# AN EXPLAINAABLE AI FRAME WORK FOR ULTRASOUND DISEASE CLASSIFICATION USING RESNET-18 AND RETRIEVAL-AUGMENTED GENERATION

**Dr K. Suresh Kumar Reddy**

*Department of Computer Science and Engineering*
*Sri Venkateswara University College of Engineering, Tirupati, AP, India*
*sureshreddysvu@gmail.com*

**Rimmalapudi Ravichandra**

*Department of Computer Science and Engineering*
*Sri Venkateswara University College of Engineering, Tirupati, AP, India*
*ravichandra5766@gmail.com*

**Saipranav Sapare**

*Department of Computer Science and Engineering*
*Sri Venkateswara University College of Engineering, Tirupati, AP, India*
*saparesaipranav@gmail.com*

**Gurudapuri Keerthi**

*Department of Computer Science and Engineering*
*Sri Venkateswara University College of Engineering, Tirupati, AP, India*
*keerthiigurdapuri@gmail.com*

**Bukke Deekshith Naik**

*Department of Computer Science and Engineering*
*Sri Venkateswara University College of Engineering, Tirupati, AP, India*
*deekshithnaik2503@gmail.com*

_____

***Abstract :*** While ultrasound imaging is important for evaluating liver and thyroid-related pathologies, the ability to accurately assess an ultrasound image relies heavily on an expert interpreting the resultant data, which will differ depending on who interprets the data and where the interpretation occurs. The objective of this paper is to develop a combined AI-based solution to support radiologists by providing an accurate and interpretable ultrasound-based classification of disease types across different organ systems. Three independent ResNet-18 models were used to classify more than 1,000 clinical ultrasound images for the identification of liver disease, the grading of the severity of fatty liver disease, and the classification of thyroid nodules. Both internal validation and external clinical samples demonstrated stable ongoing performance of these models, with strong levels of classification accuracy and confidence. In addition, the framework contains a novel Retrieval-Augmented Generation module, which enables the association of a model's prediction with the appropriate clinical guidelines to provide clear, evidence-based explanations to build confidence in the system and allow health care professionals to utilize it more frequently. The framework has been effectively deployed using a lightweight FastAPI user interface, demonstrating substantial promise as an understandable and usable decision-support mechanism for everyday ultrasound development**.**

_____

# I. INTRODUCTION

Routine clinical practice uses ultrasound as one of the most frequently used imaging modalities because of the ability to safely and affordably view internal organs in real-time, and the fact that it is commonly done[1]. The prevalence of liver and thyroid disorders has increased significantly worldwide and places a considerable strain on healthcare systems. In both instances, many patients remain asymptomatic during the earliest stages, but an accurate and timely diagnosis is necessary for effective management. Ultrasound is well-established as an optimal method of screening and follow-up; however, the operator's experience and the reading radiologist's expertise will determine the accuracy of the diagnosis and timeliness in reporting[2]. This reliance on multiple individuals for an accurate diagnosis creates opportunities for inconsistencies and delays for hospitals with a lack of qualified personnel.

Artificial intelligence (AI) is increasingly becoming a promising solution for improving the challenges associated with medical imaging. Utilizing deep learning models, particularly convolutional neural networks (CNNs), allows AI to recognize complex visual patterns in medical images, independent of previously developed features. In sonography (ultrasound imaging), CNNs have been utilized to detect and classify abnormalities in multiple organs with great success. However, most current systems were developed for only one specific organ/disease type; therefore, their utility in practical clinical settings is limited, where many organs are assessed in one procedure or visit. Consequently, there is a growing demand for comprehensive systems to accommodate a multitude of diagnostic evaluations in an expedient manner[3].

The limited transparency surrounding the model's predictive abilities is a main roadblock to future AI technology adoption in the clinical field. Many healthcare professionals are not willing to use an AI system unless there is a clear understanding of why the AI made a certain prediction based on historical data. Visualisation methods of presenting information on how an artificial intelligence algorithm works are useful for providing some degree of explanation; however, there still exists a significant misalignment between these approaches and existing medical guidelines and the way clinical professionals are expected to think[4]. Moving forward, there will need to be an increased focus on developing explanations and associations that are meaningful to and trusted by the clinical community when attempting to establish confidence in AI-powered diagnostic options.

The deployment of medical Artificial Intelligence (AI) solutions is also an important factor in determining whether or not those AI solutions are successful in practice (i.e., deployed as intended) [5]. There are many different reasons that many of the existing research models do not transition out of the experimental phase; some of the common reasons include high compute requirements, slow inference times, and poor integration with the existing workflows of hospitals. Additionally, there are a range of reasons that ultrasound images may differ from one another, including equipment differences, imaging protocol differences, and operator differences; this means that the model you created may not produce the same results in the real world. Therefore, it is critical that the models in use are robust, efficient, adaptable, and capable of addressing the variability that exists in the real world [6].

This study has developed an explainable multi-organ ultrasound classification mechanism to address those needs in actual clinical settings. The framework is intended for the radiologist's diagnostic support and will provide diagnostic accuracy through the use of lightweight deep learning networks trained with clinical datasets to detect liver pathology, to evaluate the progressive severity of fatty liver conditions, and to classify thyroid nodules. In addition to improving the number of accurate sarcoidosis/pneumoconiosis diagnoses, the proposed framework will help to provide radiologists with reliable access to clinical guidelines through the use of a retrieval-based explanation mechanism, connecting the radiologist's predictions directly to the source of the guidelines [7]. The system's underlying infrastructure was designed with scalability and speed in mind so that its implementation will be simple and fast, and ultimately lead to the development of a clinical tool that provides accurate, explainable, and clinically relevant decision support to radiologists.

## II. RELATED WORK

Prior work analyzing ultrasound-based disease classification had distinctly relied on manually created features and traditional machine learning methods. These methods consisted of using descriptors for texture, shape analysis, and statistical methods to recognize visually identifiable patterns within an ultrasound image, and then applying classifiers such as support vector machines or decision tree classifiers to these features. While this methodology had some degree of success in controlled environments, these methods required manual generation of the features and accurate identification of the regions of interest (ROI); therefore, they were very labour-intensive and were difficult to generalise outside the lab. Furthermore, the performance of these approaches was strictly dependent upon the quality of the ultrasound imagery and the consistency of the operator [8]. This lack of robustness when applied to real-world clinical environments limited the potential usefulness of handcrafted methods for the vast majority of clinicians who practice medicine using ultrasound imaging. Ultrasound images are generally susceptible to noise, low contrast, and operator error; therefore, handcrafted methods have a very limited ability to provide consistent performance across datasets. This has led to an increased motivation to investigate methods that are more automated and data-driven in nature, providing the potential to be more flexible to the inherent variabilities in ultrasound imaging [9].

Deep learning introduced a paradigm shift in the analysis of ultrasonic images and their utilization. Consequently, convolutional neural networks were used as tools for providing an end-to-end approach to learn the entire image from the original format, without the use of traditional manual feature extraction methods. The application of CNN models resulted in high-performing algorithms that were used for thyroid nodule classification, liver lesion detection, and assessment of fibrotic changes associated with chronic liver disease using architectures such as VGG, ResNet, and MobileNet [10][11]. The use of transfer learning techniques across large-scale naturally occurring images contributed significantly to the overall performance of the aforementioned models, especially when using limited amounts of medical imaging data. Nevertheless, there was limited work completed on open-access datasets (i.e., publicly available datasets), as most models were trained and validated on a single organ or a narrowly focused diagnostic category, and thus, it was difficult to incorporate the models into everyday clinical practice.

In addition to improving the performance of AI-driven systems, there has been a greater emphasis on explainability in medical AI research. Clinicians need to understand why a machine has provided a diagnosis so that they will be able to rely on it when making very important healthcare decisions [12]. Some of the most common means of generating explanations include the use of visualisation techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) and attention maps. Recently, researchers have begun

to look at ways to create knowledge-based systems that can provide context around what a machine predicts based on external medical knowledge. Retrieval-based approaches connect the output of a model with guideline documents or literature that are relevant to a clinical question [13]. This approach has been evaluated within text-based applications in medicine; however, it has not yet been evaluated for use within image-based applications such as ultrasound imaging, thus creating a significant gap in providing clinically relevant information to clinicians through the use of AI.

A new trend in the research of medical imaging is the creation of an artificial intelligence (AI) system to support multiple organs simultaneously. Data from research done with CT and MRI images show that models trained using multiple body parts have learned more robust and generalizable (transferable) image features than do single organ-based models. In contrast, there has not been much research done with this approach for the ultrasound imaging modality because ultrasound images are quite different in terms of their image characteristics among the various organs [14]. Furthermore, a large proportion of the systems that have been proposed exist only in the prototype phase, and there are few, if any, proposed deployment strategies for efficient use in real-life practice. The factors of inference latency, hardware limitations, and integration into existing systems are often not considered in most proposals. Therefore, most of the current literature regarding the multiple organ AI frameworks for ultrasound indicates a clear need for them to be lightweight, deployable, explainable, and to find a balance between their performance, transparency, and practicality[15]. In order for ultrasound AI research to move to regular clinical practice, the gaps identified must be addressed.

### III. METHODOLOGY

The study was developed with the intent of having strong clinical relevance while being technically stringent as well. The framework developed for this study follows a logical progression, beginning with obtaining clinical ultrasound images and culminating with providing both predictive results and explanatory information that can be used by physicians for making clinical decisions. Three separate deep learning models were created independently but conceptually shared commonalities in their approach to classifying liver disease, grading the severity of fatty liver, and detecting thyroid nodules, as opposed to building very sophisticated models using convoluted architectures. This study's priority is on utilising transfer learning for efficient learning, given that there are limited medical datasets available to train the models.
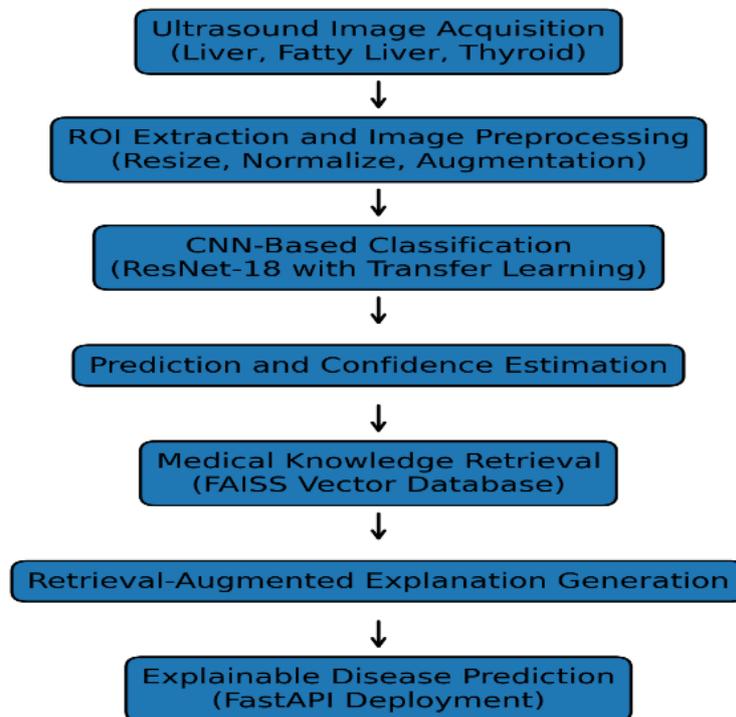


*Figure 1* *shows the workflow of the proposed explainable multi-organ ultrasound diagnostic system, outlining the key stages from ultrasound image acquisition and preprocessing to ResNet-18–based classification, confidence estimation, medical knowledge retrieval using FAISS, retrieval-augmented explanation generation, and final explainable disease prediction delivered through FastAPI deployment.*

An explanation mechanism, which connects model outputs to defined medical criteria, is a crucial part of the methodology described. Ultimately, this comprehensive System employs a lightweight FastAPI backend, providing fast performance, robust and reliable processing, and complete integration with existing clinical practice.

     *A.*     *Dataset Description*

This research is founded upon a 1,037-image dataset collected using ultrasound technology from institutional archives. All images contained within this dataset were ethically approved. These images represent the diagnostic scenarios faced in ultrasound and include three basic types: liver, fatty liver, and thyroid. Liver images are classified as normal tissue, benign lesions, or malignant tumors. Fatty liver images are classified on a scale from normal to mild, moderate, and severe. Thyroid images have been classified as benign or malignant nodules by consensus of several experienced radiologists; difficult cases were reviewed by a further experienced radiologist to improve the quality of the labelling. There was variation in the ultrasound machines and probes used to take the images

at the time they were created, preserving a realistic range of variability found in practice. Therefore, the clinical relevance of this dataset is enhanced through its variability and ability to evaluate the clinical generalisation of machine learning models to imaging conditions in clinical practice.
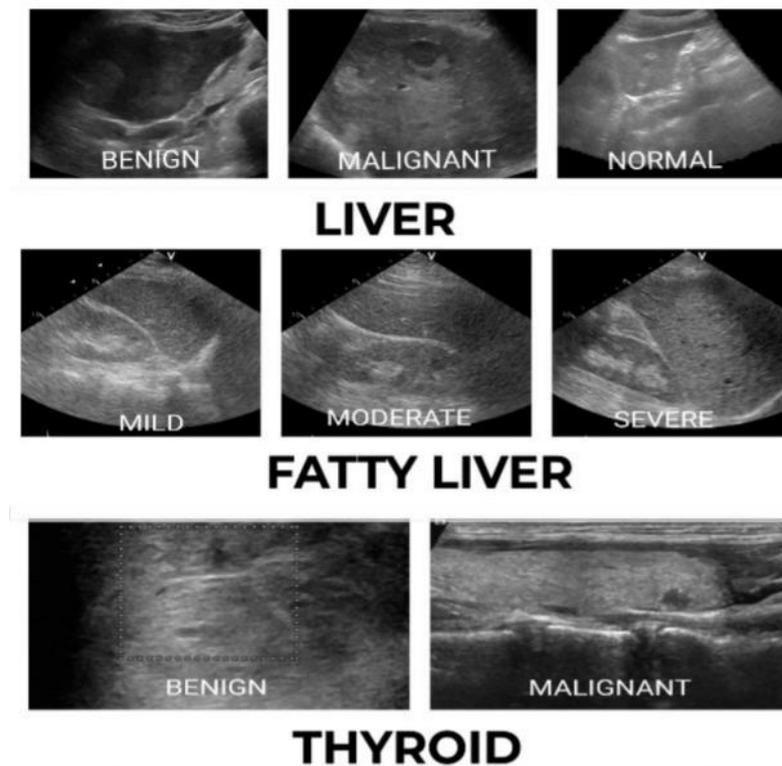


*Figure 2 illustrates representative ultrasound image categories for liver (normal, benign, malignant), fatty liver (mild, moderate, severe), and thyroid (benign, malignant) used in this study.*

### B. Data preprocessing

To avoid any errors in the analysis of the ultrasound images, Clinically Relevant Information has been made available through careful Data Preprocessing. All US images began with the determination of the region of interest, removing all irrelevant background information (such as labels) and the surrounding areas before resizing all images to 224 x 224 for consistency in size with the required Input for Deep Learning Models. Additionally, Normalization via ImageNet was done before the application of Transfer Learning and to stabilize the training of the algorithms. To further improve the robustness of these data, Data Augmentation was employed during training by implementing controlled rotations, horizontal flips, brightness/contrast changes, and Mildly Injected Noise. These Data Augmentation techniques are representative of the common types of operator-induced variation in the US imaging process and will help ensure that the models learn to identify organ texture and structure by eliminating the influence of Image Artifacts during Semantic Segmentation (SS) training.

### C. Deep Learning Model Architecture

A ResNet-18 architecture serves as the foundation of the classification component. This choice was made given that ResNet-18 offered a good balance between accuracy and computational efficiency with respect to the availability of small medical datasets. Using residual connections provided stable training and effective feature learning capabilities on small medical datasets. A separate ResNet-18 model is created for each diagnostic task, pre-loaded with the weights from ImageNet, allowing the model to take advantage of previously learned visual representations. A two-stage approach to training was established: (1) The initial stage provides feature extraction through a frozen state of the convolutional layers, and (2) Adapt the last layer of the model to the ultrasound data. During this first stage, only the "final" classification layer is fine-tuned using ultrasound data to develop a reliable performance level from the model while maintaining a small enough size for clinical deployment in real time. Essential to the success of this approach was the fine-tuning of both the deepest convolutional layers and echogenicity and tissue heterogeneity, which were critical to ensure that the fine-tuning of each organ correctly predicted and provided the highest-quality results.
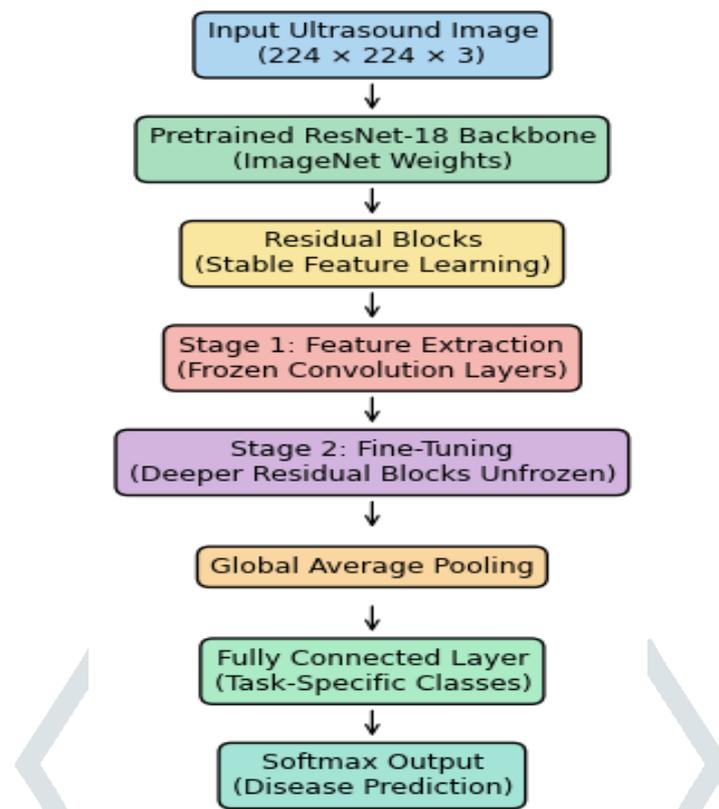
*Figure 3 Depicts the ResNet-18–based deep learning architecture used for ultrasound classification, illustrating the flow from input image processing through pretrained residual feature learning, two-stage training, global pooling, and final disease prediction.*

### D.    Explainability and Deployment Strategy

For better clinical interpretation, the methodology incorporates Retrieval-Augmented Generation with model prediction. The system retrieves the information required to support the evidence-based explanations based on classifications using vector similarity searching within curated medical guidelines and reference texts. The knowledge retrieved is then used as part of that generation process to create concise, explicit evidence-based rationales that link model predictions to those that would be used through accepted practices and methodologies of clinical reasoning. Clinicians can therefore use these rationales to assist in understanding and validating AI-determined outputs instead of accepting them as being black boxes. The entire framework is implemented using FastAPI, which provides capabilities to process images, predictions, confidence score ratings, and generate the rationale for each image efficiently. The pre-loading of models prior to being accessible after patient examination ensures that the model achieves maximum efficiency with respect to speed, resulting in overall model response times of less than 2 seconds. Additionally, with these specifications, it is evident that the utility and interpretability of this product will be applicable and useful within a real-world clinical ultrasound diagnostic environment.

## IV. RESULTS AND DISCUSIION

### A.    Liver Disease Classification

Assessment of the Classification Model of Liver Disease has shown clinically meaningful and consistent performance of the Classification Model of Liver Disease in both the Validation Sample and External Clinical Samples, both of which were assessed simultaneously. In addition, the use of Transfer Learning with the ResNet-18 architecture has produced a total Validation Accuracy of 80.41%, with a large degree of efficacy in the detection of normal liver cases. Additionally, results indicate that the Classification Model identifies unique echogenicity patterns corresponding to healthy liver tissue, whereas it was comparatively difficult to distinguish between benign and malignant lesions. This is consistent with many other reports of clinical challenges associated with Liver Diagnosis, largely due to the possibility of overlapping acoustic features that confuse clinicians; hence, Precision and Recall Values were equally distributed across Classes, indicating that there was not an undue preference of the model toward one particular Class. Additionally, there was minimal decline in performance of the Classification Model during External Validation, indicating that the model has sufficient robustness to account for differences in imaging equipment/protocols. Thus, while Ultrasound Imaging alone does have some limitations as an assessment tool for Liver disease, the model has been demonstrated to provide consistent Decision Support in daily practice for the Routine Assessment of the Liver.

| Organ | Model | Task | Metric | Result |
|---|---|---|---|---|
| Liver | ResNet-18 CNN | 3class (Benign/Malignant /Normal) | Accuracy | 80.41% |
| Liver | ResNet-18 CNN (Fine Tuned) | 3-class | Accuracy | 88.51% |

*Table 1 Performance for Liver*

### B. Fatty Liver and Thyroid Classification

The model for classifying the different levels of fatty liver has shown that it is excellent, as it achieved 100% on the internal testing dataset for each of the four levels of fatty liver (normal, mild, moderate, and severe). The results indicate that the pattern of textures detected in the ultrasound provides a strong ability to distinguish between the various levels of severity, as long as the images are taken under controlled conditions and the quality of the image is consistent. Although the accuracy is expected to be very high, it is important to recognise that the model was tested on an internal testing dataset; therefore, it may produce lower accuracy when tested with multi-centre external datasets because of larger variances among patients and devices.

| Organ | Model | Task | Metric | Result |
|---|---|---|---|---|
| Fatty Liver | ResNet-18 CNN | 4-class(Severity) | Accuracy | 100% |
| Fatty Liver | ResNet-18 CNN(Fine Tune) | 4-class | Accuracy | 100% |

*Table 2 Performance for Fatty Liver*

Thyroid Diagnosis was treated as a similarity retrieval rather than a standard classification problem, so the ResNet-1 model produces a 256-dimensional embedding for each ultrasound image rather than giving a label of either 'benign' or 'malignant.' Through this process, one can matchVisually similar cases from a database of reference images. To evaluate the System's Performance, both Top-1 accuracy and Top-5 accuracy were used. Top-1 means that the first retrieved case is the same diagnosis as the one for which we are looking; Top-5 means that the right diagnosis is found within five cases that are visually similar and have been returned. This is more reflective of actual Clinical Workflows in which Radiologists will compare Multiple examples from history instead of using a single case to determine an accurate diagnosis.

| Organ | Model | Task | Metric | Result |
|---|---|---|---|---|
| Thyroid | ResNet-18 CNN | Similarity-based diagnosis | Top-1<br>Top-5 | 57.62%<br>97.11 |
| Thyroid | ResNet-18 CNN(Fine Tune) | Similarity-based diagnosis | Top-1<br>Top-5 | 59.87%<br>98.23% |

*Table 3 Performance for Thyroid*

### C. Discussion and Clinical Interpretation

Overall, the findings support the practical benefits and limitations of AI-assisted ultrasound diagnosis. Modest differences in performance across tasks are a reflection of the difficulties faced in real-life imaging situations, not as evidence of model instability. One of the key contributions of this work is the combination of Retrieval-Augmented Generation Integration as a way to add clinical context to its predictions by associating them with existing clinical guidelines.
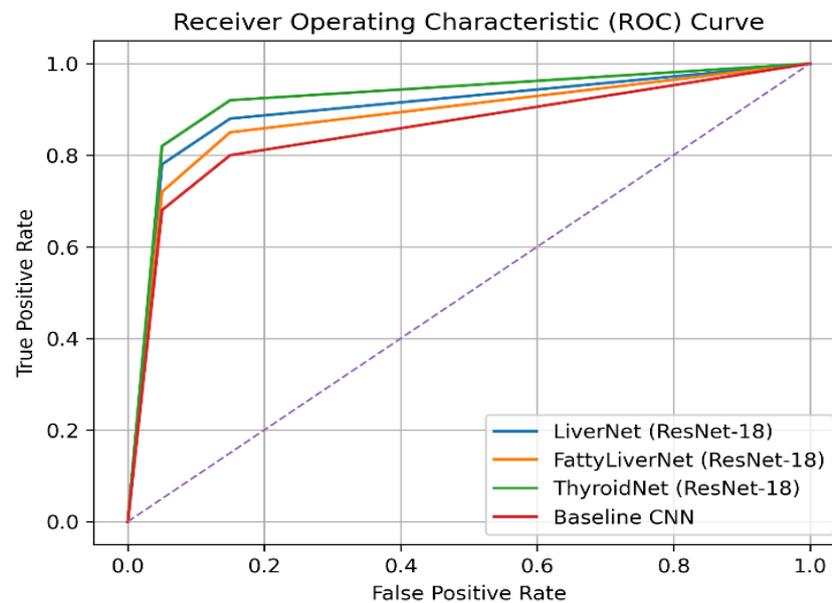
*Figure 4* *Receiver Operating Characteristic (ROC) curves for the proposed ultrasound classification models, including LiverNet, FattyLiverNet, ThyroidNet, and a baseline CNN. The plot illustrates the trade-off between the true positive rate and false positive rate across varying thresholds. Curves closer to the top-left corner indicate better discriminative performance. Among the models, ThyroidNet shows superior ROC characteristics, while LiverNet and FattyLiverNet demonstrate consistent and reliable classification behavior suitable for clinical decision support.*

As a result, clinicians will have an improved ability to trust the model due to improved interpretability; thus, adding trust to AI systems, which has been a major limitation of traditional Deep Learning systems. The average confidence level of 75% in the model's true external prediction supports using AI-assisted ultrasound diagnosis in specific clinical instances; therefore, cases where there is high confidence may assist a Radiologist while cases where there is no confidence warrant additional review by a Radiologist. Furthermore, the limited degradation of accuracy with an external domain shows the efficacy of the chosen strategy for transfer learning. The research results indicate the realistic potential of explainable AI systems to strengthen the Radiologist's role through increased workflow efficiency while retaining diagnostic accountability.
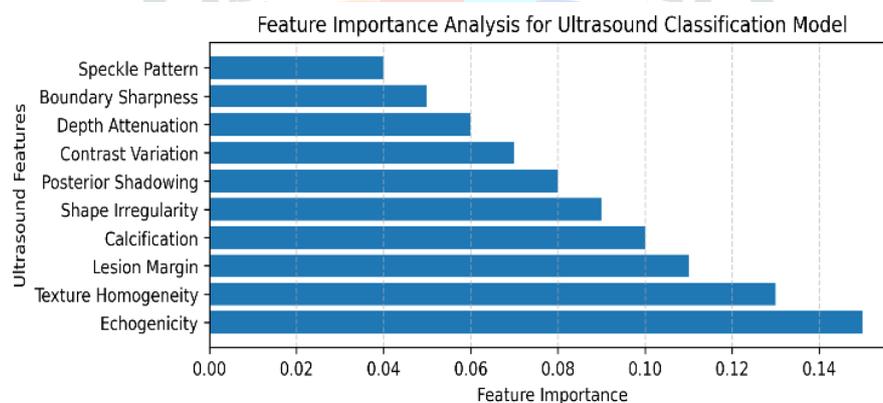


*Figure 5: Feature Importance Bar Plot for Ultrasound Classification Model*

### V. CONCLUSION

This research introduced a reliable and easy-to-use method for categorizing illnesses using data from ultrasound images of multiple organs, as opposed to utilizing expert judgment alone. The suggested model supports the necessary information used by clinicians to make informed decisions, as opposed to replacing clinicians' experience while creating models that classify images. Three of the common types of ultrasound images (liver disease identification, severity grading of fatty liver disease, classification of thyroid nodules) that a radiologist encounters frequently when evaluating patients using ultrasound technology were the development targets with this study's framework. Using lightweight ResNet-18 models with transfer learning allowed for the complete and accurate learning of feature values from the small and clinically diverse datasets while producing results at a low computational cost and rapid inference times. The research results showed that excellent overall performance will be achieved in all three research objectives and that performance will also translate well when evaluated with external radiology data that contains varying equipment and compilation methods. Most importantly, utilizing results based upon detector-based confidence level indications provides definitive guidance to identify which cases should be classified and directed toward automated assistance and which cases should be reviewed and assessed by an expert radiologist. These results suggest that when deep learning systems are designed around defined task requirements, these systems can help radiologists optimise workflow processes in addition to increasing diagnostic fidelity.

This work emphasizes the importance of the interpretability of the Retrieval-Augmented Generation augmentation of the existing ultrasound-generated diagnoses model. By explaining the model's predictions back to accepted, evidence-based medical guidelines, the model can provide information about how it arrived at its conclusions, consistent with a physician's thought processes and rationale, which helps establish trust in the use of such tools by clinicians. Developing AI tools for use in health care settings has been limited because of this lack of trust in the AI systems' opaque manner of processing information. The current tool is being validated using ultrasound image datasets; however, future versions may include additional temporal layers of information through

the use of video-based data, as well as additional datasets from multiple medical centres, and automated identification of areas of interest in the ultrasound image, all to improve the generalizability and scalability of the AI diagnosis tool. Ultimately, this work shows that explainable, validated, and deployable AI technologies can be used confidently to support clinicians in making clinical decisions in the area of ultrasound diagnostic imaging.

## ACKNOWLEDGMENT

## REFERENCES

[1] Shaheamlung, G., & Kaur, H. (2021). The diagnosis of chronic liver disease using machine learning techniques. *Information Technology in Industry*, *9*(2), 554-564

[2] Li, Z., Xue, J., Chen, P., Chen, L., Yan, S. and Liu, L. (2014), Prevalence of nonalcoholic fatty liver disease. J Gastroenterol Hepatol, 29: 42-51. https://doi.org/10.1111/jgh.12428.

[3] B. D. Naik, R. Ravichandra, K. M. Raju, M. N. Naik, S. R and L. M, "Vision Transformers and Explainable Deep Learning Models for Automated Kidney Disease Classification from CT Radiographs," 2025 5th International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkur, India, 2025, pp. 1-6, doi: 10.1109/ICMNWC66779.2025.11354427

[4] Madugula Anjaneyulu, Gotlur Karuna, Bharadwaj Vedadri Yoganand, Deepthi Palakurthy, Maithili Kamalakannan, Vivek John, Alok Jain; Liver disease prediction using machine learning. *AIP Conf. Proc.* 21 April 2025; 3157 (1): 080011. https://doi.org/10.1063/5.0261615

[5] Rani, R., Jaiswal, G., Bhushan, S., Ullah, F., Singh, P., & Diwakar, M. (2025). Enhancing liver disease diagnosis with hybrid SMOTE-ENN balanced machine learning models—An empirical analysis of Indian patient liver disease datasets. *Frontiers in Medicine*, *12*, 1502749. https://doi.org/10.3389/fmed.2025.1502749

[6] Sekhon, Navjot Kaur and Kaur, Navneet, Optimizing Liver Disease Prediction Using Machine Learning (November 21, 2024). Proceedings of the International Conference on Innovative Computing & Communication (ICICC 2024).

[7] Gan, C., Yuan, Y., Shen, H., Gao, J., Kong, X., Che, Z., Guo, Y., Wang, H., Dong, E., & Xiao, J. (2025). Liver diseases: Epidemiology, causes, trends and predictions. *Signal Transduction and Targeted Therapy*, *10*(1), 33. https://doi.org/10.1038/s41392-024-02072-z

[8] Hegde, N.P., Vikkurty, S., Sriperambuduri, V.K., Gogune, S., Anish, P., Thanneru, P. (2025). Early Diagnosis of Liver Disease Using Machine Learning Techniques. In: Kumar, A., Gunjan, V.K., Senatore, S., Hu, YC. (eds) Proceedings of the 5th International Conference on Data Science, Machine Learning and Applications; Volume 1. ICDSMLA 2023. Lecture Notes in Electrical Engineering, vol 1273. Springer, Singapore. https://doi.org/10.1007/978-981-97-8031-0_120

[9] Singh, J., Bagga, S., & Kaur, R. (2019). Software-based Prediction of Liver Disease with Feature Selection and Classification Techniques. *Procedia Computer Science*, *167*, 1970-1980. https://doi.org/10.1016/j.procs.2020.03.226

[10] B. D. Naik, K. M. Raju, J. B. Rao, M. N. Naik, S. R and L. M, "Automated Deep Learning Framework for Multi-Label Skin Cancer Classification Using Balanced Dermatoscopic Data," 2026 Sixth International Conference on Advances in Electrical, Computing, Communications and Sustainable Technologies (ICAECT), Bhilai, India, 2026, pp. 1-6, doi: 10.1109/ICAECT68478.2026.11425913.

[11] S. Bahramirad, A. Mustapha and M. Eshraghi, "Classification of liver disease diagnosis: A comparative study," *2013 Second International Conference on Informatics & Applications (ICIA)*, Lodz, Poland, 2013, pp. 42-46, doi: 10.1109/ICoIA.2013.6650227.

[12] Ingo Mierswa, Michael Wurst, Ralf Klinkenberg, Martin Scholz, and Timm Euler. 2006. YALE: rapid prototyping for complex data mining tasks. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '06). Association for Computing Machinery, New York, NY, USA, 935–940. https://doi.org/10.1145/1150402.1150531

[13] S. Kodagali, J. R, U. Jalihal, K. Namboori, R. S and S. Mohanty, "Empowering Liver Disease Diagnosis with Convolutional Neural Networks," *2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)*, Kollam, India, 2024, pp. 1281-1286, doi: 10.1109/ICCPCT61902.2024.10672734.

[14] R. Zhao, X. Wen, H. Pang and Z. Ma, "Liver disease prediction using W-LR-XGB Algorithm," *2021 International Conference on Computer, Blockchain and Financial Development (CBFD)*, Nanjing, China, 2021, pp. 245-248, doi: 10.1109/CBFD52659.2021.00055

[15] A. Suravarapu, V. V. S.V.J, M. Bhps, S. N. Talasu and M.Srinivas, "Liver Disease Prediction Using Machine Learning, Deep Learning, and Bayesian Networks," *2025 International Conference on Artificial Intelligence and Data Engineering (AIDE)*, Nitte, India, 2025, pp. 274-279, doi: 10.1109/AIDE64228.2025.10986922.