

Data-Driven Analysis and Prediction of Thyroid Trends in Women

Nitturu Asha Jyothi¹, M Charan Kumar Reddy², M Rithwik, P Deepak Chowdary³, G Vinay Yadav⁴
^{1,2,3,4}*Department of Computer Science, GITAM University Bangalore Campus, India-561203*

Abstract—Endocrine health challenges, specifically those involving the thyroid gland, represent a major global medical concern that predominantly impacts the female population. While clinical practitioners rely heavily on ultrasound imaging for diagnosis, the process is often hampered by subjective interpretation and the varying experience levels of medical staff. To mitigate these inconsistencies, we engineered an automated diagnostic framework that utilises advanced neural networks to categorise thyroid scans into two distinct groups: healthy and pathological.

Our methodology employs a streamlined, pre-trained image recognition model known for its high-performance benchmarks and minimal computational footprint. We trained this system on a comprehensive library of 5,000 verified ultrasound samples. By integrating rigorous pre-processing techniques—such as image normalisation and artificial data expansion—we ensured the model remains robust across various clinical scenarios. Finally, we deployed the tool as a web-accessible application, providing medical professionals with a responsive and reliable secondary opinion to assist in early disease identification.

I. INTRODUCTION

Thyroid health serves as a primary indicator of overall physiological balance, regulating everything from metabolic rates to heart function. When the gland malfunctions, it often leads to serious conditions like hyperthyroidism or nodules, which disproportionately affect women. While ultrasound is the standard for non-invasive screening, relying solely on human interpretation introduces the risk of subjective error.

To bridge this gap, we propose an intelligent screening system powered by deep learning. Our approach uses a refined MobileNetV2 architecture to classify ultrasound scans as either Normal or Abnormal. By integrating this technology into an accessible web platform, we provide healthcare workers with a

consistent, secondary diagnostic tool to improve early detection and patient outcomes.

II. RELATED WORK

Historically, thyroid diagnosis has relied on a mix of blood tests and radiologist-led ultrasound sessions. While effective, these manual processes are vulnerable to inconsistencies depending on the specialist's experience. In response, the medical community has turned toward automated diagnostic systems.

Recent breakthroughs in Convolutional Neural Networks (CNNs) have redefined medical image analysis. While architectures like ResNet and VGG16 are common, they are often computationally heavy. Our study adopts MobileNetV2 for its unique balance of speed and precision. This lightweight framework allows for efficient, real-time analysis without sacrificing the accuracy needed for clinical decision support.

III. METHODOLOGY

We built our classification pipeline around the TN5000 dataset, which contains 5,000 expertly labelled images. To prepare the data, we resized all scans to 224×224 pixels and normalised the pixel values to stabilise the training process. We also utilised data augmentation—including rotations and brightness shifts—to help the model generalise better to new data.

The core of our system is a MobileNetV2 model acting as a feature extractor. We added custom dense layers with a sigmoid activation function to handle the final binary classification. We trained the network using the Adam optimiser to minimise binary cross-entropy loss. Finally, we deployed the trained model through a

Streamlit application, creating a user-friendly interface for instant image analysis.

IV. RESULTS

Our experiments compared the MobileNetV2-based model against traditional machine learning and standard CNN baselines.

TABLE I. *Comparison of Methods*

Method	Accuracy	Time (ms)
Proposed (MobileNetV2)	74.2%	14.5
Baseline A (Traditional ML)	68.7%	19.8
Baseline B (Basic CNN)	71.3%	17.6

The results show that our proposed system achieves the highest accuracy (74.2%) while maintaining the fastest processing time (14.5 ms). This combination makes it particularly well-suited for high-volume clinical environments.

V. DISCUSSION

The findings of this research confirm that deep learning architectures can significantly modernise the way we approach thyroid diagnostics. By implementing the mobilenetv2 model, we successfully achieved a performance profile that balances high-level diagnostic accuracy with the lean computational requirements essential for real-world medical software.

A primary takeaway from this study is the superiority of automated feature extraction over traditional methodologies. In older machine learning frameworks, researchers had to manually define which image characteristics were important, a process prone to human oversight. Our model, however, autonomously identifies the subtle textures and structural irregularities within ultrasound scans that signal a disorder. Furthermore, transfer learning proved to be the "engine" of our success; by building upon a pre-trained network, the system gained a sophisticated understanding of image patterns that would have been impossible to learn from a smaller medical dataset alone.

While our confusion matrix indicates that most scans were classified with high precision, we did observe occasional misclassifications. These typically occurred in instances of low-quality imaging or where the visual "signature" of a normal thyroid closely resembled an early-stage abnormality. Additionally, the current system is a binary classifier, meaning it flags the presence of an issue but does not yet specify the exact pathology, such as hypothyroidism or hyperthyroidism.

To move beyond these hurdles, our future roadmap includes:

- Dataset expansion: integrating larger and more diverse image libraries to improve the model's ability to generalise across different patient demographics.
- Advanced architectures: testing even more robust frameworks like Efficient Net or ResNet to see if they offer a significant jump in accuracy.
- Explainable AI (XAI): incorporating tools that show doctors *why* the AI made a certain choice, which builds the trust necessary for clinical adoption.

VI. CONCLUSION

This study has successfully developed and validated a deep learning-based system tailored for the early detection of thyroid disorders. By utilising the MobileNetV2 architecture and the TN5000 dataset, we have created a framework that distinguishes between Normal and Abnormal thyroid states with both speed and consistency. Our results prove that these models can effectively identify abnormalities while maintaining the low latency required for a seamless user experience.

The highlight of this project is its practical application: a Streamlit-based web tool that allows healthcare providers to upload ultrasound images and receive immediate diagnostic support. This deployment moves AI out of the theoretical realm and into a space where it can actively assist in faster, more accurate preliminary screenings.

We emphasise that this technology is designed to serve as a decision-support tool—a "digital second opinion"—rather than a replacement for the nuanced judgment of a radiologist. However, it marks a significant step forward in the role of Artificial

Intelligence within endocrine health. By continuing to refine our datasets and optimising our neural networks, we aim to transform this system into a standard resource for real-world medical environments, ultimately leading to earlier intervention and better health outcomes for women worldwide

ACKNOWLEDGMENT

The authors would like to express their heartfelt gratitude to the GITAM School of Computer Science and Engineering, Bengaluru Campus, for providing the vital resources and academic environment that made this research possible. We owe a significant debt of gratitude to our project supervisor, Ms. Nitturu Asha Jyothi, Assistant Professor, for her tireless mentorship. Her insightful critiques and constant encouragement were essential in refining our methodology and ensuring the success of this study. We also wish to thank Prof. Dr. A. Vadivel, Head of the Department, for his steady professional support and leadership throughout our project work. Finally, we extend our sincere thanks to our faculty members, peers, and colleagues whose diverse contributions—whether direct or indirect—helped us bring this project to fruition

REFERENCES

- [1] F. Niyaz et al., "Deep Learning Frameworks for the Diagnostic Identification of Thyroid Irregularities in Ultrasound Scans," *IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 2025.
- [2] P. Kumari et al., "Automated Classification of Thyroid Pathologies using Advanced Convolutional Neural Networks," *International Conference on Electrical Electronics and Computing Technologies (ICEECT)*, 2024.
- [3] C. K. Viknesh et al., "Evaluating Deep Learning Models for the Structural Analysis of Thyroid Ultrasound Data," *International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*, 2024.
- [4] W. Xu et al., "Contemporary Methodologies in Deep Learning for Modern Medical Image Classification," *International Congress on Image*

and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2024.

- [5] X. A. Kesarkar and K. V. Kulhalli, "Machine Learning-Driven Approaches for Predicting Thyroid Disease Patterns," *International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 2021.
- [6] S. Gaurkhede et al., "Assessing the Efficacy of Transfer Learning in Clinical Image Diagnostic Tasks," *International Conference on Sentiment Analysis and Deep Learning (ICSADL)*, 2025.
- [7] L. Jenitha Mary et al., "AI-Powered Detection of Thyroid Abnormalities via Convolutional Neural Architectures," *International Conference on Innovative Trends in Information Technology (ICITIIT)*, 2025.
- [8] N. S. H. Sankabathula et al., "Integrating Computational Intelligence for High-Fidelity Healthcare Image Diagnostics," *International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)*, 2024.