random zoom. These transformations were applied only to the training images and were not used for validation or testing samples.

(b) Feature representation for SMOTE: After preprocessing and augmentation, each image was transformed into a four-channel representation consisting of the normalized grayscale image, Sobel-X edge map, Sobel-Y edge map, and LBP texture map. These four channels form a tensor representation that captures complementary structural and textural characteristics of thyroid nodules.

For oversampling, the four-channel tensors were flattened into classical feature vectors. SMOTE was then applied to these feature vectors to generate synthetic samples for minority classes. Importantly, SMOTE was applied only to this classical feature representation and not to quantum feature vectors, deep embeddings extracted by ViT-B16 or EfficientNet-B3, or the fused feature representation.

(c) Cross-validation protocol: Model evaluation was performed using five-fold cross-validation along with a hold-out test set. To ensure fair evaluation and avoid data leakage, both data augmentation and SMOTE were applied strictly within the training portion of each fold.

For each fold, the following procedure was followed:

(i) The dataset was first split at the patient level into training, validation, and test sets.
(ii) Within each cross-validation fold, the training subset was further divided into fold-specific training and validation partitions.
(iii) Data augmentation was applied only to the images in the fold-specific training set.
(iv) Four-channel feature extraction was performed on the augmented training images.
(v) SMOTE was applied to the resulting training feature vectors to balance minority classes
(vi) The balanced training data were then used for quantum feature extraction, deep feature extraction (ViT-B16 and EfficientNet-B3), feature fusion, and model training.

Validation and test samples remained completely untouched by augmentation or SMOTE. This protocol ensures that synthetic samples are generated only from training data and that no information from validation or test sets is introduced into the training process. Table 4 refers to the class distribution: benign and malignant. Table 5 depicts the distribution of the class based on TI-RADS.

Table 4. Class distribution

Class	Training data	Training data after preprocessing	Testing data
Benign	39	210	9
Malignant	210	525	51
Total	249	735	60

The increase in the number of training samples after preprocessing is due to data augmentation and SMOTE-based oversampling, which were applied only to the training dataset. Data augmentation generated additional variations of the original ultrasound images, while SMOTE created synthetic samples for underrepresented classes. These techniques were used exclusively on the training set to improve class balance and model robustness, while the validation and test sets retained their original data distributions. Table 6 summarizes the tumor classification performance of the proposed hybrid model and baseline models. The values reported in this table correspond to the results obtained on the independent hold-out test set. In addition, five-fold cross-validation results are reported as mean ± standard deviation to demonstrate the robustness and stability of the proposed model.

Table 5. Class distribution based on TI-RADS level after preprocessing

TI-RADS level	Training data	Training data after augmentation	Testing data
2	25	135	6
3	14	75	3
4a	67	168	18
4b	59	147	13
4c	50	125	10
5	34	85	10
Total	249	735	60

Abbreviation: TI-RADS: Thyroid Imaging Reporting and Data System.

3.3. Model architecture

The proposed model adopts a hybrid quantum–classical deep learning architecture that integrates domain-specific handcrafted features filtered through a variational quantum circuit with deep semantic representations derived from two state-of-the-art pretrained backbones—ViT-B16 and EfficientNet-B3. This fusion aims to harness the complementary strengths of quantum feature encoding for local structure and classical deep learning for global semantic understanding. Figure 4 depicts the flowchart of the proposed hybrid architecture, and Figure 5 depicts the diagrammatic representation of the hybrid quantum–classical model architecture.

(a) Overview of the architecture: The architecture comprises three parallel branches:

  (i) Quantum feature extraction branch: Operates on the 4-channel handcrafted image tensors (normalized intensity, Sobel-X, Sobel-Y, and LBP). Each image is divided into non- overlapping 2 × 2 patches per channel. Each 4-dimensional patch vector is encoded into a 4-qubit quantum state using parameterized single-qubit rotations followed by entangling CNOT gates. Post-measurement, the expectation values of Pauli-Z operators are aggregated to generate a compact, discriminative quantum-enhanced feature vector.[36,37]

  (ii) Transfer learning feature extraction branch: The same image is resized to 224 × 224 and passed through two pretrained models:

  - ViT-B16: Captures global dependencies via transformer-based attention mechanisms.[38,39]
  - EfficientNet-B3: Extracts localized features with compound scaling.[40,41]

Both models are fine-tuned end-to-end, and their penultimate layer outputs are concatenated to form a joint semantic feature representation.

  (iii) Fusion and classification branch: The outputs from the quantum and deep feature extractors are concatenated to form a comprehensive multimodal representation. This is passed through:

  - A fully connected dense layer with rectified linear unit (ReLU) activation.
  - A dropout layer (rate = 0.3) to prevent overfitting.
  - Two parallel softmax classification heads for tumor classification (benign versus malignant) and TI-RADS level prediction (classes 2, 3, 4a, 4b, 4c, 5).

(b) Mathematical representation. Let:

$q \in R^{d_q}$ denotes the quantum-filtered feature vector,

$v \in R^{d_v}$ and $e \in R^{d_e}$ be the semantic features from ViT-B16 and EfficientNet-B3, respectively.

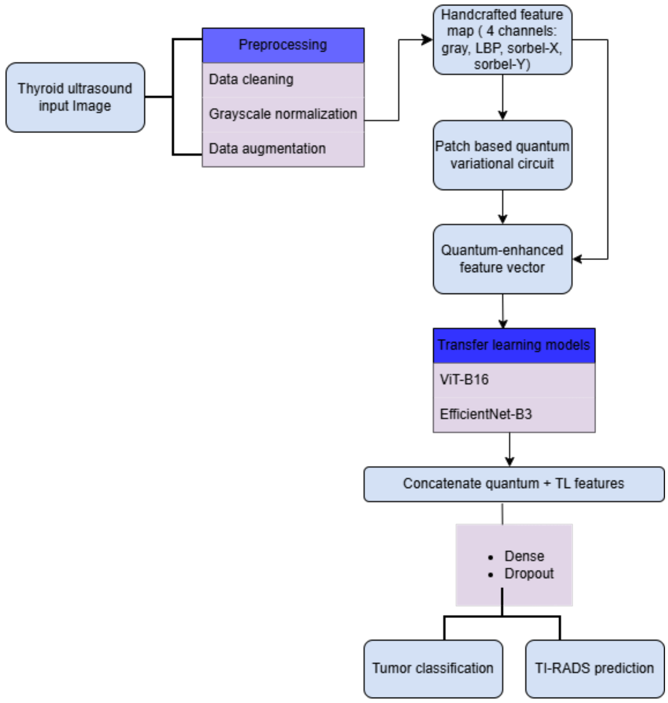

Figure 4. Flowchart of the proposed hybrid architecture: Integration of quantum-filtered handcrafted features and deep semantic representations from ViT-B16 and EfficientNet-B3 for dual-task classification

Abbreviations: EfficientNet-B3: Efficient network–B3 variant; LBP: Local binary pattern; TI-RADS: Thyroid Imaging Reporting and Data System; TL: Transfer learning; ViT-B16: Vision transformer–based, 16 × 16 patch size.

The fused feature vector is constructed by concatenating the quantum feature vector q, the ViT-B16 embedding v, and the EfficientNet-B3 feature vector e, as shown in **Equation 10**:

$$z = [q \,||\, v \,||\, e] \tag{10}$$

where || represents the concatenation operation.

The classification head processes the fused feature vector using fully connected layers with ReLU activation and dropout regularization, producing predictions for tumor classification and TI-RADS categories as shown in **Equations 11–14**:

$$h = \mathrm{ReLU}\,(W_1 z + b_1) \tag{11}$$

$$h_{drop} = \mathrm{Dropout}\,(h, p = 0.3) \tag{12}$$

$$\hat{y}_{tumor} = \mathrm{Softmax}\,(W_2 h_{drop} + b_2) \tag{13}$$

$$\hat{y}_{TI\text{-}RADS} = \mathrm{Softmax}\,(W_3 h_{drop} + b_3) \tag{14}$$

(c) Loss function: The model is trained using a multi-task learning objective that combines tumor classification loss and TI-RADS prediction loss, as defined in **Equation 15**:

$$L_{total} = \lambda_1 \cdot L_{tumor} + \lambda_2 \cdot L_{TI\text{-}RADS} \tag{15}$$

where L_{tumor} and $L_{TI\text{-}RADS}$ are the categorical cross-entropy losses, and λ_1 and λ_2 are task weighting coefficients, empirically set as $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$.

This formulation ensures balanced learning between the tumor classification and TI-RADS level prediction tasks.

3.4. Training procedure

The model was trained on a composite input comprising handcrafted quantum-enhanced features and pretrained backbone outputs. Categorical cross-entropy was used as the loss function for both binary (tumor classification) and multi-class (TI-RADS) tasks. The optimizer used was Adam with an initial learning rate of $1e^{-4}$ and a cosine learning rate scheduler with warm restarts. Training was conducted for 100 epochs with early stopping (patience = 15). Dropout regularization (rate = 0.3) was applied before the classification head to prevent overfitting. Evaluation was performed using 5-fold cross-validation and a hold-out test set to ensure reliable generalization. The classification tasks are optimized using the categorical cross-entropy loss defined in **Equation 16**:

$$L_{CE} = -\sum_{i=1}^{C} y_i \log\left(\widehat{y_i}\right) \tag{16}$$

where y_i is the ground truth and $y\hat{}_i$ is the predicted probability for class i.

Algorithm 1 presents a hybrid quantum–classical learning approach for thyroid nodule and TI-RADS classification. It integrates handcrafted and deep features with quantum feature extraction from ultrasound images. Features from ViT-B16, EfficientNet-B3, and quantum circuits are fused for dual classification. The model is trained using cross-entropy loss and optimized with the Adam optimizer.

3.5. Evaluation metrics

To evaluate the performance of the proposed hybrid model on both tumor classification and TI-RADS risk, we measured multiple classification metrics commonly used in medical imaging and diagnostic systems. These include accuracy, precision, recall (sensitivity), specificity, F1 score, and area under the receiver operating characteristic curve (AUC–ROC). Together, these metrics provide a balanced evaluation across both correct positive predictions and the avoidance of false alarms, which is essential in clinical applications like thyroid nodule classification.

All metrics were computed separately for both binary (benign versus malignant) and multi-class (TI-RADS 2–5) classification tasks. Macro-averaging was used in the multi-class setting to treat each class equally, regardless of class imbalance. Performance scores were averaged over five-fold cross-validation and additionally reported on an independent test set to ensure generalizability.

Algorithm 1. Quantum–classical hybrid learning for thyroid classification

Require: Ultrasound images I, TI-RADS labels $Y_{TI\text{-}RADS}$, tumor labels Y_{tumor}

Ensure: Predicted classes $\hat{Y}_{TI\text{-}RADS}$, $\hat{Y}_{tumor}$

For each image, I_i do
Normalize image using Z-score normalization
Apply data augmentation
Extract Sobel-X, Sobel-Y and LBP features
Construct a four-channel tensor
End for
For each 2 × 2 patch do
Encode features into a quantum circuit
Apply variational rotations
Apply CNOT entanglement
Measure expectation values
End for
Fuse quantum features with ViT-B16 and EfficientNet-B3 embeddings
Train classifier using cross-entropy loss

For the multi-class TI-RADS classification task, performance is also evaluated using macro-accuracy. In this study, macro-accuracy is defined as the mean per-class accuracy across all TI-RADS categories. Let C denote the total number of classes and Acc_i represent the accuracy of class i. The macro-accuracy metric, which computes

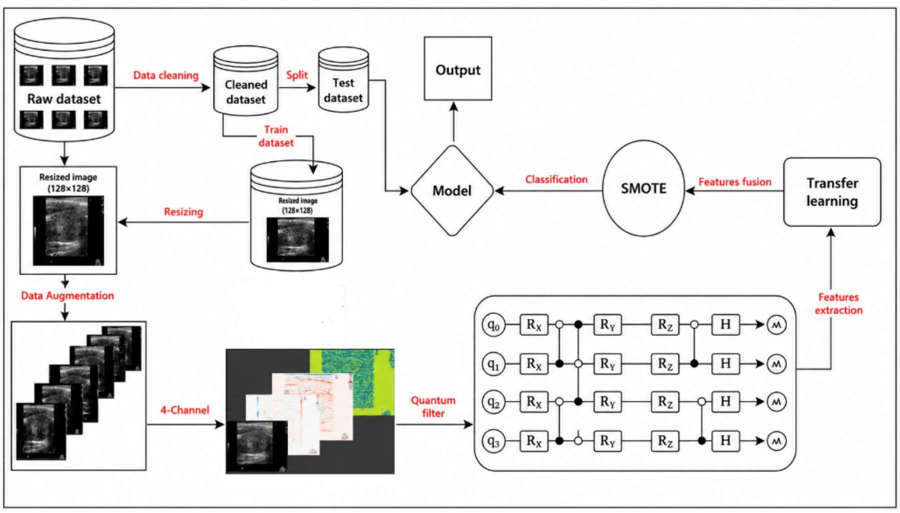

Figure 5. Proposed hybrid quantum–classical model architecture
Abbreviation: SMOTE: Synthetic minority over-sampling technique.

the average of per-class accuracies, is defined as shown in **Equation 17**:

$$MacroAccuracy = (1/\mathrm{C})\sum_{i=1}^{C} Acc_i \quad (17)$$

This metric assigns equal importance to each class, thereby providing a balanced evaluation in the presence of class imbalance.

(i) Accuracy measures the proportion of correctly predicted instances among the total number of samples, as defined in **Equation 18**:

$$\mathrm{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (18)$$

Although accuracy provides an overall indication of performance, it does not fully capture the model's behavior under class imbalance, which is common in medical datasets.

(ii) Precision quantifies how many of the predicted positive cases are actually positive, as defined in **Equation 19**:

$$\mathrm{Precision} = \frac{TP}{TP+FP} \quad (19)$$

This is critical in reducing false alarms, which can otherwise lead to unnecessary medical interventions.

(iii) Recall: In the context of thyroid nodules, high recall helps ensure malignant or suspicious nodules are not missed during diagnosis, as defined in **Equation 20**:

$$\mathrm{Recall} = \frac{TP}{TP+FN} \quad (20)$$

(iv) Specificity: Measures the ability of the model to correctly reject benign (negative) cases, as defined in **Equation 21**:_bookmark29

$$\mathrm{Specificity} = \frac{TN}{TN+FP} \quad (21)$$

In a medical context, this reduces the likelihood of false positives, which could otherwise lead to unwarranted follow-up testing or biopsies.

(v) F1-Score: Provides a balanced trade-off between precision and recall by computing their harmonic mean, which is defined as shown in **Equation 22**:

$$F1 = \frac{2 \cdot \mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \quad (22)$$

It is especially useful when dealing with class imbalance or when false negatives and false positives carry different costs.

(vi) AUC–ROC: Captures the model's ability to discriminate between classes across different thresholds. It is computed as the area under the ROC curve, which plots true positive rate against false positive rate, and is defined as shown in **Equation 23**:

$$AUC = \int_0^1 TPR(x)dx \quad (23)$$

A higher AUC indicates stronger discrimination between positive and negative classes. In clinical settings, it serves as a summary measure for diagnostic accuracy.

4. Results

The proposed model was evaluated using two complementary evaluation strategies: an independent hold-out test set and five-fold cross-validation. The hold-out test set was used to report the final unbiased performance of the trained model. In addition, five-fold cross-validation was performed on the training dataset to evaluate the stability and robustness of the proposed approach. For cross-validation experiments, performance metrics are reported as the mean and standard deviation across the five folds (mean ± standard deviation). Throughout the results section, hold-out and cross-validation results are reported separately to ensure a clear and consistent evaluation. The proposed hybrid model outperformed the baseline models on both tumor classification and TI-RADS multi-class prediction tasks. Figure 6 presents a performance comparison of the proposed model for tumor classification and TI-RADS multi-class prediction, using accuracy, precision, recall, F1-Score, and AUC–ROC.

The hybrid quantum–classical model achieved a tumor classification accuracy of 94.2% and a TI-RADS macro-accuracy of 91.8%. Precision, recall, and F1-score values were similarly high across both tasks, indicating robust generalization and sensitivity to clinically relevant patterns.

To further validate model reliability, we analyzed confusion matrices and ROC curves for each classification task. These visualizations confirmed consistent performance across all classes, including lower-prevalence categories such as TI-RADS 2 and 3. Figure 7 presents the confusion matrix for the tumor classification task (benign

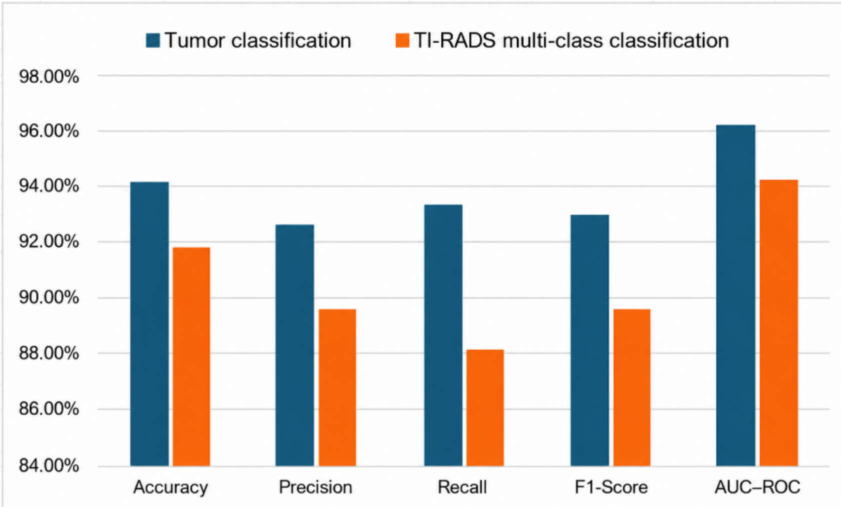

Figure 6. Performance comparison of the proposed model across tumor classification and TI-RADS multi-class classification based on accuracy, precision, recall, F1-Score, and AUC–ROC
Abbreviations: AUC–ROC: Area under the receiver operating characteristic curve; TI-RADS: Thyroid Imaging Reporting and Data System.

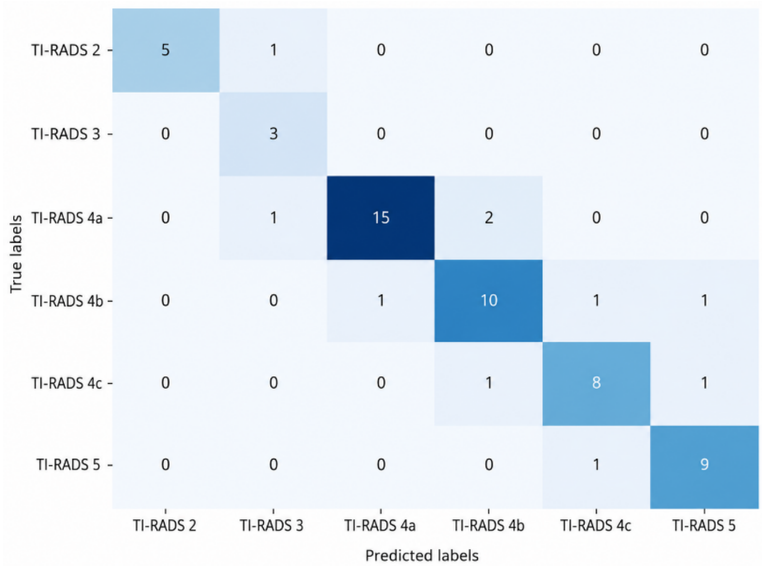

Figure 7. Confusion matrix for the tumor classification task (benign versus malignant) evaluated on the independent hold-out test set

versus malignant) obtained on the independent hold-out test set. The matrix illustrates the distribution of correctly and incorrectly classified samples and provides insight into the diagnostic performance of the proposed model. Figures 8 and 9 present the tumor model's performance over 100 epochs, showing training and validation accuracy (Figures 8) and training and validation loss (Figure 9). Figures 10 and 11 illustrate the TI-RADS model, depicting training and validation accuracy (Figures 10) and loss (Figure 11) over 100 epochs. Table 6 summarizes the performance comparison between the baseline and proposed hybrid models, while Figure 12 provides a graphical representation of the results in Table 6.

In addition to overall accuracy, clinically relevant evaluation metrics such as sensitivity, specificity, and AUC–ROC were analyzed for the tumor classification task. Sensitivity measures the model's ability to correctly detect malignant nodules, while specificity measures its

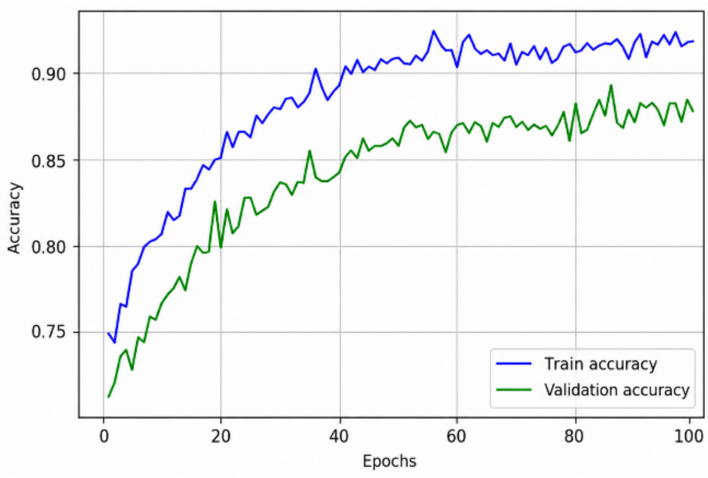

Figure 8. Tumor training versus validation accuracy over 100 epochs

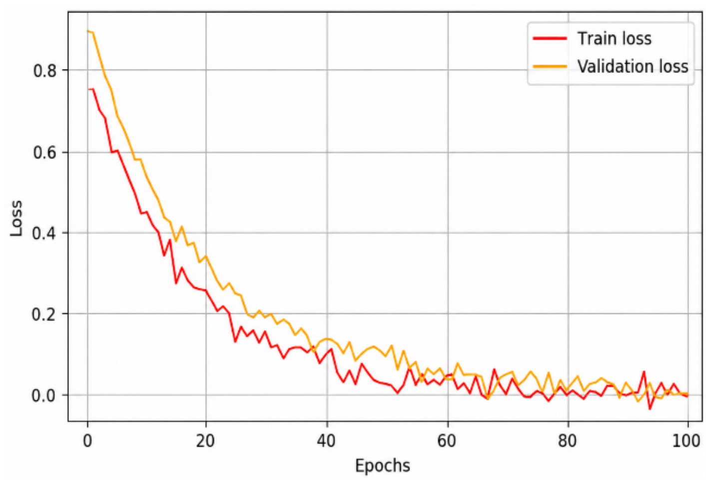

Figure 9. Tumor training versus validation loss over 100 epochs

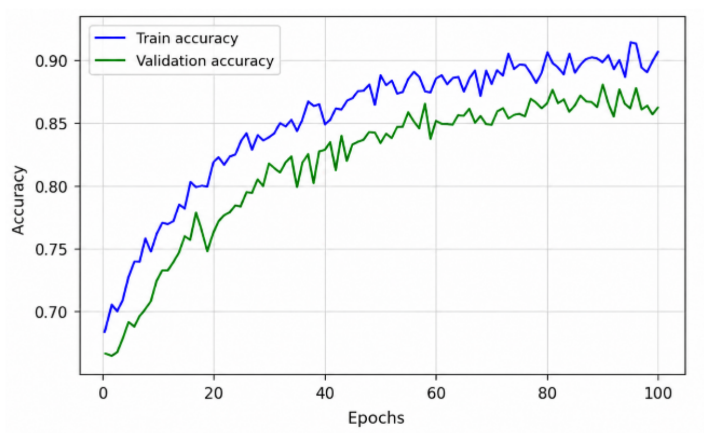

Figure 10. Thyroid Imaging Reporting and Data System (TI-RADS) training versus validation accuracy over 100 epochs

Table 6. Performance comparison of baseline and proposed hybrid models for tumor classification

Model	Accuracy	Precision	Recall	F1-score
DenseNet121	0.8673	0.8642	0.8673	0.8600
ResNet50	0.8571	0.8236	0.8571	0.8367
InceptionV3	0.7143	0.6802	0.7143	0.6950
VGG16	0.6327	0.6081	0.6327	0.6167
Proposed hybrid model	0.9420	0.9498	0.9427	0.9369

Abbreviations: DenseNet121: Dense convolutional network with 121 layers; InceptionV3: Inception network, version 3; ResNet50: Residual network with 50 layers; VGG16: Visual geometry group 16-layer network.

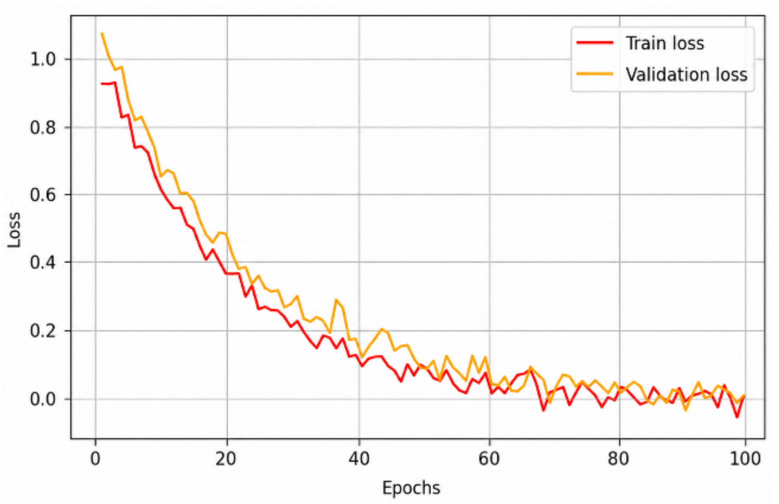

Figure 11. TI-RADS training versus validation loss over 100 epochs

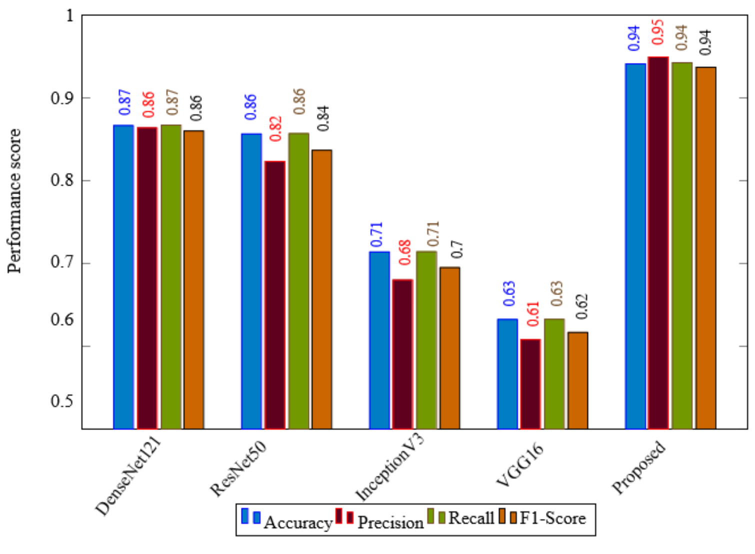

Figure 12. Graphical representation of tumor classification performance metrics across models
Abbreviations: DenseNet121: Dense convolutional network with 121 layers; InceptionV3: Inception network, version 3; ResNet50: Residual network with 50 layers; VGG16: Visual geometry group 16-layer network.

ability to correctly identify benign nodules. These metrics are particularly important in medical diagnosis since false negatives (missed malignant nodules) can have serious clinical consequences. The confusion matrix and ROC curve further illustrate the diagnostic capability of the proposed model.

4.1. Ablation study

To evaluate the contribution of each component in the proposed hybrid quantum–classical framework, an ablation study was conducted. The objective was to analyze how different architectural modules influence the final classification performance. Four model configurations were evaluated:

(i) EfficientNet-B3 only: Only deep convolutional features extracted using EfficientNet-B3.
(ii) ViT-B16 only: Only transformer-based semantic features extracted using the vision transformer (ViT-B16).
(iii) Classical fusion (ViT + EfficientNet): Fusion of EfficientNet-B3 and ViT-B16 features without the quantum filtering module.
(iv) Proposed hybrid model: Full model integrating quantum variational filtering with ViT-B16 and EfficientNet-B3 feature fusion.

Table 7 summarizes the performance of each configuration. The results demonstrate that combining transformer and CNN features improves classification performance compared to individual architectures. Furthermore, integrating the quantum filtering module further improves feature representation and captures complex structural patterns in ultrasound images.

The ablation results indicate that each component contributes to the overall performance improvement. In particular, the quantum filtering module enhances feature diversity and improves classification accuracy when combined with deep learning feature extractors. These findings are further analyzed and interpreted in the following Discussion section.

Table 7. Ablation study of the proposed architecture

Model configuration	Tumor accuracy (%)	TI-RADS macro accuracy (%)
EfficientNet-B3 only	89.4	86.7
ViT-B16 only	90.8	87.9
ViT-B16 + EfficientNet-B3	92.6	89.5
Proposed hybrid model	94.2	91.8

5. Discussion

The experimental results demonstrate that the proposed hybrid quantum–classical framework improves thyroid ultrasound classification compared with conventional deep learning approaches. By combining quantum variational filtering with deep semantic features extracted from ViT-B16 and EfficientNet-B3, the model captures both localized structural patterns and global contextual information in ultrasound images. This complementary representation contributes to improved diagnostic performance. The proposed method achieved 94.2% accuracy for tumor classification and 91.8% macro-accuracy for TI-RADS prediction, outperforming baseline models such as a 121-layer dense convolutional network, a 50-layer residual network, an Inception v3 network, and VGG16. These results suggest that integrating transformer-based global attention, convolutional feature extraction, and quantum-enhanced patch representations yields richer feature learning than individual architectures. The ablation study further confirms that fusing ViT-B16 and EfficientNet-B3 improves performance, while adding the quantum filtering module yields further gains by enhancing feature diversity and representational capacity.

6. Conclusion and future work

In conclusion, this study presents a quantum-infused hybrid deep learning framework for the dual-task classification of thyroid ultrasound images, effectively integrating quantum-filtered handcrafted features with semantic embeddings from ViT-B16 and EfficientNet-B3. The proposed model addresses key challenges in medical image analysis, including limited data, class imbalance, and high-dimensional image complexity, achieving robust performance in both binary tumor classification and multi-class TI-RADS prediction. By leveraging a multi-stage preprocessing pipeline, adaptive augmentation, and SMOTE-based oversampling, the framework demonstrates improved generalization and diagnostic reliability. Future work will focus on extending the model to incorporate multimodal clinical data, enhancing interpretability through explainable AI techniques, and optimizing computational efficiency for real-time deployment. Additionally, validating the model on larger, multi-institutional datasets and exploring implementation on quantum hardware represent important next steps toward clinical integration and scalability.

Acknowledgment

We are thankful for the assistance we have received from the Government of India's Ministry of Electronics and Information Technology.

Funding

This work was made possible by the Visvesvaraya PhD Scheme scholarship.

Conflict of interest

The authors declare no competing interests.

Author contributions

Conceptualization: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Data curation: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Formal analysis: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Investigation: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Methodology: Smitirekha Behuria
Project administration: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Software: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Resources: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Validation: Anjan Bandyopadhyay, Mingqiang Wang, Saurav Mallik
Writing–original draft: Smitirekha Behuria, Sujata Swain, Anjan Bandyopadhyay
Writing–review & editing: Anjan Bandyopadhyay, Mingqiang Wang, Saurav Mallik

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The datasets used during the current study are available in DDTI.[32]

References

1. Mnatsakanian A, Al-Khalili Y. Anatomy, Head and Neck, Thyroid muscles. In: *StatPearls*. Treasure Island, FL, USA: StatPearls Publishing; 2023.

2. Tessler FN, Middleton WD, Grant EG, *et al.* ACR Thyroid Imaging, Reporting and Data System (TI-RADS): White Paper of the ACR TI-RADS Committee. *J Am Coll Radiol.* 2017;14(5):587-595.

   doi: 10.1016/j.jacr.2017.01.046

3. Tessler FN, Middleton WD, Grant EG. Thyroid Imaging Reporting and Data System (TI-RADS): A User's Guide. *Radiology*. 2018;287(1):29-36.

   doi: 10.1148/radiol.2017171240

4. Lixandru-Petre IO, Dima A, Musat M, *et al.* Machine Learning for Thyroid Cancer Detection, Presence of Metastasis, and Recurrence Predictions—A Scoping Review. *Cancers*. 2025;17(8):1308.

   doi: 10.3390/cancers17081308

5. Ha EJ, Lee JH, Mak N, *et al.* A Deep Learning-Based Artificial Intelligence Model Assisting Thyroid Nodule Diagnosis and Management: Pilot Results for Evaluating Thyroid Malignancy in Pediatric Cohorts. *Thyroid*. 2025;35(6):652-661.

   doi: 10.1089/thy.2024.0627

6. Fontana E, Rudolph MS, Duncan R, Rungger I, Cîrstoiu C. Classical simulations of noisy variational quantum circuits. *npj Quantum Inf.* 2025;11(1):84.

   doi: 10.1038/s41534-024-00955-1

7. Liu CY, Chen SYC, Chen KC, Huang WJ, Chang YJ. Programming Variational Quantum Circuits with Quantum-Train Agent. In: Proceedings of the 2025 International Conference on Quantum Communications, Networking, and Computing (QCNC). New York, NY, USA: IEEE; 2025:544-548.

   doi: 10.1109/qcnc64685.2025.00091

8. Cao CL, Li QL, Tong J, *et al.* Artificial intelligence in thyroid ultrasound. *Front Oncol.* 2023;13:1060702.

   doi: 10.3389/fonc.2023.1060702

9. Khan T. Application of Two-Class Neural Network-based classification model to predict the onset of Thyroid Disease. In: Proceedings of the 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence). New York, NY, USA: IEEE; 2021:114-118.

   doi: 10.1109/confluence51648.2021.9377097

10. Srivastava R, Kumar P. Optimizing CNN based model for thyroid nodule classification using data augmen- tation, segmentation and boundary detection techniques. *Multimed Tools Appl.* 2023;82(26):41037-41072.

    doi: 10.1007/s11042-023-15068-8

11. Sun J, Wu B, Zhao T, *et al.* Classification for thyroid nodule using ViT with contrastive learning in ultrasound images. *Comput Biol Med.* 2023;152:106444.

    doi: 10.1016/j.compbiomed.2022.106444

12. Wang Y, Gan J. Benign and Malignant Classification of Thyroid Nodules Based on ConvNeXt. In: Proceedings of the 2022 3rd International Conference on Control, Robotics and Intelligent System. New York, NY, USA: Association for

Computing Machinery; 2022:56-60.

doi: 10.1145/3562007.3562017

13. Zhang J, Zhang H, Jiang P, *et al.* High efficiency classification of thyroid cytopathological images based on knowledge distillation and vision transformer. *Sci Rep*. 2025;15(1):30059.

doi: 10.1038/s41598-025-15728-9

14. Zheng Z, Su T, Wang Y, *et al.* A novel ultrasound image diagnostic method for thyroid nodules. *Sci Rep*. 2023;13(1):1654.

doi: 10.1038/s41598-023-28932-2

15. Aboudi N, Khachnaoui H, Moussa O, Khlifa N. Bilinear Pooling for Thyroid Nodule Classification in Ultrasound Imaging. *Arab J Sci Eng*. 2023;48(8):10563-10573.

doi: 10.1007/s13369-023-07674-3

16. Gong H, Chen J, Chen G, Li H, Li G, Chen F. Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules. *Comput Biol Med.* 2023;155:106389.

doi: 10.1016/j.compbiomed.2022.106389

17. Xiao X, Zhou Y, Zhu Y, Li Y, Qi T, Wang W. A multimodal prompt-tuning method of ultrasound diagnosis for thyroid nodule. *Front Med*. 2025;12:1686374.

doi: 10.3389/fmed.2025.1686374

18. Fang M, Ma Q, Xu B. Automated multi-class classification of thyroid nodules in ultrasound imaging using transformer-based segmentation and hybrid feature learning. *J Radiat Res Appl Sci.* 2025;18(4):101782.

doi: 10.1016/j.jrras.2025.101782

19. Zhou YT, Yang TY, Han XH, Piao JC. Thyroid-DETR: Thyroid nodule detection model with transformer in ultrasound images. *Biomed Signal Processing Control.* 2024;98:106762.

doi: 10.1016/j.bspc.2024.106762

20. Liu Q, Li Y, Hao Y, *et al.* Multimodal ultrasound multistage classification of PTC cervical lymph node metastasis via DualSwinThyroid. *Front Oncol*. 2024;14:1349388.

doi: 10.3389/fonc.2024.1349388

21. Rashed A, Medhat T, Elgarayhi A. Enhancing automatic diagnosis of thyroid nodules from ultrasound scans leveraging deep learning models. *Sci Rep*. 2025;15(1):40364.

doi: 10.1038/s41598-025-25780-0

22. Chi J, Chen J hui, Wu B, Zhao J, Wang K, Yu X, *et al.* A Dual-Branch Cross-Modality-Attention Network for Thyroid Nodule Diagnosis Based on Ultrasound Images and Contrast-Enhanced Ultrasound Videos. *IEEE J Biomed Health Inform*. 2025;29(2):1269-1282.

doi: 10.1109/jbhi.2024.3472609

23. Jiang S, Zhang N, Li C, Tong L, Yang X. Application of deep learning based on convolutional neural network model in multimodal ultrasound diagnosis of unexplained cervical lymph node enlargement. *Front Oncol*. 2025;15:1542265.

doi: 10.3389/fonc.2025.1542265

24. Shen P, Yang Z, Sun J, *et al.* Explainable multimodal deep learning for predicting thyroid cancer lateral lymph node metastasis using ultrasound imaging. *Nat Commun*. 2025;16(1):7052.

doi: 10.1038/s41467-025-62042-z

25. Maheshwari D, Ullah U, Marulanda PAO, *et al.* Quantum Machine Learning Applied to Electronic Healthcare Records for Ischemic Heart Disease Classification. *Hum Cent Comput Inf Sci.* 2023;13(06):1-8.

doi: 10.22967/HCIS.2023.13.006

26. Sajeev V, Hima Vyshnavi AM, Krishnan Namboori PK. Thyroid Cancer Prediction Using Gene Expression Profile, Pharmacogenomic Variants And Quantum Image Processing In Deep Learning Platform-A Theranostic Approach. In: Proceedings of the 2020 International Conference for Emerging Technology (INCET). New York, NY, USA: IEEE; 2020:1-5.

doi: 10.1109/incet49848.2020.9154041

27. Gomes Ataide EJ, Ponugoti N, Illanes A, Schenke S, Kreissl M, Friebe M. Thyroid Nodule Classification for Physician Decision Support Using Machine Learning-Evaluated Geometric and Morphological Features. *Sensors*. 2020;20(21):6110.

doi: 10.3390/s20216110

28. Liu T, Xie S, Zhang Y, Yu J, Niu L, Sun W. Feature selection and thyroid nodule classification using transfer learning. In: Proceeding of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). New York, NY, USA: IEEE; 2017:1096-1099.

doi: 10.1109/isbi.2017.7950707

29. Kwon SW, Choi IJ, Kang JY, Jang WI, Lee GH, Lee MC. Ultrasonographic Thyroid Nodule Classification Using a Deep Convolutional Neural Network with Surgical Pathology. *J Digit Imaging*. 2020;33(5):1202-1208.

doi: 10.1007/s10278-020-00362-w

30. Zhu Y, Fu Z, Fei J. An image augmentation method using convolutional network for thyroid nodule classification by transfer learning. In: Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC). New York, NY, USA: IEEE; 2017:1819-1823.

doi: 10.1109/compcomm.2017.8322853

31. Hang Y. Thyroid Nodule Classification in Ultrasound Images by Fusion of Conventional Features and Res-GAN

Deep Features. Singh D, ed. *J Healthc Eng.* 2021;2021:1-7.

doi: 10.1155/2021/9917538

32. Pedraza L, Vargas C, Narváez F, Durán O, Muñoz E, Romero E. An open access thyroid ultrasound image database. In: Romero E, Lepore N, eds. Proceedings of the SPIE 9287,10th International Symposium on Medical Information Processing and Analysis, 92870W. SPIE; 2015.

doi: 10.1117/12.2073532

33. Xu C, Sun Y, Zhang Y, *et al.* Stain Normalization of Histopathological Images Based on Deep Learning: A Review. *Diagnostics*. 2025;15(8):1032.

doi: 10.3390/diagnostics15081032

34. Kim YS, Kim MK, Fu N, Liu J, Wang J, Srebric J. Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artificial neural network models. *Sustain Cities Soc.* 2025;118:105570.

doi: 10.1016/j.scs.2024.105570

35. Liaw LCM, Tan SC, Goh PY, Lim CP. A histogram SMOTE-based sampling algorithm with incremental learning for imbalanced data classification. *Inf Sci.* 2025;686:121193.

doi: 10.1016/j.ins.2024.121193

36. Hou M, Wu Y, Zhang S. Two-Party Quantum Private Comparison with Pauli Operators. *Axioms*. 2025;14(8):549.

doi: 10.3390/axioms14080549

37. Hirano Y, Fujii K. Locality-Aware Pauli-Based Computation for Local Magic State Preparation. In: Proceedings of the 2025 IEEE International Conference on Quantum Computing and Engineering (QCE). New York, NY, USA: IEEE; 2025:670-680.

doi: 10.1109/qce65121.2025.00078

38. Pavithra KC, Kumar P, Geetha M, *et al.* Transformer-Based DME Classification Using Retinal OCT Images Without Data Augmentation: An Evaluation of ViT-B16 and ViT-B32 With Optimizer Impact. *IEEE Access*. 2025;13:180781-180798.

doi: 10.1109/access.2025.3620945

39. Younis EM, Ibrahim IA, Mahmoud MN, Albarrak AM. Hybrid of VGG-16 and FTVT-b16 Models to Enhance Brain Tumors Classification Using MRI Images. *Diagnostics*. 2025;15(16):2014.

doi: 10.3390/diagnostics15162014

40. Navaneethakrishnan M, Robinson Joel M, Kalavai Palani S, Gnanaprakasam GJ. EfficientNet-deep quantum neural network-based economic denial of sustainability attack detection to enhance network security in cloud. *Netw Comput Neural Syst.* 2024;36(4):1360-1384.

doi: 10.1080/0954898x.2024.2361093

41. Altaf AR, Altaf A, Rehman FU. Modified EfficientNet-B0 Architecture Optimized with Quantum-Behaved Algorithm for Skin Cancer Lesion Assessment. *Diagnostics*. 2025;15(24):3245.

doi: 10.3390/diagnostics15243245