© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715262

Federated Learning Framework for Integrating Multi Device IoT Data in Healthcare Applications

RAGUL K B[1], KUMARAN S[2], ARSHAD AHAMAD K[3], ROHINI C[4]
[1, 2, 3]*Final year UG, Department of Computer Science and Engineering, Velammal Engineering College, Chennai, India*
[4]*Assistant Professor, Department of Computer Science and Engineering, Velammal Engineering College, Chennai, India*

Abstract- ***There has been an increase in the volume of distributed, heterogeneous healthcare data with the advent of Internet of Things (IoT) technology-based medical devices. However, aggregating large volumes of sensitive patient information at a central server poses several privacy, security, and regulatory hurdles. This paper proposes a decentralized Federated Learning (FL) architecture that can accommodate multi-device IoT-based medical data integration without sharing sensitive patient information between healthcare institutions. Each institution maintains their unique model, which can be used for training on their premises, and then the models are integrated into a single federated model for collective use. A novel framework is proposed, which can accommodate collaborative training between several nodes of healthcare institutions, along with local processing of multimodal data obtained from IoT medical devices. The multimodal deep neural network architecture has been used to combine different types of input, including medical images and structured demographic or clinical attributes, through parallel feature extraction branches followed by feature-level fusion. Each healthcare institution has a unique model, which is trained on their respective data, and then the model parameters are sent to a central server for aggregation through the Federated Averaging (FedAvg) algorithm. Cross-validation of the models has also been done to check the robustness of the models, which involves training on several nodes and then validation on an external dataset. The results obtained through the federated multimodal framework indicate high classification accuracy, thus ensuring the viability of a secure, scalable, and privacy-preserving collaborative intelligence system for IoT-based healthcare environments.***

I. INTRODUCTION

In hospital environments, there are often standard methods of diagnosing thyroid. For example, there is an ultrasound test for patients. Sometimes, there can also be the addition of a thermal imaging test. This can help identify strange heat signals. After reviewing the age, gender, and past medical visits, the physician can then make a decision. All of these factors are important to the overall result. Yet, they are often evaluated individually, relying too heavily on personal review and individual tools.

What is another obstacle? Hospitals today manage vast amounts of medical information, thanks to internet-enabled devices such as ultrasound equipment, thermal sensors, and patient records. This information has vast potential for improvement in diagnosing patients, especially with the use of powerful machine learning. Yet, there is one major issue: they cannot share patient information because of privacy, moral, and internal issues. So, valuable information remains fragmented between organizations.

Patient information is routed through one server for central learning. Yet, more accurate predictions can be made through this route. However, there is also the risk of an information leak or illegal viewing. When sorting through the information to diagnose thyroid nodules, none of these risks apply. Security needs to be increased, especially when the outcome can mean life or death.

As this project was initiated, there was one thought: can hospitals learn from a powerful tool of diagnosing patients, yet keep their patient information private? Instead of sending their information to the tool, can they instead send the tool to their information?

In the face of this problem, a different approach arises in the form of a shared learning method designed to sort thyroid nodules using different data types. Rather than hubs, the system brings together three different sources of data collected through health-focused internet of things tools: thermal visuals, ultrasound frames, and personal data. There are different ways forward, and each medium has its own different processing system to turn the data into a usable pattern. From these patterns, a combined view arises. This view is what decides the different cases: normal, benign, or cancer.

In the proposed system, each hospital has the multimodal model by itself and learns from its own data. Only the parameters of the model reach the server. This is done through a method known as federated averaging (fedavg). Learning happens collectively, and no sensitive patient data leaves the hospital.

Next, the paper continues in the same fashion: section ii looks to the past and revisits the efforts of the past in federated learning and the connection to multimodal medical diagnosis. Section iii looks to the future and explains the proposed system and approach. Section iv explains the proposed test method and how the different hospitals will be compared. Results and system behavior are explained in section v. Naturally, future work is explained in section vi, the final section.

II. LITERATURE REVIEW

With the recent developments in artificial intelligence, thyroid disease diagnosis has been greatly improved through imaging, lab information, and physiological signals. However, there are challenges in combining these heterogeneous modalities in a privacy-preserving collaborative environment. This section discusses the existing literature in federated thyroid imaging, multimodal learning, and non-imaging thyroid prediction systems.

A. Federated Learning for Thyroid Ultrasound Imaging

The first example of a specific research project undertaken with the aim of collecting or utilizing thyroid ultrasound images with the application of federated learning was when Lee et al.[1] came up with a collaborative effort with six clinics with 8,457 scans, which they shared through a common network; rather than collecting all the data in one place, each location was able to train their own convolutional neural networks, such as VGG19, ResNet50, etc., with the application of the workflow known as Federated Averaging (FedAvg). The results were surprisingly good, indicating that distributed learning was just as good as conventional learning methods, but also allowed for the same results without the compromise of patient information confidentiality. Further research indicated that sharing of knowledge and insights through networks helped in providing support in analyzing images of the thyroid gland [1].

Recently, Xiang et al. [2] applied federated learning in the grouping of images of thyroid nodules, which were then segmented with the application of the UNet architecture with a multi-attention guide. The application of the architecture was able to show that segmentation of images from different datasets was possible, especially with difficult-to-segment medical images. However, the authors only concentrated on the application of ultrasound images with the algorithm, rather than more complex artificial intelligence algorithms with more than one source of input.

As can be seen by the research done by Savelonas et al.[3], there are many different tools that can be used to help with this process involving AI and the images produced by ultrasound on the thyroid gland, including the use of segmentation and a three-dimensional representation. Though there are many different methodologies that can be used to accomplish this, including the use of a doppler tracking system on many different types of ultrasound machines, federated learning is specifically discussed in great detail in this article, though there were no studies found that used multiple different types of data and came from multiple different networks that have practical applications in the real world [3].

Though there have been many different advancements in this field, the majority of the literature that can be found on federated learning and the diagnosis of thyroid nodules specifically uses images and binary results that come from an ultrasound [1-3]. There is a great deal of missing literature on the significant amount of information that exists on human physiology and the differences that exist between each individual.

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715262

B. Federated Learning in Medical Imaging and Disease Prediction

Federated learning has also been widely used in healthcare applications. Furthermore, it is being used for all types of applications in the medical field. The application of federated learning in medical fields was recently reviewed by Moshawrab et al. (2021) [4]. The authors in their review provide an overview of the current state of federated learning in disease prediction. They also discuss the aggregation strategies, data heterogeneity, differential privacy, and secure communication mechanisms in disease prediction. Nazir et al. (2021) also reviewed the different types of federated learning that can be used in medical image analysis. They also provide guidelines on the best practices of dealing with non-IID data distribution in different institutions.

Both these studies clearly indicate that federated learning can provide both accuracy and privacy; however, the implementation of these studies is based on single image modality data rather than using multimodality fusion. The development of a privacy-preserving federated learning model for the detection of thyroid disease using tabular clinical datasets with different techniques of feature selection and class balancing was recently investigated in a preliminary study in a preprint article [6]. Although the above study has covered all aspects of model development in the context of clinical structured data, no image modality types or fusion were used in the study.

C. Non-Imaging and Multimodal Thyroid Prediction

Some researchers have also studied the prediction of thyroid disease with lab results and demographics. Hu et al. [7] used machine learning algorithms to develop models to diagnose thyroid dysfunction with normal lab results. Noh et al. [8] studied the application of different machine learning algorithms in predicting hypothyroidism and thyrotoxicosis with biochemical and demographic information.

Shin et al. [9] also studied the application of heart rate signals in predicting thyrotoxicosis. This shows that other physiological signals can also contribute to thyroid disease prediction. Chaganti et al. [10] and Ji et al. [11] used machine learning algorithms on the thyroid disease dataset of the UCI machine learning repository. The authors focused on clinical attributes. The results of these studies were good in classification but were conducted in centralized environments. The authors did not consider the application of privacy-preserving distributed learning.

Thermal imaging has also been studied for thyroid disease prediction. Chantasartrassamee et al. [12] studied the application of AI in infrared imaging in the analysis of thyroid nodules. Zhang et al. [13] also proposed a system that combines ultrasound imaging with infrared imaging in thyroid disease classification. These studies also used centralized environments in training the models.

D. Research Gaps

As per the above literature, it is clear that there are some important gaps to be filled. The existing literature on federated learning-based approaches to thyroid diagnosis has mostly focused on ultrasound images. There is a lack of consideration of other approaches, such as thermal images and demographic factors. Although there have been some approaches to multimodal classification of thyroid images, they have mostly been centralized. There is a lack of scalability and privacy. Moreover, there has been a lack of validation of approaches that combine three encoders: thermal, ultrasound, and demographic. There has also been a lack of validation of approaches that utilize multimodal datasets from various hospitals.

E. Positioning of the Proposed Work

In order to address the above-mentioned limitations, this paper proposes a federated multimodal framework for thyroid nodule classification, which includes three parallel encoders:

- thermal image encoder
- ultrasound image encoder
- demographic feature encoder

In this framework, each hospital trains the entire system, but the model parameters are updated through the fedavg algorithm. This framework is different from the earlier studies, which were based on either segmentation or classification of single-modality data. This framework considers various sources of data and multi-class classification of thyroid nodules in a federated, private environment.

III. PROPOSED SYSTEM

This paper introduces a federated multimodal learning framework that is used to classify thyroid nodules in different hospitals. The framework uses the combination of thermal images, ultrasound images, and demographic information from the hospitals, without exposing the raw data from the patients. A federated server is also introduced, which collects the parameters from the models and uses them for collaboration during the training phase, with the aim of performing an effective multi-class classification.

A. Standard Federated Learning Mechanism

Federated Learning (FL) is a technique used for training a shared model with the help of several clients. All the clients collaborate with each other under the supervision of a central server. Under a conventional Federated Learning scenario, the server initializes the model parameters and sends them to all the clients. Then, each of the clients trains the model with their respective private data set and uses backpropagation for updating the model parameters. Instead of sending their respective data, the clients send back their updated model parameters to the central server.

The server aggregates the received model parameters with the help of the Federated Averaging (FedAvg) algorithm. This algorithm aggregates the model weights of the clients based on the size of their respective data sets. Mathematically, the updated global model can be expressed as:

$$w^{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_k^t$$

where K is the number of participating clients, n_k is the number of samples at client k, and n is the total number of samples. The aggregated global model is then sent back to all clients for the next round of training. This back-and-forth communication continues until it stabilizes. The distinctive aspect of this architecture is the enforcement of strict data locality. The raw training data does not leave the client institution. This approach ensures patient privacy and minimizes legal and ethical risk.

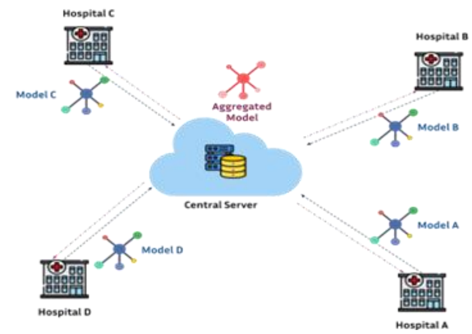

Fig 1 Federated Learning Architecture

B. Multimodal Data Representation at Each Hospital

The proposed system has three types of patient data for each of the hospitals:

1. Thermal Images that provide surface temperature pattern data, which could indicate abnormal blood flow due to thyroid nodules.
2. Ultrasound Images that provide detailed information about the shape of thyroid tissue.
3. Demographic Attributes that provide structured data about age, gender, and other clinical signs.

The data is of different types. The data has different dimensions, statistical patterns, and meanings. We cannot directly combine the input data. We define encoders for each of them to transform the input data into a condensed feature representation.

C. Local Multimodal TNF Architecture

The local model proposed consists of three parallel encoder branches and a fusion and classification module.

The thermal image encoder uses a convolutional neural network (CNN) to encode the thermal images and extract features from the spatial temperature patterns. Convolutional neural networks consist of convolutional layers that extract hierarchical features from the images. Activation functions and pooling operations are also employed to reduce the spatial dimensionality and retain the important features.

The ultrasound image encoder also uses a CNN to extract features from the ultrasound images. This CNN is optimized to extract structural features from the images.

The demographic encoder uses fully connected layers to encode the tabular data. This changes the normalized demographic data into dense features. Non-linear activation functions are also employed in the demographic encoder to learn the complex relationships between the clinical variables and the diagnostic outcomes.

Let the feature extraction functions for thermal images, ultrasound images, and demographic data be denoted as $f_T(x_T)$, $f_U(x_U)$, and $f_D(x_D)$, respectively. The corresponding latent feature representations can be expressed as:

$$\mathbf{z}_T = f_T(\mathbf{x}_T), \quad \mathbf{z}_U = f_U(\mathbf{x}_U), \quad \mathbf{z}_D = f_D(\mathbf{x}_D).$$

These encoded vectors are concatenated to form a unified multimodal representation:

$$z = [z_T \mid\mid z_U \mid\mid z_D]$$

The fused feature vector z is then passed through the fully connected layers forming the classification head. The last layer applies a softmax activation function to generate probabilities for the three classes: Normal, Benign, and Malignant.

The entire architecture with all three encoders and the fusion-based classifier is the local fused model in each hospital.

D. Federated Training of the Multimodal Fused Model

During federated training, the full multimodal fused model is trained locally at each hospital using its own private dataset. For the classification task, the model is optimized using the categorical cross-entropy loss function.

$$\mathcal{L} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

where C = 3 is the number of classes, and y_i is the predicted probability.

After training for a set number of epochs, each hospital sends their updated model parameters to the federated server. Note that the server is not sent intermediate representations of features or patient information; only the updated model weights. The server then aggregates the weights to update the global multimodal model.

The updated global model is then sent back to all the hospitals, ensuring that all the hospitals are updated simultaneously.

E. Communication Protocol and Model Update Flow

The federated communication process
The federated communication process involves a synchronized round-based method. At the beginning of every round, the server sends the latest global model to all the hospitals. Each hospital undertakes a local training process individually using its dataset, which can have non-IID properties.

The hospitals send their model updates to the server. The aggregation process involves creating an updated global model, considering the shared knowledge across hospitals. There is no direct communication between hospitals. Data is not shared in its raw form at any stage of the process.

The federated communication process is scalable as more hospitals can be added to the federated network without modifying the system.

F. Real-Time Seat Locking Mechanism

The entire system consists of numerous hospital nodes and a central federated server. Every hospital has two logical components:

1. Input layer: This consists of thermal images, ultrasound images, and demographic attributes
2. Local multimodal fused model: This consists of three encoders and a classification head.

Every hospital sends model updates to the central server. The server aggregates the model using fedavg. The server then generates a new global multimodal model. The model is then sent back to each hospital. This architecture does not require inter-hospital communication. The hospitals remain independent.

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715262

This architecture combines multimodal feature fusion with federated optimization. The architecture addresses the shortcomings of past works that only concentrated on centralized single-modality systems

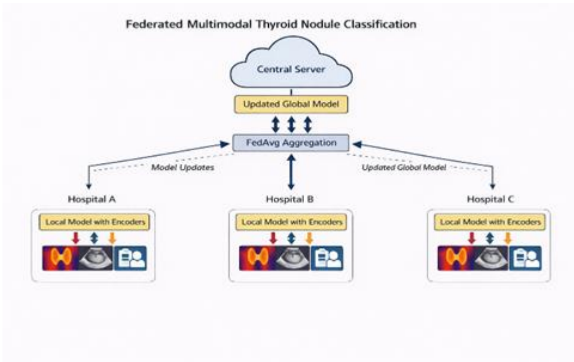

Fig 2 System Architecture

IV. IMPLEMENTATION

This section details the technical implementation of the our platform, including technology stack selection, development tools, and integration strategies.

A. Local Multimodal Model Implementation

The multimodal classification model that was proposed has been implemented using PyTorch, which is a deep learning framework. The local architecture of each of the hospital nodes is similar. There are three separate encoders for the analysis of the thermal and ultrasound images. The classification module is used for fusion. The images are analyzed using convolutional neural networks (CNNs). The CNN encoders consist of stacked convolutional layers with ReLU activation functions, batch normalization, and max pooling. The convolutional feature maps are then flattened to produce fixed-length embedding vectors.

The structured tabular attributes of the demographic data are analyzed using a fully connected neural network. First, the input features are normalized to facilitate stable optimization. The dense layers of the demographic encoder consist of non-linear activation functions to identify the underlying pattern of clinical attributes.

The output of each of the encoders is fused to produce a multimodal feature vector. The fused vector is then analyzed using fully connected layers to produce the classification module. The final output layer has softmax activation to produce probabilities for each of the three classes: Normal, Benign, and Malignant. The classification model is optimized using Adam optimizer and categorical cross-entropy loss functions. The model is trained locally by each of the hospitals for a specified number of epochs.

B. Federated Training and Aggregation Implementation

The federated learning process was carried out using the Flower framework with PyTorch as the base. The federated server manages the communication rounds with the hospitals. At the beginning of every communication round, the federated server sends the model parameters to the hospitals. The hospitals, in turn, train their local model using their private data.

Once the model training at every hospital is done, they send their model parameters back to the federated server. The federated server uses the federated averaging (FedAvg) method to compute the weighted sum of the model parameters. The weighted sum of the model parameters is computed using the number of samples in every hospital. The combined model parameters form the new global model, which is sent to all the hospitals for the next round of communication. The communication rounds are carried out for a certain number of rounds. No images or patient data are exchanged between hospitals. The model weights are exchanged instead. The model works with non-IID data. The multimodal model architectures are kept in sync.

V. RESULTS AND DISCUSSION

It is important to note that the current experimental evaluation was conducted using synthetic multimodal data. It is important to note that because of the regulatory and ethical issues, it was not possible to collect the required multi-center hospital data for the purpose of the current experimental evaluation. Therefore, the synthetic data was used for the purpose of simulating the representative thermal images, ultrasound images, and demographic variables. It is important to note that the aim of the current simulation-based experimental evaluation was to examine the architectural feasibility, stability of the federated optimization procedure, and efficacy of the multimodal fusion procedure.

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715262

A. Overall Classification Performance

The federated multimodal model that was proposed was able to achieve a classification accuracy of 91% for 100 samples that were used to test the model. The model was also able to achieve good predictive values for the Normal and Malignant classes, resulting in a recall of 100% for both classes.

The results for the Benign class suggest a recall of 73.53%, which could be attributed to the overlapping properties that exist between the Benign and Malignant classes in the synthetic dataset that was used to train the model.

The results suggest a good balance between the different classes, which can be attributed to the fact that the model was able to achieve a macro average F1-score of 0.9092 and a weighted F1-score of 0.9085.

B. Detailed Classification Metrics

To further understand the performance of the model, precision, recall, and F1-score were computed for each of the classes. The Normal class had a perfect precision and recall of 1.000. This indicates that there was no overlap between the features in the synthetic data distribution. The Malignant class was very robust with a recall of 1.000 and precision of 0.7857. This indicates that there were small false positive cases. The Benign class had a low recall of 0.7353. This indicates that there was overlap in the feature patterns. The averages indicate that the federated multimodal architecture was well-balanced in performance. The precision, recall, and F1-score confirm the success of combining thermal, ultrasound, and demographic encoders in a federated learning setup.

Classification Report:

	precision	recall	f1-score	support
Normal	1.0000	1.0000	1.0000	33
Benign	1.0000	0.7353	0.8475	34
Malignant	0.7857	1.0000	0.8800	33
accuracy			0.9100	100
macro avg	0.9286	0.9118	0.9092	100
weighted avg	0.9293	0.9100	0.9085	100

Fig3 Classification Report

C. Interpretation Under Synthetic Data Constraints

The results of the performance metrics have to be considered in the context of the simulation of synthetic data. The data was created in a controlled setting where there were no real-world external data noise. The feature distribution was relatively more structured. This led to stable training convergence as well as high accuracy values.

However, in real-world implementation with real-world datasets in hospitals, there could be some variations in performance based on non-iid data, image variations, and demographic variations. Therefore, although the accuracy of 91% was obtained, real-world implementation with real-world datasets in hospitals is needed to assess the real-world datasets' reliability.

The real-world result of this study reveals that the federated multimodal tnf framework was technically stable, scalable, and able to achieve high accuracy in classification.

VI. CONCLUSION AND FUTURE WORK

This study developed and tested a federated multimodal deep learning framework for the classification of thyroid nodules across distributed healthcare centers. The goal of the study was to develop an intelligent diagnostic system that could effectively combine the diverse data modalities from thermal imaging, ultrasound imaging, and demographic attributes without compromising the privacy of the patients. Unlike the conventional methods that require the aggregation and consolidation of the medical data from the distributed centers, the proposed framework employed the federated learning paradigm to facilitate the distributed optimization process.

The design of the framework consists of three specialized encoder modules per node in the distributed hospitals. The thermal encoder focuses on the surface physiological temperature patterns, the ultrasound encoder focuses on the structural and morphological features of the thyroid nodules, and the demographic encoder focuses on the structured attributes from the clinical data. The features from the

DOI: https://doi.org/10.64388/IREV9I9-1715262

different modalities are combined and fed into the classification head. The important thing to note is that the entire system is treated as a single entity during the federated learning process, and the entire system is responsible for the learning of the combined features from the different modalities.

The performance evaluation of the proposed system was conducted using the artificially generated multimodal datasets because the authors did not have access to the real-world clinical data from the distributed centers. The performance evaluation results showed that the proposed system could effectively classify the thyroid nodules with an accuracy of 91%, perfect recall for the normal and malignant classes, and relatively lower recall values for the benign classes. The f1-score values obtained from the macro-average and the weighted-average metrics also showed balanced performance across the classes, indicating the ability of the federated multimodal framework to learn the features and remain stable during the optimization process.

At the same time, the patterns of performance also shed light on the model's behavior. For instance, the perfect recall of the malignant class implies a very sensitive response to the most important pathological features, which is always desirable in a screening model. On the other hand, the reduced recall of the benign class implies the presence of feature overlap among the intermediate class boundaries. This again underlines the inherent complexity of multi-class thyroid diagnostics, even in a synthetic world.

It is, therefore, important to note that the above results should be understood in the context of synthetic data simulation. In a real world, there would be many other factors such as imaging artifacts, device variations, annotator biases, and data imbalance, which would affect the final imaging data. In that context, the above accuracy should be understood only as a benchmark of the proposed architecture's feasibility.

In the future, the proposed framework can be extended in several important ways. For one, the proposed framework should be validated using actual multi-center hospital data to test the robustness of the proposed architecture and its ability to generalize well. In that context, the proposed framework should be able to handle the non-iid data condition more effectively using advanced aggregation schemes such as fedprox.

Another important extension of the proposed framework would be the improved multimodal fusion strategies. In the current implementation, the proposed framework uses a simple concatenation of the individual modalities. In the future, the proposed framework can be improved by using attention-based multimodal fusion strategies to dynamically adjust the relative weights of the individual modalities such as thermal, ultrasound, and demographic.

Third, the interpretability component is also crucial and should be improved. This is because the integration of interpretable ai methods, such as gradient-based visualization for the imaging encoder and feature attribution methods for demographic inputs, would be beneficial and would increase the confidence and trust level among clinicians.

Fourth, the privacy component would also be beneficial and could be improved by integrating methods such as differential privacy, secure multi-party computation, and encrypted aggregation.

Finally, the framework could also be improved and generalized from the current three-class classification to risk stratification scoring, malignancy probability estimation, and disease progression analysis over time. The integration with electronic health records and real-time iot diagnostic systems would also increase the utility and applicability of the proposed framework.

In conclusion, the study presents a scalable and privacy-aware federated multimodal framework for thyroid nodule classification using thermal imaging, ultrasound imaging, and demographic information. The proposed system enhances the current state-of-the-art in collaborative medical ai systems and maintains the confidentiality and privacy of the data between institutions and hospitals. The proposed framework is also beneficial and provides a strong foundation and foundation for the development and implementation of further research in the field.

REFERENCES

[1] H. Lee, Y. J. Chai, H. Joo, et al., "Federated learning for thyroid ultrasound image analysis to

protect personal information: Validation study in a real health care environment," *JMIR Medical Informatics*, vol. 9, no. 5, e25869, 2021.

[2] Z. Xiang, et al., "Federated learning via multi-attention guided UNet for thyroid nodule segmentation," *Neurocomputing*, 2025.

[3] M. Savelonas, et al., "An overview of AI-guided thyroid ultrasound image analysis: Three-dimensional imaging, Doppler, federated learning, and future directions," *Knowledge*, vol. 9, no. 10, 2025.

[4] M. Moshawrab, et al., "Reviewing federated machine learning and its use in disease prediction," *Sensors*, vol. 23, no. 4, p. 2112, 2023.

[5] S. Nazir, et al., "Federated learning for medical image analysis with deep neural networks: A survey," *Diagnostics*, vol. 13, no. 9, p. 1532, 2023.

[6] "Privacy-preserving federated learning for thyroid disease detection: A comparative study on feature selection and class balancing," ResearchGate Preprint, 2025.

[7] M. Hu, et al., "Development and preliminary validation of a machine learning approach to diagnose thyroid dysfunction from routine laboratory tests," 2022.

[8] J. Y. Noh, et al., "An evaluation of the efficacy of machine learning in predicting thyrotoxicosis and hypothyroidism," *Internal Medicine*, 2024.

[9] K. Shin, et al., "A machine learning-assisted system to predict the occurrence of thyrotoxicosis using heart-rate data from wearable devices," 2023.

[10] R. Chaganti, et al., "Thyroid disease prediction using selective features and machine learning," 2022.

[11] S. Ji, et al., "SSC: A self-stack ensemble model for thyroid disease prediction," *PLOS ONE*, 2024.

[12] C. Chantasartrassamee, et al., "AI-enhanced infrared thermography for thyroid nodules," 2024.

[13] N. Zhang, et al., "Multi-modal thyroid nodule classification using ultrasound and infrared imaging," *BMC Bioinformatics*, 2023.