IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 January 2024, accepted 12 February 2024, date of publication 20 February 2024, date of current version 26 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3367806

RESEARCH ARTICLE

QuCNet: Quantum-Inspired Convolutional Neural Networks for Optimized Thyroid Nodule Classification

GUNTI SWATHI[1], ALI ALTALBE[2,3], AND R. PRASANNA KUMAR[1]

[1]Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Chennai 601103, India
[2]Department of Computer Engineering, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[3]Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding author: R. Prasanna Kumar (r_prasannakumar@ch.amrita.edu)

This work was supported by Prince Sattam bin Abdulaziz University under Project PSAU/2023/01/29091.

ABSTRACT The escalating incidence of thyroid cancer over the past decade underscores the imperative for effective classification and early detection of thyroid nodules. An automated system for this purpose could significantly aid physicians, particularly in expediting diagnostic processes. However, attaining this objective has proven challenging, primarily attributed to the constrained dataset size for medical images and the laborious process of feature extraction. This research addresses these challenges by thoroughly exploring the importance of extracting meaningful features for tumor detection and introducing a quantum-based convolutional neural network. The proposed approach employs a quantum filter transformation for intricate feature extraction, coupled with a classical neural network for the classification of thyroid nodules in ultrasound images. The classification process involves two distinct categorizations: distinguishing between nodules that are benign or malignant, and identifying the specific suspicious class to which the nodule is attributed. The amalgamation of both classifiers yields a comprehensive characterization of thyroid nodules, showcasing notable accuracy. For tumor classification, the model achieves an accuracy of 97.63%, precision of 97.72%, and recall of 97.64% on a test dataset containing 127 images. Similarly, for suspicious level classification, the model attains an accuracy of 93.87%, precision of 94.58%, and recall of 93.88% on a test dataset containing 98 images. These results surpass the performance of existing models, marking a significant advancement in the field of thyroid nodule classification. The proposed model represents a promising and innovative methodology that could offer valuable support mechanisms in clinical settings, facilitating the rapid classification and diagnosis of thyroid cancer.

INDEX TERMS Thyroid nodule classification, quanvolutional neural networks, quantum filter transformation, tumor classification, benign, malignant, TI-RADS, feature extraction, DDTI.

I. INTRODUCTION

In the contemporary healthcare landscape, the classification of thyroid nodules holds utmost importance. The thyroid gland, situated below the epiglottis, is pivotal for regulating a range of physiological functions by producing essential hormones. These functions encompass the regulation of neural and cerebral activities as well as the facilitation of optimal development and operation of essential organs [1]. Any anomalies or distortions in the thyroid can significantly impair its capacity to perform these functions. Thyroid nodules are commonly encountered, with a higher incidence among women and older individuals [2]. While most of these nodules are benign in nature, the potential for malignancy underscores the need for a precise and consistent evaluation process. To address the subjectivity and variability associated with diagnosing thyroid nodules, the medical community has adopted the Thyroid Imaging Reporting and Data System

The associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian.

(TIRADS) approach [3]. This standardized system plays a critical role in stratifying the risk and discerning thyroid nodules as either benign or malignant, relying on visual attributes observed through ultrasound imaging.

Traditionally, a range of machine learning models, including SVM, Naive Bayes, Logistic Regression, and Decision Trees [4], were employed to differentiate between benign and malignant thyroid nodules but struggled to capture intricate patterns. In recent years, deep learning models like pretrained convolutional neural networks (CNNs), including VGG16 [5], GoogLeNet [6], ResNet18 [7], transformer models, and attention mechanisms, emerged as promising solutions. However, challenges like limited generalizability and time-consuming feature extraction from small ultrasound datasets hinder their clinical applicability.

In this research, we employed a quanvolutional neural network (QuCNet) that integrates quantum data processing with classical computation to enhance the efficiency and accuracy of thyroid nodule classification. This approach tackles the complexities posed by high-dimensional data within ultrasound images of thyroid glands while also striving to enhance the generalization of small datasets, thereby improving classification performance, even in data-constrained situations.

Our proposed work's primary contributions are summarized below:

- Introducing a diagnostic support system for thyroid cancer that utilizes quantum-inspired convolutional neural networks. This system is specifically designed to be effective in the classification of thyroid nodules, even when data availability is constrained.
- The proposed model utilizes a quantum filter transformation (QFT), extracting four-channel quantum features from images. This unique approach enhances the system's ability to focus on various regions of the image, providing more detailed and nuanced information crucial for accurate classification.
- In comparison of our proposed QuCNet with established models (Vgg16, Resnet50, InceptionV3, Densenet121) on the Digital Database of Thyroid Ultrasound Image Dataset, our QuCNet model outperforms existing machine learning and deep learning models. The integration of intricate feature extraction using QFT, along with classical convolutional networks, proves to be highly effective for tumor classification in thyroid ultrasound images.

The remaining portion of this paper is organized in the following manner: Section II gives a concise overview of pertinent studies on thyroid nodule classification. Following this, Section III provides in-depth insights into the dataset, and its preprocessing techniques are detailed in Section IV. Then, Section V presents the methodology, which elaborates on the proposed framework. Progressing further, Section VI delineates the experimental analysis. Lastly, Sections VII and VIII encapsulate the findings, delve into a discussion on the utilized performance metrics, draw conclusions, and provide insights into potential avenues for future research.

II. RELATED WORKS

Given the significance of classifying the nature and early suspicion of thyroid nodule, substantial attention has been dedicated to this endeavor. In the past, various machine learning models have been employed to address this objective [8].

Gomes et al. [3] harnessed geometric and morphological features in addition to conventional machine learning to improve the classification of thyroid nodules. This approach surpassed existing techniques and held promise as an automated diagnostic system for validating TIRADS guidelines. Saini et al. [4] employed four classification models (Naive Bayes, Logistic Regression, KStar, and Decision Tree - J48) for predicting thyroid disease using machine learning. Their findings demonstrated that the J48 model outperformed all other models.

Later, deep learning methods have been explored to overcome the constraints associated with traditional models. Liu et al. [9] presented a technique that combines deep features from a CNN model with conventional features like scale-invariant feature transform and histogram of oriented gradient. This integration forms a hybrid feature space, aiming to improve the classification process. Kwon et al. [5] used the VGG16 model to create a nodule classification system addressing subjectivity and variability concerns while applying data augmentation and cross-validation to reduce overfitting. Chi et al. [6] developed an automated diagnosis system utilizing a fine-tuned GoogleLeNet model for feature extraction from pre-processed images and, consequently, employed a Random Forest classifier for the final classification. Hussain et al. [10] used an ensemble classifier that combined Gabor features and multiple classifiers to classify mass regions of interest (ROIs) from mammograms into benign and malignant categories.

Zhu et al. [7] utilized image preprocessing to extract ROI, data augmentation using traditional methods for data preparation, and a pre-trained residual network for classifying the nodules. Hang et al. [2] integrated deep features with traditional ones, improved image quality using methods like histogram equalization, and extracted multiple features. They then utilized the Res-GAN model for classification. Khan et al. [11] employed a two-class neural network classifier in combination with feature selection using mutual information and used the synthetic minority oversampling technique for class balancing to create a machine learning-based model for thyroid disease classification, achieving strong performance.

One of the recent research by Srivastava et al. [12] developed an optimized CNN model for identifying thyroid nodules. This work encompassed the incorporation of diverse architectures, such as Visual Geometry Group-16, ResNet-50, DenseNet, and AlexNet. Also, hyper-parameter tuning was employed to experimentally determine optimal values for learning rate and dropout factors. Jiawei et al. [13] employed

TC-ViT, a transformer-based model incorporating contrast learning. With an emphasis on TI-RADS, this application sought to enhance the classification of thyroid nodules. Wang et al. [14] utilized an algorithm based on ConvNeXt, combining artificial marker removal, data expansion, and an attention mechanism for classifying thyroid nodules. Feres et al. [15] utilized two models for thyroid nodule classification: CNNs (VGG16, EfficientNetB0, ResNet50) and ViTs (ViT_B16, Hybrid ViT), trained with Softmax and SVM classifiers. Results indicated that the hybrid ViT with SVM surpassed all other models in effectiveness for tumor classification.

Baima et al. [16] introduced a Dense Nodal Swin-Transformer (DST) method for diagnosing thyroid nodules with ultrasound images. It employed image segmentation, feature map construction across four stages, and a dense connection mechanism within each stage block. This approach enhanced diagnostic performance by effectively utilizing multi-layer features. The Lightweight Global Attention Module (LGAM) was developed by Yu et al. [17] to distinguish between benign and malignant thyroid nodules in ultrasound images. The incorporation of the attention mechanism demonstrated notable effectiveness in enhancing the diagnostic capabilities of the model. Zheng et al. [18] combines a multiscale localization network, leveraging data-driven size and aspect ratio patterns, with an advanced two-way classification network processing ultrasound images. The enhanced feature extraction and fusion of deep features, shallow features, and nodule aspect ratio contribute to a precise differentiation between benign and malignant nodules.

Aboudi et al. [19] employs a unique bilinear convolutional method, combining outputs from two CNN models using outer products, to classify thyroid nodules in 447 ultrasound images. Results show that BCNN algorithms outperform CNN architectures, offering a valuable second opinion for accurate and objective nodule classification. Thyroid Region Prior Guided Feature Enhancement Network (TRFE+) was employed by Gong et al. [20] to overcome challenges during ultrasound segmentation of thyroid nodules. The method incorporated a novel multi-task learning framework, adaptive gland region feature enhancement, and normalization for channel dimension during training, ultimately improving accuracy. The Brief Efficient Thyroid Network (BETNET), a deep CNN model for the location and classification of thyroid nodules in ultrasound images, was developed by Zhu et al. [21]. This network, built on the VGG-19 architecture with 16 convolutional layers, 3 fully connected layers, and a softmax layer, exhibited superior diagnostic performance. The preprocessing procedures, which included shrinking and removing black areas, helped BETNET recognize and classify thyroid nodules more effectively.

At present, quantum computing is considered a pioneering solution for addressing computational challenges that traditional hardware finds insurmountable. Its advanced speed and computational prowess have proven highly successful, when combined with machine learning [22]. A study conducted by Maheshwari et al. [23] has utilized optimized quantum support vector machine and hybrid quantum multi-layer perceptron algorithms to classify cardiovascular diseases. These models exhibited computational efficiency, rendering them apt for real-time healthcare applications, and showcased competitive performance when compared to more intricate architectures in recent research. Xiong et al. [24] employed Qiskit to simulate quantum circuits and developed VQNet, a hybrid neural network that combines quantum and classical architectures. They applied this model to classify the MNIST and CIFAR-10 datasets, showcasing the superior efficiency and accuracy of quantum machine learning over classical methods.

Conventional methods, relying on classical methods for feature extraction, face challenges in the medical imaging domain due to the limited size of datasets and resource-intensive processes. This constraint becomes more pronounced in scenarios where the nuanced details of pathology, such as subtle texture variations in thyroid nodules, require a more sophisticated approach. To address these challenges, our study proposes a quantum-classical approach utilizing quantum filter transformation to capture intricate features. This innovative strategy presents a promising solution for overcoming the limitations of conventional methods, facilitating a more comprehensive characterization of Thyroid nodule ultrasound imaging features.

III. DATA DESCRIPTION

The classification of medical images is of paramount importance, as it necessitates the use of real-time data, which is often scarce and invaluable. We have leveraged the Digital Database of Thyroid Ultrasound Images (DDTI) [25], a publicly available dataset that provides a comprehensive collection of B-mode ultrasound images. Notably, this dataset includes descriptions and annotations of suspicious thyroid lesions, all of which have been classified by a panel of more than two radiologists. DDTI undergoes regular updates to enhance its content and provide more diverse data. At present, it encompasses a total of 478 images from 399 medical cases. Among these images, 460 are categorized into two main groups: Benign and Malignant, as illustrated in the Figure 1.

Moreover, experts in the medical imaging domain have classified the thyroid nodules in all 478 images using the Thyroid Imaging Reporting and Data System (TI-RADS). TI-RADS employs a numerical scoring system based on the attributes observed within five ultrasound categories [26] found in a thyroid nodule. Suspicious characteristics are allocated extra points, and the total TI-RADS score is determined by the sum of these feature points. The TI-RADS scores range from Benign (2) to probably benign (3), Low suspicion of malignancy (4a), Intermediate suspicion of malignancy (4b), Moderate Suspicion of malignancy (4c), and High suspicion of malignancy (5). These classifications are visually illustrated in Figure 2. An overview of the dataset's statistics is presented in Figure 3.

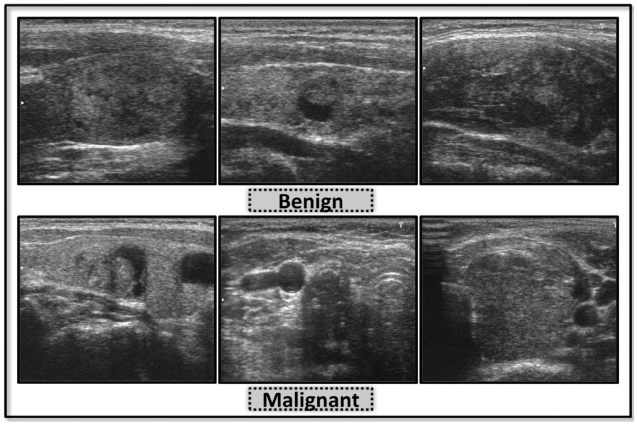


FIGURE 1. Visual Representation of Tumor Classification for Ultrasound Images in the Dataset.

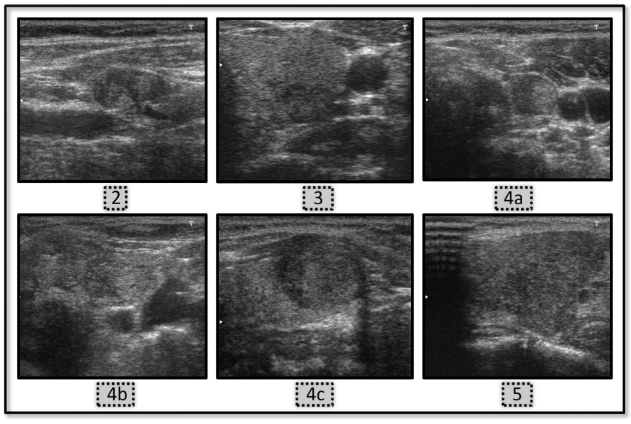


FIGURE 2. Visual Representation of TI-RADS Classification for Ultrasound Images in the Dataset.

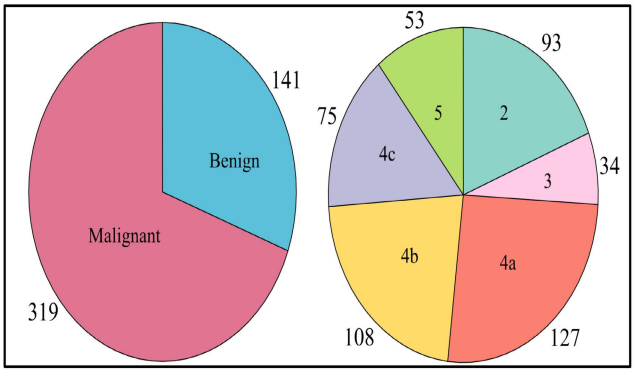


FIGURE 3. Two- level Distribution of Class Instances in the DDTI.

IV. PREPROCESSING

In order to ensure the uniformity and compatibility of the image data for quantum filters, we have implemented a comprehensive preprocessing pipeline encompassing four key steps: resizing, normalization, data augmentation, and oversampling. Initially, the raw input images were resized to a consistent dimension of 128 × 128 pixels to facilitate optimal division into a 4-channel configuration by the quantum kernel. Following resizing, normalization techniques, specifically mean subtraction and standard deviation scaling, were employed to normalize the pixel values of the images.

These initial preprocessing steps are applied to both datasets and are subsequently divided into training and testing sets using an 80-20 split. Recognizing the limited size of our dataset, we incorporated data augmentation strategies on the training set that include zooming in and out within a range of 0.99 to 1.01, brightness adjustments with a range of 0.8 to 1.2, and horizontal flipping. Additionally, in order to tackle the issue of class imbalance in the dataset, we utilized the Synthetic Minority Over-sampling Technique (SMOTE) [27]. SMOTE generates synthetic instances for the underrepresented class by interpolating between existing samples, ensuring equitable and precise predictions across all classes. The class counts before and after the preprocessing pipeline are presented in Table 1 and Table 2.

TABLE 1. Tumor class distribution in the training dataset before and after preprocessing, and in the unprocessed test dataset.

Classes	Train dataset	Train dataset after Preprocessing	Test dataset
Benign	102	1848	39
Malignant	231	1216	88
Total	333	3064	127

TABLE 2. Suspicious level class distribution in the training dataset before and after preprocessing, and in the unprocessed test dataset.

Classes	Train dataset	Train dataset after Preprocessing	Test dataset
2	74	1184	19
3	19	608	5
4a	101	1616	26
4b	84	1344	21
4c	62	992	16
5	40	980	11
Total	380	6724	98

V. METHODOLOGY

In this current work, we have employed a quantum convolution neural network or a Quanvolutional neural network [28], which falls in the category of Classic-quantum computation. This model is an enhancement of the classical CNNs with an added quanovlutional filter. Quanvolutional filters utilize random quantum circuits to extract features from input data by transforming specific localized spatial subsegments of the data. These circuits can exhibit either structured or random configurations.

In our study, we deliberately adopted random quantum circuits over circuits with specific structural characteristics.

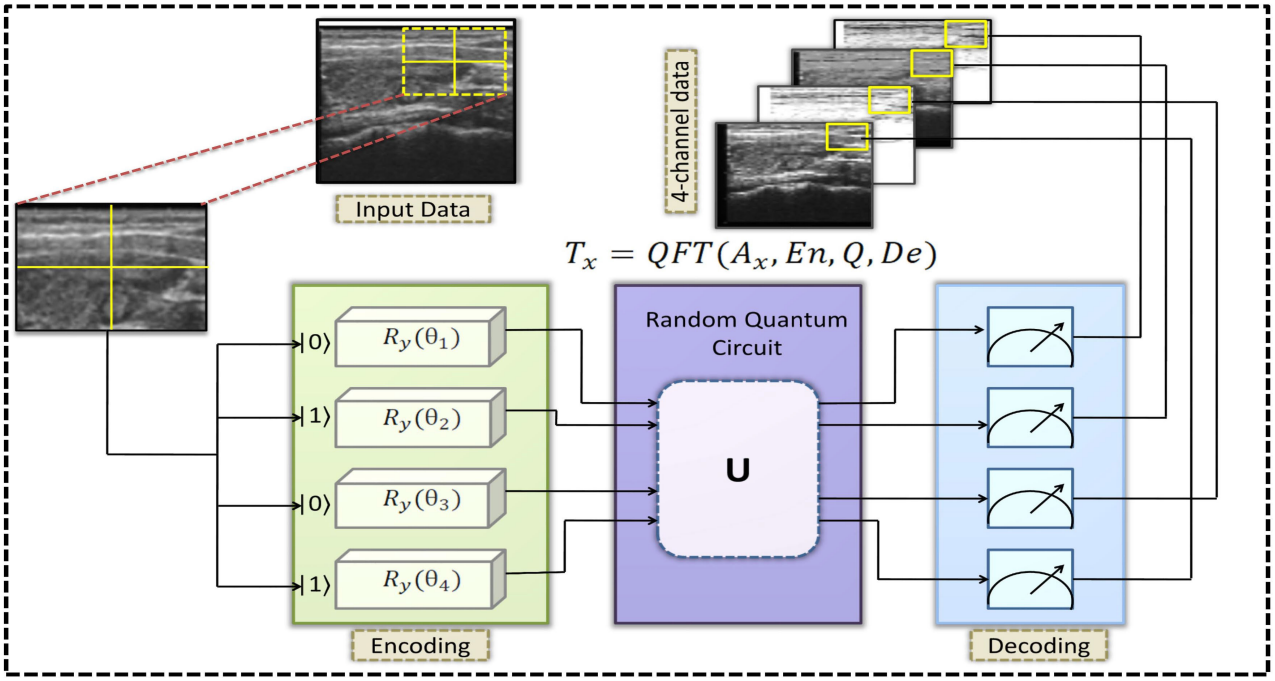


FIGURE 4. Semantic Representation of the Quantum Filter Transformation Process: A Three-Stage Conversion from Input Image to 4-Channel Output.

This strategic choice was guided by considerations of computational simplicity, a rationale stemming from the relatively modest size of our dataset. Furthermore, the utilization of random quantum circuits has yielded notable benefits in terms of accelerated training. At a broader conceptual level, quanvolutional filters serve as a transformative intermediary that takes input data and produces an output scalar value through the manipulation of a 2D matrix of scalar elements using a universal quantum computing circuit. This transformation involves a three-stage process encompassing encoding, the application of a quantum random circuit, and decoding.

A. ENCODING

During the encoding phase, the generation of quantum data is initiated by transforming classical data, marking a pivotal step in the computational process. In our research, we utilized the $R_x(\Theta)$ gate, a single-qubit rotation gate that rotates around the x-axis by an angle of Θ radians. Notably, the x-axis on the Bloch sphere corresponds to the Pauli X gate [29], which toggles the qubit between the $|0\rangle$ and $|1\rangle$ states. For each filter, it assigns a specific encoding function denoted as En. This encoding function is applied to the 2-dimensional subsegment, A_x, obtained from the input image through the filter. Consequently, the initial quantum state, I_x, is formulated as,

$$I_x = En(A_x) \tag{1}$$

B. QUANTUM RANDOM CIRCUITS

In the subsequent step, the encoded quantum state I_x, derived from the encoding process is subjected to a quantum circuit to fully harness the capabilities of quantum computing. This circuit, denoted as the unitary operator U, is instrumental in carrying out quantum computations and can be generated through a trainable variational quantum circuit, or through the utilization of a random circuit (Q).We have opted for the use of a random circuit denoted as Q, primarily due to its computational suitability for the relatively modest size of our dataset.The outcome of this computational process results in the quantum output state, O_x, which can be expressed as,

$$O_x = Q(I_x) = Q(En(A_x)) \tag{2}$$

C. DECODING

The culminating step in our quantum data processing pipeline involves the transformation of the quantum state into a scalar output. The prominent method for this decoding process is quantum measurement. During measurement, the quantum state collapses into one of its fundamental states determined by the associated probability distribution. This ensures the consistency and reliability of the quanvolutional filter output. To achieve this, we employ a decoding function (De), which allows us to obtain the Terminal decoded state (T_x). This decoded state represents a scalar value and can be expressed as,

$$T_x = De(O_x) = De(Q(En(A_x))) \tag{3}$$

The comprehensive three-stage process known as Quantum Filter Transformation (QFT), as illustrated in Figure 4, can be formally defined as follows,

$$T_x = QFT(A_x, En, Q, De) \tag{4}$$

In the implementation the initial input of 128 × 128 pixels image is divided into segments of 2 × 2 pixels, and each segment is processed by the QFT. The obtained 4 expectation values after the QFT is mapped into four different channels. Each channel finally forms a 64 × 64 pixel image. This process would resemble a classical CNN correspond to a convolution with a 2 ×2 kernel and a stride equal to two.

The QFT down samples the resolution and introduces local distortion in the image, while preserving the global shape. This downsampling operation enables the model to effectively focus on capturing essential patterns and structures at various perspective levels, thereby facilitating the extraction of global features from the input. This distinctive approach contrasts with classical CNNs, which yield a singular output from each convolutional operation. The incorporation of QFT provides QuCNet with a performance edge, augmenting its ability to achieve precise image classification.

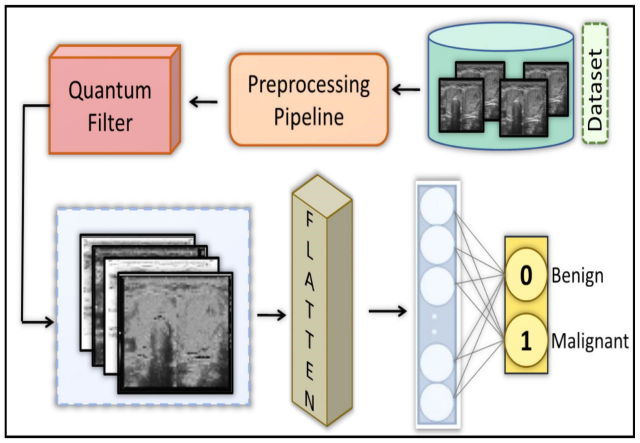


FIGURE 5. Comprehensive Workflow of the QuCNet Architecture for Tumor Classification.

D. MODEL

For the proposed classification tasks, we have developed two distinct models, both sharing a common QuCNet framework but differing in complexity. The initial step in both classification tasks involves applying the QFT operation to a subsection of the pre-processed image as

$$Zq_k{}^{(i,i)} = QFT(A_{(i,i)}, En, Q_k, De) \tag{5}$$

This computation, represented by (5), is performed for all subsegments of the image within a region determined by the filter size (k). The quantum-computed segments are then organized into k respective channels, as illustrated in (6).

$$Channel_k = \left[Zq_k{}^{(1,1)}, Zq_k{}^{(1,2)} \cdots Zq_k{}^{(m,n)}\right] \tag{6}$$

Here, the variables m and n range from 1 to the height/k and width/k, respectively, for the first channels. The resulting channelled state given in (7), represents the final computed state of all sub-segments in their corresponding channels.

$$Z_{chanelled} = [Channel_1, Channel_2, Channel_3, Channel_4] \tag{7}$$

These operations are consistent for both classification tasks, with the subsequent architecture exhibiting different complexities.

The tumour classification model, responsible for distinguishing between Benign and Malignant classes, adheres to a simple structure as presented in Figure 5. This model comprises the following architectural elements: Input layer for image ingestion, QFT generating a 4-channel output, concatenation, and flattening of these channels, and finally, a dense layer for output generation. The Z-channelled state, as represented in (7), undergoes flattening before being passed to a dense layer.

$$Z_{flattened} = flattened\left(Z^2_{CNN}\right) \tag{8}$$

$$Z^1_{dense} = f^1\left(W^1 \cdot Z_{flattened} + b^1\right) \tag{9}$$

$$Z^2_{dense} = f^1\left(W^2 \cdot Z^1_{dense} + b^2\right) \tag{10}$$

Here, W and b represent the weights and biases in the network, respectively, serving as trainable parameters that are updated with each training epoch. The f^1 denotes the ReLU activation function, Subsequent to the aforementioned steps, (11) describes the output layer (Y) with f^2 as the SoftMax activation function. Y contains the probability distribution for Benign and Malignant classes.

$$Y = f^2\left(W^0 \cdot Z^2_{dense} + b^0\right) \tag{11}$$

On the other hand, the suspicious classification model is more intricate, featuring a design characterized by the following components: Input layer for image input, QFT for 4-channel output, two pairs of Conv2D (convolutional) and MaxPool (max-pooling) layers, a flattening layer, and two dense layers for class prediction. This architecture is designed to handle multi-class classification tasks and is depicted in Figure 6.

Similar to the tumor classification model, the initial step involves applying QFT to the input image.The $Z_{channelled}$ state obtained from (7) serves as input for two sets of convolutional and pooling layers. These layers, as depicted in (12) and (13), perform convolutional and max-pooling operations, respectively. This design enables the extraction of features and reduction of spatial dimensions, enhancing the model's ability to discern intricate patterns.

$$Z^1_{CNN} = max\_poll\left(Conv2D(Z_{chanelled})\right) \tag{12}$$

$$Z^2_{CNN} = max\_poll\left(Conv2D(Z^1_{CNN})\right) \tag{13}$$

Here,

$$Conv2D((X_{(i,j)})) = f\left(\sum_{y=1}^{Y}\sum_{z=1}^{Z} W_{y,z} \cdot X_{i+y-1,j+z-1} + b\right) \tag{14}$$

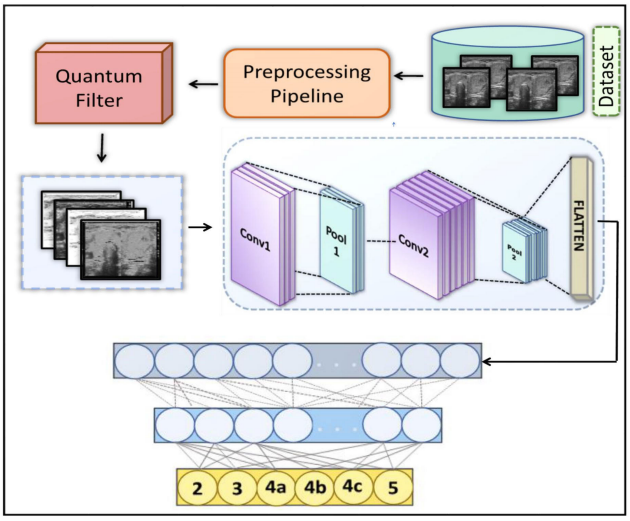


FIGURE 6. Comprehensive Workflow of the QuCNet Architecture for Suspicious Level Classification.

Here, y and z represent the dimensions of convolutional filters. f serves as the activation function, W stands for the weight, and X corresponds to the input image. Following these operations, the architecture of the suspicious level classification aligns with the tumor classification model.

$$max - poll\left(X_{i,j}\right) = max_{y=1}^{p} max_{z=1}^{p} X_{(j-1)p+y,(j-1)p+z} \quad (15)$$

Here, p is the pooling window size.

The output from the Z_{cnn} is flattened and processed through two dense layers to provide the final class prediction. This process is encapsulated in (8), (9), (10), and (11), representing the subsequent steps in the model's architecture for suspicious classification. These dense layers, akin to those in tumor classification, contribute to determining the level of suspicion associated with the image.

VI. EXPERIMENT ANALYSIS

In consideration of the current limitations of the availability of efficient and dependable quantum computers, we have pragmatically harnessed the Qiskit library, provided by IBM [30], to simulate quantum circuits on classical computing devices. Every experiment was conducted on a consistent computing environment utilizing a Tesla P100 Graphics Processing Unit, supported by a 25 GB RAM memory capacity.

The quantum processing is characterized by a random quantum circuit, where the initial stage consists of an embedding layer comprising local rotations scaled by a factor of π. In our experimental setup, we opted for a conservative approach by configuring the kernel size of the quantum convolutional filter to 2. This choice was made to manage computational complexity, which was increasing exponentially with increase in kernel size. The number of layers (n_layers) of quantum circuit repeated in quantum phase is set to 2, corresponding to the results of its variation

TABLE 3. Impact of the (*n_layers*) parameter on performance metrics of tumor classification.

n_layers	Accuracy	Precision	Recall	F1 Score
1	0.9528	0.9526	0.9528	0.9524
2	**0.9764**	**0.9772**	**0.9764**	**0.9761**
3	0.9449	0.9449	0.9449	0.9442
4	0.9370	0.9366	0.9370	0.9365

TABLE 4. Impact of the (*n_layers*) parameter on performance metrics of suspicious level classification.

n_layers	Accuracy	Precision	Recall	F1 Score
1	0.8776	0.9050	0.8776	0.8697
2	**0.9387**	**0.9458**	**0.9388**	**0.9329**
3	0.9082	0.9235	0.9082	0.8964
4	0.8469	0.8894	0.8469	0.8306

represented in Table 3 and 4. The model's performance, especially for the suspicion level classifier, has deteriorated with an increase in the number of layers. In the classical computation phase, both the models compiled using Adam optimizer [31]. Throughout the training process, the binary cross-entropy loss function and categorical cross-entropy loss function was employed on tumour classification and suspicious classification respectively, to quantify the dissimilarity between the true class labels and the predicted class probabilities, with the overarching objective of minimizing this dissimilarity. The binary cross-entropy loss function is given by.

$$BCE\left(a, \hat{a}\right) = -\left(a \cdot log\left(\hat{a}\right) + (1-a) \cdot log\left(1-\hat{a}\right)\right) \quad (16)$$

In (16) a is actual label, representing the label 0 or 1 corresponding to benign and malignant and $\hat{a}$ is the predicted probability. The categorical cross-entropy loss function is given by,

$$CCE\left(a, \hat{a}\right) = -\sum_{i=0}^{n-1} a_i \cdot log\left(\hat{a}_i\right) \quad (17)$$

In (17) a is a one-hot encoded vector representing true classes of the suspicion scale, $\hat{a}$ is the predicted probability distribution and n is the number of classes.

The internal layers of both models were furnished with the Rectified Linear Unit (ReLU) [32] activation function. The (18) gives the mathematical representation of ReLU. The incorporation of ReLU served the essential purpose of introducing non-linearity into the models, thus effectively mitigating the vanishing gradient problem.

$$RELU\ (x) = max\ (0, x) \quad (18)$$

In the output layer, the Softmax activation function was employed to transform the real-valued vector obtained from the penultimate layer into a probability vector. (19) illustrates

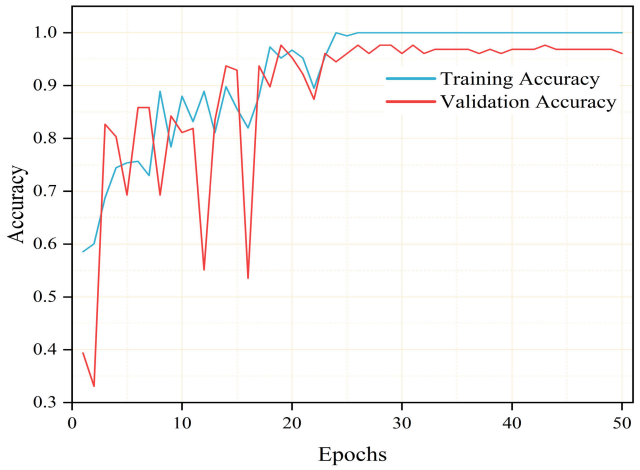


FIGURE 7. Accuracy plot of Tumor Classification over 50 epochs.

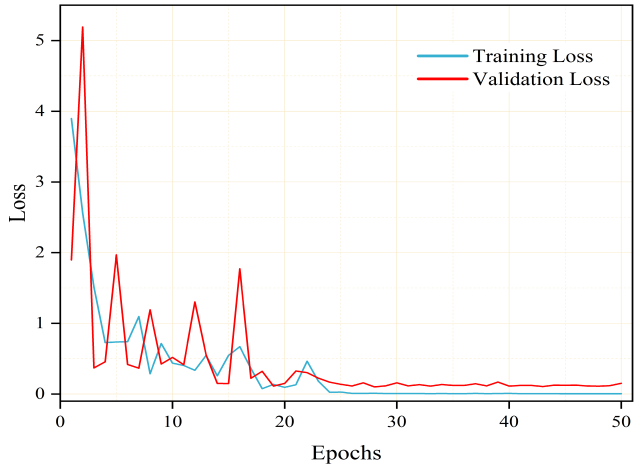


FIGURE 8. Loss plot of Tumor Classification over 50 epochs.

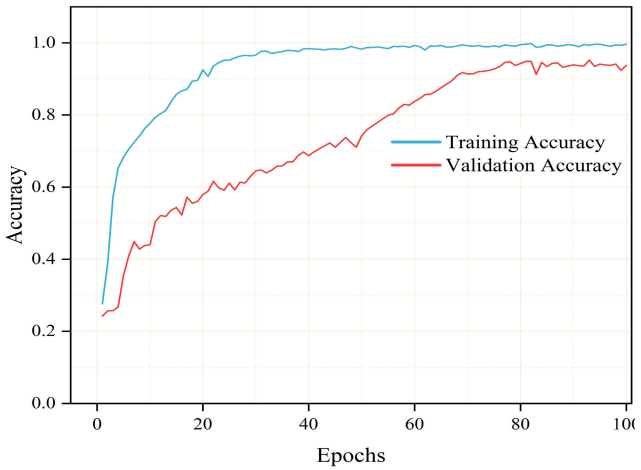


FIGURE 9. Accuracy plot of Suspicious Level Classification over 100 epochs.

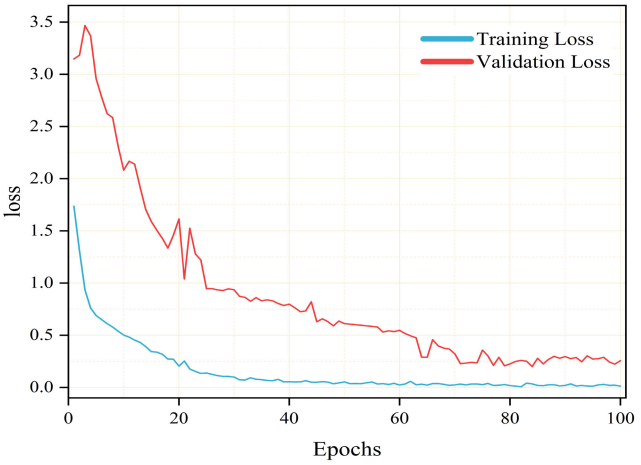


FIGURE 10. Loss plot of Suspicious Level Classification over 100 epochs.

the normalization process applied, guaranteeing that the resulting values fall within the range of 0 and 1.

$$S\,(y)\,i = \frac{exp\,(yi)}{\sum_{j=i}^{n} exp\,(yi)} \tag{19}$$

Moreover, to facilitate dynamic adjustments in the learning rate and mitigate the risk of overfitting, the "ReduceLROnPlateau" callback function along with "EarlyStopping" was incorporated into the training regimen. Specifically, in instances where the loss metric demonstrated no discernible improvement over a consecutive span of four epochs, a reduction factor of 0.5 was applied to the learning rate. This adjustment ensured that the learning rate did not descend below the predefined minimum threshold of 10^{-6}. The "EarlyStopping" mechanism was configured with a patience parameter set to three, coupled with the implementation of model checkpointing to save the best weights during the training process. These hyperparameter settings were consistently applied to both the classification models.

The tumour classifier underwent a 50-epoch training regimen, and early stopping was employed to save the optimal model weights at the 26th epoch. The training process for this model took approximately 8.15 hours. In contrast, the suspicion classification model underwent an extended training period of 100 epochs, with early stopping ensuring the preservation of the best model at the 73rd epoch. This extended training duration took approximately 16 hours. Comprehensive training and validation plots for both classifiers on the augmented datasets are depicted in Figure 7, 8, 9 and 10.

VII. RESULT AND DISCUSSION

This section evaluates the effectiveness of both the classifiers by presenting their results on the test set and making a comparative analysis with the established pretrained CNN models.

A. PERFORMANCE METRICS

Evaluating the effectiveness of a machine learning model is a crucial step in its development. This involves assessing the model's performance using various metrics and pro-

vide insights into the model's quality by making accurate predictions based on the given data. In the context of our Thyroid Nodule Classification, we employed a range of performance metrics to evaluate the predictions made by our model. These metrics serve as critical gauges, allowing us to gauge how well the model is functioning and to make informed assessments about its overall performance. This section provides a concise overview of all the metrics employed in this evaluation process.

1) CONFUSION MATRIX

A classification model's performance on test data is visually summarized through a confusion matrix, which displays accurate and inaccurate predictions. This matrix evaluates metrics such as Recall, Precision, Accuracy, and F1_score. The essential components include correctly predicted positives (True Positives), correctly predicted negatives (True Negatives), inaccurately predicted positives (False Positives), and inaccurately predicted negatives (False Negatives).

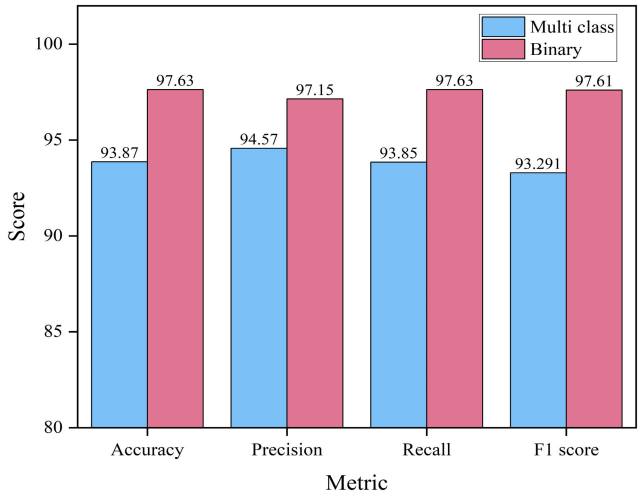


FIGURE 11. Performance Metric Results for Tumor Classifier and Suspicious Level Classifier.

2) VALIDATION LOSS AND ACCURACY

Validation loss and accuracy are crucial metrics assessing a model's generalization. Like standard loss and accuracy, they are evaluated after each epoch using a validation set. Validation loss measures the disparity between predicted and actual values, aiming for minimal error, while accuracy emphasizes the correct classification percentage on new data. This meticulous evaluation ensures a comprehensive understanding of the model's effectiveness beyond its training data.

3) PRECISION

Precision is employed to overcome the limitations of Accuracy by determining the accuracy of positive predictions. This metric is computed as the ratio of True Positives to the total positive predictions.

$$Precision(P) = \frac{TP}{TP + FP} \tag{20}$$

TABLE 5. Detailed analysis of tumor classifier performance across different classes.

Class	Precision	Recall	F1 Score
Benign	1.00	0.92	0.9
Malignant	0.97	1.00	0.98

TABLE 6. Detailed analysis of suspicious level classifier performance across different classes.

Class	Precision	Recall	F1 Score
2	1.00	0.95	0.97
3	1.00	0.40	0.57
4a	0.90	1.00	0.95
4b	0.88	1.00	1.00
4c	1.00	1.00	1.00
5	1.00	0.82	0.90

4) RECALL

Much like the Precision metric, Recall seeks to measure the portion of actual positives correctly identified. It is computed as the division of True Positives by the total number of positives, encompassing both correct positive predictions and inaccurately identified negatives (False Negatives).

$$Recall(R) = \frac{TP}{TP + FN} \tag{21}$$

5) F1_SCORE

The F1_score evaluates a model's predictive ability by analyzing its class-wise performance, focusing on True Positives and False Negatives. The metric ranges from 0 to 1, where a value of 1 signifies perfect Precision and Recall, and a value of 0 indicates either zero Precision or Recall. The score is calculated by dividing True Positives by Total Positives.

$$F1\_Score = 2 \times \frac{P \times R}{P + R} \tag{22}$$

Figure 7 and Figure 8 provide a comprehensive overview of the tumour classification model's learning process, indicating a discernible improvement in accuracy up to the 26th epoch, followed by stabilization of curve. The model's optimal configuration, discerned through rigorous training, has been checkpointed, and its efficacy is evaluated on a previously unseen test-set comprising 127 images. Specifically, these images have not been exposed to the model during any phase of training, ensuring a proper evaluation scenario.

The overall performance metric scores of the Tumour classifier are illustrated in Figure 11, demonstrating a

TABLE 7. Performance metrics of various models on benign and malignant image classification, including precision, recall, and F1 score. The 'Support' column represents the number of instances in each class.

Models	Benign			Malignant		
	Precision	Recall	F1 Score	Precision	Recall	F1 Score
Densnet121	0.88	0.74	0.81	0.89	0.95	0.92
Resnet50	0.74	0.67	0.70	0.86	0.90	0.88
InceptionV3	0.57	0.54	0.55	0.68	0.68	0.68
Vgg16	0.40	0.46	0.43	0.74	0.69	0.72
Proposed Model	**1.00**	**0.92**	**0.96**	**0.97**	**1.00**	**0.98**
Support	39			88		

TABLE 8. Comparative performance evaluation of proposed model and baseline cnn architectures for tumour classification.

Models	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
Propopsed Model	**0.9387**	**0.9458**	**0.9388**	**0.9329**

commendable accuracy of 97.63% and F-1 score of 97.61%. These metrics collectively affirm the classifier's ability to consistently differentiate between benign and malignant classes, underscoring the stability and reliability of its predictions. A more detailed insight into the classifier's individual class metrics is meticulously outlined in Table 5. Additionally, the confusion matrix in Figure 13 elucidates the model's predictions for each class, offering a granular perspective on its performance.

Similarly, Figure 9 and Figure 10 depict the learning process of the suspicious level classification model, demonstrating a discernible improvement in accuracy up to the 73rd epoch, followed by stabilization of the curve. The model's optimal configuration, established through rigorous training, is then evaluated on a previously unseen test set comprising 98 images, ensuring a proper evaluation scenario. The overall performance metric scores of the suspicious level classifier, as presented in Figure 11, showcase a commendable accuracy of 93.87% and an F-1 score of 93.29%. These metrics affirm the classifier's consistent differentiation between classes, highlighting the stability and reliability of its predictions. Table 6 provides a detailed insight into the individual class metrics of the classifier, while the confusion matrix in Figure 14 elucidates the model's predictions for each class, providing a granular perspective on its performance.

B. COMPARITIVE ANALYSIS AGAINST EXISTING CLASSIFICATION MODELS

In our comparative analysis of model performance, we employed four well-established pre-trained benchmark CNN models: DenseNet121 [33], ResNet50 [34], InceptionV3 [35], and VGG16 [36]. These models underwent training on both augmented datasets for tumour classification and

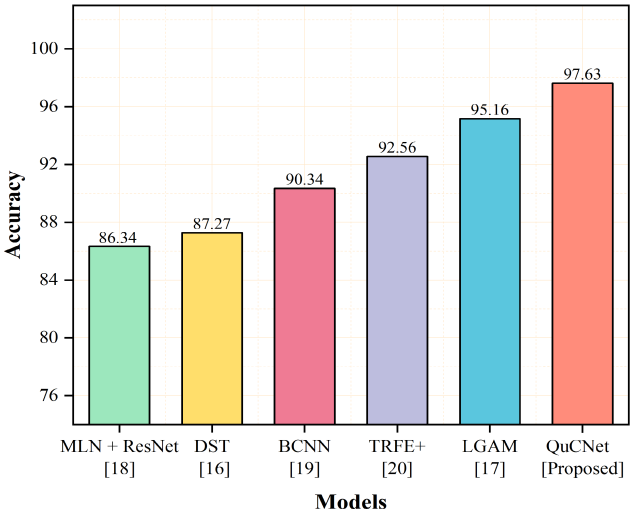
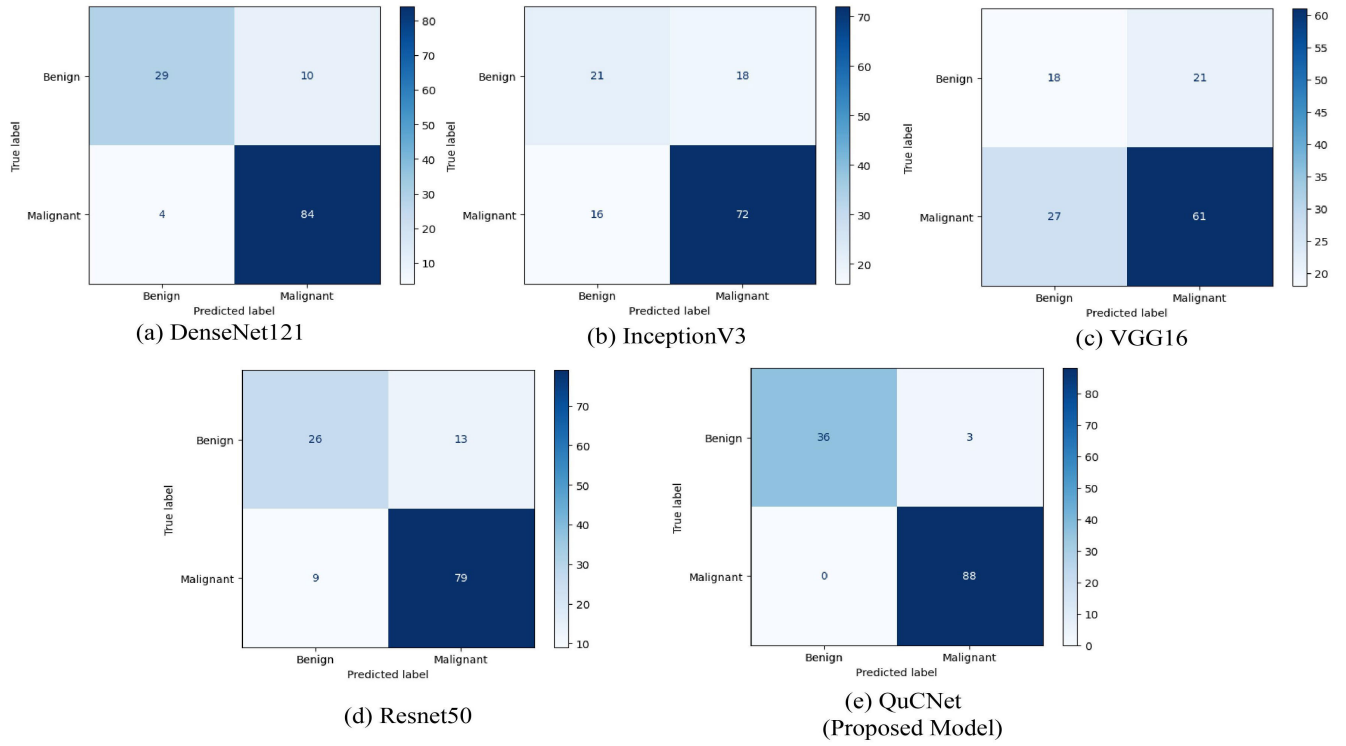


FIGURE 12. Accuracy Comparison of the Proposed QuCNet Model and Existing Approaches.

suspicious level classification. We maintained a consistent training and testing environment for these benchmark models, aligning with the conditions applied to our proposed models. The pre-trained weights of all models were employed, and training involved removing the top layer.

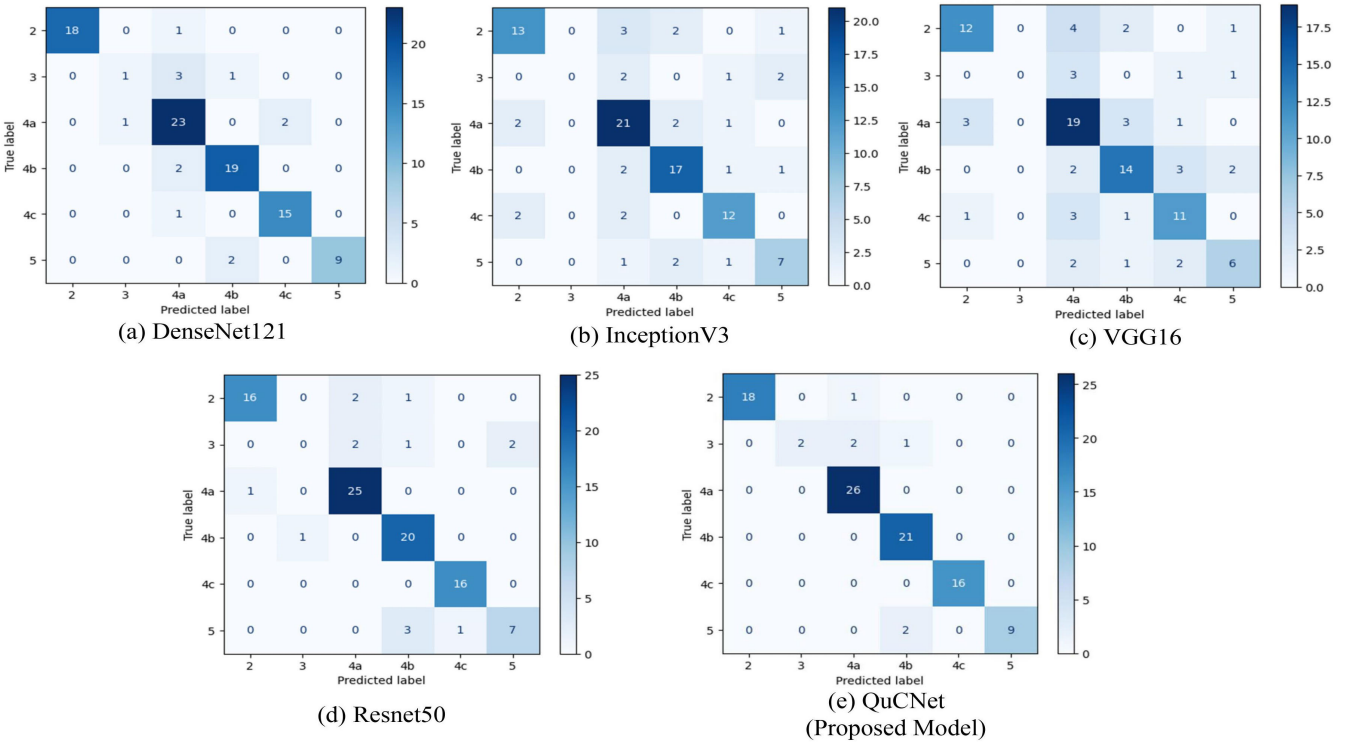
The classification performance on both the tasks are in order, VGG16 < InceptionV3 < ResNet50 < DenseNet121 < QuCNet. Table 8 and Table 10, represents the detailed overall performance metrics including Accuracy, Precision, Recall and F1-Scores of tumour classification and suspicious level classification respectively. Notably, The DenseNet121 has provided an accuracy of 88.98% for the tumour classification and 86.73% for the suspicious level classification. This was the highest among the four CNN models and the closest model to the proposed model on the both classification tasks. In contrast, VGG16 performed poorly for both tasks, demonstrating a substantial 30% difference in accuracy compared to the proposed model. InceptionV3 and ResNet50 exhibited mediocre performance on both classification tasks, lacking notable highlights or drawbacks. To understand the low-level details of model performance, the confusion matrices for tumor classification are presented in Figure 13 and suspicious level classification in Figure 14. Furthermore,



(a) DenseNet121 (b) InceptionV3 (c) VGG16

(d) Resnet50 (e) QuCNet (Proposed Model)

FIGURE 13. Confusion Matrices of Tumor Classification models: DenseNet121, InceptionV3, ResNet50, VGG16, and QuCNet.



(a) DenseNet121 (b) InceptionV3 (c) VGG16

(d) Resnet50 (e) QuCNet (Proposed Model)

FIGURE 14. Confusion Matrices of Suspicious Level Classification models: DenseNet121, InceptionV3, ResNet50, VGG16, and QuCNet.

we undertook an exhaustive class-level analysis utilizing precision, recall, and F-1 score metrics, delineated in Table 7 and Table 9. These outcomes reveal the QuCNet model's robust performance across individual classes. Notably, precision

TABLE 9. Performance metrics of various models on suspicious level classification, including precision, recall, and F1 score. The 'Support' column represents the number of instances in each class.

Class	Metrics	Proposed Model	Densenet121	Resnet50	InceptionV3	Vgg16	Support
2	Precision	**1.00**	1.00	0.94	0.76	0.75	**19**
	Recall	**0.95**	0.95	0.84	0.68	0.63	
	F1 Score	**0.97**	0.97	0.89	0.72	0.69	
3	Precision	**1.00**	0.50	0.00	0.00	0.00	**5**
	Recall	**0.40**	0.20	0.00	0.00	0.00	
	F1 Score	**0.57**	0.29	0.00	0.00	0.00	
4a	Precision	**0.90**	0.77	0.86	0.08	0.58	**26**
	Recall	**1.00**	0.88	0.96	0.81	0.73	
	F1 Score	**0.95**	0.82	0.91	0.74	0.64	
4b	Precision	**0.88**	0.86	0.80	0.74	0.67	**21**
	Recall	**1.00**	0.90	0.95	0.81	0.67	
	F1 Score	**1.00**	0.88	0.87	0.77	0.67	
4c	Precision	**1.00**	0.88	0.94	0.75	0.61	**16**
	Recall	**1.00**	0.94	1.00	0.75	0.69	
	F1 Score	**1.00**	0.91	0.97	0.75	0.65	
5	Precision	**1.00**	1.00	0.78	0.64	0.60	**11**
	Recall	**0.82**	0.82	0.64	0.64	0.55	
	F1 Score	**0.90**	0.90	0.70	0.64	0.57	

scores for both tumor and suspicious level classifications surpass 0.88 for all classes, with several achieving a perfect precision of 1.

From a holistic standpoint, the proposed model demonstrates efficacy even at the class level, consistently achieving the maximum value in each row. Despite the modest dataset size, both models exhibit commendable accuracies exceeding 90%. Noteworthy is the classification of class "3" in the QuCNet model, which, despite a minimal number of samples, exhibits a precision and recall of 1.00 and 0.4, respectively. Although there are some incorrect predictions, this performance is notably superior to the next-best pretrained benchmark model, which has a precision of 0.5 and recall of 0.2, with no other model achieving correct classifications in this particular class.

Upon comparison with our proposed model, these benchmark models exhibited limitations in accurately classifying images, largely stemming from the limited size of the dataset. Despite leveraging image augmentation techniques to enhance their classification capabilities, these models proved insufficient in the end for the pretrained CNN's. Their limitations were attributed to the constrained learning capabilities of convolution and max-pooling operations, which struggled to effectively recognize and assign importance to crucial features. Our proposed QuCNet addresses these limitations by transforming the image into four separate input channels, offering heightened attention to various regions within the image. Furthermore, the results solidifies that the QuCNet model is very effective with the low data classification and has enhanced performance compared to the other models.

TABLE 10. Comparative performance evaluation of proposed model and baseline CNN architectures for suspicious level classification.

Models	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
Propopsed Model	**0.9387**	**0.9458**	**0.9388**	**0.9329**

Limited attention has been devoted to suspicion level classification, in contrast to the plethora of established works focusing on tumor classifications. In Figure 12, we present a comprehensive comparative analysis of the QuCNet model alongside recent existing works. The findings reveal that DST and ResNet incorporated MLN achieved modest accuracies of 87.27% and 86.34%, respectively. Conversely, the BCNN, TRFE+, and LGAM models exhibited slight improvements with accuracies of 90.34%, 92.56%, and 95.16%, respectively. Notably, the proposed QuCNet achieved maximum performance, boasting an accuracy of 97.63%. The results presented in Table 9 and Figure 12 collectively suggest that leveraging quantum computation in conjunction with classical data yields notable benefits, particularly when dealing with limited data. In order to assess the robustness of our tumor classification model, we conducted evaluations using a small-scale dataset titled as the Algerian Ultrasound Images Thyroid Dataset (AUITD) [37]. This dataset is sourced from hospitals in Setif city, Algeria, consists of 352 images meticulously labelled by medical professionals.

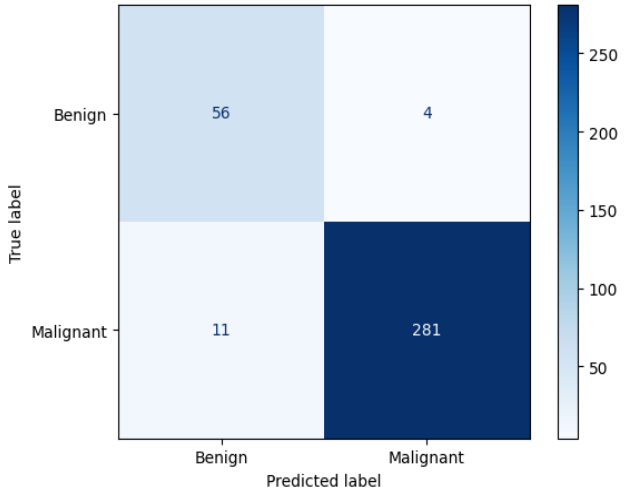


FIGURE 15. Confusion matrix of Proposed Tumor Classifier on AUITD.

AUITD comprises 60 images categorized as Benign class and 292 as Malignant class. Notably, the dataset lacks a Suspicions level classification, prompting our evaluation to solely focus on the Tumour classification model. The proposed QuCNet-based Tumor classification model demonstrated notable performance, achieving an accuracy of 95.74%, Precision of 96.04%, Recall of 95.74%, and F1_Score of 95.83%. The confusion matrix obtained by the testing process is presented in Figure 15. These outcomes serve as a testament to the robustness of the proposed classifier, highlighting its potential applicability in real-world medical analysis scenarios. While the QuCNet exhibits commendable performance in classification tasks, a bottleneck arises from the internal challenges associated with quantum operations in the architecture. The current limitation lies in the unavailability of reliable quantum computers, leading to prolonged computation times when employing hybrid computing for training purposes. Acknowledging the swift advancements in quantum computing technology, the proposed model stands poised to serve as a foundational in the quantum applications.

VIII. CONCLUSION

In conclusion, the classification of Thyroid nodules stands as a pivotal undertaking within the realm of medical image analysis. This study has pioneered the application of quantum-inspired Convolutional Neural Networks to effectively categorize thyroid nodules into benign and malignant classes, as well as assign them to distinct stages of suspicion across six classes. Despite the constraints of a limited dataset, the integration of quantum filter transformation has significantly enhanced the precision of image classification, resulting in accuracy rates of 97.63% and 93.87% for tumour and suspicious level classification, respectively. These achievements surpass the performance of established models like DenseNet121, ResNet50, InceptionV3 and VGG16 in the field, underscoring the efficacy of the proposed approach.

The distinctive feature extraction capabilities of our model enable a new level of precision in diagnosing thyroid abnormalities even with the limited data. In practical applications, this precision serves as a helping hand for doctors, aiding them in confidently characterizing the nature of tumors. In the real-world scenario, the continuous evolution of technology and the growing volume of medical images pose a challenge that demands extensive computational power. Quantum computing, renowned for its high computation speed, holds promise in addressing these challenges once a reliable quantum machine is successfully developed. This study serves as a foundational step towards the utilization of quantum-based approaches in thyroid nodule detection.

Looking ahead, future research should explore the integration of 3D imaging techniques to provide a more comprehensive representation of nodule structure. Furthermore, leveraging advanced quantum computers could significantly reduce the computational costs and time required for model training, leading to faster results and more efficient training processes. Additionally, investigating object segmentation techniques is essential for achieving a clearer and more nuanced understanding of thyroid nodules. By addressing these avenues for improvement, the potential for even more refined and accurate classification models becomes evident. This research contributes to the current understanding of thyroid nodule classification and lays the foundation for further advancements in Quantum medical image analysis, ultimately enhancing the diagnostic capabilities of thyroid pathology.

REFERENCES

[1] *How Does the Thyroid Gland Work?* Accessed: 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/

[2] Y. Hang, "Thyroid nodule classification in ultrasound images by fusion of conventional features and res-GAN deep features," *J. Healthcare Eng.*, vol. 2021, pp. 1–7, Jul. 2021.

[3] E. J. Gomes Ataide, N. Ponugoti, A. Illanes, S. Schenke, M. Kreissl, and M. Friebe, "Thyroid nodule classification for physician decision support using machine learning-evaluated geometric and morphological features," *Sensors*, vol. 20, no. 21, p. 6110, Oct. 2020.

[4] A. Saini, K. Guleria, and S. Sharma, "Machine learning approaches for early identification of thyroid disease," in *Proc. World Conf. Commun. Comput. (WCONF)*, Jul. 2023, pp. 1–6.

[5] S. W. Kwon, I. J. Choi, J. Y. Kang, W. I. Jang, G.-H. Lee, and M.-C. Lee, "Ultrasonographic thyroid nodule classification using a deep convolutional neural network with surgical pathology," *J. Digit. Imag.*, vol. 33, no. 5, pp. 1202–1208, Oct. 2020.

[6] J. Chi, E. Walia, P. Babyn, J. Wang, G. Groot, and M. Eramian, "Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network," *J. Digit. Imag.*, vol. 30, no. 4, pp. 477–486, Aug. 2017.

[7] Y. Zhu, Z. Fu, and J. Fei, "An image augmentation method using convolutional network for thyroid nodule classification by transfer learning," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, Dec. 2017, pp. 1819–1823.

[8] C.-L. Cao, Q.-L. Li, J. Tong, L.-N. Shi, W.-X. Li, Y. Xu, J. Cheng, T.-T. Du, J. Li, and X.-W. Cui, "Artificial intelligence in thyroid ultrasound," *Frontiers Oncol.*, vol. 13, May 2023, Art. no. 1060702.

[9] T. Liu, S. Xie, Y. Zhang, J. Yu, L. Niu, and W. Sun, "Feature selection and thyroid nodule classification using transfer learning," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 1096–1099.

[10] M. Hussain, "Ensemble classifier for benign-malignant mass classification," *Int. J. Comput. Vis. Image Process.*, vol. 3, no. 1, pp. 66–77, Jan. 2013.
[11] T. Khan, "Application of two-class neural network-based classification model to predict the onset of thyroid disease," in *Proc. 11th Int. Conf. Cloud Comput., Data Sci. Eng.*, Jan. 2021, pp. 114–118.
[12] R. Srivastava and P. Kumar, "Optimizing CNN based model for thyroid nodule classification using data augmentation, segmentation and boundary detection techniques," *Multimedia Tools Appl.*, vol. 82, no. 26, pp. 41037–41072, Nov. 2023.
[13] J. Sun, B. Wu, T. Zhao, L. Gao, K. Xie, T. Lin, J. Sui, X. Li, X. Wu, and X. Ni, "Classification for thyroid nodule using ViT with contrastive learning in ultrasound images," *Comput. Biol. Med.*, vol. 152, Jan. 2023, Art. no. 106444.
[14] Y. Wang and J. Gan, "Benign and malignant classification of thyroid nodules based on ConvNeXt," in *Proc. 3rd Int. Conf. Control, Robot. Intell. Syst.*, Aug. 2022, pp. 56–60.
[15] F. Jerbi, N. Aboudi, and N. Khlifa, "Automatic classification of ultrasound thyroids images using vision transformers and generative adversarial networks," *Sci. Afr.*, vol. 20, Jul. 2023, Art. no. e01679.
[16] N. Baima, T. Wang, C.-K. Zhao, S. Chen, C. Zhao, and B. Lei, "Dense Swin transformer for classification of thyroid nodules," in *Proc. 45th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2023, pp. 1–4.
[17] L. Yu, S. Qu, Z. Cong, and X. An, "Ultrasound image classification of thyroid nodules based on attention mechanism," *J. Phys., Conf. Ser.*, vol. 2637, no. 1, Nov. 2023, Art. no. 012048.
[18] Z. Zheng, T. Su, Y. Wang, Z. Weng, J. Chai, W. Bu, J. Xu, and J. Chen, "A novel ultrasound image diagnostic method for thyroid nodules," *Sci. Rep.*, vol. 13, no. 1, p. 1654, Jan. 2023.
[19] N. Aboudi, H. Khachnaoui, O. Moussa, and N. Khlifa, "Bilinear pooling for thyroid nodule classification in ultrasound imaging," *Arabian J. Sci. Eng.*, vol. 48, no. 8, pp. 10563–10573, Aug. 2023.
[20] H. Gong, J. Chen, G. Chen, H. Li, G. Li, and F. Chen, "Thyroid region prior guided attention for ultrasound segmentation of thyroid nodules," *Comput. Biol. Med.*, vol. 155, Mar. 2023, Art. no. 106389.
[21] J. Zhu, S. Zhang, R. Yu, Z. Liu, H. Gao, B. Yue, X. Liu, X. Zheng, M. Gao, and X. Wei, "An efficient deep convolutional neural network model for visual localization and automatic diagnosis of thyroid nodules on ultrasound images," *Quant. Imag. Med. Surg.*, vol. 11, no. 4, pp. 1368–1380, Apr. 2021.
[22] L. Wei, H. Liu, J. Xu, L. Shi, Z. Shan, B. Zhao, and Y. Gao, "Quantum machine learning in medical image analysis: A survey," *Neurocomputing*, vol. 525, pp. 42–53, Mar. 2023.
[23] D. Maheshwari, U. Ullah, P. Marulanda, A. García-Olea, I. Gonzalez, J. Merodio, and B. Zapirain, "Quantum machine learning applied to electronic healthcare records for ischemic heart disease classification," *Hum.-Cent. Comput. Inf. Sci*, vol. 13, p. 17, Feb. 2023.
[24] H. Xiong, X. Duan, Y. Yu, J. Zhang, and H. Yin, "Image classification based on quantum machine learning," in *Proc. 5th Int. Conf. Intell. Control, Meas. Signal Process. (ICMSP)*, May 2023, pp. 891–895.
[25] L. Pedraza, C. Vargas, F. Narváez, O. Durán, E. Muñoz, and E. Romero, "An open access thyroid ultrasound-image database," *Proc. SPIE*, vol. 9287, pp. 188–193, Jan. 2015, doi: 10.1117/12.2073532.
[26] F. N. Tessler, W. D. Middleton, E. G. Grant, J. K. Hoang, L. L. Berland, S. A. Teefey, J. J. Cronan, M. D. Beland, T. S. Desser, M. C. Frates, L. W. Hammers, U. M. Hamper, J. E. Langer, C. C. Reading, L. M. Scoutt, and A. T. Stavros, "ACR thyroid imaging, reporting and data system (TI-RADS): White paper of the ACR TI-RADS committee," *J. Amer. College Radiol.*, vol. 14, no. 5, pp. 587–595, May 2017.
[27] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.
[28] M. Henderson, S. Shakya, S. Pradhan, and T. Cook, "Quanvolutional neural networks: Powering image recognition with quantum circuits," *Quantum Mach. Intell.*, vol. 2, no. 1, pp. 1–9, Jun. 2020.
[29] X. Gao, M. Krenn, J. Kysela, and A. Zeilinger, "Arbitrary D-dimensional Pauli X gates of a flying qudit," *Phys. Rev. A, Gen. Phys.*, vol. 99, no. 2, Feb. 2019, Art. no. 023825, doi: 10.1103/physreva.99.023825.
[30] M. Treinish, "Qiskit/qiskit-metapackage: Qiskit 0.44.0," Version 0.44.0, Zenodo, Jul. 2023, doi: 10.5281/zenodo.8190968.
[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[32] A. Fred Agarap, "Deep learning using rectified linear units (ReLU)," 2018, *arXiv:1803.08375*.
[33] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269.
[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[35] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
[36] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.
[37] A. Maroua. *Algerian Ultrasound Images Thyroid Dataset: Auitd*. Accessed: 2023. [Online]. Available: https://www.kaggle.com/datasets/azouzmaroua/algeria-ultrasound-images-thyroid-dataset-auitd

GUNTI SWATHI is currently pursuing the bachelor's degree with the Amrita School of Computing, Chennai. She has amassed four years of valuable experience in the field of computer science. During this time, she has distinguished herself as an adept expert, showcasing proficiency in various areas of computer science, with a specific focus on machine learning and deep learning. She is actively engaged in various research projects, demonstrating her commitment to exploring and contributing to the cutting-edge advancements in the field.

ALI ALTALBE received the M.Sc. degree in information technology from Flinders University, Australia, and the Ph.D. degree in information technology from The University of Queensland, Australia. He is currently an Associate Professor with the Department of IT, King Abdulaziz University, Jeddah, Saudi Arabia.

R. PRASANNA KUMAR is currently the Chairperson and an Associate Professor with the Amrita School of Computing, Chennai. With 22 years of experience in the field of teaching, he has established himself as an Expert in several areas of computer science. His research interests include data analytics, machine learning, the theory of computation, compiler design, and python programming. He is also supervising eight Ph.D. research scholars and he has successfully guided eight M.E. projects. Additionally, he serves as a Doctoral Committee Member for ten Ph.D. research scholars. He has been invited as a Guest Lecturer and a resource person for Faculty Development Programs, and he has also contributed as a program committee member and the session chair in various national and international conferences.