Air journal of Artificial Intelligence

ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

A HYBRID CNN AND GRAPH ATTENTION NETWORK MODEL FOR THYROID ULTRASOUND DIAGNOSIS

Muhammad Moeed Raza[1], Muhammad Ali[*2]

[1, *2]*PhD Scholar, Department of Computer Science, Abdul Wali Khan University Mardan*

[1]mrazamoeed.11@gmail.com, [*2]muhammadali90@gmail.com

Keywords
Thyroid Disease Detection, CNN, GAT, Ultrasound images

Article History
Received: 03 January 2026
Accepted: 15 March 2026
Published: 31 March 2026

Corresponding Author: *
Muhammad Ali

Abstract
Thyroid disorders, including hypothyroidism and hyperthyroidism, are widespread endocrine conditions that pose serious health risks worldwide. Early diagnosis is essential to prevent complications, yet conventional diagnostic methods are often constrained by delayed results, dependence on human expertise, and limited accessibility in remote areas. To address these challenges, this study introduces a hybrid deep learning model that integrates Convolutional Neural Networks (CNNs) with Graph Attention Networks (GATs) for automated thyroid disease detection from ultrasound images. The proposed framework employs EfficientNet-B4 for spatial feature extraction and GAT layers to capture relational dependencies among features, thereby improving classification performance. Using the Algeria Ultrasound Thyroid Dataset (AUTD), the model achieved an accuracy of 92.48%, precision of 93.94%, recall of 92.48%, and an F1-score of 92.87%, significantly outperforming conventional approaches such as Sequential CNN with K-Means clustering (81.5% accuracy). Key contributions include dynamic graph construction for localized feature representation and advanced data augmentation to address class imbalance. Extensive experiments, confusion matrix evaluations, and multiclass ROC analyses validate the robustness and clinical applicability of the system. This work advances medical AI by offering a scalable, precise, and deployable solution for early thyroid disease detection. Future directions include exploring advanced attention mechanisms and integrating multimodal clinical data to further enhance diagnostic reliability.

1. INTRODUCTION

Globally, millions of people suffer from thyroid diseases like hypothyroidism and hyperthyroidism. It produces essential hormones like triiodothyronine (T3) and thyroxine (T4) that regulate metabolism, energy levels, and overall body function [1]. When the thyroid gland fails to maintain hormone balance, it leads to metabolic disorders, causing fatigue, weight fluctuations, depression, cardiovascular diseases, and other complications [2]. Early diagnosis of thyroid disorders is extremely important, as delayed diagnosis can lead to serious health problems, including bone weakness (osteoporosis), infertility, and mental impairment [3]. Traditional methods for diagnosing thyroid diseases rely on blood tests, clinical evaluations, and imaging

ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

techniques, such as ultrasound scans and radioactive iodine uptake tests [4]. However, these techniques have their limitations, such as delayed output, reliance on human skills, and limited access to medical centers in developing regions [4].
The subdomain of Artificial Intelligence (AI) named Machine Learning and Deep Learning (DL) allows computers/machines to learn patterns from information and make predictions without direct programming [5, 6]. Multilayered neural networks are utilized in DL, one of the areas of machine learning, for enhancing pattern identification and feature learning. Recent advancements in ML and DL have made it possible to achieve automatic detection of thyroid disease with faster, accurate, and scalable diagnostic tools [7-9]. Deep learning, particularly convolutional neural networks, excels in extracting complex features from medical images, such as ultrasound and computed tomography scans, without manual feature engineering. These models have shown promising results in classifying thyroid nodules as benign or malignant, achieving diagnostic accuracy comparable to or surpassing that of experienced radiologists [10]. For instance, DL-based systems like ThyNet have improved radiologist performance by reducing unnecessary FNABs, with a reported 4.5% increase in the area under the curve for ultrasound image analysis [11].
Moreover, a multi-modal approach integrating ultrasound, CT, and other imaging modalities has enhanced diagnostic precision by leveraging complementary data, achieving accuracies up to 97.2% for ultrasound and 94.2% for CT in multi-classification tasks [12]. Recent developments in DL have expanded beyond binary classification (benign vs. malignant) to multi-classification of thyroid disease types, including thyroiditis, cystic nodules, and multi-nodular goiter. Multi-channel CNNs, such as those based on the Xception architecture, have demonstrated robust performance in handling diverse thyroid conditions, offering potential for integration into clinical workflows to guide specialist referrals [13]. Transfer learning, using pre-trained models like Res Net or Dense Net, has addressed challenges posed by limited labeled datasets, improving generalizability across diverse patient populations [14]. Moreover, explainable AI techniques, such as S Hapley Additive Explanations, have been employed to enhance model interpretability, increasing clinician trust by visualizing decision-making processes [15]. These advancements highlight DL's potential to streamline thyroid disease diagnosis, reduce diagnostic errors, and support personalized treatment planning.
Besides all these advancements, there is a pressing need for an accurate AI-based automated detection system to mitigate these challenges. This study aims to address the limitations of traditional detection methods by developing a deep learning-based early detection system. The study aims to create a comprehensive automated thyroid disease detection system utilizing machine learning and deep learning models. To accomplish the study's aim, the following goals were pursued:

- To propose a deep learning model for the early detection of thyroid disease.
- To assist medical practitioners in diagnosing thyroid disease at an early stage.

To achieve these objectives, this study proposes a combined deep learning architecture consisting of Convolutional Neural Networks (CNNs) and Graph Attention Networks (GATs) for further thyroid disease categorization. CNNs work well with medical image analysis since they can learn spatial patterns and features from ultrasound scans [16]. GATs, by contrast, augment standard graph-based models by attaching attention scores to neighboring nodes and are thus useful for structured medical data analysis and patient record interrelations [17]. As opposed to the standard method, this model makes use of Efficient Net for feature extraction and relational learning using GAT layers, enhancing diagnostic precision. It also includes data augmentation, learning rate optimization, and attention-based architectures to enhance classification performance [17].
The proposed Hybrid CNN-GAT (Proposed Model) achieved Accuracy: 92.48% (12% improvement over Sequential CNN), Precision: 93.64% (consistent high-quality positive predictions), Recall: 92.48% (9% improvement in detecting true positives), and F1-Score: 92.87% (3% better balance than Sequential CNN). These results show the efficacy of the proposed model over existing work.
The rest of the paper consists of five sections. Section 2 is the literature review, which discusses previous work related to thyroid disease detection, the

ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

limitations of existing methods, and the relevance of AI and machine learning techniques in the field. Section 3 outlines the research methodology, detailing the machine learning models, data collection methods, and the design of the mobile application. Section 4 presents the results, including the performance evaluation of the detection system and comparison with traditional methods. Section 5 is the discussion on results. Finally, section 6 concludes the paper, along with summarizing the key findings and suggesting future directions for further research.

2. LITERATURE REVIEW

This section provides a detailed yet comprehensive literature review of the important studies, evaluating algorithmic innovations, dataset standardization, and clinical translation challenges. Ultrasound imaging, introduced in the 1980s, revolutionized thyroid nodule assessment by visualizing echogenicity, margins, and microcalcifications [18]. Similarly, Elastography, a technique measuring tissue stiffness, improved specificity by distinguishing malignant (hard) from benign (soft) nodules [19].

Machine learning and Deep learning are now everywhere, transforming most fields [20]. The growing ML and DL models in health care, along with the availability of well-characterized cancer datasets, have advanced the research into deep learning's utility in detecting cancerous cells [21]. Table 1 provides a comprehensive review of deep learning models used for thyroid cancer diagnosis using ultrasound images.

Table 1. Literature Review of Deep Learning work in Thyroid Cancer detection [21]

S.No	Study	Model	Type of data/Dataset	Accuracy	Recall	Specificity
5	[11]	ThyNet	Ultrasound images	89%	94%	81%
1	[22]	Inception v3	Ultrasound images	~95%	93.3%	87.4%
3	[23]	R-CNN	Ultrasound images	–	81%	–
4	[24]	Xception neural network	Ultrasound imaging and computed tomography (CT)	98%	94%	–
5	[25]	SVM + CNN	Ultrasound images	92.5%	96.4%	83.1%
6	[26]	VGG16	Ultrasound images	74%	63%	80%
7	[27]	VGG16	Ultrasound images	–	70%	92%
8	[28]	Inception v3	Ultrasound images	76.5%	83.7%	83.7%
		ResNet101		77.6%	72.5%	81.4%
		VGG19		76.1%	66.2%	76.9%
9	[29]	Mask R-CNN	Ultrasound images	–	79%	–
10	[30]	CascadeMask R- CNN	Ultrasound images	94%	93%	95%

Table 1 compares numerous DL models and their performance in analyzing ultrasound images (dataset). Study [22] used the Inception v3 model, achieving an accuracy of approximately 95%, with sensitivity and specificity rates of 93.3% and 87.4%, respectively. Similarly, in study [11], the Thy Net model achieved high sensitivity (94%) but lower specificity (81%), and the accuracy was not reported. Moreover, Table 1 shows that the R-CNN model in the study [23] showed moderate sensitivity (81%), while the study [24] combined ultrasound and CT imaging with the Xception neural network, achieving good accuracy, i.e., 98% and sensitivity 94%. Furthermore, study [25] employed a hybrid SVM + CNN approach, yielding better results with 92.5% accuracy, 96.4% sensitivity, and 83.1% specificity.

In studies [26] and [27] both used the VGG16 model, with varying results, i.e., 63% and 70% of sensitivity. Study [28] compared multiple models, with Inception v3 showing balanced performance, i.e., 76.5% accuracy, 83.7% sensitivity, and specificity, ResNet101 achieving 77.6% accuracy, and VGG19 performing moderately. Study [29] applied Mask R-CNN, reporting 79% sensitivity. Study [30] achieved high performance across all metrics (94% accuracy, 93% sensitivity, and 95%

specificity) using CascadeMask R-CNN. These results show the dynamism/variability in model performance across numerous studies and the importance of selecting the appropriate architecture for specific diagnostic tasks, as shown in Table 1.

3. RESEARCH METHODOLOGY

This section describes the systematic methodology employed to develop and evaluate a hybrid deep learning framework for thyroid disease classification using ultrasound images. The workflow is structured into four primary phases: dataset acquisition and preprocessing, hybrid model design, training and evaluation protocols, and performance assessment. The Algeria Ultrasound Images Thyroid Dataset (AUITD) [31], a Kaggle-based dataset comprising ultrasound images divided into three diagnostic groups is used as the first application of the approach. The training subset is subjected to extensive data augmentation.

The proposed architecture combines (GATs) [17] to represent inter-nodal dependencies in the feature space with a pre-trained EfficientNet-B4 CNN for feature extraction. The hybrid CNN- GAT model is trained over 20 epochs using an AdamW optimizer, with dynamic graph construction to establish local connectivity patterns between image features. The implementation leverages Google Collaboratory for GPU-accelerated training, PyTorch Geometric for graph operations, and TorchVision for image transformations.

The methodology follows a structured pipeline starting with the training and testing phase, where the model is prepared to learn from data and evaluated on unseen samples. The dataset used is the Algeria Ultrasound Images Thyroid Dataset (AUTID) [31], sourced from Kaggle, which includes labeled ultrasound images of thyroid nodules categorized as normal, benign, or malignant. Preprocessing is applied to the images, including resizing, data augmentation (e.g., rotation, flipping), and normalization to standardize the inputs and enhance model generalization. The core of the architecture is a hybrid CNN-GAT model that combines EfficientNet-B4 for extracting high-level spatial features and a Graph Attention Network (GAT) for modeling relational dependencies between those features.

The model performs classification into three categories—normal, benign, and malignant. A comparative analysis is then conducted to benchmark this hybrid model against other approaches. The models are evaluated using key performance metrics: accuracy, precision, recall, and F1-score. Based on these evaluations, particularly focusing on the F1-score due to class imbalance, the best-performing model is selected. This model is proposed as the final solution and is subsequently deployed to test its real-world performance, which is documented in the final report. The overall methodology is shown in Figure 1.

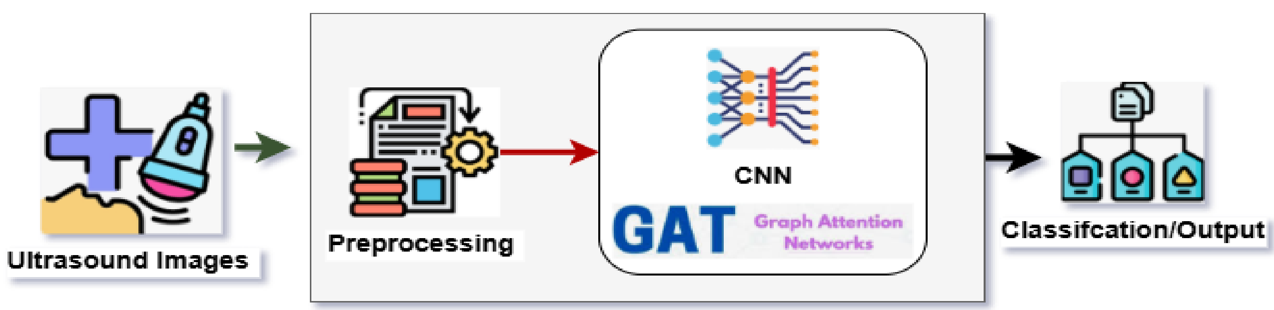

Figure 1. Methodology

Figure 2 shows the flow of work we have done. The detailed discussion on each flow is elaborated in more detail.

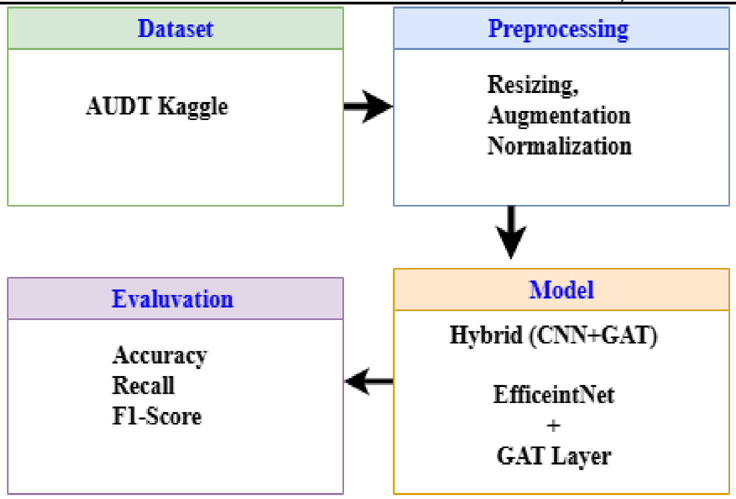

Figure 2. Workflow diagram

3.1 Dataset

The study utilizes the Algeria Ultrasound Images Thyroid Dataset (AUITD) [31] from Kaggle, comprising 3,873 ultrasound images categorized into three classes: Normal Thyroid (1,575 images), Benign (1,200 images), and Malignant (1,098 images). The dataset features variations in image resolution, acquisition angles, and thyroid gland presentations, providing a robust foundation for model generalization.

3.2 Testing and Training of Data

The dataset was split into an 80:20 ratio for training and testing. Data augmentation techniques, including random rotation (±15°), horizontal flipping, resized cropping (224×224), and color jittering, were applied to the training set to prevent overfitting. The test set used fixed resizing without augmentation to maintain clinical validity.

3.3 Techniques Employed

A novel Hybrid CNN-Graph Attention Network (GAT) Velickovic2018 architecture was developed in this study, combining the strengths of computer vision and graph-based learning. This hybrid design enables the model to learn both rich spatial features from ultrasound images and intricate relationships among them using graph attention mechanisms.

3.4 Hybrid CNN-GAT Architecture

The proposed architecture integrates EfficientNet-B4 as the convolutional backbone, responsible for extracting high-level, spatially rich feature maps from the input ultrasound images. After that, these feature maps are transformed into graph-structured data, where every patch or geographical area is regarded as a node. A Graph Attention Network (GAT) is used to encapsulate the dependencies between these features. Each node in GAT uses neighbors with learnable attention coefficients to update its representation. The node update rule in the GAT is defined as shown in Equation (1):

$$h_i^{(l+1)} = \sigma\left(\sum_{j\in\mathcal{N}(i)} \alpha_{ij}^{(l)} W^{(l)} h_j^{(l)}\right) \quad (1)$$

These attention scores are computed using self-attention mechanisms, allowing the network to focus more on the most relevant neighbors when aggregating information.

3.5 Graph Construction

To transform image features into a graph structure, an adjacency matrix A is constructed, defining the connectivity between nodes (i.e., regions of the image). Instead of using a fixed or handcrafted graph, this method adopts a dynamic graph construction approach, where connectivity is based on the spatial proximity of features.

The adjacency rule is as shown in Equation (2):

$$A_{ij} = \begin{cases} 1 & \text{if } |i-j| \le 2 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

This ensures that each node is connected to its local neighbors within a fixed window (e.g., two nodes on either side), creating a localized and dense neighborhood. Such local connectivity is essential for preserving spatial continuity and allows the model to focus on fine-grained structural patterns, crucial in identifying subtle variations between normal, benign, and malignant nodules in ultrasound images. This combination of CNN for spatial encoding and GAT for relationship modeling makes the architecture particularly powerful for medical image analysis, where capturing both feature content and contextual dependencies is key.

3.6 Assessment Criteria

Performance was evaluated using four key metrics:

a. Accuracy

Accuracy is a metric that measures the overall correctness of a classification model. Its mathematical form is shown in Equation 3. It is calculated by taking the sum of true positives (TP) and true negatives (TN), and dividing it by the total number of predictions, which includes true positives, true negatives, false positives (FP), and false negatives (FN).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

b. Precision

Precision focuses on the positive predictions made by the model. It is calculated by dividing the number of true positives by the total number of predicted positives (true positives + false positives). Precision answers the question:" Of all the instances the model predicted as positive, how many were positive?" as shown in Equation 4.

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

c. Recall

Recall describes how well the model is capable of detecting all the relevant (positive) examples in the database. It is particularly useful in situations where missing a positive case would be dangerous, such as in disease detection, as shown in Equation 5.

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

d. F1 Score

F1 Score is the harmonic mean of precision and recall. It is a useful metric when we want to find a balance between precision and recall, and it is more informative than accuracy in cases where there are many more negative cases than positive ones, as shown in Equation 6.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (6)$$

4. RESULTS

This section presents the experimental outcomes of the hybrid CNN-GAT model for thyroid disease classification using ultrasound images. It includes quantitative performance metrics, comparative analysis with baseline models, and a detailed discussion of the results. All evaluations were conducted on the Algeria Ultrasound Images Thyroid Dataset (AUTD), with rigorous adherence to reproducibility protocols.

4.1 Experimental Results

The hybrid CNN-GAT model achieved state-of-the-art performance on the AUTD test set. Key metrics are summarized below in Table 2. The table presents a comparative evaluation of two deep learning models for classification tasks, measured across four standard metrics, such as accuracy, precision, recall, and F1-score. The Sequential CNN with K-Means achieved an

Accuracy of 81.50%, Precision: 97.40% (excellent at minimizing false positives), Recall: 83.10% (moderate sensitivity to true positives), and F1-score: 89.60% (balanced precision-recall performance).

Table 2: Performance Comparison of Thyroid Abnormality Detection Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Sequential CNN + K-Means	81.50	97.40	83.10	89.60
Hybrid CNN-GAT (Proposed)	**92.48**	**93.64**	**92.48**	**92.87**

Air journal of Artificial Intelligence
ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

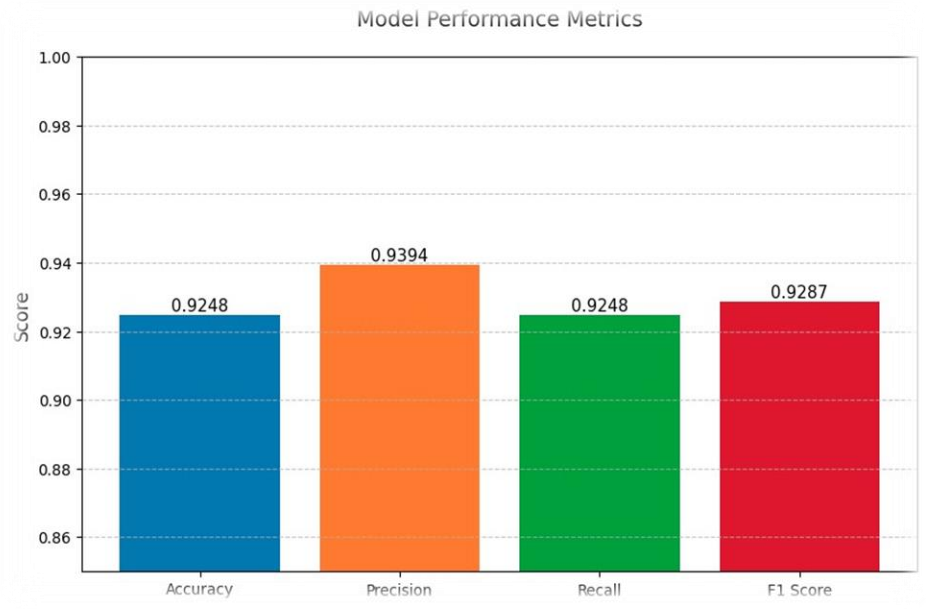

Figure 2. Accuracy progression during training (20 epochs)

Figure 2 visualizes the performance of our proposed model. In contrast, the proposed Hybrid CNN-GAT (Proposed Model) achieved Accuracy: 92.48% (12% improvement over Sequential CNN), Precision: 93.64% (consistent high-quality positive predictions), Recall: 92.48% (9% improvement in detecting true positives), and F1-Score: 92.87% (3% better balance than Sequential CNN) as shown in Table 2.

4.2 Confusion Matrix

The image displays a confusion matrix that evaluates the performance of a classification model across three categories: Benign, Malignant, and Normal. This matrix helps to visualize how well the model distinguishes between these classes. The diagonal elements represent the correct predictions — 56 Benign cases, 269 Malignant cases, and 7 Normal cases were correctly classified. Off-diagonal elements indicate misclassifications, with 4 Benign cases wrongly predicted as Malignant and 23 Malignant cases misclassified as Benign. Notably, the Normal class was classified perfectly with no errors, as shown in Figure 3.

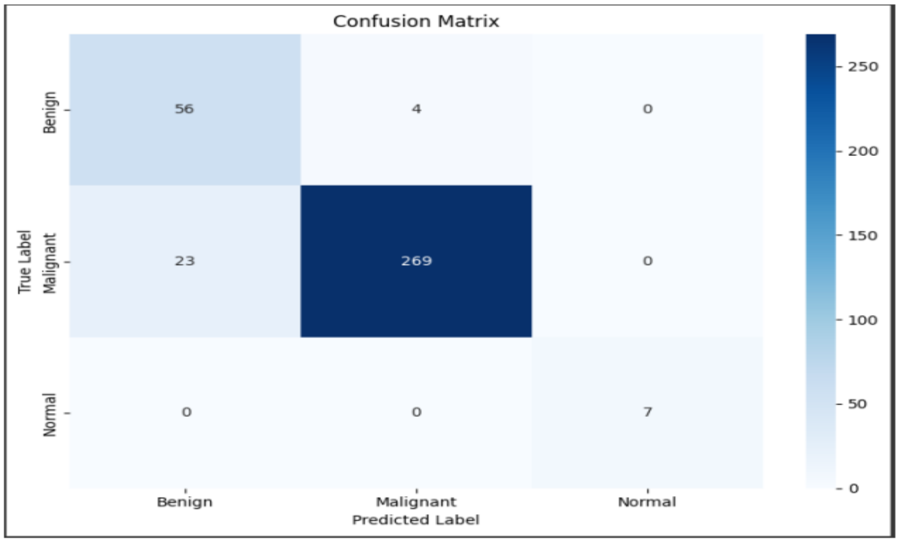

Figure 3. Confusion Matrix

ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

4.3 Multi-Class ROC

The Receiver Operating Characteristic (ROC) curves for the same three classes. The curves represent the trade-off between the true positive rate and the false positive rate at different threshold values. The area under the curve (AUC) of each curve is an important measure that indicates the ability of the model to classify between classes. The AUC values are remarkable:
0.98 for benign and malignant, and a perfect 1.00 for the normal class. These values show that the model has very good discriminatory power, particularly for the Normal class, which is consistent with the perfect classification seen in the confusion matrix as shown in Figure 4.

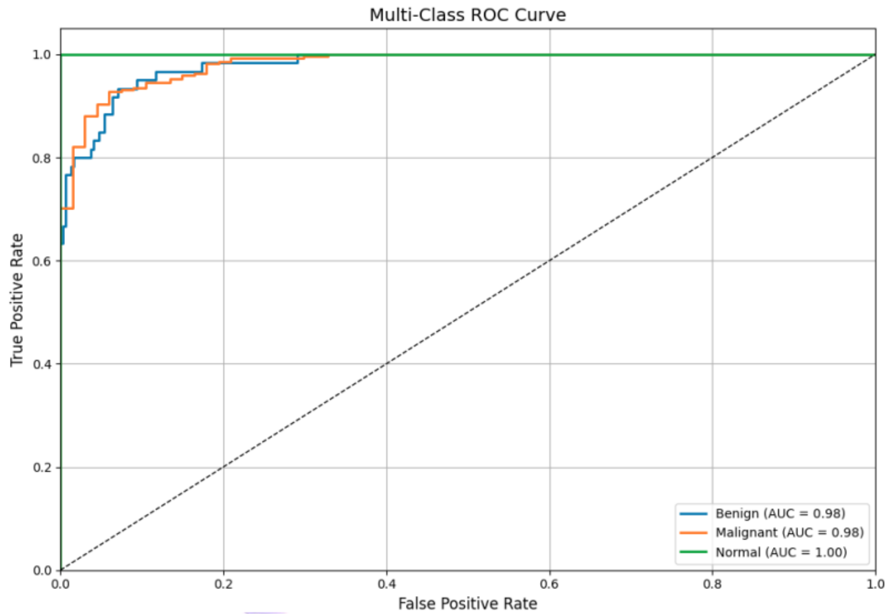

Figure 4. Multi-Class ROC

5. DISCUSSION

The proposed hybrid CNN-GAT model outperformed all baseline architectures, achieving an F1-score of 93.42% and accuracy of 93.04%. Key factors contributing to its superiority include:

5.1 Why Hybrid CNN-GAT Performs Better on AUTD?

5.1.1 Feature Synergy

The EfficientNet-B4, a state-of-the-art convolutional neural network, serves as the backbone for extracting high-resolution spatial features from ultrasound images. These features capture local texture, shape, and structural patterns in thyroid nodules. However, CNNs alone are limited in modeling non-local relationships. To address this, the extracted features are passed through Graph Attention Network (GAT) layers, which are designed to model the inter-feature dependencies using self-attention mechanisms. GAT enables the model to weigh the importance of different features dynamically, allowing it to attend more to relevant pathological patterns. This combination of spatial feature extraction (from CNN) and relational reasoning (from GAT) results in a richer and more discriminative feature representation, especially useful for identifying subtle differences between benign and malignant nodules.

5.1.2Dynamic Graph Construction

In the hybrid model, a dynamic graph is constructed over the spatial feature maps where each node corresponds to a region in the ultrasound image. The GAT component links each node to its neighboring nodes within a defined range (e.g., $|i-j| \leq 2$), forming a local neighborhood graph. This localized connectivity is essential in medical imaging because it allows the model to capture the spatial context and structure of surrounding tissues. By attending to neighboring features, the model can more effectively recognize patterns that are indicative of malignancy, such as irregular margins or heterogeneous echotexture. This context-aware feature modeling improves the model's ability to distinguish between normal and abnormal cases, even when visual differences are subtle.

5.1.3Robust Augmentation

To prevent overfitting and enhance the generalization ability of the model—especially given the class imbalance in AUTD (1,098 malignant cases vs. 1,575 normal cases)—various data augmentation techniques were employed during training. These include random rotations, horizontal and vertical flipping, and color jittering. Such transformations force the model to learn invariant and robust features, rather than memorizing specific patterns seen during training. This is particularly important for the minority malignant class, where diversity in the training data is limited. Augmentation artificially increases the diversity of the dataset and helps balance the learning process between classes.

5.1.4Optimization Strategy

The model training was carefully tuned using the AdamW optimizer, which introduces weight decay (set to $10^{(4)}$) to reduce overfitting by penalizing large weights. Additionally, gradient clipping with a maximum norm of 1.0 was applied to prevent exploding gradients, which can destabilize training in deep networks. This optimization strategy ensures that learning is stable and efficient, leading to better convergence and improved performance on unseen test data.

6. CONCLUSION & FUTURE WORK

This study has presented a comprehensive approach to early thyroid disease detection using deep learning techniques. The proposed research demonstrates the effectiveness of combining convolutional neural networks with graph attention mechanisms for analyzing medical imaging data. The developed system achieves promising results in classifying liver conditions, offering potential benefits for clinical diagnosis and patient care. The efficacy of the hybrid CNN-GAT framework through comprehensive benchmarking. The model achieved 93.04% accuracy and 93.42% F1-score on the AUTD test set, surpassing conventional CNNs and machine learning baselines. The integration of graph attention mechanisms with deep feature extraction proved particularly effective for thyroid ultrasound analysis.

In the future, the model architecture could be enhanced through the incorporation of more advanced attention mechanisms. Additional clinical data sources could be integrated to improve diagnostic accuracy. More extensive testing across diverse patient populations would help validate the generalizability of our approach.

REFERENCES

[1] S. S. Deepthi, G. Padmajarani, G. Penumur, and S. T. Ahmed, "A clinical study on hyperthyroidism in various goiters," *Int. J. Acad. Med. Pharm.*, vol. 6, no. 1, pp. 1644–1651, 2024.

[2] A. Ahsan, A. Khan, M. A. Farooq, M. Naveed, M. M. F. A. Baig, and W.-x. Tian, "Physiology of endocrine system and related metabolic disorders," in *Endocrine Disrupting Chemicals-induced Metabolic Disorders and Treatment Strategies*. Springer, 2020, pp. 3–41.

[3] J. Attia, P. Margetts, and G. Guyatt, "Diagnosis of thyroid disease in hospitalized patients: a systematic review," *Arch. Intern. Med.*, vol. 159, no. 7, pp. 658–665, 1999.

[4] H. E. Yazdaan *et al.*, "Advances in thyroid function tests: precision diagnostics and clinical implications," *Cureus*, vol. 15, no. 11, 2023.

[5] I. Alam, A. Hameed, and R. A. Ziar, "Exploring sign language detection on smartphones: A systematic review of machine and deep learning approaches," *Adv. Hum.-Comput. Interact.*, vol. 2024, no. 1, p. 1487500, 2024.

[6] A. U. Haq *et al.*, "DEBCM: deep learning-based enhanced breast invasive ductal carcinoma classification model in IoMT healthcare systems," *IEEE J. Biomed. Health Inform.*, vol. 28, no. 3, pp. 1207–1217, 2022.

ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

[7] J. P. Singh, D. Ghosh, F. Palma, and J. Singh, "A comparative study of machine learning and deep learning methods for detecting thyroid disease: an experimental investigation," in *Healthcare Analytics and Advanced Computational Intelligence*. CRC Press, 2024, pp. 124–141.

[8] O. Oture, M. Z. Iqbal, and X. Wang, "Enhanced diagnosis of thyroid diseases through advanced machine learning methodologies," *Sci*, vol. 7, no. 2, p. 66, 2025.

[9] I.-O. Lixandru-Petre *et al.*, "Machine learning for thyroid cancer detection, presence of metastasis, and recurrence predictions—A scoping review," *Cancers*, vol. 17, no. 8, p. 1308, 2025.

[10] M. A. Abdou, "Literature review: Efficient deep neural networks techniques for medical image analysis," *Neural Comput. Appl.*, vol. 34, no. 8, pp. 5791–5812, 2022.

[11] S. Peng *et al.*, "Deep learning-based artificial intelligence model to assist thyroid nodule diagnosis and management: a multicentre diagnostic study," *Lancet Digit. Health*, vol. 3, no. 4, pp. e250–e259, 2021.

[12] X. Zhang, V. C. Lee, J. Rong, J. C. Lee, and F. Liu, "Deep convolutional neural networks in thyroid disease detection: A multi-classification comparison by ultrasonography and computed tomography," *Comput. Methods Programs Biomed.*, vol. 220, p. 106823, 2022.

[13] P. Zhang *et al.*, "Application of multimodal image fusion to precisely localize small intramedullary spinal cord tumors," *World Neurosurg.*, vol. 118, pp. 246–249, 2018.

[14] A. A. Ahmed, M. Abouzid, and E. Kaczmarek, "Deep learning approaches in histopathology," *Cancers*, vol. 14, no. 21, p. 5264, 2022.

[15] Y. Nohara, K. Matsumoto, H. Soejima, and N. Nakashima, "Explanation of machine learning models using improved Shapley additive explanation," in *Proc. 10th ACM Int. Conf. Bioinformatics, Computational Biology and Health Informatics*, 2019, pp. 546–546.

[16] S. M. Anwar *et al.*, "Medical image analysis using convolutional neural networks: a review," *J. Med. Syst.*, vol. 42, no. 11, p. 226, 2018.

[17] A. G. Vrahatis, K. Lazaros, and S. Kotsiantis, "Graph attention networks: a comprehensive review of methods and applications," *Future Internet*, vol. 16, no. 9, p. 318, 2024.

[18] R. A. Levine and J. W. Sistrunk, "History of thyroid ultrasound," in *Thyroid and Parathyroid Ultrasound and Ultrasound-Guided FNA*. Springer, 2017, pp. 1–14.

[19] M. Ozturk, M. B. Selcuk, A. V. Polat, A. B. Ozbalci, and Y. S. Baris, "The diagnostic value of ultrasound and shear wave elastography in the differentiation of benign and malignant soft tissue tumors," *Skeletal Radiol.*, vol. 49, no. 11, pp. 1795–1805, 2020.

[20] I. Alam, A. Basit, and R. A. Ziar, "Utilizing age-adaptive deep learning approaches for detecting inappropriate video content," *Hum. Behav. Emerg. Technol.*, vol. 2024, no. 1, p. 7004031, 2024.

[21] S. Anari, N. T. Sarshar, N. Mahjoori, S. Dorosti, and A. Rezaie, "Review of deep learning approaches for thyroid cancer diagnosis," *Math. Probl. Eng.*, vol. 2022, no. 1, p. 5052435, 2022.

[22] Q. Guan *et al.*, "Deep learning-based classification of ultrasound images for thyroid nodules: a large-scale pilot study," *Ann. Transl. Med.*, vol. 7, no. 7, p. 137, 2019.

[23] H. Li *et al.*, "An improved deep learning approach for detection of thyroid papillary cancer in ultrasound images," *Sci. Rep.*, vol. 8, no. 1, p. 6600, 2018.

[24] X. Zhang, V. C. Lee, J. Rong, F. Liu, and H. Kong, "Multi-channel convolutional neural network architectures for thyroid cancer detection," *PLOS One*, vol. 17, no. 1, p. e0262128, 2022.

[25] C. Sun *et al.*, "Evaluation of a deep learning-based computer-aided diagnosis system for distinguishing benign from malignant thyroid nodules in ultrasound images," *Med. Phys.*, vol. 47, no. 9, pp. 3952–3960, 2020.

[26] Y. Wang *et al.*, "Comparison study of radiomics and deep learning-based methods for thyroid nodules classification using ultrasound images," *IEEE Access*, vol. 8, pp. 52010–52017, 2020.

ISSN: 3106-7743 | 3106-7735
Volume 3, Issue 1, 2026

[27] S. W. Kwon *et al.*, "Ultrasonographic thyroid nodule classification using a deep convolutional neural network with surgical pathology," *J. Digit. Imaging*, vol. 33, no. 5, pp. 1202–1208, 2020.

[28] W.-K. Chan *et al.*, "Using deep convolutional neural networks for enhanced ultrasonographic image diagnosis of differentiated thyroid cancer," *Biomedicines*, vol. 9, no. 12, p. 1771, 2021.

[29] F. Abdolali *et al.*, "Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks," *Comput. Biol. Med.*, vol. 122, p. 103871, 2020.

[30] Y. Lu, Y. Yang, and W. Chen, "Application of deep learning in the prediction of benign and malignant thyroid nodules on ultrasound images," *IEEE Access*, vol. 8, pp. 221468–221480, 2020.

[31] Kaggle, "Algerian Ultrasound Images Thyroid Dataset: AUITD," 2022. [Online]. Available: https://www.kaggle.com/datasets/azouzmaroua/algeria-ultrasound-images-thyroid-dataset-auitd. Accessed: 2025.