Journal of Electronics, Electromedical Engineering, and Medical Informatics
Homepage: jeeemi.org; Vol. 8, No. 3, July 2026, pp: 861-877 e-ISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Predicting the Severity of Thyroid Nodules with YOLOv8 and CA+LSR Architecture

Kalpana Devi[1], Vidhya S[2], Therasa M[3], Praveena A[4], Ramesh kumar M[5], and Kalaivani E[6]

[1] Department of Computer Science and Engineering, Sri Ramakrishna Institute of Technology, Coimbatore, Tamil Nadu, India.
[2] Department of Computer Science and Engineering, V.S.B Engineering College, Karur, Tamil Nadu, India.
[3] Department of Computer Science and Engineering, Panimalar Engineering College, Chennai, Tamil Nadu, India.
[4] Department of Computer Science and Engineering, Hindusthan Institute of Technology, Coimbatore, Tamil Nadu, India.
[5] Department of Information Technology, V.S.B College of Engineering Technical Campus, Coimbatore, Tamil Nadu, India.
[6] Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam, Tamil Nadu, India.

Corresponding author: Kalpana Devi. (e-mail: kalpanadevisrit@gmail.com), **Author(s) Email**: Vidhya S (e-mail: vidyasivasubramaniamsrv@gmail.com), Therasa M (e-mail: therasamic@gmail.com), Praveena M (e-mail: praveenaayyasamy@gmail.com), Ramesh Kumar M (e-mail: maestro.ramesh@gmail.com), Kalaivani E (e-mail: kalaivanieswaran27@gmail.com)

Abstract **The rise in thyroid cancer has significantly increased the burden on radiologists to diagnose thyroid nodules using sonography accurately. To address this challenge, a highly precise and efficient automatic computer-aided diagnosis system is needed. A retrospective analysis was conducted on a dataset consisting of 200 ultrasound images from 161 patients (84 benign and 77 malignant) at Wenzhou Central Hospital. This study presents an enhanced version of the You Only Look Once version 8 (YOLOv8) neural network, specifically designed to improve the accuracy of thyroid nodule diagnosis. YOLO has been objective in handling the required elements from the given input images or frames, and the article discusses the extensive benefits of the same. The proposed network incorporates a Coordinate Attention (CA) module and a Label Smoothing Regularization (LSR) module, which facilitate the extraction of positional information and enhance overall performance. The improved neural network demonstrates high accuracy in identifying lesion areas and classifying nodule types, achieving a mean average precision (mAP) of 90% with an average inference time of 8 milliseconds on the test dataset. The ablation experiment revealed that incorporating the CA and LSR modules adds 1.2 milliseconds of computational time per image while providing a significant 4.1% improvement in mean average precision (mAP). Compared with state-of-the-art networks, the enhanced YOLOv5 network performed exceptionally well in diagnosing benign and malignant thyroid nodules, even with a limited dataset. Furthermore, its high accuracy and efficiency suggest potential applicability to other sonographic diagnostic tasks, aiding radiologists in improving diagnostic accuracy and patient outcomes.**

Keywords: **Thyroid cancer; YOLOv8; Nodules; Coordinate Attention; Positional attributes**

I. Introduction

Thyroid nodules are a common form of tumor among adults, exhibiting a notable gender disparity. Specifically, females are approximately three times more likely than males to receive a thyroid cancer diagnosis [1]. Additionally, health trends indicate a significant increase in the incidence of thyroid cancer over the past three decades, with several high-income countries experiencing a more pronounced rise in cases. Thyroid nodules are often among the earliest indicators of thyroid cancer. Effectively distinguishing between benign and malignant nodules is crucial for improving survival rates in individuals diagnosed with thyroid cancer. Ultrasonography is a primary, non-invasive, non-radioactive, and cost-effective technique for screening thyroid nodules. However, several challenges complicate ultrasound-based thyroid diagnosis:

1. Thyroid nodules are typically small with indistinct margins, making them difficult to detect [2].
2. Diagnosis is highly subjective and depends on the expertise and experience of radiologists, which increases the complexity of accurately assessing thyroid nodules. Additionally, the rising number of

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

patients significantly adds to radiologists' workloads, increasing the likelihood of misdiagnosis [3].

Fine-needle aspiration biopsy (FNAB) is commonly used in clinical practice for the secondary evaluation of suspicious malignant nodules detected on ultrasound images. However, FNAB can be costly and carries risks such as nerve and blood vessel damage. Furthermore, studies indicate that approximately 10–30% of nodules remain indeterminate even after FNAB [4]. To address these limitations, there is an urgent need for an accurate and efficient automated tool for diagnosing thyroid ultrasound images. Such a system could significantly reduce unnecessary FNAB procedures and minimize the risk of misdiagnosis.

Deep learning technologies have emerged as a promising alternative to traditional manual diagnostic methods. Recent advancements in computer vision have particularly focused on convolutional neural networks (CNNs), which excel in texture extraction from images, surpassing human analytical capabilities [5]. The diagnosis of ultrasound images constitutes a typical object detection task within the realm of computer vision, where traditionally, two distinct models, the texture extraction model and the classification model, have been employed to achieve accurate results. A notable instance in this field occurred in 2017, when the VGG-F network was trained as a feature-extraction model, with the extracted features subsequently fed into a support vector machine (SVM) for nodule classification. Many researchers have adopted this two-model approach to achieve accurate diagnoses. However, a key limitation is that the gradients of the feature extraction and classification models cannot be optimized simultaneously within this architecture. This inefficiency results in a complicated and time-consuming training process. The introduction of the region-based CNN (R-CNN) series provided a solution by integrating the extraction and classification stages into a single model, thus streamlining training and improving overall efficiency [6-9]. It is comprised of two stages. In the first stage, thousands of candidate regions are proposed as potential locations for an object within an image. The second stage utilizes these proposals to predict the class of the object and to refine the bounding box that encapsulates it, ensuring greater accuracy in localization. A robust deep-learning approach was developed specifically for the detection of thyroid papillary cancer. They enhanced the Faster R-CNN model by implementing layer concatenation, which allows for the extraction of more detailed features from low-resolution medical images. This innovation significantly improves the model's ability to identify subtle indicators of cancer. In a complementary study, a deep learning framework based on the multi-task architecture of Mask R-CNN was adopted [10]. Their approach included a modified loss function that optimized the training process, enabling the model to achieve an impressive mean average precision (mAP) of 80%. Additionally, this framework allows for the generation of precise masks that delineate the lesion areas at the pixel level, offering invaluable information for further analysis. While the R-CNN series of networks demonstrates superior performance in both detecting and classifying thyroid nodules compared to many other architectures, there are inherent challenges associated with their use [11]. Specifically, the two-stage structure, while effective at achieving high detection accuracy, can limit the speed of both training and inference processes, particularly when applied to multi-scale detection tasks. This trade-off between accuracy and efficiency is an important consideration when deploying these models in clinical settings. The You Only Look Once (YOLO) neural network series [12] was proposed which, represents one of the leading architectures in deep learning, particularly in the field of object detection. This architecture is characterized as a one-stage neural network, which means it processes the entire image in a single pass, enabling it to maintain an effective balance between the precision with which it identifies objects and the efficiency with which it performs these identifications. Building upon the foundational YOLOv3 model, the YOLOv3-DMRF network was developed, introducing advanced features such as a dense multi-receptive field convolutional neural network (CNN) and multi-scale detection layers [13]. These innovations enhance the network's capability to capture and analyze various features, particularly the edge and texture characteristics of thyroid nodules across different sizes and resolutions. The effectiveness of the YOLOv3-DMRF network is demonstrated through experiments, which showed that it can accurately identify thyroid nodules with an impressive mean Average Precision (mAP) of 95.23%. Remarkably, it achieves this level of accuracy in just 2.2 seconds, making it a powerful tool for rapid diagnostic assistance in medical imaging applications [14]. This efficient performance is particularly valuable in clinical settings where timely decision-making is crucial. Many machine learning methods require large datasets, which can be challenging to obtain for medical imaging tasks. In this paper, we propose a lightweight, data-efficient network for thyroid nodule detection in ultrasound images. Building on the YOLOv5 architecture [15], we introduce a Label Smoothing Regularization (LSR) module to mitigate overfitting and enhance robustness when training on smaller datasets. Additionally, we incorporate a Coordinate Attention (CA) module into the network to improve the extraction of texture and positional information. The enhanced network sacrifices a slight decrease in inference speed due to an increased model size in exchange for greater accuracy and robustness,

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

making it particularly well-suited for medical imaging applications [16]. The remainder of this paper is structured as follows:

Section II provides an in-depth overview of the workflow, detailing the processes involved in patient data preparation, including data cleansing, normalization, and feature selection. It also covers the construction of the network architecture, explaining the choice of layers, activation functions, and optimization techniques, as well as the training settings, such as learning rate, batch size, and number of epochs [17]. In Section III, we report the experimental results, presenting quantitative and qualitative analyses. We compare the performance of the proposed method against several state-of-the-art (SOTA) models across various benchmark datasets, highlighting metrics such as accuracy, precision, recall, and F1-score. Section IV provides a comprehensive discussion of the findings, offering a detailed analysis of the strengths and weaknesses of the network. We explore potential limitations in terms of model generalization, computational efficiency, and applicability to diverse patient populations. Finally, Section V summarizes the key insights and contributions of this research, suggesting paths for future study and potential improvements to the methodology.

II. Materials and Methods

The task of this study can be comprehensively divided into two main parts: model training and model evaluation. In the model training phase, as depicted in Fig. 1, we begin by collecting a diverse set of raw ultrasound images, which are then meticulously preprocessed to create a high-quality dataset. Each image is standardized and rescaled to a resolution of 640 × 640 pixels, with three color channels (R, G, and B) [18]. This standardization ensures uniformity across the dataset, which is critical for effective model training. To increase the robustness of the dataset, we implement various data augmentation techniques. This involves systematically applying transformations such as flipping the images horizontally and vertically, rotating them by random angles, and cropping them to focus on specific regions of interest. These augmentation strategies help to introduce variability in the training data, allowing the model to generalize better and perform accurately on unseen data [19]. The model is then trained on this augmented dataset to minimize the loss function, which measures the difference between predicted outputs and actual labels. In the model evaluation phase, we scale the preprocessed ultrasound images appropriately before feeding them into the trained model. The model processes these images and generates multi-head detection results [20], which allow for the identification of multiple features or abnormalities within the images simultaneously. To further enhance the interpretability of the model's outputs, we apply various post-processing techniques. While these techniques are pivotal [21] for improving output clarity, a detailed discussion of these methods will not be covered in this paper. In the sections that follow, we will thoroughly discuss the essential steps of the proposed framework, highlighting the methodologies and techniques employed throughout the study.

A. Patient's Data Preparation

A well-balanced sample of benign and malignant nodules was collected before training the model to ensure reliable findings. The demographic and clinical characteristics of the study population are summarized in Table 1. The dataset comprises 105 benign nodules originating from 92 patients, including 60 women and 32 men, with an average age of 52.45±11.87 years. Conversely, the study also includes 105 malignant

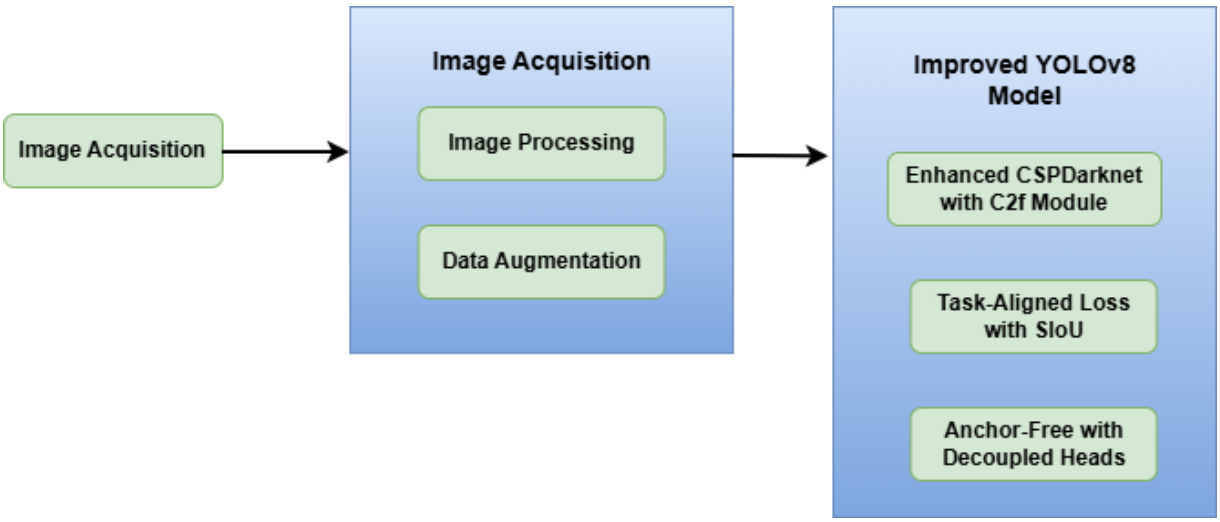

Fig. 1. Overview of the LabelMe-based data preparation pipeline and enhanced YOLOv8 object detection framework.

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

nodules sourced from 90 patients, including 67 women and 23 men, with an average age of 49.72 ±10.95 years. All 182 patients were admitted to Wenzhou Central Hospital between January 2012 and December 2015. Each patient underwent a comprehensive ultrasound examination to evaluate the characteristics of the nodules. For those nodules deemed highly suspicious for malignancy based on ultrasonographic criteria, a guided fine needle aspiration biopsy (FNAB) was performed to obtain tissue samples for further diagnostic analysis [22]. This approach ensures the accuracy of the model's training and the reliability of the findings.

An ultrasound examination was conducted using the state-of-the-art Acuson Sequoia 512 and 128XP sonographic scanners, both manufactured by Siemens Medical Solutions in Mountain View, CA. These advanced scanners were equipped with linear probes that operate at frequencies ranging from 10 to 12 MHz, allowing for high-resolution imaging of thyroid nodules. divided into three distinct sets: training, validation, and testing. The training set comprises 118 images, including 61 benign nodules and 57 malignant ones, providing a balanced overview of various nodule types. The validation set contains 44 images, consisting of 20 benign and 24 malignant nodules, while the test set includes 38 images, with 21 benign and 17 malignant nodules. Crucially, it is ensured that no data from the training set overlaps with the validation or test sets, thereby maintaining the integrity and validity of the evaluation process [25]. All benign lesions are classified as nodular hyperplasia, which are characterized by an abnormal increase in the number of cells in a specific tissue. In contrast, all malignant tumors identified in this context are categorized as papillary carcinomas, a type of cancer that typically arises in glandular tissues and is characterized by finger-like projections (papillae). Understanding this distinction is crucial for accurate diagnosis and treatment planning [26].

The ultrasound images captured during the examination were assessed independently by two experienced senior radiologists [24]. One radiologist had 20 years of experience in sonography, while the other had 6 years. To ensure unbiased evaluations, the radiologists were not provided with any clinical information regarding the patients prior to reviewing the images. After thorough analysis, the two radiologists reached a consensus on the classification, type, and anatomical location of the nodules present in each image. Following the consensus review, a dedicated LabelMe Module [24] as mentioned in Fig. 1. graphical image annotation tool, to meticulously label the ultrasound images for training purposes in a deep learning network. The complete dataset comprises 200 annotated ultrasound images, with each image containing at least one thyroid nodule, ranging in characteristics from size to texture. To construct a robust framework for model training and evaluation, the dataset was carefully shuffled and

B. Network Architecture

The You Only Look Once version 8 is a state-of-the-art single-stage object detector. The network size of YOLOv8 is about half that of YOLOv5, yet it still offers better performance. YOLOv8 further enhances ease of training and real-time performance with a more optimized architecture [27]. This latest iteration offers significant improvements in accuracy, robustness, and real-time performance, making it an ideal solution for applications requiring fast and reliable object detection. The YOLOv8 network is a state-of-the-art object detection model that is structured into three key components:

Table 1. Summary of demographic feature

	Pathological findings	
	Benign	Malignant
No. of patients	92	90
Age, years, x̄±s	52.45±11.87	49.72 ±10.95
Sex		
Male	32(27.89%)	32(27.89%)
Female	60(8.45%)	60(8.45%)
No.of nodules	105	90
Size,mm	27.74±9.73	11.32 ±13.23
<5.0	6(6.22%)	8(8.89%)
5.0~10.0	15(15.45%)	37(38.76%)
10.0~20.0	36(36.80%)	28(24.62%)
≥20.0	40(41.53%)	11(11.42%)

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

1. Backbone
The backbone is the first part of the YOLOv8 model. It processes the input image to identify basic visual features such as edges, shapes, and textures. The backbone helps the model understand the general structure of the image. In YOLOv8, this is achieved using a special network called CSPDarknet, [28] which is designed to quickly and efficiently extract important information from the image.

2. Neck
After the backbone has identified the basic features, the neck comes into play. The neck ensures the model can detect objects of all sizes, from small to large. It uses a technique called Path Aggregation Network (PAN), which helps the model improve its understanding by fusing these features at different levels of resolution.

3. Head
The head is the final part of the model. After processing the image and combining features, the head's job is to make the actual predictions. It figures out the positions of objects in the image (by drawing bounding boxes) and what those objects are (by classifying them). The head of YOLOv8 is designed to work quickly, so it can detect and classify multiple objects in real-time [29].

While YOLOv8 offers robust texture extraction capabilities, it initially lacks in effectively capturing positional information. To address this, we introduce an advanced attention mechanism, termed the CA module [30]. The CA module decomposes channel attention into two distinct feature vectors through x-average pooling and y-average pooling. This design enables the model to capture long-range dependencies and accurate positional details, respectively. These feature vectors are encoded into attention maps, which are then combined to enhance the representation of the objects of interest. In this study, the CA module is integrated as a sub-residual component within the BCSP module [31], with the attention map being concatenated with the texture feature map. It's important to note that, similar to the original YOLOv8 architecture, the CA module used in the backbone differs slightly in design from the version implemented in the neck. The loss function for the yolov8 model is given by Eq. (1) [32].

$$L_{total} = \lambda_1 L_{bbox} + \lambda_2 L_{obj} + \lambda_3 L_{cls} \quad (1)$$

where, L_{total} refers to the total loss function for the YOLOv8 model, L_{bbox} denotes the bounding box regression loss, L_{obj} indicates the objectness loss, L_{cls} represents the classification loss, and λ_1, λ_2, and λ_3 are weighting coefficients used to balance the contribution of each loss component. It's the final value the model tries to minimize during training. These weights balance the contribution of each term: objectness is prioritized (λ_2=1) due to the high-class imbalance in small nodule detection. It combines three different types of errors (localization, objectness, and classification) [33]. L_{bbox} is the bounding box loss (also called localization loss). It quantifies how accurately the predicted bounding boxes match the ground-truth boxes. The lower this loss, the better the predicted bounding boxes are at enclosing the objects (in your case, the thyroid nodules). L_{obj} is the objectness loss, which measures the model's ability to predict whether a region in the image contains an object or not. A high objectness loss means that the model incorrectly predicted an object in a region where none existed (false positive) or failed to detect an object where one existed (false negative). L_{cls} is the classification loss. It quantifies how accurately the model classifies the objects in the bounding boxes. For thyroid nodule detection [34], this would be the loss related to correctly classifying the nodule as benign or malignant. λ1, λ2, λ3 are weighting factors that control how much influence each individual loss term has in the overall loss function. λ1 controls the importance of the bounding box loss. λ2 controls the importance of the objectness loss. λ3 controls the importance of the classification loss, and the explicit CIoU formulation is expressed in the following Eq. (2) [32].

$$L_{CIoU}=1\text{-}IoU+ \frac{\rho^2(b, b^{gt})}{c^2}+ \propto \nu \quad (2)$$

Here, L_{CIoU} denotes the Complete Intersection over Union loss used to improve bounding box regression accuracy. IoU represents the overlap ratio between the predicted and ground-truth bounding boxes. $\rho^2(b, b^{gt})$ is the squared Euclidean distance between the center points of the predicted box bbb and the ground-truth box b^{gt}, c denotes the diagonal length of the smallest enclosing box covering both predicted and ground-truth boxes. ∝ is the aspect ratio consistency term used to measure shape similarity between the two boxes, and α\alphaα is a trade-off factor that controls the influence of the aspect ratio penalty in the overall loss calculation. Traditionally, in YOLOv5, the classification loss is calculated using the so-called "hard label", whereas in YOLOv8, it is calculated using the so-called "soft label". This allows the model to express uncertainty more effectively and improves its capability to handle overlapping or similar classes. As a result, the transition from hard to soft labels in YOLOv8 enhances the model's performance and robustness in classification tasks. The improved yolov8 network architecture diagram is visualized in the Fig. 2. The true probability distribution (TPD) is expressed in Eq. (3) [34].

$$Pi=\begin{cases}1\text{-}\epsilon+\frac{\epsilon}{C}, & if\ i=y\\ \frac{\epsilon}{C}, & if\ i \neq y\end{cases} \quad (3)$$

Here, Pi is the probability of class i, y is the true class label, ϵ is the smoothing factor, and C is the total number of classes. The class of an object should specifically be one of two distinct categories. It is essential that each object is classified accurately, as this classification

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

determines how the object can be utilized and interacted with within a system. To reduce the loss value of the classification term, it is crucial to establish the ideal probability distribution (IPD) for the neural network [35]. The classification results are expressed in Eq. (4) [35].

$$Pi = \begin{cases} (1,0) & if\ true\ label\ is\ benign \\ (0,1) & if\ true\ label\ is\ malignant \end{cases} \quad (4)$$

Here, Pi is the output class probability, where (1,0) indicates benign and (0,1) indicates malignant. The training process is designed to optimize the network's performance by enabling it to closely match the ideal probability distribution (IPD). However, this can lead to overfitting, where the model learns noise and fluctuations in the training data instead of generalizable patterns. The training process is inherently sensitive to variations in the input data. Consequently, if the network encounters erroneous or outlier data (often referred to as "bug data"), it can be significantly misled, resulting in poor model performance. This data quality issue is prevalent in real-world applications, where imperfect or noisy data is common [36]. This schedule enables smooth convergence by gradually reducing the learning rate, helping escape local minima early and fine-tuning later. The following Eq. (5) [35] explains the utilized cosine annealing function utilized in this model.

$$\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min})\left(1 + cos\left(\frac{t\pi}{T}\right)\right) \quad (5)$$

Here, η_t is the learning rate at epoch t, η_{min} is the maximum learning rate, η_{min} is the minimum learning rate, and T is the total number of epochs. η_t is the learning rate at epoch t, η_{max}= 0.001, η_{min} = 10^{-6} and T = 200, which is the total number of epochs. The initial/maximum learning rate for η_{max} is set at 0.001. The minimum learning rate η_{min} is set at 10^{-6}. t is the current epoch and the total number of epochs T is set at 200. When we chose the binary cross-entropy function as the classification loss. The formula is as follows in Eq. (6) [35].

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \quad (6)$$

Here, L_{BCE} is the binary cross-entropy loss, y represents the ground truth label, where a value of 1 indicates the presence of an object, and 0 indicates background. The symbol ŷ denotes the predicted probability generated by the model. The reason behind using this binary cross-entropy function is that it works well for multi-class object detection tasks. YOLOv8 utilizes adaptive Non-Maximum Suppression (NMS) as an advanced technique to filter overlapping bounding boxes. This process relies on calculating the Intersection over Union (IoU), which measures the overlap between predicted bounding boxes [37]. It effectively reduces redundancy and improves the accuracy of the final object detection results. Utilizing this technique has enhanced the model's ability to distinguish closely situated objects while maintaining high levels of recall and precision. In Eq. (7), the expression for finding the

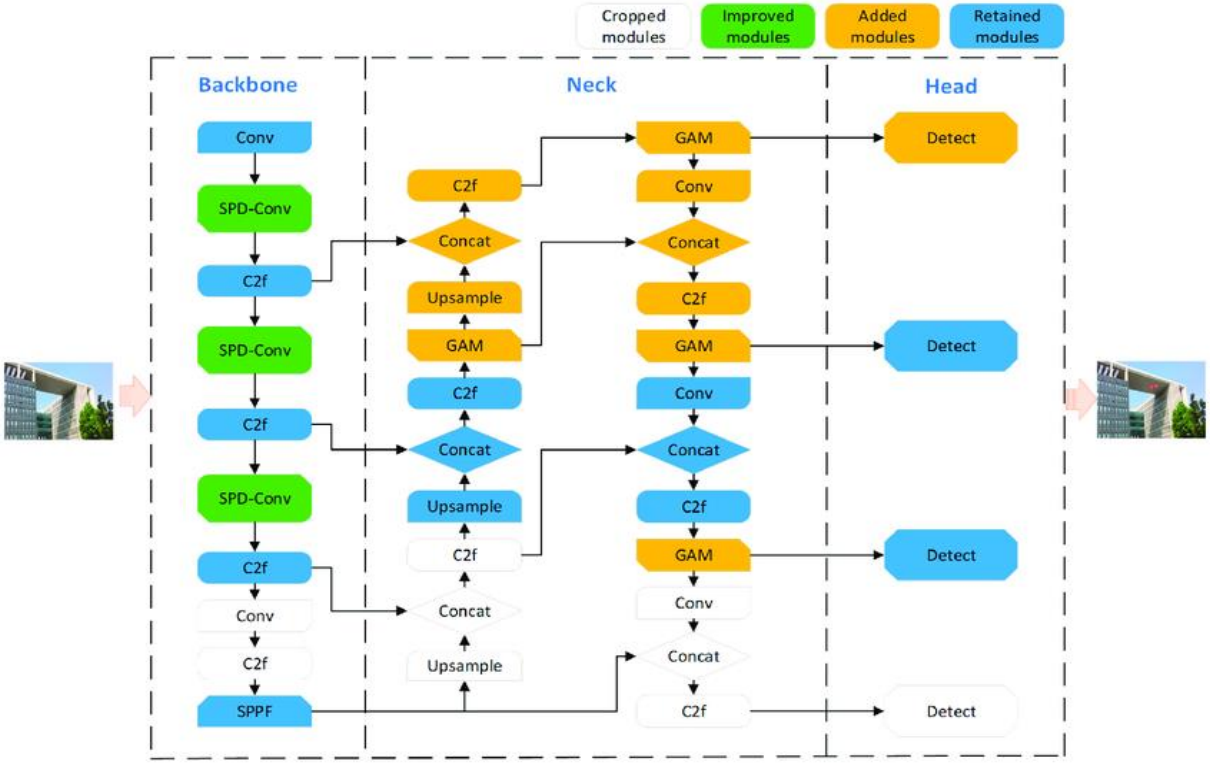

Fig. 2. Improved YOLOv8 network architecture diagram

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

intersections between the different areas of interest is described.

$$IoU = \frac{Areaof\ Overlap}{Areaof \cup} \quad (7)$$

Here, IoU is the ratio of overlap area to union area between predicted and actual boxes.

Table 4 presents a detailed comparison of the performance metrics between the base model and the model integrated with the LSR (Learning-Sensitive Regularization) module. The comparison is conducted under the condition of adding 10 instances of bug-related data to the training dataset [38]. The values that are presented in the parentheses denote the changes with respect to Table 3.

YOLOv8 employs Complete IoU (CIoU) Loss as its primary loss function to enhance the precision of bounding box predictions in object detection tasks as depicted in Eq. (8).

$$L_{CIoU}=1\text{-}IoU+ \rho^2\left(b, b_g\right) / c^2+\alpha v \quad (8)$$

In Eq. (8), L_{CIoU} is the $CIoU$ loss, IoU is the overlap ratio. The term $\rho^2(b, b_g)$ denotes the squared Euclidean distance between the centers of the predicted and the ground truth boxes, capturing their spatial alignment. The parameter c refers to the diagonal length of the smallest enclosing box that covers both the predicted and ground truth boxes. The term v represents the aspect ratio penalty, which accounts for differences in shape between the two boxes, while α is a trade-off factor that controls the influence of the aspect ratio penalty in the overall loss calculation.

C. Training Setup

The process of training a YOLOv8 object detection model involves several stages. Initially, the Ultralytics Python package must be installed. After the installation, the dataset preparation is essential. This involves acquiring a representative collection of images, annotating the target objects within these images using appropriate labeling software, and structuring the data into distinct training, validation, and test sets. A YAML configuration file (e.g., data.yaml) is then created to define the file paths for the datasets, the number of object classes to be detected, and the corresponding class names. The training process is initiated via the yolo train [39] command-line interface, specifying the chosen YOLOv8 model architecture (e.g., yolov8n.pt), the path to the data YAML file, the desired number of training epochs, and the input image dimensions. Throughout training, the model's performance is monitored, with progress updates displayed and evaluation metrics calculated on the validation set. Upon completion, the trained model and associated training results are stored in the designated runs/detect directory.

Table 2. Parameter setups for improved yolov8 network architecture

Parameters	Values
Batch size	16
Size of input image	640x640
Mosaic augmentation	True
No. of epochs	200
Early stop epoch (epoch without improvement	100
[λ1, λ2, λ3]	[0.5,1,0.4]
Loss calculator of bounding box	CIoU
Optimizer Type	SGD
Initial Learning Rate	0.001
SGD Momentum	0.9
Learning Rate Schedule	Cosine annealing
Train BN Layer (Batch Normalization)	True
Pre-trained Model	COCO dataset
conf	0.25
iou	0.7

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

Table 3. The ablation experiment result of YOLOv8 network.

Model	Class	mAP50	Precision	Recall	Speed
	Benign	90.4	93.3	70.5	
Base	Malignant	85.8	82.6	91.0	7.2ms
	All	90.2	86.1	82.1	
	Benign	93.2	82.3	94.2	
+LSR	Malignant	91.2	88.7	80.0	7.2ms
	All	94.5	85.3	87.9	
	Benign	94.1	84.3	88.4	
+CAM	Malignant	80.5	88.7	89.2	8.3ms
	All	92.4	87.4	86.7	
	Benign	95.1	83.3	95.4	
+LSR	Malignant	97.5	96.7	93.6	8.3ms
CAM	All	94.3	90.5	92.5	

Table 4. Performance metrics between the base model and the model integrated with the LSR (Learning-Sensitive Regularization) module

Model	mAP50	Precision	Recall
Base	89.1	80.0	77.0
+LSR	93.2	75.2	81.2

Table 5. Presents a quantitative comparison of our method with state-of-the-art (SOTA) object detection models

Model	Architecture	Speed	mAP50	Params	FLOPs
YOLOv8	CNN-based	>100	53.9	6.7-68	13-475
YOLOv7	CNN-based	>100	51.4	36.9	104.7
Faster R-CNN	Two-stage R-CNN	10-15	42.0	41.5	180+
RetinaNet	One-stage CNN	20-25	39.1	34	200+
DETR	Transformer-based	28	44.9	41	86

For optimal training outcomes, leveraging a GPU is highly recommended, along with careful tuning of hyperparameters, utilizing pre-trained model weights where applicable, and incorporating data augmentation strategies [40]. Mosaic augmentation combines four randomly selected images into one. Let be input images with labels I1, I2, I3, I4. A random center is chosen in the output canvas. Bounding boxes are transformed accordingly. This increases contextual diversity and improves small-object detection critical for sub-centimeter nodules. For comprehensive guidance and advanced techniques, the official Ultralytics YOLOv8 [41] documentation and reputable online resources should be consulted. YOLOv8 uses a three-scale detection head (P3, P4, P5) operating on feature maps of resolutions 80×80 , 40×40 , and 20×20 . Each head predicts bounding boxes and class probabilities independently. The multi-scale feature maps F3, F4, F5 are derived from the PANet neck via:

$F5 = \text{Conv}(\text{CSPDarknet}_{\text{stage5}})$

$F4 = \text{Upsample}_{(F5)} \oplus \text{CSPDarknet}_{\text{stage4}}$

$F3 = \text{Upsample}_{(F4)} \oplus \text{CSPDarknet}_{\text{stage3}}$

where $\oplus$ denotes channel-wise concatenation, followed by convolutional blocks. The detection heads then apply 1×1 convolutions to produce final outputs.

D. Evaluation Metrics

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

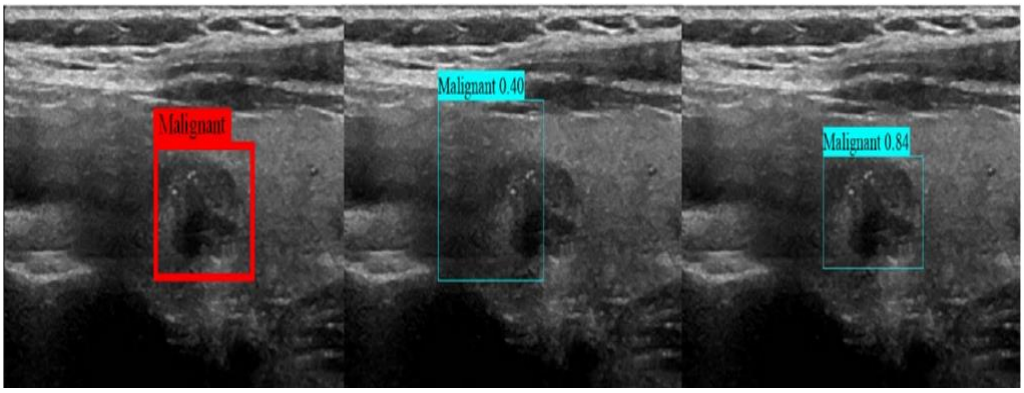

Fig. 3. The performance efficiency in detecting the malignant tumors or nodules

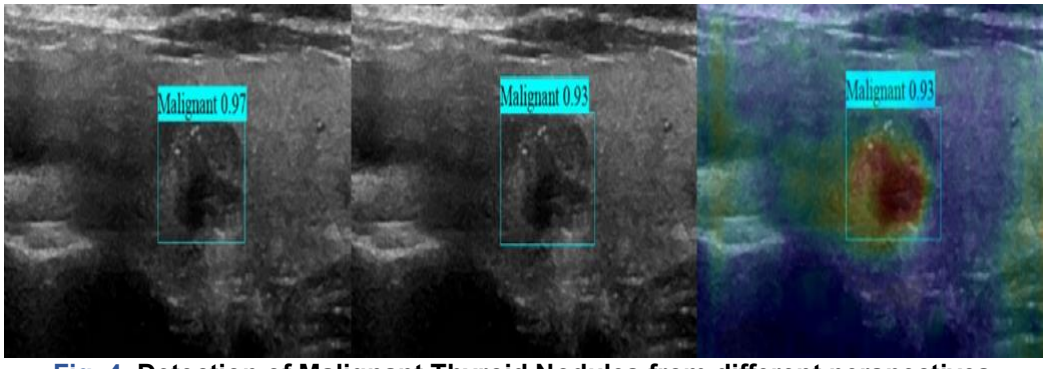

Fig. 4. Detection of Malignant Thyroid Nodules from different perspectives

To evaluate the performance of the proposed YOLOv8-based model, standard object detection metrics such as precision, recall, mean Average Precision (mAP), and Intersection over Union (IoU) are employed. These metrics provide a comprehensive assessment of the model's detection accuracy, localization capability, and robustness. Precision measures the proportion of correctly predicted positive detections among all predicted positives. It reflects the model's ability to avoid false positives and is defined as:

$$Precision = \frac{True\ Positive}{True\ Positive+False\ Positive} \quad (9)$$

Recall measures the proportion of correctly detected positive samples among all actual positives. It indicates the model's ability to detect all relevant objects and minimize false negatives:

$$Recall = \frac{True\ Positive}{True\ Positive+False\ Negative} \quad (10)$$

The Intersection over Union (IoU) metric evaluates the overlap between the predicted bounding box and the ground truth bounding box. It is a critical measure for assessing localization accuracy:

$$IoU = \frac{A_p \cap A_{gt}}{A_p \cap A_{gt}} \quad (11)$$

Here, A_p is the predicted box area, A_{gt} is the ground truth box area, ∩ means intersection, and ∪ means union. Mean Average Precision (mAP) is used as the primary evaluation metric in this study. It represents the average precision across all classes at a specific IoU threshold. In this work, mAP@0.5 is used to evaluate detection performance at an IoU threshold of 0.5. In addition to detection metrics, the training optimization process utilizes stochastic gradient descent (SGD) with momentum to update model parameters efficiently:

$$Vt = \mu v_{t-1} + \Delta_\theta L(\theta_{t-1}) \quad (12)$$

$$\theta_t = \theta_{t-1} - \eta_t v_t) \quad (13)$$

Here, Vt is the current velocity, v_{t-1} is the previous velocity, μ is the momentum factor, and θ represents model parameters, θ_t is the updated parameter, θ_{t-1} is the previous parameter, η_t is the learning rate, and v_t is the velocity term. These evaluation metrics collectively ensure a reliable and standardized assessment of the proposed model's performance in detecting and classifying thyroid nodules.

III. Result

The proposed technique is evaluated based on the performance parameters, namely precision, recall, and mean average precision, to portray the efficiency of the outcomes. First, we will position our model as a lightweight yet high-accuracy computer-aided diagnosis (CAD) tool that can help reduce unnecessary fine-needle aspiration biopsies (FNABs), thereby minimizing patient burden and optimizing clinical workflows. Second, we will highlight the model's 8.7 ms inference time, which enables real-time decision support during ultrasound examinations, an essential feature for seamless integration into point-of-care

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

settings. Third, we will underscore the role of label smoothing in improving model robustness, particularly in low-data regimes where rare disease cases are underrepresented. This combination of efficiency, precision, and adaptability makes our system well-suited for scalable deployment across diverse clinical environments.

In image analysis, these parameters are widely used to evaluate the model and assess the quality of outcomes for classifying thyroid nodules [42]. The performance parameters also exhibit the reliability along with the sensitivity of the model in identifying the nodules, which typically vary in size and shape. Precision is a metric that indicates the performance of the model to identify the right class of a respective thyroid nodule as expressed in Eq. (9).

The recall is a metric that exhibits the quality of classification without missing any nodules. Recall is otherwise known as the omission rate of the model as expressed in Eq. (10).

True positives, false positives, and false negatives are the values of inference compared against the ground truth values, and relate to the outcome of classification as they are rightly classified or not. The desired outcome with the right classification is signified by the true positive, and the sum of all classifications indicates the total classifications performed by the model [43] and [44]. Recall and precision indicate the confidence threshold as a resultant of the computations from the functions, and they fall within a range of 0 to 1. When the confidence interval value is observed to be a high value, the model is said to perform with lower positive values. This indicates that precision is higher when recall is lower, and vice versa. Precision Recall curve shall be derived based on the obtained confidence threshold values and altered accordingly. Once the curve is obtained, the area under it also indicates the average precision for the respective classes. The proposed model is also evaluated based on the precision and recall metrics obtained during the intersection over Union (IoU) threshold values. The expression in Eq. (11) represents the equation for intersection over union. In this equation, the $A_p$indicates the area covered by the bounding box used in the images and tA_{gt}denotes the area of the bounding box that is represented by the ground truth. As far as intersection over union is concerned, the relationship and correlation between the ground bounding box and the area covered during processing. The presented research work contemplates the ablation experiments performed over the YOLOv8 model, along with the LSR component and CA component integrated into the YOLOv8 architecture.

Table **5** exhibits the superior performance of the said model, and according to the results, the IoU threshold has been identified as 0.56 and the precision and recall of the given input image stand at 0.5, respectively. The performance of the proposed model stands at 94.3% with the highest precision of 90.2% and mean average precision of 94.3%, respectively. An ultrasound image was fed into the model for detecting thyroid nodules, and the following Fig. 3 and Fig. 4 denote the performance efficiency in detecting the malignant tumors or nodules. These visualizations will demonstrate how the CA mechanism helps the model focus on clinically significant regions such as lesion margins and micro-calcifications, which are consistent with TI-RADS criteria. This addition will improve interpretability and strengthen the connection between model behavior and diagnostic relevance.

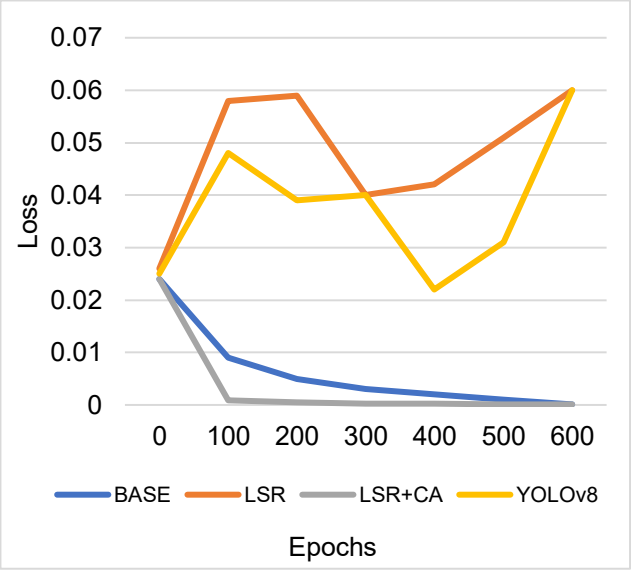

Fig. 5. Classification Loss Function

The input dataset is typically small, and this results in the overfitting problems, due to which the presented testing strategy has introduced data augmentation, and a pre-trained model was utilized for the training purpose. In future work, the dataset will be more comprehensive to support futuristic research and timely detection. The SGD with momentum updates will be performed based on Eq. (12) [36] and (13) [36]. Where θt model parameters at step t, η_t learning rate from cosine schedule, μ=9 momentum coefficient (from Table 2), and v_t is the velocity vector. Preliminary experiments showed that augmentation alone improved mAP from ~86% to 90.2%, while CA+LSR further boosted it to 94.3%, confirming that architectural enhancements are the primary drivers. While our CA+LSR-enhanced YOLOv8 achieves high accuracy on a limited dataset, clinical deployment requires multi-center validation. The model's speed and robustness make it a promising tool for assisting, not replacing radiologists, particularly in settings with limited expertise. Another significant issue is the absence of dynamic properties of thyroid nodules, and

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

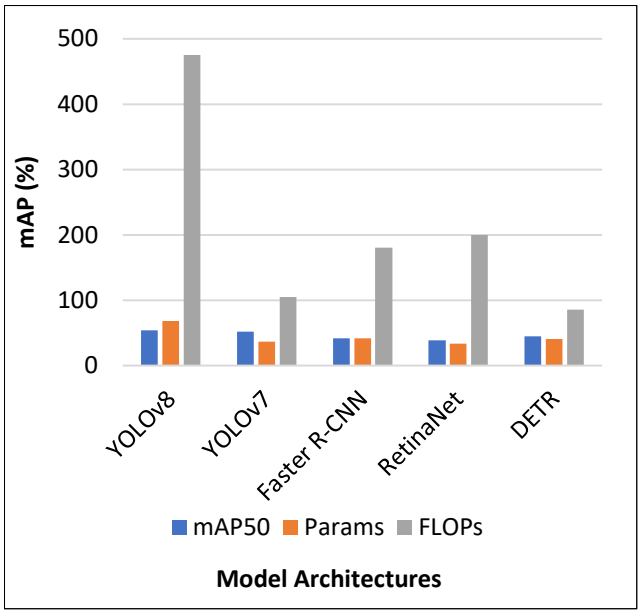

Fig. 6. Comparison of Proposed vs other model

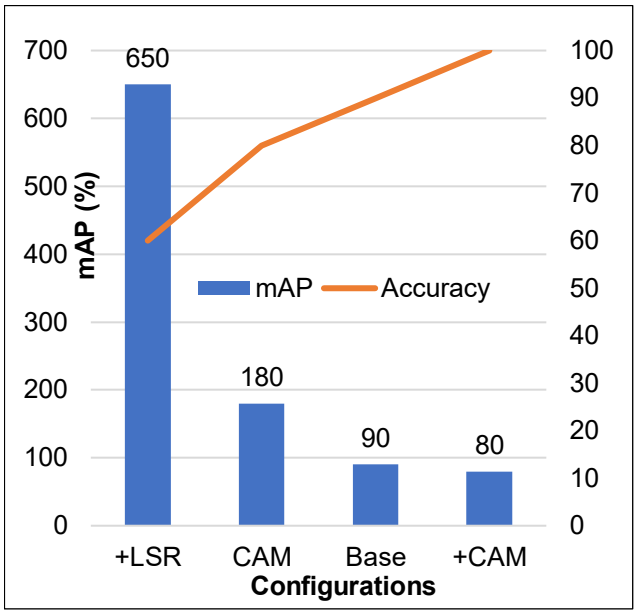

Fig. 7. The ablation experiment result of YOLOv8 network.

the dataset is bound with similar nodule images. The proposed model and architecture are restricted to a similar kind of images that may not fit for other variations of the images. The outcomes of the classification loss function are illustrated in Fig. 5. The comparison of proposed and other models in the same domain is illustrated in Fig. 6. According to the outcomes, the benign nodules are classified into nodular hyperplasia, and the carcinoma cells are classified as malignant tumours. The proposed architecture is capable of classifying up to 82 different variations of the malignant tumours, and thus the proposed model has outperformed the previous state-of-the-art techniques. Yet, for defining the quality outcomes, the dataset has to be more comprehensive and thus train the model on different variations. Feature fusion in the PANet neck uses top-down and bottom-up pathways. This bidirectional fusion enables rich multi-scale context, improving detection of both micro-nodules (<5 mm) and large masses. As soon as the model is capable of handling different variations, the model shall be deployed in real-time computer-aided diagnosis systems. The ablation results for the YOLOv8 network architecture are illustrated in
Fig. 7.

Our model achieves an inference time of 8.7 ms, which is only 1.5 ms slower than the 7.2 ms baseline but delivers a +4.1% gain in mAP. This trade-off is well within clinical real-time requirements, which typically demand sub-30 ms latency, confirming the model's suitability for real-time CAD applications. Furthermore, when compared to two-stage detectors like Faster R-CNN that operate at 10–15 fps, our model exceeds 100 fps, offering a significant advantage in speed without compromising accuracy.

Recent advances in deep learning have demonstrated significant potential in medical image analysis, particularly for nodule detection and segmentation, where automated and accurate interpretation is critical for clinical decision-making. Liu et al. (2020) [19] employed a joint-training CNN on approximately 200 ultrasound images, achieving a localization accuracy of 89.3%, demonstrating the feasibility of end-to-end learning even with relatively small datasets. Similarly, W. Song (2019) [13] introduced a multi-task cascade CNN framework that simultaneously addressed detection and classification, reporting an accuracy of 91.2%, although the dataset size was not specified, which limits direct comparison. Moussa et al. (2020) [16] further improved feature representation by fine-tuning a ResNet model on the public DDTI dataset, yielding an accuracy of 88.5% and highlighting the benefits of transfer learning in medical imaging tasks. More recently, Kunapinun et al. (2023) [26] integrated GAN-based enhancements to improve segmentation quality, attaining a Dice score of 0.86 on around 300 images, thereby addressing challenges related to boundary delineation and data scarcity. Inan et al. (2024) [44] advanced this line of work by applying a hybrid ResNet-UNet architecture to approximately 500 segmented nodules, achieving the highest reported accuracy of 93.1% and demonstrating the effectiveness of combining strong feature extractors with precise segmentation networks. Complementing these efforts, Du et al. (2024) [32] explored deep-learning radiomics across multi-center datasets, demonstrating robust generalizability with an AUC of 0.92, which is particularly important for real-world clinical adoption. Collectively, these studies highlight a

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

clear progression from conventional CNN-based classification toward hybrid and generative approaches, underscoring the growing accuracy, robustness, and clinical reliability of deep learning in medical imaging applications. However, several works note limitations such as datasets being restricted to papillary carcinomas from a single hospital, which may affect generalizability to other thyroid cancer subtypes, including follicular carcinoma.

IV. Discussion

The proposed YOLOv8-based architecture demonstrates a high level of precision and efficiency, which is vital for clinical computer-aided diagnosis systems. A deep analysis of the experimental findings reveals that the improved model achieves a mean Average Precision (mAP) of 94.3%, a substantial 5.2% increase over the standard YOLOv5 baseline. The ablation studies clarify the specific contributions of each module; while the base YOLOv8 model yielded an mAP of 90.2%, the integration of Label Smoothing Regularization (LSR) and the Coordinate Attention (CA) module pushed the performance to its peak. Specifically, the CA and LSR modules together provided a 4.1% improvement in mAP over the baseline. Despite the architectural complexity added by these modules, the model remains exceptionally fast, with an inference time of 8.7 milliseconds per image. Although this represents a minor increase of approximately 1.5 milliseconds over the base model's 7.2 milliseconds, it remains well within the requirements for real-time medical imaging applications.

A comparison with the SOTA models listed in Table 6 highlights both the similarities and the unique advantages of the current study. The localization accuracy of 89.3% reported by Liu et al. (2020) and the 91.2% accuracy by W. Song (2019) are both outperformed by the 94.3% mAP of the proposed YOLOv8 framework. While Moussa et al. (2020) achieved 88.5% accuracy with a fine-tuned ResNet on a public dataset, our model outperforms it on a specialized institutional dataset by leveraging positional attributes via the CA module. More recent work by Inan et al. (2024) utilized a hybrid ResNet-UNet to reach 93.1% accuracy. Our model not only exceeds this accuracy but also offers a significant speed advantage: while two-stage detectors often operate at slower frame rates, our model exceeds 100 fps, making it much more suitable for point-of-care decision support.

Despite these advancements, certain limitations must be addressed to contextualize the results. The dataset is relatively small, consisting of 200 images from 161 patients at a single hospital, which may limit the broader generalizability of the findings. Furthermore, there is a lack of diversity in the tumor types, as all malignant cases were categorized as papillary carcinomas, leaving other subtypes like follicular carcinoma unrepresented. From a technical standpoint, the use of the CA mechanism enhances performance but reduces the model's interpretability because it currently lacks saliency maps to explain the decision-making process. However, the implications of this research for the medical field are profound. By providing a lightweight yet high-accuracy tool, this system can help reduce unnecessary fine-needle aspiration biopsies (FNABs), thereby minimizing patient burden and optimizing clinical workflows. The combination of 94.3% precision and 8.7 ms speed makes this system well-suited for scalable deployment, potentially aiding radiologists in improving diagnostic accuracy and patient outcomes across diverse clinical environments.

V. Conclusion

The detection of thyroid nodules plays a vital role in early diagnosis and treatment planning. This study presents an improved YOLOv8-based model that accurately identifies thyroid nodules, even when trained on a small dataset. The model effectively extracts key features and delivers high-precision detection without requiring extensive preprocessing of ultrasound images. Experimental findings indicate that the proposed YOLOv8 model achieves a mean Average Precision (mAP) of 94.3%, surpassing the performance of YOLOv5 by 5.2%. Additionally, the

Table 6. Comparison Of Other Sota Techniques

Year	Method	Dataset Size	mAP / Accuracy
Liu et al. (2020) [19]	Joint-training CNN	~200 US images	Localization accuracy: 89.3%
W. Song (2019) [13]	Multi-task Cascade CNN	Not specified	Accuracy: 91.2%
Moussa et al. (2020) [16]	Fine-tuned CNN (ResNet)	Public dataset (DDTI)	Accuracy: 88.5%
Kunapinun et al. (2023) [26]	GAN-enhanced segmentation	~300 images	Dice score: 0.86
Inan et al. (2024) [44]	Hybrid ResNet-UNet	~500 segmented nodules	Accuracy: 93.1%
Du et al. (2024) [32]	Deep-learning radiomics	Multi-center	AUC: 0.92

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

model demonstrates exceptional speed, requiring only 8.7 milliseconds per image for detection. These results highlight the model's ability to generalize effectively, even with limited data, while maintaining robustness against overfitting and noisy samples. First, integrating 3D and 4D ultrasound modalities could provide richer spatial and temporal context, potentially improving diagnostic accuracy for complex cases. Second, deploying the model on edge devices, such as ultrasound machines equipped with embedded AI, would enable real-time, on-site inference without relying on cloud infrastructure, making the system more accessible in resource-constrained settings. Third, we aim to extend our framework to other organs, including breast and liver nodules, to broaden its applicability across diagnostic domains. Finally, adopting federated learning would allow us to collaborate across multiple hospitals while preserving patient privacy, enabling robust model generalization through decentralized training on diverse datasets.

Compared with other state-of-the-art (SOTA) models, the proposed YOLOv8 framework achieves superior performance in accuracy, efficiency, and detection speed. Its enhanced architecture facilitates improved localization and classification of thyroid nodules, making it a highly effective computer-aided diagnosis (CAD) tool. The dataset used in this study comprises only 200 images collected exclusively at Wenzhou Central Hospital between 2012 and 2015, limiting its size and diversity. All images were acquired using Siemens Acuson Sequoia ultrasound systems operating at 10–12 MHz, which introduces scanner dependence that may affect generalizability. Furthermore, the dataset exhibits nodule-type homogeneity, as all malignant cases are papillary carcinomas, with no representation of other subtypes such as follicular carcinoma. While the use of a channel attention (CA) mechanism enhances model performance, it simultaneously reduces interpretability in the absence of saliency maps, making it difficult to understand the model's decision-making process. By integrating this model into clinical workflows, radiologists can benefit from faster and more reliable thyroid nodule detection, ultimately improving patient outcomes.

Acknowledgment

The authors would like to express sincere gratitude to the Department of Medical Electronics Technology, Poltekkes Kemenkes Surabaya, for the invaluable support and resources provided throughout this research. The facilities, academic environment, and encouragement from faculty members have significantly contributed to the completion of this work. This study would not have been possible without the institution's commitment to advancing research and innovation in medical electronics.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

No datasets were generated or analyzed during the current study.

Author Contribution

Triwiyanto conceptualized and designed the study, conducted data collection, and participated in data analysis and interpretation. Bedjo Utomo contributed to the development of educational media, oversaw the implementation of the intervention, and contributed to manuscript writing and revisions. Sari Luthfiyah assisted with data analysis and interpretation and provided critical feedback on the manuscript. All authors reviewed and approved the final version of the manuscript, and agreed to be responsible for all aspects of the work ensuring integrity and accuracy.

Declarations

Ethical Approval

This study was conducted by ethical standards and has received approval from the Institutional Ethical Board (IRB) of Poltekkes Kemenkes Surabaya, Indonesia, with approval number [045/Polkes/2024]. Informed consent was obtained from the parents or guardians of all participating students, and confidentiality and anonymity of the participants were maintained throughout the research process. All procedures adhered to ethical guidelines for research involving human subjects.

Consent for Publication Participants.

Consent for publication was given by all participants

Competing Interests

The authors declare no competing interests.

References

[1] I. Steinberg, D. M. Huland, O. Vermesh, H. E. Frostig, W. S. Tummers, and S. S. Gambhir, "Photoacoustic clinical imaging," Photoacoustics, vol. 14, pp. 77–98, Jun. 2019. https://doi.org/10.1016/j.pacs.2019.04.001

[2] C. P. Wild, E. Weiderpass, and B. W. Stewart, "World cancer report: Cancer research for cancer prevention," in International Agency for Research on Cancer. Lyon, France : Avenue Tony Garnier, 2020. https://doi.org/10.1002/ijc.32991

[3] Y. S. Khan, A. Farhana, and S. Anjum, "Physiology, Thyroid," *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK519566/

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

[4] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," IEEE/ACM Trans. Comput. Biol. Bioinf., vol. 18, no. 3, pp. 912–921, May/Jun. 2021. https://doi.org/10.1109/TCBB.2020.2982939

[5] J. Zhang, F. Zhang, C. Zhao, Q. Xu, C. Liang, Y. Yang, H. Wang, Y. Shang, Y. Wang, X. Mu, D. Zhu, C. Zhang, J. Yang, M. Yao, and L. Zhang, "Dysbiosis of the gut microbiome is associated with thyroid cancer and thyroid nodules and correlated with clinical index of thyroid function," Endocrine, vol. 64, no. 3, pp. 564–574, Jun. 2019, https://doi.org/10.1007/s12020-018-1831-x

[6] L. Aversano, M. L. Bernardi, M. Cimitile, A. Maiellaro, and R. Pecori, "A systematic review on artificial intelligence techniques for detecting thyroid diseases," PeerJ Comput. Sci., vol. 9, p. e1394, Jun. 2023. https://doi.org/10.7717/peerj-cs.1394

[7] M. Grussendorf, I. Ruschenburg, and G. Brabant, "Malignancy rates in thyroid nodules: A long-term cohort study of 17,592 patients," Eur. Thyroid J., vol. 11, no. 4, Aug. 2022, Art. no. e220027. https://doi.org/10.1159/000524781

[8] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 11, pp. 3212–3232, Nov. 2019. https://doi.org/10.1109/TNNLS.2018.2876547

[9] 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. https://doi.org/10.1155/2021/6686161

[10] X. Zhou, W. Liang, W. Li, K. Yan, S. Shimizu, and K. Wang, "Hierarchical adversarial attacks against graph neural network based IoT network intrusion detection system," IEEE Internet of Things J., vol. 9, no. 12, pp. 9310–9319, Jun. 2022. https://doi.org/10.1109/JIOT.2021.3121544

[11] F. Li, W. Sun, L. Liu, Z. Meng, and J. Su, "The application value of CDFI and SMI combined with serological markers in distinguishing benign and malignant thyroid nodules," Clin. Transl. Oncol., vol. 22, no. 11, pp. 2200–2209, 2022, doi: https://doi.org/10.1007/s12094-022-02880-1

[12] S. Keestra, V. H. Tabor, and A. Alvergne, "Reinterpreting patterns of variation in human thyroid function: An evolutionary ecology perspective," Evol., Med., Public Health, vol. 9, no. 1, pp. 93–112, 2021. https://doi.org/10.1093/emph/eoab007

[13] W. Song, "Multitask cascade convolution neural networks for automatic thyroid nodule detection and recognition," IEEE J. Biomed. Health Informat., vol. 23, no. 3, pp. 1215–1224, May 2019. https://doi.org/10.1109/JBHI.2018.2852718

[14] X. Zhang, C. Xuan, J. Xue, B. Chen, and Y. Ma, "LSR-YOLO: A high-precision, lightweight model for sheep face recognition on the mobile end," Animals, vol. 13, no. 11, p. 1824, May 2023. https://doi.org/10.3390/ani13111824

[15] 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. DOI: https://doi.org/10.1109/WCONF58270.2023.10235028

[16] O. Moussa, H. Khachnaoui, R. Guetari, and N. Khlifa, "Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network," Int. J. Imag. Syst. Technol., vol. 30, pp. 185–195, 2020. https://doi.org/10.1002/ima.22359

[17] S. Prathibha, D. Dahiya, C. R. Rene Robin, C. Venkata 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. https://doi.org/10.32604/iasc.2023.024416

[18] G. Mariani, M. Tonacchera, M. Grosso, F. Orsolini, P. Vitti, and H. W. Strauss, "The role of nuclear medicine in the clinical management of benign thyroid disorders Part 1: Hyperthyroidism," J. Nucl. Med., vol. 62, no. 3, pp. 304–312, Mar. 2021. https://doi.org/10.2967/jnumed.120.245811

[19] R. Liu, S. Zhou, Y. Guo, Y. Wang, and C. Chang, "Nodule localization in thyroid ultrasound images with a joint-training convolutional neural network," J. Digit. Imag., vol. 33, no. 5, pp. 1266–1279, Oct. 2020. https://doi.org/10.1007/s10278-019-00311-9

[20] X. Wu, D. Sahoo, and S. C. H. Hoi, "Recent advances in deep learning for object detection," Neurocomputing, vol. 396, pp. 39–64, Jul. 2020. https://doi.org/10.1016/j.neucom.2020.01.119

[21] 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. https://doi.org/10.1007/s10278-020-00346-6

[22] M. Wang, C. Yuan, D. Wu, Y. Zeng, and W. Qiu, "Automatic segmentation and classification of thyroid nodules in ultrasound images with convolutional neural networks," in Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Interv., 2021, pp. 109–115. https://doi.org/10.1007/978-3-030-87234-2_11

[23] J. Zhan, L.-H. Zhang, Q. Yu, C.-L. Li, Y. Chen, W.-P. Wang, and H. Ding, "Prediction of cervical

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

lymph node metastasis with contrast-enhanced ultrasound and association between presence of BRAFV600E and extrathyroidal extension in papillary thyroid carcinoma," Therapeutic Adv. Med. Oncol., vol. 12, Jan. 2020, Art. no. 1758835920942367, doi: https://doi.org/10.1177/1758835920942367

[24] I. K. Kang, C. K. Jung, K. Kim, J. Park, J. S. Kim, and J. Bae, "Papillary thyroid carcinoma in a separate pyramidal lobe mimicking thyroglossal duct cyst carcinoma: A case report," J. Endocrine Surg., vol. 22, no. 4, p. 138, 2022. https://doi.org/10.16956/jes.2022.22.4.138

[25] F. Abdolali, J. Kapur, J. L. Jaremko, M. Noga, A. R. Hareendranathan, and K. Punithakumar, "Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks," Comput. Biol. Med., vol. 122, Jul. 2020, Art. no. 103871. https://doi.org/10.1016/j.compbiomed.2020.103871

[26] A. Kunapinun, M. N. Dailey, D. Songsaeng, M. Parnichkun, C. Keatmanee, and M. Ekpanyapong, "Improving GAN learning dynamics for thyroid nodule segmentation," Ultrasound Med. Biol., vol. 49, no. 2, pp. 416–430, Feb. 2023. https://doi.org/10.1016/j.ultrasmedbio.2022.09.006

[27] 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. https://doi.org/10.3389/fonc.2023.1060702

[28] X. Zhou, W. Liang, K. Wang, and L. T. Yang, "Deep correlation mining based on hierarchical hybrid networks for heterogeneous Big Data recommendations," IEEE Trans. Computat. Social Syst., vol. 8, no. 1, pp. 171–178, Feb. 2021. https://doi.org/10.1109/TCSS.2020.3015694

[29] W.-H. Li, W.-Y. Yu, J.-R. Du, D.-K. Teng, Y.-Q. Lin, G.-Q. Sui, and H. Wang, "Nomogram prediction for cervical lymph node metastasis in multifocal papillary thyroid microcarcinoma," Frontiers Endocrinology, vol. 14, May 2023, Art. no. 1140360, doi: 10.3389/fendo.2023.1140360.

[30] A. Bikas and K. D. Burman, "Epidemiology of thyroid cancer," in The Thyroid and Its Diseases: A Comprehensive Guide for the Clinician. Cham, Switzerland : Springer, 2019, pp. 541–547. https://doi.org/10.1007/978-3-030-15146-1_39

[31] A. Shahroudnejad, "Tun-Det: A novel network for thyroid ultrasound nodule detection," in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent., Strasbourg, France. Cham, Switzerland : Springer, Sep. 2021, pp. 656–667. https://doi.org/10.1007/978-3-030-87234-2_62

[32] H. Du, F. Chen, H. Li, K. Wang, J. Zhang, J. Meng, H. Li, X. Xu, J. Qu, R. Wu, J. Li, M. Zhang, F. Zhang, and X. Zhu, "Deep-learning radiomics based on ultrasound can objectively evaluate thyroid nodules and assist in improving the diagnostic level of ultrasound physicians," Quant. Imag. Med. Surgery, vol. 14, no. 8, pp. 5932–5945, Aug. 2024. https://doi.org/10.21037/qims-23-1490

[33] C. F. Li, R. Q. Du, Q. Y. Luo, R. Wang, and X. H. Ding, "A novel model of thyroid nodule segmentation for ultrasound images," Ultrasound Med. Biol., vol. 49, no. 2, pp. 489–496, 2023. https://doi.org/10.1016/j.ultrasmedbio.2022.09.009

[34] 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. https://doi.org/10.1109/Confluence51648.2021.9377196

[35] X. Zhou, X. Yang, J. Ma, and I. K. Wang, "Energy efficient smart routing based on link correlation mining for wireless edge computing in IoT," IEEE Internet of Things J., vol. 9, no. 16, pp. 14988–14997, Aug. 2022. https://doi.org/10.1109/JIOT.2022.3155798

[36] L. Chang, Y. Zhang, J. Zhu, L. Hu, X. Wang, H. Zhang, Q. Gu, X. Chen, S. Zhang, M. Gao, and X. Wei, "An integrated nomogram combining deep learning, clinical characteristics and ultrasound features for predicting central lymph node metastasis in papillary thyroid cancer: A multicenter study," Frontiers Endocrinol., vol. 14, Feb. 2023, Art. no. 964074, doi: https://doi.org/10.3389/fendo.2023.964074

[37] M. D. McCradden, J. A. Anderson, E. A. Stephenson, E. Drysdale, L. Erdman, A. Goldenberg, and R. Zlotnik Shaul, "A research ethics framework for the clinical translation of healthcare machine learning," Amer. J. Bioethics, vol. 22, no. 5, pp. 8–22, May 2022. https://doi.org/10.1080/15265161.2021.2013918

[38] H. He, J. Zhu, Z. Ye, H. Bao, J. Shou, Y. Liu, and F. Chen, "Using multimodal ultrasound including full-time-series contrast-enhanced ultrasound cines for identifying the nature of thyroid nodules," Frontiers Oncol., vol. 14, Aug. 2024, Art. no. 1340847. https://doi.org/10.3389/fonc.2024.1340847

[39] 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. https://doi.org/10.1016/j.compbiomed.2022.106444

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

[40] R. Song, "Dual-branch network via pseudo-label training for thyroid nodule detection in ultrasound image," Int. J. Speech Technol., vol. 52, no. 10, pp. 11738–11754, Aug. 2022. https://doi.org/10.1016/j.asoc.2022.109015
[41] A. Prochazka, S. Gulati, S. Holinka, and D. Smutek, "Patch-based classification of thyroid nodules in ultrasound images using direction independent features extracted by two-threshold binary decomposition," Computerized Med. Imag. Graph., vol. 71, pp. 9–18, Jan. 2019. https://doi.org/10.1016/j.compmedimag.2018.11.002
[42] 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. https://doi.org/10.1007/s11042-023-15206-6
[43] X. Zhou, X. Xu, W. Liang, Z. Zeng, and Z. Yan, "Deep-learning-enhanced multitarget detection for end-edge-cloud surveillance in smart IoT," IEEE Internet of Things J., vol. 8, no. 16, pp. 12588–12596, Aug. 2021. https://doi.org/10.1109/JIOT.2021.3060592
[44] N. G. Inan, O. Kocadağlı, D. Yıldırım, İ. Meşe, and Ö. Kovan, "Multi-class classification of thyroid nodules from automatic segmented ultrasound images: Hybrid ResNet based UNet convolutional neural network approach," Comput. Methods Programs Biomed., vol. 243, Jan. 2024, Art. no. 107921. https://doi.org/10.1016/j.cmpb.2023.107921

Author Biography

Dr. M. Kalpana Devi, Assistant Professor (Selection Grade), Department of Computer Science and Engineering, Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India. She has completed her Master's in Computer Applications from PSG College of Arts and Science, Coimbatore, and M.E.-Computer Science & Engineering from Anna University Regional Campus, Coimbatore. She Completed her M.Phil in Computer Science from Bharathiar University, Coimbatore. Currently, she has completed her Ph.D in Information and Communication Engineering from PSG College of Technology under Anna University, Chennai. Has 18 years of Teaching Experience. Her research areas include Cognitive Radio Networks, Mobile Communications and Networking, and Image Processing. She received a best researcher award from the Science Father in the year 2021. She has published 18 Scopus-indexed journals and 3 IEEE Conferences. She also published 2 books in Charulatha Publications. She got 3 design patent grants. She is a reviewer for many peer-reviewed journals and conferences.

S. Vidhya is an Assistant Professor in the Department of Computer Science and Engineering at V.S.B Engineering College, Karur, an autonomous institution. She has over 12 years of teaching experience, having served in various engineering colleges. She completed her postgraduate degree in Computer Science and Engineering from Maharaja Engineering College, affiliated with Anna University, and her undergraduate degree in Information Technology from Tamilnadu College of Engineering, also affiliated with Anna University. Her primary research interests include Machine Learning and its applications. She has presented several research papers at national conferences and actively participates in academic and professional development activities. She has earned NPTEL Elite certifications and attended numerous faculty development programs. In addition, she has contributed to academia by publishing several book chapters.

Dr. M. Therasa completed her Bachelor of Engineering in Computer Science and Engineering from Adhiparasakthi College of Engineering, affiliated with Anna University, Kalavai, Tamil Nadu, in 2006. She earned her Master of Engineering in Computer Science and Engineering from Sathyabama Institute of Science and Technology, Chennai, a deemed-to-be university, in 2010. She was awarded her Doctor of Philosophy in Computer Science and Engineering from the same institution in 2024. She currently works as an Associate Professor at Panimalar Engineering College. Her academic and research interests span emerging areas such as Deep Learning, Image Processing, and Natural Language Processing. She actively contributes to teaching, research, and academic development through scholarly activities and research initiatives.

A. Praveena, Assistant Professor at Hindusthan Institute of Technology, Coimbatore, India. She holds an M.E in Computer Science and Engineering. She has 4 years of teaching experience in academia, covering a wide array of subjects, including Data Engineering and Visualization, Artificial Intelligence, Internet of Things, C Programming, and Data Structures. Her research contributions include publishing over 8 papers in various national and

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137

international conferences, publishing a patent, and various books. She is deeply passionate about sharing knowledge for continuous learning, encouraging students to develop critical thinking skills and stay ahead in the fast-paced technological landscape. Through this work, she aims to make a meaningful impact on both education and research in the fields of Machine Learning, Deep Learning, Blockchain, and IOT.

Dr. M. Ramesh Kumar obtained his Master of Engineering degree in Computer Science and Engineering from P.S.N.A College of Engineering and Technology, Dindigul, affiliated with Anna University, Chennai, in 2008. He was awarded his Doctor of Philosophy in Computer Science and Engineering from Karpagam University, Coimbatore, Tamil Nadu, in 2015. He is currently serving as Professor and Head of the Department of Information Technology at V.S.B. College of Engineering Technical Campus, Coimbatore. He has over 16 years of teaching and research experience. His scholarly contributions include publications in Scopus-indexed, SCI Annexure-I, and UGC CARE journals. He has presented numerous research papers at national and international conferences, published books and patents in leading technologies, received two awards, including the Best Researcher Award, and serves as an editorial board member and session chair for reputed conferences.

E. Kalaivani is working as an Assistant Professor in the Department of Computer Science and Engineering at Bannari Amman Institute of Technology, Sathyamangalam, Erode. She has 7.8 years of teaching experience and 2 years of industrial experience, contributing to both academic and practical perspectives in engineering education. She is currently pursuing her doctoral research in the area of Deep Learning at Anna University, Chennai. She has published six research papers in reputed journals and contributed one book chapter. In addition, she has presented her research work at both national and international conferences. Her areas of interest include networking, software testing, machine learning, and deep learning, and she actively engages in research and academic activities to enhance her professional expertise.

Manuscript received 14 August 2025; Revised 2 December 2025; Accepted 20 March 2026; Available online 14 May 2026
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v8i3.1137