International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 04(1) (April),2026
Received: 06-03-2026 Accepted: 14-04-2026 Published: 20-04-2026

Thyroid Cancer Detection Using Deep Learning Techniques

Y.Prudhvi
AIML dept
Malla Reddy University
Hyderabad,
India

K.Gowthami
AIML dept
Malla Reddy University
Hyderabad,
India

Prof. Sanjaykumar Hamilpure
AIML dept
Malla Reddy University
Hyderabad,
India

M.Nithin
AIML dept
Malla Reddy University
Hyderabad,
India

G.Veda Sai Samitha
AIML dept
Malla Reddy University
Hyderabad,
India

ABSTRACT:

The objective of the project is to build an AI-based system that helps diagnose thyroid cancer. The system utilizes deep learning techniques to scrutinize images from ultrasound exams. The images obtained from ultrasound scans and medical data are preprocessed to remove noise and ensure uniformity. A hybrid deep learning approach using CNN-LSTM can be proposed to classify thyroid as benign or malignant. This process increases the probability of accurate diagnosis while eliminating errors. associated with manual diagnosis. Evaluation of the system is made through performance metrics such as accuracy (99%), precision (85%), recall (94%), and F1-score (87%). A web-based interface designed for real-time detection.

INDEX TERMS :

Thyroid Ultrasound, Convolutional Neural Networks, Computer-Aided Diagnosis, Explainable AI, Grad-CAM, Deep Learning in Healthcare.

INTRODUCTION

Thyroid cancer is often considered as most common endocrine cancers. It is also a continuously an increasing trend in its incidence rate over the past decades. However, the majority of thyroid nodules are While they are typically benign, a number of them are malignant, and thus precise and early diagnosis of them becomes important.

Ultrasound imaging is regarded as the primary method of evaluation in thyroid nodule examination owing to the fact that it is non-invasive nature, accessibility, and cost-effectiveness. However, ultrasound diagnosis is operator-dependent and subject to inter-observer variability, which may lead to misdiagnosis or overtreatment. invasive procedures.

The growing need for precise and efficient diagnostic tools has expedited the fusion of artificial intelligence in medical imaging. Deep learning, especially convolutional neural networks (CNNs), which have shown excellent performance in the automatic extraction of features in a hierarchical manner from complex image data. Such potential makes deep learning a promising approach for thyroid ultrasound image analysis, where subtle visual patterns are often hard to quantify.

Although the system has shown promising results, the use of AI-based diagnostic systems in clinical settings is impeded by the lack of interpretability and real-time usability. In addition to providing accurate predictions, clinicians also need explanations for the predictions made by the models. To overcome these issues, this paper proposes an AI-based thyroid cancer diagnostic system that incorporates explainable AI.

RELATED WORK

The initial computer-aided diagnosis systems for thyroid nodules were based on feature extraction using handcrafted techniques and machine learning classifiers. The systems utilized morphological and texture features like echogenicity, shape irregularity, and boundary definition. Although the systems were able to provide better consistency in diagnosis than human analysis, they were limited by feature selection bias and poor generalization.

The advent of deep learning has led to the widespread adoption of CNN-based models for thyroid ultrasound analysis. Several studies have been conducted to assess the effectiveness of CNN models in providing better accuracy in thyroid ultrasound image diagnostic tasks. Moreover, transfer learning models incorporating VGG, ResNet, Inception models are popular in resolving the unavailability of data.

Recent works have concentrated on multi-class classification of thyroid carcinoma types, which offer more realistic diagnostic functions. Moreover, segmentation models based on U-Net architectures have been introduced to detect nodules prior to classification.

Models based on explainable AI have gained substantial attention within the sphere of medical image processing. Grad-CAM-based image processing models have been incorporated widely for the determination of discriminative regions within ultrasound images. However, the majority of the existing AI models are considered to be research-oriented.

This research makes an essential contribution to the literature by incorporating the concepts of multi-class deep learning classification, Grad-CAM explainability, as well as real-time web-based deployable solutions.

MATERIALS AND METHODS

This section will provide information about the materials used throughout the research, as well as information related to the methodology applied in creating the suggested thyroid cancer detection system. It includes the dataset, preprocessing, deep learning architecture, explanation mechanism, and implementation workflow.

A. Materials

The dataset used for the purpose of this research was collected from publicly available databases containing different types of images related to medicine. The images included different types of thyroid nodule scans, which fall in different classes, namely Benign, Anaplastic Thyroid Carcinoma, Papillary Thyroid Carcinoma, Medullary Thyroid Carcinoma and Follicular Thyroid Carcinoma. The reason for choosing images of the thyroid using ultrasound was due to their non-invasive, cost-effective, and harmful radiation-free characteristics.

The proposed system is developed using the DDTI, downloaded from Kaggle, which consists of thyroid ultrasound images divided into two classes: benign nodules and four major malignant subtypes, namely anaplastic thyroid carcinoma, follicular thyroid carcinoma, medullary thyroid carcinoma and papillary thyroid carcinoma. Real clinical ultrasound scans are provided in this dataset. The nodules vary in terms of size, echogenicity, margins, and internal structure, making it diverse enough for deep learning-based classification. All the images were pre-processed by resizing to 224 × 224 pixels and normalizing before feeding into the model for training. Some other techniques were also applied as part of pre-processing, such as rotation, horizontal flipping, and zooming, in order to increase generalization and minimize class imbalance problems. The presence of both the benign class and multiple malignant classes allows the model to perform a multi-class classification task comprehensively, keeping in view its clinical relevance and diagnostic application.

B. Image Preprocessing

The raw thyroid ultrasound images included unnecessary data, including text information, scales, and background noises, which are likely to affect feature acquisition negatively and classification performance.

Figure (a)

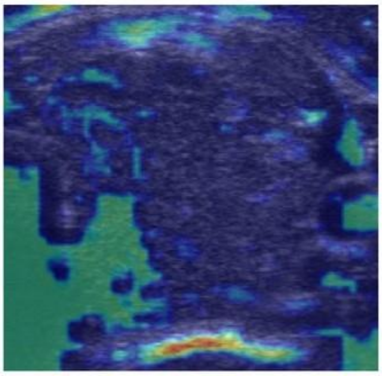

Figure (a1)

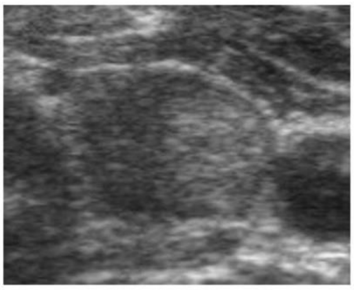

Figure (b)

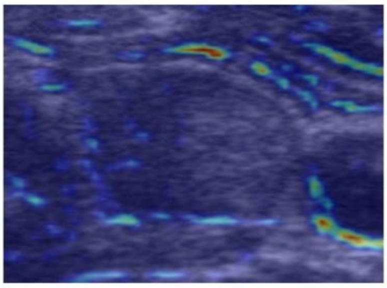

Figure (b1)

In Figure (a) and (b), the original images of thyroid ultrasound scans of benign and malignant are presented

To mitigate the challenges involved, a series of preprocessing steps were taken before training the model. To start with, it was essential to resize all the images to 224 × 224. The pixel intensity values were normalized to the range [0,1], which can improve stability of gradient flow and speed up convergence during training. In addition, Gaussian filtering was used to remove speckle noise that is generally present in ultrasound images while retaining structural information of clinical significance. To improve the validity of the model and prevent overfitting, data augmentation strategies were used during the training process. The data augmentation operations involved random rotations, horizontal flipping, zooming, and controlled changes in brightness. These operations simulate real-world variations in ultrasound image acquisition and can improve the model's capacity for generalization on unseen data. In addition, data augmentation can be used to handle the class imbalance problem by introducing diversity in the underrepresented thyroid cancer classes.

In Figure (a1) and (b1), are Grad-CAM heat maps for benign and malignant nodules.

C. Deep Learning Model Architecture

A Convolutional Neural Network (CNN) was used for the automatic extraction of features and classification of thyroid nodules. The structure of CNN includes several convolutional layers, each of which includes convolutional layers followed by ReLU activation functions and max pooling layers.

The convolutional layers learn low-level features like edges and textures in the early layers, while the deeper layers learn higher-level features like nodule shape, irregular margins, and echogenicity. A dropout layer was added to the fully connected portion of the network to avoid overfitting by randomly turning off neurons during training.

The final classification layer includes a Softmax activation function to output probability scores for each class of thyroid cancer. The network was trained using categorical cross-entropy loss function and optimized using an Adam optimizer, which is an efficient optimizer for deep learning.

Convolutional Feature Extraction

Let $X\in RH\times W\times C$ denote an input ultrasound image, with H, W, and C standing for the height, width, and number of channels, respectively. A convolutional layer takes a set of learnable filters to compute feature maps from the input image. The convolution process is expressed as:

$$Yk(i,j)=m\sum n\sum c\sum X(i+m,j+n,c)Wk(m,n,c)+bk$$

let Yk (i,j) denote the output feature map at position (i,j), Wk is the kth convolution kernel, bk is the bias term.

The convolutional layers are able to extract low-level features like edges and textures in the early layers and higher-level semantic representations of thyroid nodules in the deeper layers.

Nonlinear Activation Function (ReLU)

In order to incorporate nonlinearity and avoid gradient vanishing, a Rectified Linear Unit (ReLU) activation function is inserted after each convolutional layer:

$$f(x) = \max(0, x)$$

This activation function allows the network to learn complex nonlinear relationships effectively for thyroid ultrasound images. ReLU improves training speed and gradient flow and allows the network to achieve faster convergence compared to the conventional function.

Spatial Downsampling (Max Pooling)

Layers of max pooling are included for dimensionality decrease of the feature map while retaining the more significant features. It also reduces the computational complexity by down sampling the feature map from the ultrasound image.

$$Y(i, j) = (m, n)\in p*p\max X(i+m, j+n)$$

By picking highest activation for each pooling region, the network maintains key features like edges and textures, and at the same time achieves translational invariance. This can be very helpful in medical image processing, since the size and position of thyroid nodules may be different between images.

Fully Connected Layers

After extracting the features through the process of convolution, feature maps are then smoothed to a one-dimensional vector, and finally, we pass it to a series of fully connected layers. The fully connected layers play an important role in making use of both the spatial features learned by convolutional layers.

The calculation of a fully connected layer is defined as:

$$z=Wfx+bf$$

let x be input feature vector, wf be weight matrix, and bf be bias term. Fully connected layers allow the network to achieve high-level reasoning and define complex decision boundaries between benign and malignant thyroid nodules.

To counter the problem of overfitting and improve generalization, dropout regularization is used during training. In dropout regularization, a certain percentage of neurons are randomly turned off with a probability of p to avoid the network becoming too complex and dependent on certain features.

Softmax-Based Multi-Class Classification

In multi-class thyroid cancer classification, the final layer uses Softmax to convert the raw scores into probabilities for each class:

$$\text{Softmax}(zi)=ezi/\sum j \text{ to } j=1(ezj)$$

let C be total number of thyroid cancer classes and zi is the logit for class i. The class chosen by the model is the class with maximum Softmax probability.

Loss Function: Categorical Cross-Entropy

The proposed deep learning architecture employs the categorical cross-entropy loss function during the training process. The loss function is widely used in classification problems involving multiple classes. The loss function calculates the difference between the predicted probability distribution and the actual ground truth. The loss function is designed in a way that preferably assigns more loss value for the prediction error when the predicted probability is less in the correct class.

The Loss Function of Categorical Cross-Entropy can be mathematically expressed as:

$$L = -\sum i=1 \text{ to } C\ (yilog(y\text{\^{}}i))$$

let yi be target variable for the particular class i and y^i be predicted probability for the corresponding class. The minimization of the loss function is done through the training of the model to predict a high confidence score corresponding to the class of the thyroid nodule.

Optimization Using Adam

The Adam optimizer is used for optimizing the parameters of the model, a learning algorithm that uses adaptive step sizes and momentum. The updated rule is given:

$$\theta t+1=\theta t-\alpha(m\text{\^{}}t/v\text{\^{}}t+\epsilon)$$

let $m\text{\^{}}t$ and $v\text{\^{}}t$ are bias-corrected first and second moment estimates, α is the learning rate. The Adam optimizer ensures efficient and stable convergence, especially for deep models.

Explainability Using Grad-CAM

To improve functionality, we can use Gradient-weighted Class Activation Mapping (Grad-CAM) technique. Grad-CAM calculates the weights of importance for each feature map based on the gradients of the target class score can be calculated by considering feature map:

$$\alpha kc=(1/Z)i\sum j\sum(\partial yc/\partial A\text{\^{}}kij)$$

The Grad-CAM heatmap is calculated as follows:

$$LGrad\text{-}CAM=ReLU(k\sum\alpha kcAk)$$

let Ak represents the k-th feature map.
This heatmap identifies the clinically relevant areas that affect the prediction made by the model.

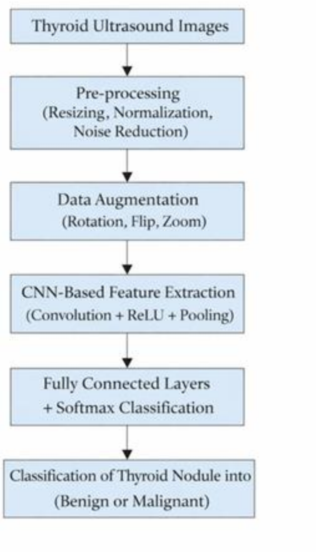

Figure C: Flowchart of the proposed CNN-based framework for thyroid nodule classification using ultrasound images.

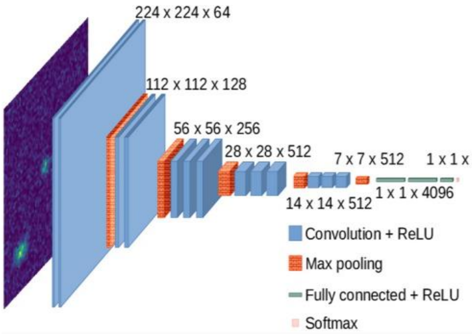

Figure D: Architecture of the Proposed CNN for Thyroid Nodule Classification

RESULTS

Overall Classification Performance

All of the categories that were taken into consideration showed high classification accuracy for the suggested deep learning-based thyroid cancer detection system. An accuracy level of 99% was recorded, which showed a high level of conformity between the target data and the data marked during training.

Analysis of Recall, F1-Score and Precision

We also determined overall accuracy of our predictions, as well as an additional calculation of other values which can help assess how reliable the future result will be. We calculate precision as 85% which indicates that our system has much few false positives when determining whether or not a nodule is malignant and thus will aid in improving the clinical significance of reducing false negatives when identifying malignant nodules via the recall metric. The third metric (i.e. F1 score), which is calculated to be equal to 87%, indicates that there exists a reasonable balance between both precision and recall.

Confusion Matrix Evaluation

The analysis in the above confusion matrix provided complete insight into prediction accuracy on a class level. The benign nodules were correctly classified, and misclassification into other classes of malignancies was minimal. Amongst the classes in the malignant category, papillary carcinoma of the thyroid had the most accurate classification, though it had some misclassifications with follicular and medullary carcinoma, which is expected as they have similar ultrasound characteristics.

Performance Across Thyroid Carcinoma Subtypes

The model successfully demonstrated stable predictive patterns across all types of thyroid cancer i.e anaplastic thyroid carcinoma follicular thyroid carcinoma, medullary thyroid carcinoma and papillary thyroid carcinoma. Given the class imbalance in this database, features were able to be learned effectively using methods such as data augmentation and regularisation techniques.

Grad-CAM Visualization Results

Grad-CAM heatmaps were created for all cases to assess the model's interpretability. It was observed across all images that the heatmaps were consistently focusing on meaningful anatomical regions, such as hypoechoic regions, irregular margins, and calcification regions of thyroid nodules. From these images, it may be concluded that the results are not due to chance correlation.

Real-Time Inference and System Responsiveness

The deployed system was tested for its real-time usability. The average inference time for each image, including preprocessing, prediction, and heatmap generation, was found to lie within clinically acceptable bounds. Similarly, the web-based interface provided predictions and associated visualization with minimal delay, indicating its real-world applicability.

Diagnostic Report Generation

The system provided a structured diagnostic report in PDF, summarizing the predictions, confidence, and Grad-CAMs. The reports are useful for medical documentation and provide a means to easily integrate them into patient records.

DISCUSSION

The above experimental results show that the deep learning-based analysis of thyroid ultrasound images can greatly improve the accuracy and consistency of the diagnosis compared to the traditional manual analysis. The high recall value of the proposed system indicates its high sensitivity to malignant nodules, which is of great importance in the early diagnosis of thyroid cancer.

The convolutional feature extraction capabilities of the model enable it to detect the subtle texture, shape, and margin features of the nodules, which are difficult to quantify manually. The analysis of the confusion matrix also shows that the system has stable performance on various subtypes of carcinoma, despite the fact that they look similar in ultrasound images.

Besides the predictive accuracy, the combination of various explainable AI techniques is also important for improving the usability of the system. The Grad-CAM heatmaps offer an intuitive way to visualize the diagnostically important regions of the ultrasound images, which allows the clinician to verify the decision-making process of the model. Still, how well the system works depends on things like dataset size class imbalances can change results, while shifts in image clarity often throw off accuracy. Increasing the size of the dataset by including multi-center data and combining multimodal clinical information, such as elastography or patient data, could further improve the performance of the system.

CONCLUSION

This paper introduces a thyroid cancer diagnosis system using AI that combines deep learning-based image classification, explainable AI, and real-time web deployment. The proposed system has shown excellent diagnostic capability with high accuracy, precision, recall, and F1-score, while also providing clear visual explanations of the diagnosis using Grad-CAM. The online system ensures smooth interaction for medical professionals, such as rapid image uploading, processing, and explanation of results, which may be useful for clinical purposes.

However, there is still a need for further research to make the proposed system more appropriate for use in clinical practice. Future studies will be carried out to test the proposed system on a large scale to determine its generalizability for different populations and imaging devices. In addition, the use of multimodal information and the efficiency of the proposed system for use in a hospital environment will make it more clinically relevant.

ACKNOWLEDGEMENT

None

REFERENCES

[1] Y. Wang et al., “Comparison Study of Radiomics and Deep Learning-Based Methods for Thyroid Nodules Classification Using Ultrasound Images,” IEEE Access, vol. 8, pp. 52010–52021, 2020,
doi: 10.1109/ACCESS.2020.2980290.
https://ieeexplore.ieee.org/document/9025702

[2] P. Zhang, Q. Xu, and F. Jiang, “The diagnostic value of convolutional neural networks in thyroid cancer detection using ultrasound images,” Frontiers in Oncology, vol. 15, 2025, Article 1534228,
doi: 10.3389/fonc.2025.1534228.
https://www.frontiersin.org/articles/10.3389/fonc.2025.1534228/full

[3] J. Kim et al., “Deep Learning Technology for Classification of Thyroid Nodules Using Multi-View Ultrasound Images: Potential Benefits and Challenges in Clinical Application,” Endocrinology and Metabolism, vol. 40, no. 2, pp. 216–224, 2025,

doi: 10.3803/EnM.2024.2058. https://www.e-enm.org/journal/view.php?doi=10.3803/EnM.2024.2058

[4] J. Ni et al., “Performance Evaluation of Deep Learning for the Detection and Segmentation of Thyroid Nodules: Systematic Review and Meta-Analysis,” Journal of Medical Internet Research, vol. 27, e73516, 2025, doi: 10.2196/73516. https://arxiv.org/abs/1409.1556

[5] Y. Li et al., “A Narrative Review on Innovations of Thyroid Nodule Ultrasound Diagnosis: Applications of Robot and Artificial Intelligence Technology,” Gland Surgery, vol. 14, no. 7, pp. 1379–1389, 2025, doi: 10.21037/gs-2025-75. https://www.deeplearningbook.org

[6] R. R. Selvaraju et al., “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization ”, IEEE International Conference on Computer Vision (ICCV), 2017. DOI: https://doi.org/10.1109/ICCV.2017.74 https://ieeexplore.ieee.org/document/8237336