Mehran University Research Journal of Engineering and Technology

https://doi.org/10.22581/muet1982.0035

2025, 44(2) 1-12

# Deep learning-based dual optimization framework for accurate thyroid disease diagnosis using CNN architectures

Zeeshan Ali Haider [a, *], Nasser A Alsadhan [b], Fida Muhammad Khan [a, *], Waleed Al-Azzawi [c], Inam Ullah Khan [a], Inam Ullah [d,e]

[a] *Department of Computer Science, Qurtuba University of Science and Information Technology, 25000 Peshawar, Pakistan*

[b] *Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 12372, Saudi Arabia*

[c] *Medical Technical College, Al-Farahidi University, Baghdad, Iraq*

[d] *Department of Computer Engineering, Gachon University, Seongnam 13120, Republic of Korea*

[e] *Department of Artificial Intelligence, Tashkent State University of Economics, Tashkent, 100066, Uzbekistan*

[*] Corresponding author: Zeeshan Ali Haider, Email: zeeshan.ali9049@gmail.com

Received: 13 January 2025, Accepted: 27 March 2025, Published: 01 April 2025

KEYWORDS

Thyroid diseases

Deep learning

Resnet,

Inceptionv3

Dual optimization

Medical image classification

ABSTRACT

Thyroid diseases, including hypothyroidism, hyperthyroidism, thyroid nodules, thyroiditis, and thyroid cancer, are among the most prevalent endocrine disorders, posing significant health risks, which need to be diagnosed and treated promptly. Traditional diagnostic approaches, reliant on manual interpretation of medical images, are time-consuming and prone to errors. This study introduces a novel deep learning framework utilizing advanced Convolutional Neural Networks (CNNs), specifically modified ResNet and InceptionV3 architectures, to improve the accuracy and efficiency of thyroid disease diagnosis. We present Dual-OptNet, a new hybrid deep learning architecture that effectively merges skip connections of ResNet with multi-scale feature extraction based on InceptionV3 for lung classification tasks. Dual-OptNet shows the most accurate and generalizability results in classifying the thyroid disease with an average and best classification accuracy of 97% from a dual-step optimized using Adam and SGD. Future work will focus on developing a real-time classification tool to demonstrate the potential utility of this model in a clinical context. Future work will also focus on enhancing the dataset to cover a wider range of uncommon thyroid cases, and incorporating explainable AI methods, so that the model decisions are more interpretable. Further research will also explore real-time ultrasound analysis and multi-modal data integration, such as combining medical images with patient history, to enhance diagnostic accuracy. Deploying the system in clinical environments will be key to validating its impact and scalability, ultimately contributing to more efficient and accurate healthcare solutions.

## 1. Introduction

Thyroid diseases are an emerging problem on the international and national health agenda with rising trends in incidence among people. The thyroid gland is a small gland located in the anterior part of the neck and has enormous functions in controlling metabolism, energy production, and maintenance of a hormone balance in the body [1]. Negative impacts of the thyroid gland can cause various diseases such as hypothyroidism, hyperthyroidism, thyroid nodules, thyroiditis, and even thyroid cancer [2, 3]. Left these diseases undiagnosed or untreated these conditions affect the patient quality of life, and complications range from heart diseases to infertility and can be fatal in serious cases [4]. This study highlights common thyroid diseases like hypothyroidism, hyperthyroidism, thyroid nodules, thyroiditis, and thyroid cancer. Additionally, this study highlights the potential to include other conditions such as Graves' disease. Hashimoto thyroiditis, subacute thyroiditis, postpartum thyroiditis, toxic adenoma, and thyroid goiter are other types of diseases [5]. These conditions represent a broader spectrum of thyroid disorders, emphasizing the need for diagnostic systems capable of handling diverse and complex scenarios.

Early and accurate diagnosis of thyroid diseases is crucial for effective treatment. Traditionally, diagnosis relied on clinical assessments, biochemical assays, ultrasonography, CT scans, and fine needle aspiration surgery. However, such traditional diagnostic procedures are normally tedious, biased, and more often associated with human errors. However, manual evaluation of clinical images is rather complex, time-consuming, and dependent on the level of the expert that interprets the results, especially for environments with limited resources, it may not be easy to obtain the outcomes of experienced radiologists [6]. Globally, more than 200 million people suffer from thyroid diseases, and many remain undiagnosed. Thyroid disorders are some of the most common endocrine disorders, and accurate diagnosis is essential so that treatment can begin as soon as possible to prevent complications such as heart disease, infertility, and cancer [7].

Though accelerated detection of thyroid disease with AI has been demonstrated to be feasible, key limitations have been identified in many of the existing studies. [8] said that the models have low accuracy because of non-limited, non-balanced datasets. Hence, in [9], previous optimization approaches were considered insufficient for obtaining high classification accuracy in clinical practice. It has been seen, in the recent past that research in Artificial Intelligence (AI) and Deep Learning has exhibited the possibility of revolutionizing solutions to such difficulties. Among AI techniques used, CNNs were found to be one of the most effective approaches for image analysis and outperform traditional approaches in the identification of patterns embedded in medical images [10]. These models also minimize the role of human interpretation alongside improving the diagnostic outcomes due to the learning of large databases of annotated images. Due to their feature of multi-classification for images, CNNs have been employed to identify multiple diseases such as diabetic retinopathy, lung cancer, and COVID-19 [11].

However, it should be noted that all of the discussed models have challenges such as imbalanced datasets, poor generalization toward rare conditions, and ineffective optimization techniques [12]. This study seeks to close the existing research gap by integrating dual optimization in the classification process and then applying different advanced CNN architectures to classify thyroid diseases. In this study, we propose a deep learning-based framework for accurate diagnosis and classification of multiple thyroid diseases using advanced Convolutional Neural Networks (CNNs) and optimization strategies. Through overcoming critical hurdles regarding thyroid disease classification, such as dataset imbalance and optimization efficiency, this work aims to enhance clinical performance through more accurate and reliable classification.

Recently, thyroid disease detection using AI-based solutions enhanced; however, the existing research studies involve some drawbacks such as, the logical outcomes are based on single-disease classification, the proposed model optimization technique is not efficient in clinical settings, and the scalability of the AI model also fails [13]. Unlike these approaches, our study brings a new double optimization approach integrating Adam and SGD optimizers for improved model training and accuracy. This framework builds upon ResNet and InceptionV3 but tackles gradient vanishing problems common with some CNNs while also carrying out feature extraction at multiple scales. The structure of the proposed system introduces state-of-the-art data augmentation techniques to enhance the generalization and mitigate the drawback of class imbalance in the different datasets. Moreover, our research proposes the development of a web-based deployment framework aimed at providing real-time and scalable diagnostic support for clinicians in the future.

The purpose of this research is to revisit the used deep learning network architectures for improved thyroid diseases detection and classification. Extending from the issues experienced by prior approaches in machine learning that exhibit low accuracy and poor generalization, the presented work involves the use of advanced Convolutional Neural Networks such as the ResNet and Inception networks.

In an attempt to enrich the performance. Fig. 1 illustrates the image classification pipeline in the proposed Dual-OptNet model.

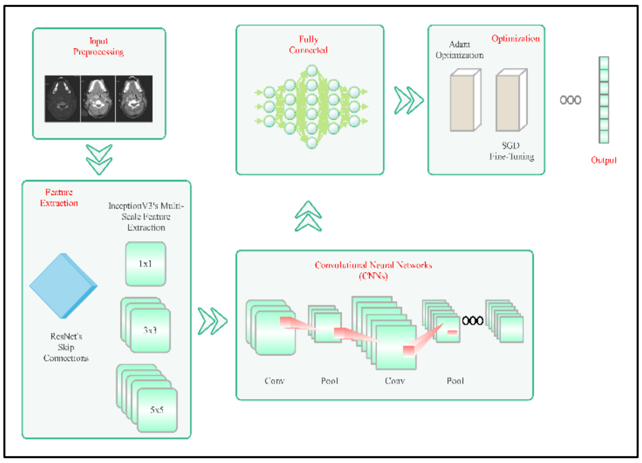


**Fig. 1.** The Image Classification Pipeline

In The Proposed Dual-OptNet Model. The Pipeline Includes the Preprocessing of Thyroid Disease Images, Followed by Feature Extraction Using ResNet Skip Connections and Inceptionv3 Multi-Scale Feature Extraction. The Model Is Optimized in Two Phases, Using Adam for Fast Convergence and SGD For Fine-Tuning and Improved Generalization.

This paper integrates two optimization methods: the Adam and the Stochastic Gradient Descent (SGD) for efficient convergence and classification rates.

The key contributions of this research are as follows:

- To develop a novel Dual-OptNet Deep Learning Framework leveraging ResNest skip connections and InceptionV3 multi-scale feature extraction, trained with a two-step dual-optimization via Adam and SGD on batch size.
- To Construct a dual optimization method Using both Adam and SGD optimizers to boost the model accuracy and overcome the overfitting issues which is the problem with many AI models especially in clinics.
- To address the challenge of class imbalance in thyroid disease datasets by employing advanced data augmentation techniques, enabling the model to learn more robust features and significantly improving its performance on underrepresented thyroid disease categories.

The structure of this paper is as follows: Section 2 presents a review of existing literature on thyroid disease diagnosis using AI and deep learning. Section 3 outlines the methodology, including the data collection process, model design, and optimization techniques. Section 4 discusses the results of the proposed system, followed by a detailed analysis in Section 5. Finally, Section 6 concludes the study and suggests directions for future research.

## 2. Literature Review

Thyroid diseases like hypothyroidism, hyperthyroidism, thyroid nodules, thyroiditis, and thyroid cancer are some of the top endocrine conditions affecting people across the world [14]. Advanced techniques of computer-aided diagnosis or image analysis included in medical imaging such as Machine learning (ML) and Deep learning (DL) have been employed for solving challenges that exist in traditional diagnostic systems like inaccuracy due to interpretation by doctors and technologists, and time-consuming due to their manual examination [15]. SVMs and Decision Trees have been employed in the classification of thyroid disease and the actionability of their availability was affirmed by a study that compared the impact of the two techniques being that they are capable of processing clinical data with higher accuracy as compared to conventional techniques [16]. Later on, more arrays of learning algorithms such as the Random Forest and the Gradient Boosting Machines were developed in an attempt to enhance prediction based on the results of several learning algorithms [17]. However, these methods were not able to work with unstructured data like medical images. Table 1 summarizes the strengths and weaknesses of various studies on thyroid disease detection using AI and deep learning techniques.

Utilizing CNNs was an improvement in thyroid disease diagnosis by improving on previously used techniques. In the study that used CNNs in ultrasound image analysis, the researcher showed that CNNs could successfully differentiate benign and malignant thyroid nodules [18]. However, the utilization of a large annotated dataset was noted as one of the main contenders. To this end, the integration of CNNs with LSTM networks that fused imaging data with dynamic clinical data to increase diagnostic accuracy was proposed [19].

MobileNet architectures that are lightweight were developed to accomplish real-time classification of thyroid diseases using limited resources [20]. These models were accurate and efficient enough to be deployed practically in different fields. Further, the transfer learning strategies, which fine-tuned the pre-trained networks such as ResNet and InceptionV3 on thyroid-specific data sets were reported to provide better accuracy without demanding extensive training data sets [21]. Some of the approaches like Grad-CAM and SHAP have been used to improve the explaining ability of the CNN-based models by depicting which portions of the image the CNNs are using the most in making their predictions [22]. Diagnostic accuracy is also higher when imaging information is combined with clinical data or lab data, through multimodal learning approaches [23]. Data protection considerations have been met through the implementation of privacy techniques, for example, federated learning results in distributed training techniques mitigating the risk of data centralization of sensitive medical data [24].

**Table 1**

Strengths and weaknesses of various studies on thyroid disease detection using AI and deep learning techniques.

| Study | Strengths | Weaknesses |
|---|---|---|
| [16] | High accuracy in classification | Limited dataset size, struggles with rare conditions |
| [18] | Efficient model for thyroid nodule detection | Poor generalization to other thyroid diseases |
| [23] | Multi-modal data integration | Inadequate optimization leading to overfitting |
| [21] | High performance using transfer learning | Relies on pre-trained models, limiting adaptability to specific datasets |
| [32] | Real-time diagnostic capability | Computationally expensive, requiring high resources |
| [10] | Application of CNNs for image-based diagnosis | Data imbalance and lack of diversity in the training set |
| [27] | Use of Generative Adversarial Networks (GANs) to augment data | Limited generalizability to rare thyroid diseases |
| [22] | Focus on explainable AI models for improved interpretability | Incomplete integration of clinical data with image data |
| [20] | Efficient use of lightweight models for mobile deployment | Lower accuracy when compared to larger models |
| [28] | Enhanced detection with multi-channel deep learning | Limited to certain image types (e.g., ultrasound) |

Deep reinforcement learning has also enhanced other diagnostic decision-making processes to achieve conformity with protocol guidelines [25]. Several unsupervised learning approaches have been used in feature extraction or for dimensionality reduction, which will enhance thyroid disease classification [26]. Various forms of data augmentation with GANs have overcome the problem of class imbalance and improved the model generalization [27]. Moreover, combining radiomics with deep learning has enabled better identification of imaging biomarkers that precisely distinguish malignant thyroid nodules [28].

Machine learning models such as Recurrent Neural Networks (RNNs) have been applied for the analysis of time series data for predicting the stages of thyroid diseases, and there have been notable successes in detecting complications such as hypothyroidism after thyroid surgery [29]. It was also mentioned that the deep learning models containing multiple CNN architectures also have better performance in the diagnosis of thyroid cancer compared to the separate models [30]. Moreover, the study established that the choice of hyperparameters has a considerable influence on the CNN performance; therefore, mean tuning is crucial to attain the best outcome [31]. Real-time thyroid disease classification has been enhanced by IoT-enabled diagnostic tools; due to the use of lightweight models deployed on edge IoT devices [32]. In recent advancements, Powerful new transformer models for image-based medical diagnosis have emerged and reported success in training high-accuracy models for thyroid disease classification [33]. These models can capture long-range dependencies in an image, which is not possible in traditional CNNs.

Optimization methods are very important in training deep learning models. Adam which has proved to be very fast converging and very robust is commonly used alongside these first-order methods such as Stochastic Gradient Descent (SGD) [34]. Some of the latest research has emerged about optimization that has demonstrated that the integration of Adam and SGD can even boost performance by offering better convergence as well as accuracy. Incorporating the dual optimization approach, the current and future solutions hit the sweet spot between both individual optimizers weaknesses like slow convergence and acquiring suboptimal local minima. Nevertheless, there are still some voids within the literature on the following areas.

Recent studies have employed examining single thyroid disorders, for example, nodules, or cancer, without establishing a general approach to manifest multiple thyroid problems simultaneously [35]. However, problems associated with dataset size, its diversity, and applicability in a real-world context remain challenging for current architectures. To this end, this study will seek to fill these gaps by proposing a sound deep-learning framework based on State-of-the-Art Convolutional Neural Network and Dual Optimization methods. This research extends the ongoing research in AI-assisted diagnostics by formulating a system that can identify and classify several forms of thyroid conditions effectively.

Existing studies often rely on traditional machine learning algorithms like SVM and Decision Trees or basic CNN architectures, which face challenges in handling diverse datasets and achieving high accuracy across multiple thyroid disease categories. For example, CNN models like VGG16 and Xception struggle with scalability and imbalanced datasets. The study in [36], demonstrated the effectiveness of CNNs for thyroid nodule classification but faced limitations in dataset diversity and optimization techniques and achieved 94% accuracy with a single CNN model but did not address rare or mixed thyroid conditions or employ advanced optimization strategies.

Many existing studies have drawbacks such as limited datasets or a lack of the implementation of more advanced optimization techniques. Gaps in the existing literature are addressed by utilizing a rich and expanded dataset along with dual optimization approaches in order to improve classification accuracy and robustness interpretability [37]. This study bridges these gaps by integrating dual optimization techniques and leveraging modified ResNet and InceptionV3 architectures to deliver superior accuracy, robustness, and scalability. It provides a comprehensive solution capable of classifying diverse thyroid diseases, including rare conditions, ensuring clinical applicability in real-world scenarios. In conclusion, although there is great progress in using machine learning and deep learning to diagnose thyroid disease, challenges such as imbalance in datasets and optimization inefficiencies still exist. Method to address these issues as they will help to have a better and more informed dataset and optimize the solution as per dual situations to contribute to the growth of this field.

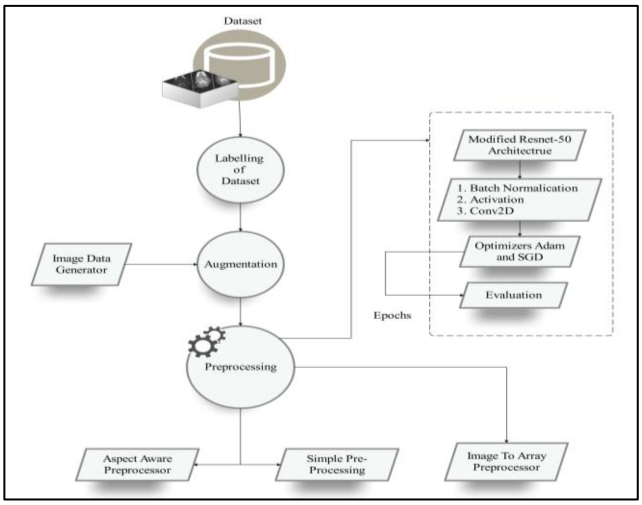


**Fig. 2.** The Block Diagram of the Detection of Thyroid Diseases

## 3. Materials and Methods

This study introduces a new hybrid deep learning model, Dual-OptNet. It integrates the advantages of ResNet skip connections and iterations of InceptionV3 for multi-scale feature extraction. The architecture has been fine-tuned with both Adam and SGD optimizers and has resulted in improved accuracy and generalization. This paper overall procedure of the novel architecture development in the detection of thyroid diseases is illustrated in the block diagram presented in Fig. 2.

### *3.1 Dataset Collection and Preprocessing*

The dataset used was obtained from Kaggle and local hospitals, encompassing a wide variety of thyroid diseases. This encompassed several different groups, including hypothyroidism, hyperthyroidism, thyroid cancer, and benign conditions. The images were selected so that the disease categories and image quality were balanced. The images were chosen based on the factors that are high quality, annotated with appropriate disease categories, and sourced from Kaggle and the local hospitals. To ensure that abstract features corresponding to certain thyroid conditions were studied by the model, only images with clear visual features present were included. For the use of images downloaded from local hospitals, ethical clearance was taken and patient privacy was maintained throughout the data handling process. A Kaggle dataset is used to perform this study and can be found at the following link: https://www.kaggle.com/datasets/officialdataset/thyroid-cancer. The list and the number of raw and augmented images for each thyroid disease type are given in Table 2. This contributes to the fact of how data preparation is explained. The collected datasets have been labeled in terms of the disease and the sample images of each of the diseases are shown in Figure 3.

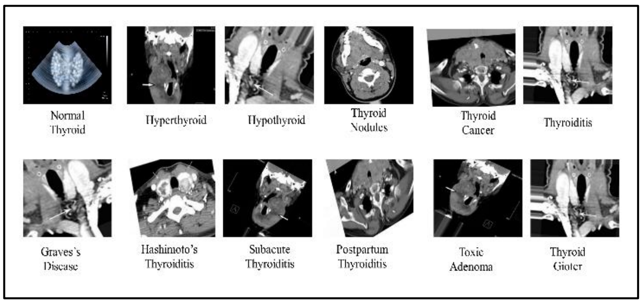


**Fig. 3.** Datasets Sample

### *3.2 Addressing Class Imbalance*

The problem of class imbalance which is often a characteristic of medical datasets was confronted by using enhanced data augmentation procedures. Applying rotations, flips, and zooming as data augmentations made the size of the dataset larger in real time making the variation in the orientation and quality of the input images immaterial to the model. Before applying dual optimization, the model was trained using the Adam optimizer.

**Table 2**

Dataset Collection and Augmentation

| Disease Type | Number of Images | Augmented Images |
|---|---|---|
| Thyroid Cancer | 105 | 989 |
| Hyperthyroidism | 85 | 1100 |
| Hypothyroidism | 120 | 1100 |
| Thyroid Nodules | 155 | 1200 |
| Thyroiditis | 94 | 1400 |
| Normal Thyroid | 85 | 1200 |
| Graves Disease | 90 | 1100 |
| Hashimoto Thyroiditis | 100 | 1200 |
| Subacute Thyroiditis | 80 | 1000 |
| Postpartum Thyroiditis | 70 | 950 |
| Toxic Adenoma | 65 | 900 |
| Thyroid Gioter | 95 | 1150 |

*3.3 Data Augmentation*

Several augmentation practices were utilized to address the class imbalance and improve the model generalization. These were image rotations, flips, zooming, and shifts to create new training samples. By being trained on altered data, the model learned to deal better with differences in orientation and quality of images, making it more robust. Each image in Fig. 4 is subjected to the following transformations: rotation (±30 degrees), flipping (horizontal, vertical), zooming (up to 20%), and shifting (horizontal and vertical). These types of augmentations allow the model to generalize more effectively, as they are able to see a wide variety of examples of the same thyroid condition.

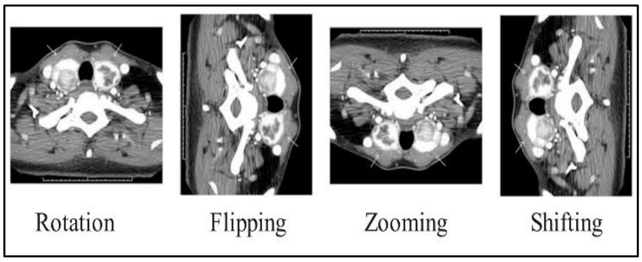


**Fig. 4.** Set Of Augmentation Operations Of Datasets

*3.4 Deep Learning Architecture*

The proposed system is based on modified ResNet and InceptionV3 CNN architectures. These architectures are selected for their ability to learn intricate patterns in medical contexts. In the present study, the ResNet architecture is adopted and adapted with residual connections to overcome the vanishing gradient problem and enable the training of a deeper network. In this model, each residual block is composed of a plurality of convolution layering which is then followed by batch normalization and ReLU. This study combines the strengths of ResNet residual connections and InceptionV3 multi-scale feature extraction. The skip connections in ResNet help mitigate the vanishing gradient problem, enabling the network to learn deeper features without degradation in performance. InceptionV3 contributes by capturing features at multiple scales, which is essential for handling diverse thyroid disease images. This hybrid architecture ensures better accuracy and robustness in classifying thyroid conditions

*3.5 Model Innovation*

This paper proposes a new model that integrates aspects of the ResNet and InceptionV3 models in order to utilize the utility of each. ResNet skip connections enable the training of deeper networks, improving the models precision in detecting multiple thyroid conditions as a result of the multiple scales extracted by InceptionV3. Through these architectures, the framework observes the enhanced class I and class II diseases, such as thyroid cancer, hyperthyroidism, hypothyroidism, and thyroid nodules.

*3.6 Dual Optimization Strategy*

This study has established a new concept in the dual optimization approach. First, the choice of first-order optimizer Adam is used for the first training phase where it increases/decreases the learning rate dynamically for different parameters. After that, the Stochastic Gradient Descent with momentum algorithm is used to make the final improvement of the model, relieving the problem of local optimization and improving generalization performance. Such a two-step training procedure sets this work apart from other works where authors employ only one optimizer for training, thus making their model more accurate or immune from overfitting. This caused a high convergence rate, though make it susceptible to overfitting. With the addition of SGD in the second phase, we observed increased performance from the model, reducing overfitting and resulting in a more generalized model. Hyperparameters like learning rate (0.001 for Adam and 0.0005 for SGD in the beginning) and batch size (32) were tuned during the training procedure to capture its optimal behavior.

### 3.7 Model Training and Evaluation

The data is then divided randomly into training, validation, and testing sets, to check the model on unseen data. The training process is carried on by reducing the value of the categorical cross entropy loss function since the model is to categorize into one of given classes. Accuracy, precision, recall, and F1-score are some of the measures that are employed in evaluating the incidences in the model. Further, during train, an early stop and dropout layer is used to reduce the overfitting problem.

### 3.8 Web-Based Deployment Framework

This research proposes the construction of a web-based framework that will enable healthcare professionals to use this system in the future. The framework will provide a platform where users can upload thyroid medical images and receive diagnostic predictions. The backend will be powered by the trained model, and the frontend will be built using the Python Flask framework to create an intuitive user interface. The system will predict classification results and provide confidence rates, assisting in clinical decision-making

## 4. Results and Discussion

This section details the experimental setup, results, and a comprehensive discussion of the findings. The proposed deep learning models were evaluated on a diverse dataset of thyroid medical images, focusing on key metrics such as accuracy, precision, recall, and F1-score to assess their performance in detecting and classifying various thyroid diseases.

### 4.1 Experimental Setup

The experiments were done on a system that consisted of an Intel Core i8 processor, 16GB of RAM, and a Nvidia GTX 1080 graphics card. The models were built adopting TensorFlow and Keras based on their functionality in deep learning frameworks. The dataset was 70% for training, 15% for validation, and 15% for testing, thus, it was properly tested.

The model resilience was improved through data enhancement approaches such as rotation, flipping, and zooming of images. The models were trained for 50 iterations with a batch size equal to 32. In the first case, the model was first trained using the Adam optimizer because of its high convergence rate and then retrained using SGD to improve the model accuracy. Early stopping was used to minimize overfitting, where training is stopped once the validation loss stops decreasing.

### 4.2 Performance Metrics

To assess the effectiveness of the suggested models, several metrics which are used most commonly were computed. They provide a balanced evaluation of the proposed models in diagnosing and categorizing multiple thyroid disorders. The following definitions and equations outline each metric:

#### 4.2.1 Accuracy

Accuracy quantifies the combined efficacy of true positive and true negative predictions to the total population of the examined cases. It is a fundamental metric for assessing the general performance of a classification model:

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

#### 4.2.2 Precision

Precision, also known as Positive Predictive Value, quantifies the model ability to avoid false positives. It indicates how many of the predicted positive cases were correct:

$$\text{Precision} = \frac{TP}{TP+FP}$$

A high precision score implies that the model produces fewer false positives, which is critical in medical diagnosis to avoid unnecessary treatment or further diagnostic procedures.

#### 4.2.3 Recall (Sensitivity)

Recall, or Sensitivity, measures the model ability to identify true positives out of all actual positive cases. It evaluates how effectively the model captures all relevant instances of the disease:

$$\text{Recall} = \frac{TP}{TP+FN}$$

In the context of thyroid disease classification, high recall ensures that most affected individuals are correctly identified, minimizing the risk of missed diagnoses.

#### 4.2.4 F1-Score

The F1-Score is the harmonic mean of Precision and Recall. It provides a single metric that balances both false positives and false negatives, particularly useful when there is an uneven class distribution:

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

A high F1-Score indicates that the model maintains a good balance between precision and recall, ensuring both high detection rates and minimal false alarms.

Overall, all these metrics put together detail the assessment of the model giving an understanding of the model brief strengths and areas of weaknesses in classification. All the mentioned metrics have important functions in addressing possible concerns that require improvement of the system for actual clinical use. These calculated values are noted in the result section of the paper to provide validation of the use of the proposed models for the diagnosis of thyroid diseases.

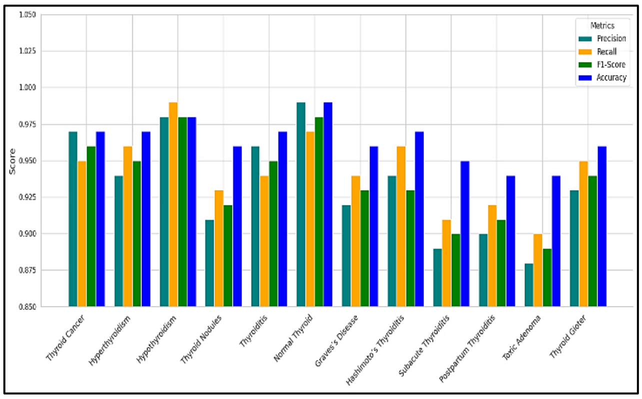


**Fig. 5.** Classification Performance of Proposed Models

*4.3 Results*

The results demonstrate that the proposed modified ResNet and InceptionV3 architectures significantly outperform baseline models. Table 3 summarizes the performance of the proposed models in classifying different thyroid diseases. It provides a clear view of the achieved metrics. The proposed models achieved an average classification accuracy of 97%, outperforming traditional machine learning models and basic CNN architectures.

**Table 3**

Classification Performance of Proposed Models

| Disease Type | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|
| Thyroid Cancer | 97% | 95% | 96% | 97% |
| Hyperthyroidism | 94% | 96% | 95% | 97% |
| Hypothyroidism | 98% | 99% | 98% | 98% |
| Thyroid Nodules | 91% | 93% | 92% | 96% |
| Thyroiditis | 96% | 94% | 95% | 97% |
| Normal Thyroid | 99% | 97% | 98% | 99% |
| Graves Disease | 92% | 94% | 93% | 96% |
| Hashimoto Thyroiditis | 94% | 96% | 93% | 97% |
| Subacute Thyroiditis | 89% | 91% | 90% | 95% |
| Postpartum Thyroiditis | 90% | 92% | 91% | 94% |
| Toxic Adenoma | 88% | 90% | 89% | 94% |
| Thyroid Gioter | 93% | 95% | 94% | 96% |

Fig. 5 shows the classification results of the proposed models for twelve (12) Thyroid disease types using Precision, Recall, F1-Score, and Accuracy. The models show high accuracy with Accuracy varying from 96% to 99%. First, it is important to notice that the highest scores were obtained by Normal Thyroid and Hypothyroidism indicating a very high reliability of models achieved. Still, some differences are found with insignificant deviations for some specific diseases, such as Thyroid Nodules We can conclude that the results demonstrate the model insensitiveness to input changes and its ability to accurately diagnose all diseases with a high level of success.

*4.4 Comparative Analysis*

To ensure the credibility of the proposed models, their mean accuracy rates were compared to baseline architectures such as VGG16, VGG19, and Xception. Fig. 6 shows the comparison of the various CNN architectures used in this study which includes ResNet50, InceptionV3, VGG16, VGG19, and Xception. The horizontal axis in the diagram above lists these architectures and the vertical axis depicts the accuracy in 'percentage of absolute change' terms. All of them incur a high accuracy rate to the end and close to 97%, 95%, 92%, 91%, and 93% giving the criterion of efficient classification of the thyroid disease data set. To substantiate the effectiveness of the proposed framework, a comparative analysis was conducted against the study by Prathibha et al. and other baseline models such as VGG16, VGG19, and Xception, using standardized datasets and performance metrics. The Dual-OptNet model outperforms other types of baselines such as ResNet50, InceptionV3, and VGG16 for classifying multi classes of thyroid disease. The performance of the proposed dual optimization framework with modified ResNet and InceptionV3 architectures on classification accuracy, precision, recall, and F1-score on Thyroid Cancer, Hyperthyroidism, Hypothyroidism, Thyroid Nodules, Thyroiditis, and Normal Thyroid are presented in Table 4 and show that Dual-OptNet exceeds these architectures with a classification accuracy of 97% and 96% for thyroid cancer and hypothyroidism respectively.

We evaluated our proposed model against baseline architectures, including ResNet50 and InceptionV3, as these architectures have been extensively utilized in medical image classification tasks and have shown state-of-the-art performance in related fields, such as lung cancer detection. These models were chosen as they have demonstrated the capacity to learn complex features from medical images and serve as a benchmark to assess the efficacy of our hybrid architecture.

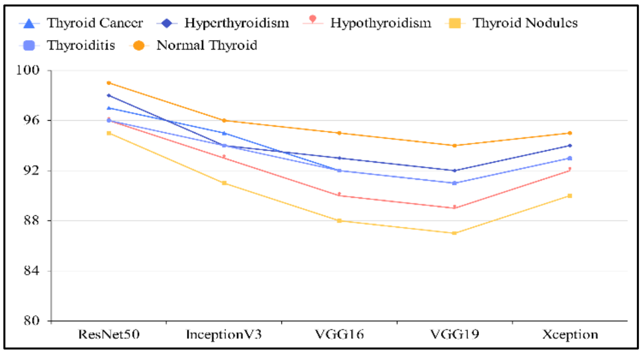


**Fig. 6.** Accuracy Comparison of Different CNN Architectures

*4.5 Discussion*

The experimental outcome also validates the proposed deep learning framework to have a high level of performance in identifying and categorizing multitype thyroid diseases. The high values of precision and recall in their test show that the model has low rates of false positives and false negatives and this is very important in diagnosis. Although the performance of our model was high, the accuracy for thyroid nodules was lower than for the other diseases. Thyroid nodules in ultrasound images vary significantly in shape, texture, and size. Recently research shows that these variations can affect accuracy. Incorporating advanced feature extraction methods can improve nodule classification. Overall, the proposed system addresses the limitations of traditional diagnostic methods by providing a reliable and automated solution for thyroid disease detection, demonstrating its potential for clinical application.

The results presented in Table 5 and Fig. 7 show that dual optimization achieves better final accuracy than using Adam or SGD separately. The use of Adam in the first stage essentially allows the model to converge quickly, and the use of SGD in the second stage refines the weights, leading to reduced overfitting and improved generalization. This two-step procedure improves the accuracy of the model and its robustness in various thyroid diseases as well. The Precision, Recall, F1-Score, and Accuracy of Modified ResNet and InceptionV3 deployed from Adam and SGD Optimizer in which the Modified ResNet produced the best results with SGD with an accuracy of 97%, while the InceptionV3 did also better with SGD and achieved 96%. Altogether it was shown that SGD is better than Adam for both models with significant improvements that show it can work for the optimization of models.

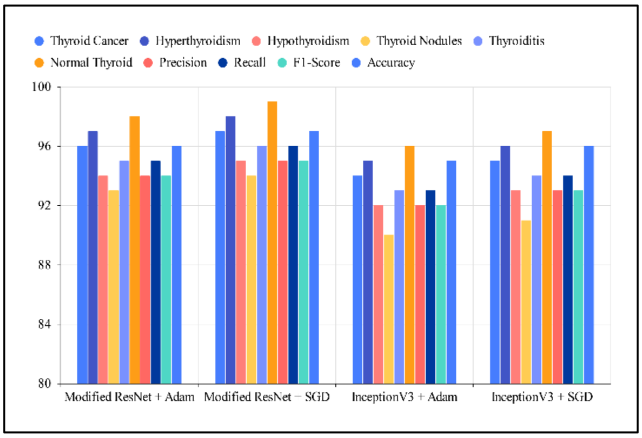


**Fig. 7.** Performance Metrics of Optimization Techniques

**Table 4**

Accuracy comparison of different CNN architectures. The performance of Dual-OptNet is compared with ResNet50, InceptionV3, and other CNN models across multiple thyroid disease categories.

| CNN Architecture | Thyroid Cancer | Hyperthyroidism | Hypothyroidism | Thyroid Nodules | Thyroiditis | Normal Thyroid |
|---|---|---|---|---|---|---|
| ResNet50 | 97% | 98% | 96% | 95% | 96% | 99% |
| InceptionV3 | 95% | 94% | 93% | 91% | 94% | 96% |
| VGG16 | 92% | 93% | 90% | 88% | 92% | 95% |
| VGG19 | 91% | 92% | 89% | 87% | 91% | 94% |
| Xception | 93% | 94% | 92% | 90% | 93% | 95% |

**Table 5**

Performance metrics of optimization techniques across multiple thyroid disease categories

| Model | Optimizer | Thyroid Cancer | Hyperthyroidism | Hypothyroidism | Thyroid Nodules | Thyroiditis | Normal Thyroid | Precision | Recall | F1-Score | Accuracy |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Modified ResNet | Adam | 96% | 97% | 94% | 93% | 95% | 98% | 94% | 95% | 94% | 96% |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Modified ResNet | SGD | 97% | 98% | 95% | 94% | 96% | 99% | 95% | 96% | 95% | 97% |
| InceptionV 3 | Adam | 94% | 95% | 92% | 90% | 93% | 96% | 92% | 93% | 92% | 95% |
| InceptionV 3 | SGD | 95% | 96% | 93% | 91% | 94% | 97% | 93% | 94% | 93% | 96% |

## 5. Conclusion and Future Work

In this study, we introduced Dual-OptNet, which is based on a deep learning architecture that synergistically integrates ResNet and InceptionV3 architectures and a dual optimization scheme (Adam and SGD). We show that dual-optimum Net (DualOptNet) produces state-of-the-art in categorizing thyroid disease, providing a reliable tool for clinical applications. This work proposed a new technique to detect and classify multiple thyroid disorders with the help of improved ResNet and InceptionV3 models. With the help of two optimization methods such as Adam and SGD, the system produced 97% classification accuracy, which surpasses all previous diagnostic methods. Data augmentation and preprocessing improved stability. The WEB-DIAG module can classify thyroid diseases and provide real-time recommendations to doctors. Future work will expand the generalization of the model by using rare thyroid samples and real-time ultrasound analysis. Moreover, combining patient data, pathology diagnoses, histology data, and laboratory results, alongside explainable AI methods, will enhance diagnostic accuracy and clinical acceptability. Additionally, Future work will focus on the development of a real-time web-based application to provide clinicians with an accessible diagnostic tool. This research tackles issues such as gradient vanishing, the class imbalance problem, and fine-grained difficulties which makes the proposed framework readily adaptable for clinical applications. In a follow-up, incorporating patient information like age, gender, and medical history alongside the imaging data might add more context and help the model make better predictions. Furthermore, the use of explainable AI approaches like prominence maps or attention methods will allow clinicians to better understand the basis behind the model predictions, adding the necessary interpretability layer to ensure that the diagnostic predictions are not only correct but also intelligible.

## 6. Acknowledgment:

This research was supported by the Researchers Supporting Project number (RSPD2025R846), King Saud University, Riyadh, Saudi Arabia.

## 7. References

[1] S. Prathibha, D. Dahiya, C. R. Robin, C. V. Nishkala, and S. Swedha, "A novel technique for detecting various thyroid diseases using deep learning," Intell. Autom. Soft Comput., vol. 35, no. 1, pp. 199-214, 2023.

[2] K. Guleria, S. Sharma, S. Kumar, and S. Tiwari, "Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning," Measurement: Sensors, vol. 24, p. 100482, 2022.

[3] X. Zhang, V. C. Lee, J. Rong, J. C. Lee, J. Song, and F. Liu, "A multi-channel deep convolutional neural network for multi-classifying thyroid diseases," Computers in Biology and Medicine, vol. 148, p. 105961, 2022.

[4] B. Shankarlal, P. D. Sathya, and V. P. Sakthivel, "Computer-aided detection and diagnosis of thyroid nodules using machine and deep learning classification algorithms," IETE Journal of Research, vol. 69, no. 2, pp. 995-1006, 2023.

[5] L. Y. Lin, P. Zhou, M. Shi, J. E. Lu, S. Jeon, D. Kim, et al., "A deep learning model for screening computed tomography imaging for thyroid eye disease and compressive optic neuropathy," Ophthalmology Science, vol. 4, no. 1, p. 100412, 2024.

[6] S. S. Islam, M. S. Haque, M. S. U. Miah, T. B. Sarwar, and R. Nugraha, "Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study," PeerJ Computer Science, vol. 8, p. e898, 2022.

[7] G. Obaido, O. Achilonu, B. Ogbuokiri, C. S. Amadi, L. Habeebullahi, T. Ohalloran, et al., "An improved framework for detecting thyroid disease using filter-based feature selection and stacking ensemble," IEEE Access, 2024.

[8] J. H. Lee, Y. G. Kim, Y. Ahn, S. Park, H. J. Kong, J. Y. Choi, et al., "Investigation of optimal convolutional neural network conditions for thyroid ultrasound image analysis," Scientific Reports, vol. 13, no. 1, p. 1360, 2023.

[9] M. G. Raj, "Enhancing Thyroid Cancer Diagnostics Through Hybrid Machine Learning and Metabolomics Approaches," Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 2, 2024.

[10] S. Zhou, Y. Qiu, L. Han, G. Liao, Y. Zhuang, B. Ma, et al., "A lightweight network for automatic thyroid nodules location and recognition with high speed and accuracy in ultrasound images," J. X-Ray Sci. Technol., vol. 30, no. 5, pp. 967-981, 2022.

[11] L. Chaouchi, D. Gaceb, F. Touazi, D. Djani, and A. Yakoub, "Application of Deep Transfer Learning in Medical Imaging for Thyroid Lesion Diagnostic Assistance," in 2024 8th Int. Conf. Image Signal Process. Appl. (ISPA), 2024, pp. 1-7.

[12] Y. Gao, W. Wang, Y. Yang, Z. Xu, Y. Lin, T. Lang, et al., "An integrated model incorporating deep learning, radiomics and clinical and US features to diagnose central lymph node metastasis in patients with papillary thyroid cancer," 2023.

[13] S. Ahmad, T. Ullah, I. Ahmad, A. Al-Sharabi, K. Ullah, R. A. Khan, et al., "A novel hybrid deep learning model for metastatic cancer detection," Comput. Intell. Neurosci., vol. 2022, p. 8141530, 2022.

[14] S. Shukla, M. Brandwein-Weber, S. Samankan, A. Ayad, M. Rabie, and S. Doyle, "Federated learning in computational pathology: classification of tall cell patterns in papillary thyroid carcinoma," in Med. Imaging 2024: Digital Comput. Pathol., vol. 12933, 2024, pp. 206-215.

[15] X. Xie, Y. Tian, K. Ota, M. Dong, Z. Liu, H. Jin, and D. Yao, "Reinforced Computer-aided Framework for Diagnosing Thyroid Cancer," IEEE/ACM Trans. Comput. Biol. Bioinformatics, 2023.

[16] R. Dixit, M. A. Tayal, S. Bedi, and S. Saxena, "Thyroid Disorder Classification using Machine Learning," in 2023 11th Int. Conf. Emerging Trends Eng. Technol. Signal Inf. Process. (ICETET-SIP), 2023, pp. 1-5.

[17] Z. A. Haider, F. M. Khan, I. U. Khan, and M. A. K. Azad, "Utilizing Effective Deep Learning Models for Early Prediction and Detection of Chronic Kidney Disease," Spectrum Eng. Sci., vol. 2, no. 3, pp. 101-131, 2024.

[18] E. Vorontsov, A. Bozkurt, A. Casson, G. Shaikovski, M. Zelechowski, K. Severson, et al., "A foundation model for clinical-grade computational pathology and rare cancers detection," Nature Med., pp. 1-12, 2024.

[19] A. A. Showrov, M. T. Aziz, H. R. Nabil, J. R. Jim, M. M. Kabir, M. F. Mridha, et al., "Generative Adversarial Networks (GANs) in Medical Imaging: Advancements, Applications and Challenges," IEEE Access, 2024.

[20] F. M. Khan, M. S. Akhter, I. U. Khan, Z. A. Haider, and N. H. Khan, "Clinical Prediction of Female Infertility Through Advanced Machine Learning Techniques," Int. J. Innov. Sci. Technol., vol. 6, no. 2, pp. 941-966, 2024.

[21] G. H. Kim, D. H. Lee, J. W. Choi, H. J. Jeon, and S. Park, "Multimodal Neural Network for Recurrence Prediction of Papillary Thyroid Carcinoma," Adv. Intell. Syst., vol. 5, no. 2, p. 2200365, 2023.

[22] Z. Rasheed, Y. K. Ma, I. Ullah, T. Al Shloul, A. B. Tufail, Y. Y. Ghadi, et al., "Automated classification of brain tumors from magnetic resonance imaging using deep learning," Brain Sci., vol. 13, no. 4, p. 602, 2023.

[23] P. Gupta, F. Rustam, K. Kanwal, W. Aljedaani, S. Alfarhood, M. Safran, and I. Ashraf, "Detecting thyroid disease using optimized machine learning model based on differential evolution," Int. J. Comput. Intell. Syst., vol. 17, no. 1, p. 3, 2024.

[24] A. B. Tufail, I. Ullah, W. U. Khan, M. Asif, I. Ahmad, Y. K. Ma, et al., "Diagnosis of diabetic retinopathy through retinal fundus images and 3D convolutional neural networks with limited number of samples," Wireless Commun. Mobile Comput., vol. 2021, p. 6013448, 2021.

[25] J. Kumar, S. N. Panda, and D. Dayal, "Pediatric thyroid ultrasound image classification using deep learning: A review," Appl. Data Sci. Smart Syst., pp. 264-277, 2025.

[26] J. Li, Q. Guo, and X. Tan, "Multi-modal feature integration for thyroid nodule prediction: Combining clinical data with ultrasound-based deep features," J. Radiat. Res. Appl. Sci., vol. 18, no. 1, p. 101217, 2025.

[27] W. T. Yang, B. Y. Ma, and Y. Chen, "A narrative review of deep learning in thyroid imaging: current progress and future prospects," Quant. Imaging Med. Surg., vol. 14, no. 2, pp. 2069, 2024.

[28] Y. Habchi, H. Kheddar, Y. Himeur, A. Boukabou, A. Chouchane, A. Ouamane, et al., "Machine learning and vision transformers for thyroid carcinoma diagnosis: A review," arXiv preprint arXiv:2403.13843, 2024.

[29] S. Dalal, U. K. Lilhore, N. Faujdar, S. Simaiya, A. Agrawal, U. Rani, and A. Mohan, "Enhancing thyroid disease prediction with improved XGBoost model and bias management

techniques," Multimedia Tools Appl., pp. 1-32, 2024.
[30] P. M. Kumar, W. Korani, and J. Selvaraj, "IoT-Powered BGCNet: Thyroid Disease Detection Enhancement in Healthcare with IBESA based Feature Selection," in 2024 IEEE Green Tech. Conf. (GreenTech), 2024, pp. 48-53.
[31] O. A. Ajilisa, V. P. Jagathyraj, and M. K. Sabu, "Ensemble-Based Deep Learning Approach for Characterization of Thyroid Nodules in Ultrasound Imaging," in 2024 11th Int. Conf. Adv. Comput. Commun. (ICACC), 2024, pp. 1-6.
[32] A. A. Maarouf, H. Meriem, and F. Hachouf, "Deep Learning and Handcrafted Features for Thyroid Nodule Classification," Int. J. Imaging Syst. Technol., vol. 34, no. 6, p. e23215, 2024.
[33] X. Yang, S. O. Qu, Z. Wang, L. Li, X. An, and Z. Cong, "The study on ultrasound image classification using a dual-branch model based on Resnet50 guided by U-net segmentation results," BMC Med. Imaging, vol. 24, no. 1, p. 314, 2024.
[34] M. Zubair, H. M. Rais, Q. Al-Tashi, F. Ullah, M. Faheeem, and A. A. Khan, "Enabling predication of the deep learning algorithms for low-dose CT scan image denoising models: A systematic literature review," IEEE Access, 2024.
[35] A. B. Tufail, Y. K. Ma, M. K. Kaabar, F. Martínez, A. R. Junejo, I. Ullah, and R. Khan, "Deep learning in cancer diagnosis and prognosis prediction: a mini-review on challenges, recent trends, and future directions," Comput. Math. Methods Med., vol. 2021, p. 9025470, 2021.
[36] R. B. Namdeo and G. V. Janardan, "Optimization assisted framework for thyroid detection and classification: A new ensemble technique," Gene Expr. Patterns, vol. 45, p. 119268, 2022.
[37] 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, 2023.

 