Original

Detection of Valve Vegetations in Native and Prosthetic Valves using Echocardiographic Radiomics and Deep Learning on Transesophageal Echocardiography Images

Farid Esmaely (PhD)[1], Pardis Moradnejad (MD)[2], Shabnam Boudagh (MD)[3], Seyed Mohammad Zamani-Aliabadi (MSc)[1], Hamid Reza Pasha (MSc)[2,3], Ahmad Bitarafan-Rajabi (PhD)[1,2]*, Leyla Ansari (PhD Candidate)[1,4]

[1]Department of Medical Physics, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[2]Cardiovascular Research Center, Rajaie Cardiovascular Institute, Tehran, Iran
[3]Echocardiography Research Center, Rajaie Cardiovascular Institute, Tehran, Iran
[4]Department of Nulear Medicine, Imam Sajjad Hospital, Yasuj university of medical science, Yasuj, Iran

ABSTRACT

Background: Infective Endocarditis (IE) is a life-threatening condition that requires rapid and accurate diagnosis. Transesophageal Echocardiography (TEE) is the gold standard for detecting valve vegetations; however, its interpretation is highly operator-dependent and particularly challenging in patients with prosthetic valves. Recent advances in artificial intelligence, especially deep learning, offer opportunities to improve diagnostic accuracy and reduce observer variability.

Objective: This study aimed to evaluate the performance of deep learning–based models for detecting vegetations in TEE images to support the diagnostic workflow of IE.

Material and Methods: In this retrospective experimental study, a Faster Region-based Convolutional Neural Network (Faster R-CNN) was implemented to localize valve vegetations in TEE images. Four model configurations were developed using DenseNet121 and ResNet50 backbones, each trained in frozen and fine-tuned modes. All models were pretrained on RadImageNet. The dataset consisted of 1,000 annotated TEE frames acquired from both native and prosthetic heart valves.

Results: The fine-tuned DenseNet121 model achieved the best performance, with a mean Average Precision (mAP) of 0.653 and an Area Under the Curve (AUC) of 0.858. Its frozen version demonstrated lower performance (mAP=0.416, AUC=0.640). The fine-tuned ResNet50 model reached a mAP of 0.593 and an AUC of 0.789, while the frozen ResNet50 showed the lowest performance (mAP=0.403, AUC=0.601).

Conclusion: Both fine-tuned DenseNet121 and ResNet50 models demonstrated effective localization of valve vegetations in TEE images, with comparable IoU performance. Although DenseNet121 showed superior classification accuracy, the similar localization results highlight the potential of both models as physician-assistive tools for enhancing IE diagnostic workflows.

Citation: Esmaely F, Moradnejad P, Boudagh Sh, Zamani-Aliabadi SM, Pasha HR, Bitarafan-Rajabi A, Ansari L. Detection of Valve Vegetations in Native and Prosthetic Valves using Echocardiographic Radiomics and Deep Learning on Transesophageal Echocardiography Images. *J Biomed Phys Eng*. 2026;16(2):139-152. do: 10.31661/jbpe.v0i0.2505-1915.

Keywords

Infective Endocarditis; Transesophageal Echocardiography; Deep Learning; Faster R-CNN; Valve Vegetations; Computer-Assisted Image Interpretation; Heart Valves

*Corresponding author:
Ahmad Bitarafan-Rajabi
Department of Medical Physics, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
E-mail: bitarafan@hotmail.com

Received: 3 May 2025
Accepted: 16 August 2025

Introduction

Infective Endocarditis (IE) remains a life-threatening condition, with vegetations representing a key lesion in both native and prosthetic valves [1]. Early and accurate detection of IE is essential to reduce complications and improve patient outcomes, as delayed diagnosis is associated with worse clinical prognosis [2]. Although diagnostic methods have advanced, current techniques, including Transthoracic Echocardiography (TTE), Transesophageal Echocardiography (TEE), blood cultures, and other imaging modalities, still have important limitations [3]. These limitations include difficulty in detecting small or prosthetic valve vegetations, operator-dependent interpretation, inter-observer variability, and susceptibility to human error. Modified duke criteria represent the diagnostic gold standard for IE and guide clinical decision-making, with echocardiography serving as a major criterion [4].

TEE is crucial for detecting endocardial involvement and valve vegetations in IE; however, its diagnostic accuracy is limited in Prosthetic Valve Endocarditis (PVE) [5,6]. According to the duke criteria, blood culture constitutes the other major criterion. Recent studies suggest that Positron Emission Tomography–Computed Tomography (PET/CT) and TEE outperform blood cultures in diagnosing IE, with diagnostic weights of 3.5 and 1.015, respectively, highlighting the central role of imaging in PVE diagnosis [7].

Misidentification of lesions on TEE can worsen clinical outcomes, as prosthetic valve artifacts, such as shadowing, reverberation, and motion noise, often obscure vegetations [8]. Despite the gold standard for IE imaging, TEE has inherent limitations, including operator dependency, inter-observer variability, and challenges in detecting small or prosthetic vegetations due to artifacts. In PVE, these limitations reduce echocardiographic sensitivity to 50%–70% [9,10]. Operator-related variability remains significant, with studies reporting only moderate-to-low inter-observer agreement, even among experienced echocardiographers, particularly for small or equivocal vegetations [11].

According to the 2023 ESC guidelines, PET/CT can aid in PVE diagnosis, though it has limited sensitivity for small vegetations and may yield false-positive results due to early postoperative inflammation or prosthetic artifacts [12,13]. Moreover, PET/CT's limited spatial resolution can lead to false negatives, especially for small vegetations or low-uptake infectious foci [10,14]. In native valve endocarditis, PET/CT sensitivity may be as low as 31%, mainly due to small lesion size and low metabolic activity [12]. Delayed or missed detection of vegetations, particularly in PVE, is associated with worse outcomes and increased morbidity and mortality [15].

The limitations of both TEE and PET/CT underscore the need for integrating Artificial Intelligence (AI) to enhance diagnostic accuracy in complex cases [10,16]. In this context, echocardiographic radiomics has emerged as a novel deep learning approach for disease detection and classification [17,18]. Deep learning automatically extracts low-level features and constructs high-level representations, enabling the detection of complex imaging patterns [19].

Recent advances, particularly in Convolutional Neural Networks (CNNs) and transfer learning, have transformed medical imaging analysis [19-21]. CNNs excel at automatically extracting intricate features from imaging data, making them highly effective for echocardiographic radiomics [22]. Transfer learning leverages pre-trained models to reduce data requirements and training time, improve accuracy, and enable near real-time analysis, thereby overcoming key limitations of conventional machine learning [21].

ResNet and DenseNet, recognized for their efficient feature extraction, are widely applied in medical imaging and have enhanced diagnostic performance across numerous tasks

[22,23]. ResNet addresses the vanishing gradient problem via shortcut connections, making it well-suited for high-resolution medical images and improving diagnostic precision [24-26]. DenseNet, in contrast, connects each layer to all subsequent layers, promoting feature reuse and efficient gradient flow, achieving strong performance with fewer parameters [26,27]. Models are often initialized with ImageNet-pretrained weights to accelerate training and improve generalization [24]. However, as ImageNet contains non-medical images, its features may not optimally transfer to clinical tasks. RadImageNet, with over 1.3 million labeled medical images, enables models to learn features better suited for medical imaging. Pretraining ResNet or DenseNet on RadImageNet strengthens feature extraction in TEE, particularly for detecting subtle patterns such as valve vegetations. Incorporating these backbones into Faster R-CNN has demonstrated strong performance in ultrasound and echocardiographic analysis, supporting its use in TEE image interpretation [26,28].

In this study, we propose a novel deep learning-based radiomics framework using Faster R-CNN for valve vegetation detection in TEE. To our knowledge, this approach has not been previously evaluated. We hypothesize that it can significantly improve sensitivity and specificity, particularly in challenging cases like prosthetic valve endocarditis, leading to faster diagnosis and better clinical outcomes.

Material and Methods

Study Plan

This retrospective study included TEE images from 287 patients with confirmed IE and 113 control subjects (August 2022–November 2024, Rajaie Center), covering all types of heart valves. IE diagnoses were established based on the modified duke criteria and independently verified by two expert echocardiographers to ensure accurate labeling of valve vegetations.

Patients were included if they had definite valve vegetations according to the modified Duke criteria, high-quality TEE images with visible lesions on any cardiac valve, and complete agreement between two expert echocardiographers. Exclusion criteria included poor-quality or incomplete images, missing clinical data, or the presence of adjacent prosthetic aortic and mitral valves, which could introduce artifact-related diagnostic limitations.

TEE Views: For the assessment of valve vegetation

The following TEE views were used to assess valve vegetations: for the mitral valve, Mid-Esophageal (ME) 2-chamber, ME 4-chamber, and ME long-axis views; for the aortic valve, ME long-axis and ME Aortic Valve (AV) short-axis views; for the tricuspid valve, ME 4-chamber and ME short-axis views; and for the pulmonary valve, the ME Right Ventricular (RV) inflow–outflow view. These views were selected to detect vegetations on both native and prosthetic valves. Figure 1 shows representative examples, with vegetations indicated by red arrows.

Selecting and Extracting Frames from the TEE Series

For each TEE view, the frame, in which the vegetation appeared largest and clearest, was selected, regardless of cardiac phase. To prevent data leakage, only one frame per view per patient was used. All images were saved in uncompressed DICOM format using Python and SimpleITK, and only frames confirmed by two expert echocardiographers were included.

Bounding Mask

Each frame was preprocessed to correct artifacts using ImageJ. Vegetation masks were created in 3D Slicer (v5.6), saved in NRRD format, and converted to binary TIFF files using Python. Fan beam masks delineated the valid ultrasound region and were combined with vegetation masks to remove irrelevant areas. Bounding boxes, defined as the smallest rectangles surrounding non-zero regions, were generated for model training (Figure 2). Frames without vegetation were left unboxed.

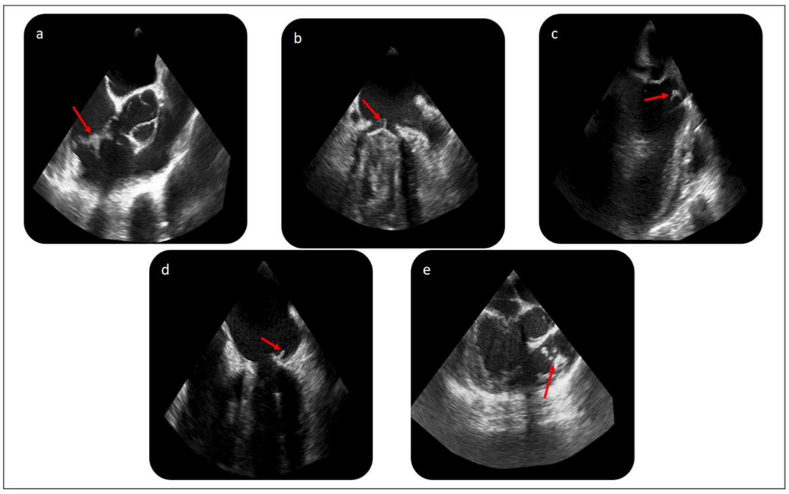

Figure 1: Representative Transesophageal Echocardiography (TEE) views with vegetations (red arrows). (**a**) Short-axis view of native tricuspid and aortic valves, (**b**) Mid-Esophageal (ME) 2-chamber view of a mechanical mitral valve, (**c**) ME long-axis view of an aortic prosthesis and native mitral valve, (**d**) ME 4-chamber view of the native mitral valve, (**e**) ME Right Ventricular (RV) inflow-outflow view of the native pulmonary valve.

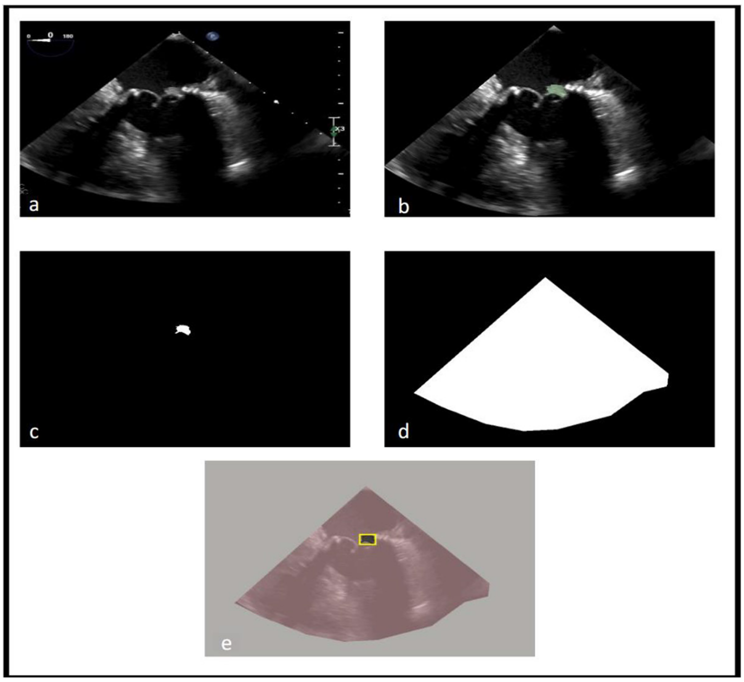

Figure 2: Visualization of mask generation and preprocessing steps. (**a**) Original 4-chamber Transesophageal Echocardiography (TEE) frame with a bioprosthetic mitral valve, (**b**) Manual vegetation segmentation using 3D Slicer, (**c**) Extracted binary vegetation mask, (**d**) Binary fan-beam mask indicating the imaging field, (**e**) Overlay of masks and ground truth bounding box (yellow) on the original frame.

Fan beam masks restricted input to the relevant region, while empty labels trained the model to ignore frames without lesions.

Data Preprocessing

Normalization

The selected TEE frames were preprocessed for classification by converting them to grayscale, applying a bilateral filter to reduce noise while preserving edges, normalizing pixel values to the range [0,1], and resizing to 1024×1024 pixels to optimize computational efficiency and segmentation accuracy [29].

Data Augmentation

To enhance dataset diversity and improve model robustness, each image was augmented fivefold using carefully selected transformations [30, 31]. These transformations included elastic deformation to simulate probe-induced shape variations during acquisition, speckle distortion to introduce the grainy texture produced by sound-wave interference, reverberation simulation to mimic artifacts arising from reflective tissue interfaces, shadow simulation to reproduce acoustic shadows caused by dense or bony structures, and multiplicative noise to simulate device-related intensity variations.

Data augmentation was performed using Albumentations, applied only to the training set, and consistently applied to both images and corresponding masks. The total number of images was calculated as:

Total Images=Original Images×(Number of Augmentations+1)

Dataset Splitting

We employed Scikit-Learn's stratified splitting to ensure balanced representation across the five TEE views. Because all images contained vegetation masks, stratification by lesion presence was unnecessary. The dataset was divided into training (70%), validation (15%), and test (15%) sets, maintaining consistent view distribution to minimize bias.

Deep Learning Model Design

To evaluate the influence of backbone architecture and training strategy, we implemented four Faster R-CNN variants, all pretrained on RadImageNet. The models differed in backbone, comparing ResNet50 with DenseNet121, and in training mode, using either frozen or fine-tuned parameters. This resulted in four configurations: ResNet50 (frozen), ResNet50 (fine-tuned), DenseNet121 (frozen), and DenseNet121 (fine-tuned). All models were trained on the same dataset under identical conditions to ensure a fair comparison.

Model Architecture

We employed Faster R-CNN with a Feature Pyramid Network (FPN) as the core detection model. Binary masks were converted into bounding boxes by enclosing all non-zero pixels, making them suitable for object detection training. The detection head included a Region Proposal Network (RPN), a classifier, and a bounding box regressor, enabling accurate localization of lesions across multiple scales.

Training Protocol

Training employed early stopping, terminating if the validation mAP did not improve for seven consecutive epochs (patience=7).

Stochastic Gradient Descent (SGD) was used with a learning rate of 0.005, momentum of 0.9, and weight decay of 0.0001, based on prior studies and preliminary tuning. Batch sizes were adjusted per model to balance GPU utilization and training stability: 16 for fine-tuned DenseNet121, 20 for fine-tuned ResNet50, 24 for frozen DenseNet121, and 32 for frozen ResNet50.

Evaluation Metrics

Only the original test images (without augmentation) were used to ensure a fair and realistic evaluation.

Two sets of metrics were employed to assess both localization and classification performance. For bounding box detection, we used mean average precision at an IoU threshold of 0.5 (mAP@0.5) [32] and Intersection over Union (IoU) [33], as defined in Equations (1) and (2), respectively. To evaluate the model's ability to discriminate between vegetation-positive and vegetation-negative frames (i.e., image-level classification), we calculated the F1 score and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC).

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \qquad \text{(Eq. 1)}$$

$$mAP@0.5 = \frac{1}{N}\sum_{i=1}^{N}\text{APi} \qquad \text{(Eq. 2)}$$

Workflow

The workflow comprised data preparation followed by Faster R-CNN training using either ResNet50 or DenseNet121 backbones. Model performance was then evaluated on an independent test set using the defined metrics and ROC analysis (Figure 3).

Results

Patient characteristics

Among the 286 reviewed cases, 196 patients were included in the radiomics analysis, 100 with native valves and 96 with prosthetic valves, to ensure balanced groups and consistent image quality. The mean age of the included patients was 55±23.8 years. A total of 227 valves with confirmed vegetation were analyzed as individual samples. In some patients, multiple valves were affected; each involved valve was treated as a separate sample to allow valve-level analysis. Table 1 summarizes vegetation sizes and sample counts by valve type.

In some cases, a single TEE view captured vegetations on two valves simultaneously, resulting in duplicate frame representations. To eliminate redundancy, 27 overlapping frames were excluded, 25 from ME long-axis views (mitral and aortic) and 2 from ME 4-chamber views (mitral and tricuspid). This process yielded 500 unique vegetation-positive frames

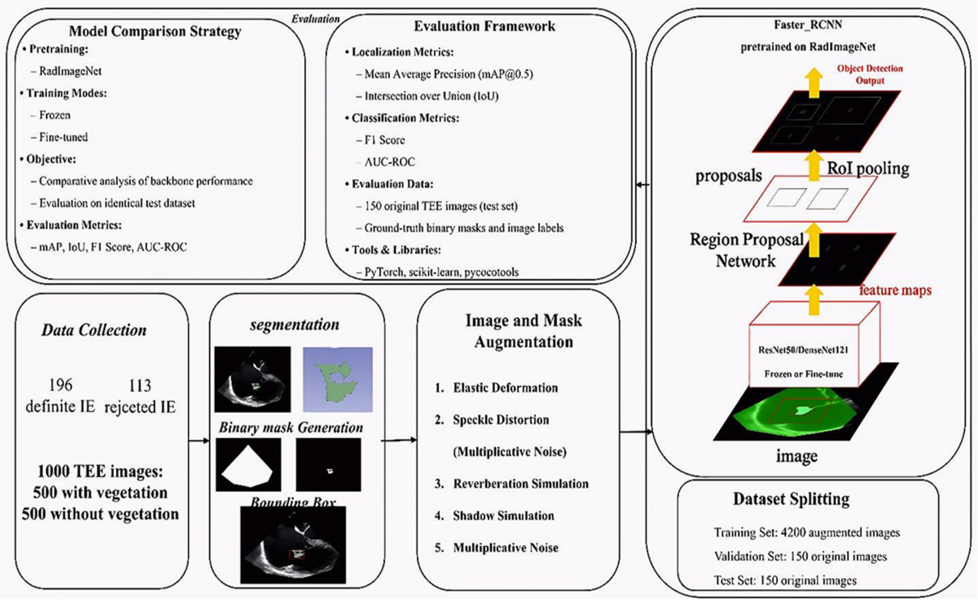

Figure 3: Workflow from data collection to evaluation using Faster Region-based Convolutional Neural Network (Faster R-CNN) with RadImageNet-pretrained backbones. (AUC-ROC: Area Under the Receiver Operating Characteristic Curve, TEE: Transesophageal Echocardiography, IE: Infective Endocarditis)

across standard TEE views. An equal number of negative frames were selected from controls, maintaining a matched view distribution. The final dataset consisted of 1,000 balanced frames. View-specific distributions are summarized in Table 2.

Model Training and Early Stopping Behavior

To prevent overfitting, early stopping based on validation mAP was implemented. As shown in Figure 4, the frozen ResNet50 plateaued early at approximately 0.470, whereas the unfrozen ResNet50 reached 0.644. The frozen DenseNet121 achieved 0.496, while the unfrozen DenseNet121 attained the highest mAP of 0.686. Although the unfrozen DenseNet121 required more training epochs, it delivered the best overall performance.

Quantitative Evaluation of Detection Performance

Models were evaluated on a held-out test set using F1 Score, mAP, and Intersection over Union (IoU, mean±SD). As shown in Table 3, unfrozen backbones outperformed

Table 1: Distribution of valve types, number of samples, and vegetation size measurements in all patients

Type of valves	Number of samples	Vegetation size (mm) (Mean±SD)
Native mitral	50	24.67±9.36
Mechanical mitral	38	11.57±7.58
Bioprosthetic mitral	5	11.70±7.14
Native aortic	50	16.04±5.37
Mechanical aortic	31	8.85±6.67
Bioprosthetic aortic	2	22.50±9.19
Native tricuspid	25	20.85±11.89
Mechanical tricuspid	4	6.08±3.54
Bioprosthetic tricuspid	2	8.00±4.24
Native pulmonary	10	18.17±7.39
Mechanical pulmonary	5	13.70±3.77
Bioprosthetic pulmonary	5	13.56±4.34

Table 2: Distribution of vegetation-positive frames across the five standard Transesophageal Echocardiography (TEE) views used for valve assessment. Each view contributed to the dataset based on its relevance to specific cardiac valves

TEE View	Valve(s) Assessed	Number of Frames
ME 2-Chamber	Mitral	93
ME 4-Chamber	Mitral or Tricuspid	122
ME Long-Axis	Aortic or Mitral but of them	151
ME AV Short-Axis	Aortic or Tricuspid	114
ME RV Inflow-Outflow	Pulmonary	20
Total		500

TEE: Transesophageal Echocardiography, ME: Mid-Esophageal, AV: Aortic Valve, RV: Right Ventricular

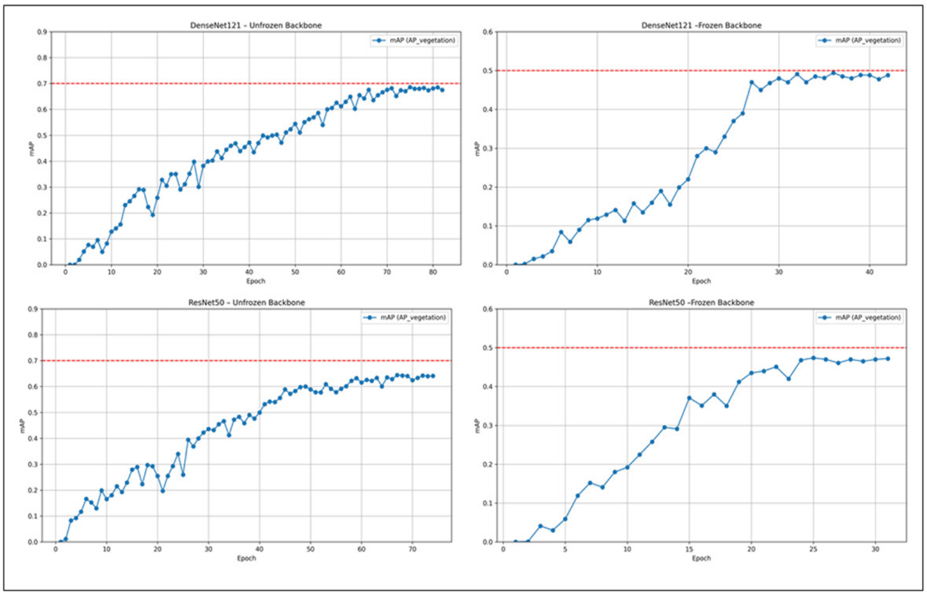

Figure 4: Validation mean Average Precision (mean AP) curves across training epochs for four model configurations: ResNet50 and DenseNet121 in frozen and unfrozen modes.

Table 3: Quantitative Results – Final F1 Score, mean Average Precision (mAP), and Intersection over Union (IoU) on the test set for four model configurations.

	DenseNet121_ Frozen	DenseNet121_ Unfrozen	ResNet50_ Frozen	Resnet50_ Unfrozen
F1 Score	0.431	0.664	0.4222	0.607
mAP (vegetation)	0.416	0.653	0.403	0.593
IoU (mean±std)	0.282±0.120	0.532±0.281	0.267±0.104	0.501±0.272

frozen counterparts, with the unfrozen DenseNet121 achieving the highest overall performance. Fine-tuning enhanced feature extraction and improved vegetation detection.

Image-level Classification

Figure 5 shows the ROC curves for all evaluated models, based on classification of test set images as positive (containing at least one vegetation bounding box) or negative. The unfrozen DenseNet121 achieved the highest AUC (0.858; 95% CI: 0.811–0.889), outperforming all other configurations. Its frozen counterpart showed reduced performance (AUC: 0.640; 95% CI: 0.570–0.680). The fine-tuned ResNet50 attained an AUC of 0.789 (95% CI: 0.761–0.816), whereas the frozen ResNet50 performed the lowest (AUC: 0.601; 95% CI: 0.500–0.650). Confidence intervals were estimated via non-parametric bootstrapping with 1,000 iterations to ensure robust evaluation.

Statistical Comparison of Model Performance

Only the unfrozen models were subjected to statistical comparison, as the frozen configurations consistently underperformed and were excluded to avoid misleading conclusions. Given the limited test set size (n=150), non-parametric analyses were applied. IoU

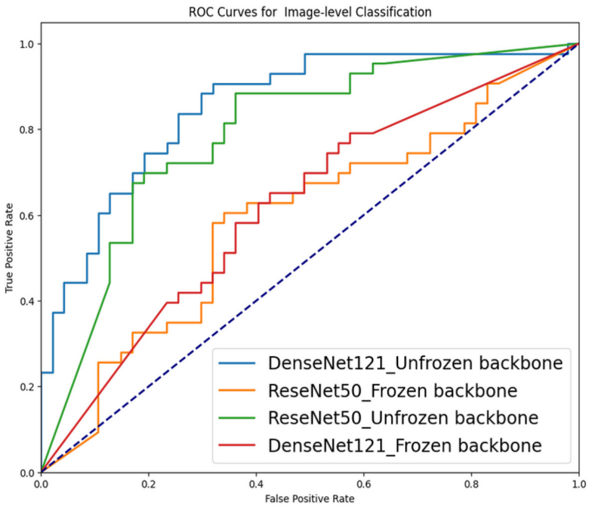

Figure 5: Receiver Operating Characteristic (ROC) curves for image-level classification on the test set.

was calculated on 75 positive test images and compared between unfrozen DenseNet121 and ResNet50 using the Wilcoxon signed-rank test. Although DenseNet121 showed a slightly higher IoU (0.532 vs. 0.501), the difference was not statistically significant (*P*-value=0.053), indicating comparable performance in vegetation localization (Figure 6).

In contrast, DenseNet121 (unfrozen) achieved a significantly higher AUC (0.858; 95% CI: 0.811–0.889) than ResNet50 (unfrozen) (0.789; 95% CI: 0.761–0.816), with minimal confidence interval overlap (~9%), reflecting superior classification performance. Overall, while both models performed similarly in localizing vegetation via bounding boxes, DenseNet121 clearly outperformed ResNet50 in classifying the presence of vegetation.

Discussion

Although TEE is widely used for diagnosing infective endocarditis, its accuracy heavily depends on operator skill and interpretation [34]. Central to clinical guidelines, TEE faces particular challenges in PVE, where acoustic shadowing can obscure key anatomical details, potentially leading to false positives or negatives. Moreover, interpretation remains subjective, contributing to inconsistent diagnoses. Factors, such as patient anatomy and suboptimal acoustic, windows can reduce image quality [35]. Moreover, manual analysis of TEE frames is time-consuming and error-prone, especially under high workloads. These challenges have motivated increasing exploration of AI-based methods to improve diagnostic consistency and detect subtle patterns, particularly in low-quality or artifact-prone images, a critical need in PVE.

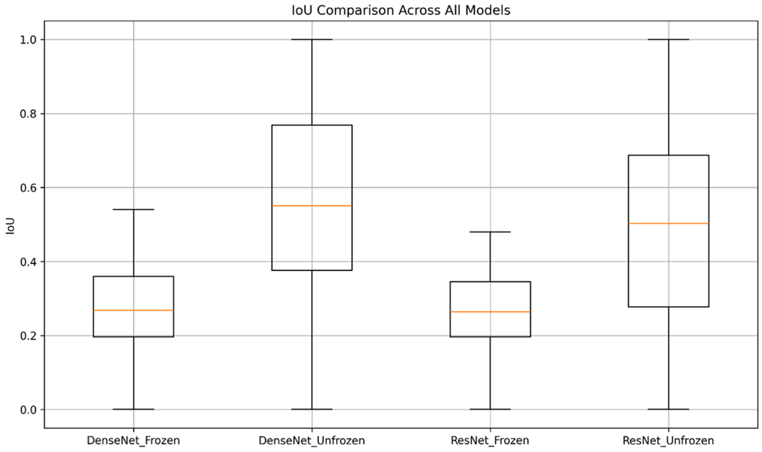

Figure 6: Boxplot of Intersection over Union (IoU) scores for four model configurations. Unfrozen DenseNet121 showed the highest median IoU, but differences between unfrozen models were not statistically significant.

In this study, we leveraged RadImageNet-pretrained ResNet50 and DenseNet121 backbones to enhance performance given our limited dataset (n=1,000). Unlike ImageNet, RadImageNet offers domain-specific features better suited for grayscale medical images, including ultrasound modalities such as TEE [36]. While newer architectures, such as Swin Transformer and EfficientNet, have gained attention, we found CNN-based backbones more appropriate for our scenario. EfficientNet, designed primarily for classification and pretrained on general-purpose ImageNet images, lacks segmentation capabilities and medical imaging pretraining [37]. Swin Transformer can perform well on smaller datasets, but studies indicate that transformers generally require larger datasets to fully realize their advantages [38]. Thus, a CNN-based approach with domain-specific pretrained weights, such as RadImageNet, provided a more stable and practical solution [36].

Our results demonstrate that fine-tuned RadImageNet-pretrained models, particularly DenseNet121, improve both detection and classification of valve vegetations in TEE images. DenseNet121 outperformed ResNet50 in classification (AUC: 0.858 vs. 0.789) with minimal confidence interval overlap (~9%), likely due to its feature-reuse architecture. Localization performance was slightly higher for DenseNet121 (IoU: 0.532 vs. 0.501), though this difference was not statistically significant (P-value=0.053). These findings highlight the effectiveness of fine-tuned RadImageNet-based CNN backbones for automated TEE analysis and enhanced diagnostic consistency in infective endocarditis.

Although RadImageNet includes multiple medical modalities, it does not contain echocardiographic data, particularly TEE. This likely explains why unfrozen models, capable of adapting more fully to TEE-specific patterns, performed better in both detection and classification [36]. While both ResNet50 and DenseNet121 showed similar mAP for lesion localization, DenseNet121 achieved a higher classification AUC. In our study, ResNet50 could detect vegetation regions adequately but lacked the deeper feature representation and cross-layer connectivity of DenseNet121 [39], which appear crucial for confident classification in echocardiographic images.

Comparisons with prior studies support our findings. While Zhu et al. [40] reported ResNet50 superiority in classifying DCIS on ultrasound, our TEE-based study favored DenseNet121, likely due to end-to-end fine-tuning and enhanced spatial modeling for low-contrast, complex echocardiographic features. He et al. [41] found Swin Transformer outperformed CNNs on ImageNet-pretrained models; however, no publicly available Swin Transformer pretrained specifically on medical imaging exists, and domain-specific pretraining often improves performance. Notably, studies in thyroid nodules [42] and oral cancer detection [43] also reported DenseNet121 outperforming ResNet50, consistent with our findings.

DenseNet121's advantages likely stem from improved gradient flow, feature reuse, and reduced parameter redundancy due to its short skip connections [44,45]. These characteristics are particularly beneficial for detecting small, low-contrast vegetations in TEE images [46,47]. While computationally demanding, DenseNet121's architecture favors low-contrast, fine-textured images, and our dual-task setup (detection and classification) further highlights its strengths. The small, irregular shape of vegetations and limited dataset size may have additionally favored DenseNet's stability [45,48].

Accordingly, this study represents the first application of deep learning-based object detection for localization and classification of valve vegetations in TEE. By leveraging domain-specific pretraining and comparing CNN architectures, we establish a performance benchmark and demonstrate DenseNet121's suitability for this complex imaging task.

Although histology remains the gold standard for diagnosing infective endocarditis, it is not always available; we relied on the modified duke criteria, which report over 90% specificity for definite cases [5,49]. Annotation was restricted to frames with clearly visible vegetations, independently verified by two fellowship-trained echocardiographers in full agreement, ensuring high-confidence labels for robust model training and evaluation.

Limitations include a relatively small number of labeled cases, preventing separate models for each valve type. Consequently, all valves were combined into a single model, which may obscure valve-specific features, though expert-validated segmentations and similar echogenicity likely minimized performance loss. Future studies should include external validation across multiple centers to assess generalizability, expand datasets, particularly for prosthetic and mechanical valves, and explore advanced architectures, such as Swin Transformers to further enhance performance.

Conclusion

Our results support the use of fine-tuned DenseNet121 for TEE image analysis. This model demonstrated strong performance in detecting and localizing vegetations and may also be applicable to identifying other lesion types in similar ultrasound data. DenseNet121 achieved a high AUC in classifying images with vegetation, highlighting its potential to assist clinical diagnosis. When used alongside the echocardiographer's report, it could help flag suspicious cases, reduce missed findings, and improve diagnostic confidence and accuracy in real-world settings.

Authors' Contribution

F. Esmaely was primarily responsible for study design, TEE image preprocessing, deep learning model implementation, evaluation, and manuscript drafting. SM. Zamani-Aliabadi contributed to data preparation, assisted with model training and performance analysis, and supported manuscript formatting. A. Bitarafan-Rajabi supervised the study, guided the methodology, and critically revised the manuscript. P. Moradnejad contributed to the clinical validation of cases, participated in patient selection, and revised the manuscript. Sh. Boudagh contributed to the interpretation of echocardiographic data and provided overall project supervision. All authors read, revised, and approved the final version of the manuscript.

Ethical Approval

The research was approved by the Ethics Committee of Iran University, with the approval ID [IR.IUMS.FMD.REC.1401.285].

Informed Consent

This study was conducted using anonymized retrospective data, and informed consent was waived by the Ethics Committee.

Funding

This study was supported by the Iran University of Medical Sciences, whose financial support is sincerely appreciated.

Conflict of Interest

None

References

1. Galeone A, Gardellini J, Di Nicola V, Perrone F, Menzione MS, Di Gaetano R, Luciani GB. Early and Long-Term Outcomes of Patients Undergoing Surgery for Native and Prosthetic Valve Endocarditis: The Role of Preoperative Neutrophil-to-Lymphocyte Ratio, Neutrophil-to-Platelet Ratio, and Monocyte-to Lymphocyte Ratio. *J Clin Med.* 2025;**14**(2):533. doi: 10.3390/jcm14020533. PubMed PMID: 39860541. PubMed PMCID: PMC11766379.
2. Fukuchi T, Iwata K, Ohji G. Failure of early diagnosis of infective endocarditis in Japan--a retrospective descriptive analysis. *Medicine (Baltimore).* 2014;**93**(27):e237. doi: 10.1097/MD.0000000000000237. PubMed PMID: 25501088. PubMed PMCID: PMC4602777.
3. Abadie BQ, Wang TKM. Contemporary Multi-mo-

dality Imaging of Prosthetic Aortic Valves. *Rev Cardiovasc Med.* 2025;**26**(1):25339. doi: 10.31083/RCM25339. PubMed PMID: 39867176. PubMed PMCID: PMC11759978.

4. Li JS, Sexton DJ, Mick N, Nettles R, Fowler VG Jr, Ryan T, et al. Proposed modifications to the Duke criteria for the diagnosis of infective endocarditis. *Clin Infect Dis.* 2000;**30**(4):633-8. doi: 10.1086/313753. PubMed PMID: 10770721.
5. d'Almeida S, Reischmann K, Andreß S, Felbel D, Stephan T, Hay B, et al. Evaluating the Duke Criteria for infectious endocarditis in a single-center retrospective study. *Sci Rep.* 2024;**14**(1):19524. doi: 10.1038/s41598-024-70196-x. PubMed PMID: 39174590. PubMed PMCID: PMC11341858.
6. Carević V, Mladenović Z, Perković-Avelini R, Bečić T, Radić M, Fabijanić D. Three-Dimensional Transesophageal Echocardiography in the Diagnosis and Treatment of Mitral Prosthetic Valve Endocarditis-A Narrative Review. *Medicina (Kaunas)*. 2021;**58**(1):23. doi: 10.3390/medicina58010023. PubMed PMID: 35056331. PubMed PMCID: PMC8779064.
7. Ten Hove D, Slart RHJA, Glaudemans AWJM, Postma DF, Gomes A, Swart LE, et al. Using machine learning to improve the diagnostic accuracy of the modified Duke/ESC 2015 criteria in patients with suspected prosthetic valve endocarditis - a proof of concept study. *Eur J Nucl Med Mol Imaging.* 2024;**51**(13):3924-33. doi: 10.1007/s00259-024-06774-y. PubMed PMID: 38904778. PubMed PMCID: PMC11527948.
8. Pamnani A, Skubas NJ. Imaging artifacts during transesophageal echocardiography. *Anesth Analg.* 2014;**118**(3):516-20. doi: 10.1213/ANE.0000000000000084. PubMed PMID: 24557099.
9. Habib G, Lancellotti P, Antunes MJ, Bongiorni MG, Casalta JP, Del Zotti F, et al. 2015 ESC Guidelines for the management of infective endocarditis: The Task Force for the Management of Infective Endocarditis of the European Society of Cardiology (ESC). Endorsed by: European Association for Cardio-Thoracic Surgery (EACTS), the European Association of Nuclear Medicine (EANM). *Eur Heart J.* 2015;**36**(44):3075-128. doi: 10.1093/eurheartj/ehv319. PubMed PMID: 26320109.
10. Saby L, Laas O, Habib G, Cammilleri S, Mancini J, Tessonnier L, et al. Positron emission tomography/computed tomography for diagnosis of prosthetic valve endocarditis: increased valvular 18F-fluorodeoxyglucose uptake as a novel major criterion. *J Am Coll Cardiol.* 2013;**61**(23):2374-82. doi: 10.1016/j.jacc.2013.01.092. PubMed PMID: 23583251.
11. Stamate E, Piraianu AI, Ciobotaru OR, Crassas R, Duca O, Fulga A, et al. Revolutionizing Cardiology through Artificial Intelligence-Big Data from Proactive Prevention to Precise Diagnostics and Cutting-Edge Treatment-A Comprehensive Review of the Past 5 Years. *Diagnostics (Basel).* 2024;**14**(11):1103. doi: 10.3390/diagnostics14111103. PubMed PMID: 38893630. PubMed PMCID: PMC11172021.
12. Delgado V, Ajmone Marsan N, De Waha S, Bonaros N, Brida M, Burri H, et al. 2023 ESC guidelines for the management of endocarditis: developed by the task force on the management of endocarditis of the European Society of Cardiology (ESC) endorsed by the European Association for Cardio-Thoracic Surgery (EACTS) and the European Association of Nuclear Medicine (EANM). *Eur Heart J.* 2023;**44**(39):3948-4042. doi:10.1093/eurheartj/ehad193. PubMed PMID: 37622656.
13. Sammartino AM, Bonfioli GB, Dondi F, Riccardi M, Bertagna F, Metra M, Vizzardi E. Contemporary Role of Positron Emission Tomography (PET) in Endocarditis: A Narrative Review. *J Clin Med.* 2024;**13**(14):4124. doi: 10.3390/jcm13144124. PubMed PMID: 39064164. PubMed PMCID: PMC11277723.
14. Pizzi MN, Roque A, Fernández-Hidalgo N, Cuéllar-Calabria H, Ferreira-González I, Gonzàlez-Alujas MT, et al. Improving the Diagnosis of Infective Endocarditis in Prosthetic Valves and Intracardiac Devices With 18F-Fluordeoxyglucose Positron Emission Tomography/Computed Tomography Angiography: Initial Results at an Infective Endocarditis Referral Center. *Circulation.* 2015;**132**(12):1113-26. doi: 10.1161/CIRCULATIONAHA.115.015316. PubMed PMID: 26276890.
15. Papadimitriou-Olivgeris M, Ledergerber B, Siedentop B, Monney P, Frank M, Tzimas G, et al. Beyond the Timeline: 1-Year Mortality Trends in Early Versus Late Prosthetic Valve Endocarditis. *Clin Infect Dis.* 2025;**80**(4):804-6. doi: 10.1093/cid/ciae392. PubMed PMID: 39067056. PubMed PMCID: PMC12043058.
16. Kusunose K. Transforming Echocardiography: The Role of Artificial Intelligence in Enhancing Diagnostic Accuracy and Accessibility. *Intern Med.* 2025;**64**(3):331-6. doi: 10.2169/internalmedicine.4171-24. PubMed PMID: 39048361. PubMed PMCID: PMC11867741.
17. Kusunose K. Radiomics in Echocardiography: Deep Learning and Echocardiographic Analysis. *Curr Cardiol Rep.* 2020;**22**(9):89. doi: 10.1007/s11886-020-01348-4. PubMed PMID: 32648059.

18. Tsampras T, Karamanidou T, Papanastasiou G, Stavropoulos TG. Deep learning for cardiac imaging: focus on myocardial diseases, a narrative review. *Hellenic J Cardiol.* 2025;**81**:18-24. doi: 10.1016/j.hjc.2024.12.002. PubMed PMID: 39662734.
19. Baeßler B, Engelhardt S, Hekalo A, Hennemuth A, Hüllebrand M, Laube A, et al. Perfect Match: Radiomics and Artificial Intelligence in Cardiac Imaging. *Circ Cardiovasc Imaging.* 2024;**17**(6):e015490. doi: 10.1161/CIRCIMAGING.123.015490. PubMed PMID: 38889216.
20. Chen J, Liu Y, Wei S, Bian Z, Subramanian S, Carass A, et al. A survey on deep learning in medical image registration: New technologies, uncertainty, evaluation metrics, and beyond. *Med Image Anal.* 2025;**100**:103385. doi: 10.1016/j.media.2024.103385. PubMed PMID: 39612808. PubMed PMCID: PMC11730935.
21. Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. *BMC Med Imaging.* 2022;**22**(1):69. doi: 10.1186/s12880-022-00793-7. PubMed PMID: 35418051. PubMed PMCID: PMC9007400.
22. Li Z, Zheng Y, Shan D, Yang S, Li Q, Wang B, et al. ScribFormer: Transformer Makes CNN Work Better for Scribble-Based Medical Image Segmentation. *IEEE Trans Med Imaging.* 2024;**43**(6):2254-65. doi: 10.1109/TMI.2024.3363190. PubMed PMID: 38324425.
23. Afifi S, Kaur R, GholamHosseini H, Monsef A, Ullah E, Baig MM, Salama Abdelhady A. Detection of Metastatic Tissues in Histopathologic Images using DenseNet-121 with Data Augmentation. *Annu Int Conf IEEE Eng Med Biol Soc.* 2025;**2025**:1-7. doi: 10.1109/EMBC58623.2025.11251669. PubMed PMID: 41336055.
24. Xu W, Fu YL, Zhu D. ResNet and its application to medical image processing: Research progress and challenges. *Comput Methods Programs Biomed.* 2023;**240**:107660. doi: 10.1016/j.cmpb.2023.107660. PubMed PMID: 37320940.
25. Alruwaili M, Mohamed M. An Integrated Deep Learning Model with EfficientNet and ResNet for Accurate Multi-Class Skin Disease Classification. *Diagnostics (Basel).* 2025;**15**(5):551. doi: 10.3390/diagnostics15050551. PubMed PMID: 40075797. PubMed PMCID: PMC11898587.
26. Cao Z, Duan L, Yang G, Yue T, Chen Q. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. *BMC Med Imaging.* 2019;**19**(1):51. doi: 10.1186/s12880-019-0349-x. PubMed PMID: 31262255. PubMed PMCID: PMC6604293.
27. Nurmaini S, Sapitri AI, Tutuko B, Rachmatullah MN, Rini DP, Darmawahyuni A, et al. Automatic echocardiographic anomalies interpretation using a stacked residual-dense network model. *BMC Bioinformatics.* 2023;**24**(1):365. doi: 10.1186/s12859-023-05493-9. PubMed PMID: 37759158. PubMed PMCID: PMC10536702.
28. DYap MH, Goyal M, Osman F, Martí R, Denton E, Juette A, Zwiggelaar R. Breast ultrasound region of interest detection and lesion localisation. *Artif Intell Med.* 2020;**107**:101880. doi: 10.1016/j.artmed.2020.101880. PubMed PMID: 32828439.
29. Zhang L, Wang X, Yang D, Sanford T, Harmon S, Turkbey B, et al. Generalizing Deep Learning for Medical Image Segmentation to Unseen Domains via Deep Stacked Transformation. *IEEE Trans Med Imaging*. 2020;**39**(7):2531-40. doi: 10.1109/TMI.2020.2973595. PubMed PMID: 32070947. PubMed PMCID: PMC7393676.
30. Snider EJ, Hernandez-Torres SI, Hennessey R. Using Ultrasound Image Augmentation and Ensemble Predictions to Prevent Machine-Learning Model Overfitting. *Diagnostics (Basel).* 2023;**13**(3):417. doi: 10.3390/diagnostics13030417. PubMed PMID: 36766522. PubMed PMCID: PMC9914871.
31. Chlap P, Min H, Vandenberg N, Dowling J, Holloway L, Haworth A. A review of medical image data augmentation techniques for deep learning applications. *J Med Imaging Radiat Oncol.* 2021;**65**(5):545-63. doi: 10.1111/1754-9485.13261. PubMed PMID: 34145766.
32. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft COCO: Common Objects in Context. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. Computer Vision – ECCV 2014. Lecture Notes in Computer Science. Cham: Springer; 2014. p. 740-55.
33. Everingham M, Eslami SA, Van Gool L, Williams CK, Winn J, Zisserman A. The pascal visual object classes challenge: A retrospective. I*nt J Comput Vis.* 2015;**111**(1):98-136. doi: 10.1007/s11263-014-0733-5.
34. Moon KB, Tattersall MC, Adoe M, Osman F, Rahko PS. Evaluation of inter-observer variability regarding aortic and mitral valve findings on transesophageal echocardiograms ordered for suspected endocarditis. *Echocardiography.* 2022;**39**(7):906-17. doi: 10.1111/echo.15400. PubMed PMID: 35733298. PubMed PMCID: PMC9541542.
35. Lau V, Priestap F, Landry Y, Ball I, Arntfield R. Diagnostic Accuracy of Critical Care Transesophageal

Echocardiography vs Cardiology-Led Echocardiography in ICU Patients. *Chest.* 2019;**155**(3):491-501. doi: 10.1016/j.chest.2018.11.025. PubMed. PMID: 30543807.

36. Mei X, Liu Z, Robson PM, Marinelli B, Huang M, Doshi A, et al. RadImageNet: An Open Radiologic Deep Learning Research Dataset for Effective Transfer Learning. *Radiol Artif Intell.* 2022;**4**(5):e210315. doi: 10.1148/ryai.210315. PubMed PMID: 36204533. PubMed PMCID: PMC9530758.
37. Boice EN, Hernandez-Torres SI, Snider EJ. Comparison of Ultrasound Image Classifier Deep Learning Algorithms for Shrapnel Detection. *J Imaging.* 2022;**8**(5):140. doi: 10.3390/jimaging8050140. PubMed PMID: 35621904. PubMed PMCID: PMC9144026.
38. Takahashi S, Sakaguchi Y, Kouno N, Takasawa K, Ishizu K, Akagi Y, et al. Comparison of Vision Transformers and Convolutional Neural Networks in Medical Image Analysis: A Systematic Review. *J Med Syst.* 2024;**48**(1):84. doi: 10.1007/s10916-024-02105-8. PubMed PMID: 39264388. PubMed PMCID: PMC11393140.
39. Huang G, Liu Z, Pleiss G, Maaten LV, Weinberger KQ. Convolutional Networks with Dense Connectivity. IEEE Trans Pattern Anal Mach Intell. 2022;44(12):8704-16. doi: 10.1109/TPAMI.2019.2918284. PubMed PMID: 31135351.
40. Zhu M, Kuang Y, Jiang Z, Liu J, Zhang H, Zhao H, et al. Ultrasound deep learning radiomics and clinical machine learning models to predict low nuclear grade, ER, PR, and HER2 receptor status in pure ductal carcinoma in situ. *Gland Surg.* 2024;**13**(4):512-27. doi: 10.21037/gs-23-417. PubMed PMID: 38720675. PubMed PMCID: PMC11074652.
41. He X, Bai XH, Chen H, Feng WW. Machine learning models in evaluating the malignancy risk of ovarian tumors: a comparative study. *J Ovarian Res.* 2024;**17**(1):219. doi: 10.1186/s13048-024-01544-8. PubMed PMID: 39506832. PubMed PMCID: PMC11539702.
42. Kim J, Kim MH, Lim DJ, Lee H, Lee JJ, Kwon HS, et al. Deep Learning Technology for Classification of Thyroid Nodules Using Multi-View Ultrasound Images: Potential Benefits and Challenges in Clinical Application. *Endocrinol Metab (Seoul).* 2025;**40**(2):216-24. doi: 10.3803/EnM.2024.2058. PubMed PMID: 39805576. PubMed PMCID: PMC12061742.
43. Okazaki S, Mine Y, Yoshimi Y, Iwamoto Y, Ito S, Peng TY, et al. RadImageNet and ImageNet as Datasets for Transfer Learning in the Assessment of Dental Radiographs: A Comparative Study. *J Imaging Inform Med.* 2025;**38**(1):534-44. doi: 10.1007/s10278-024-01204-9. PubMed PMID: 39048809. PubMed PMCID: PMC11811346.
44. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Honolulu, HI: IEEE; 2017. p. 4700-8.
45. Zhang YD, Satapathy SC, Zhang X, Wang SH. COVID-19 Diagnosis via DenseNet and Optimization of Transfer Learning Setting. *Cognit Comput.* 2021:1-17. doi: 10.1007/s12559-020-09776-8. PubMed PMID: 33488837. PubMed PMCID: PMC7812362.
46. Hou Y, Wu Z, Cai X, Zhu T. The application of improved densenet algorithm in accurate image recognition. *Sci Rep.* 2024;**14**(1):8645. doi: 10.1038/s41598-024-58421-z. PubMed PMID: 38622153. PubMed PMCID: PMC11018628.
47. Senior R, Becher H, Monaghan M, Agati L, Zamorano J, Vanoverschelde JL, et al. Clinical practice of contrast echocardiography: recommendation by the European Association of Cardiovascular Imaging (EACVI) 2017. *Eur Heart J Cardiovasc Imaging.* 2017;**18**(11):1205-1205af. doi: 10.1093/ehjci/jex182. PubMed PMID: 28950366.
48. Okonta KE, Adamu YB. What size of vegetation is an indication for surgery in endocarditis? *Interact Cardiovasc Thorac Surg.* 2012;**15**(6):1052-6. doi: 10.1093/icvts/ivs365. PubMed PMID: 22962320. PubMed PMCID: PMC3501298.
49. Philip M, Tessonier L, Mancini J, Mainardi JL, Fernandez-Gerlinger MP, Lussato D, et al. Comparison Between ESC and Duke Criteria for the Diagnosis of Prosthetic Valve Infective Endocarditis. *JACC Cardiovasc Imaging.* 2020;**13**(12):2605-15. doi: 10.1016/j.jcmg.2020.04.011. PubMed PMID: 32563658.