

Ultrasound Image Segmentation in Cardiovascular Disease: A Review of Techniques and Clinical Applications

Zuoyan Li^a, Lianfang Du^b, Fan Li^{c,*}

^a School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai, China; ^b Department of Ultrasound, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; ^c Department of Ultrasound, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Received July 18, 2025; revision requested November 16, 2025; accepted November 24, 2025

Abstract: This review provides a comprehensive overview of echocardiographic image segmentation techniques in representative clinical contexts, including valvular heart disease (VHD), congenital heart disease (CHD), and left ventricular (LV) dysfunction and heart failure. The article first outlines the standardized workflow for ultrasound (US) image segmentation, establishing a technical framework for comparing and evaluating different approaches. This is followed by a brief overview of the pathophysiological characteristics of these conditions, highlighting their anatomical segmentation requirements and clinical implications. Subsequently, the review summarizes the evolution and key contributions of traditional image processing, machine learning (ML), and deep learning (DL) methods across different disease types. Comparative analysis indicates that DL models, with their end-to-end training strategies and hierarchical feature representation, demonstrate substantial advantages in segmentation automation and accuracy. Finally, the article discusses current challenges in data acquisition, model generalizability, and clinical integration, and explores future research directions to promote the development and translational application of intelligent echocardiographic analysis.

Key words: Valvular heart disease; Congenital heart disease; Left ventricular dysfunction; Deep learning; Echocardiographic image segmentation

Advanced Ultrasound in Diagnosis and Therapy 2026; 03: 137-147

DOI: [10.26599/AUDT.2026.250063](https://doi.org/10.26599/AUDT.2026.250063)

Echocardiography (Echo) is a widely used imaging modality in clinical practice for evaluating cardiac structure and function, offering advantages such as non-invasiveness, safety, and repeatability [1]. Diagnostic interpretation based on imaging signs relies on observing cardiac motion and doppler flow characteristics to identify morphological and hemodynamic abnormalities, making it valuable for the diagnosis and treatment evaluation of various heart diseases. However, its clinical application remains limited by several factors, including operator dependence, suboptimal acoustic windows, and subjectivity in interpreting imaging signs [2].

Current medical image processing primarily focuses on image enhancement, segmentation, registration, and computer-aided diagnosis systems (CADx) [3]. Traditional approaches often rely on handcrafted rules or physical models, leading to limited accuracy and generalizability. In contrast, artificial intelligence (AI), particularly deep learning (DL) models, has demonstrated superior performance in feature extraction and pattern recognition from medical images. DL has been increasingly applied to retrospective analyses of Echo data, showing promising potential in structural segmentation [4,5], functional assessment [6,7], disease identification [8,9], and telemedicine applications [10,11], thus offer-

* Corresponding author: Department of Ultrasound, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, 241 West Huaihai Road, Shanghai, 200030, China. (Fan Li)
e-mail: medicinel@sjtu.edu.cn (F L)

ing a new direction for intelligent Echo analysis.

Among these tasks, image segmentation plays a central role in medical image processing and analysis, serving as a fundamental step for anatomical structure delineation and region identification. However, traditional segmentation methods face considerable challenges in real-world scenarios, such as ambiguous boundaries between the LV and myocardium, and significant temporal variations in cardiac morphology [12]. DL models, with their powerful representation learning capabilities, have enabled high-precision segmentation across multiple organs and imaging modalities, significantly improving the automation of image analysis [13]. AI-driven segmentation has been widely explored in Echo, computed tomography (CT), and magnetic resonance imaging (MRI), and has become a research hotspot in the pursuit of intelligent medical image analysis, showing great potential for future applications.

Different cardiovascular diseases impose distinct imaging and segmentation demands [14]. In the case of valvular disease, segmentation of heart valves is essential for assessing regurgitation and stenosis; in heart failure, accurate delineation of the left ventricle is vital for evaluating ejection fraction. By contrast, CHDs demand precise localization of structural anomalies across diverse anatomical variants. To reflect these clinical needs, this review categorizes segmentation approaches based on major disease types, aiming to bridge algorithmic advances with practical applications.

General Workflow of Ultrasound Image Segmentation

Before examining disease-specific applications, a clear understanding of the general workflow of US image segmentation is foundational to subsequent disease-specific applications [15]. As illustrated in figure 1, the general workflow involves three core stages: image preprocessing, region of interest (ROI) localization, and pixel-wise segmentation [16,17].

In the first stage, US images undergo various preprocessing steps aimed at enhancing image quality and improving downstream model performance [11]. Techniques such as filtering, denoising, and binarization are employed to improve image quality [18], while data augmentation methods including rotation, translation,

and scaling are applied to enrich the diversity of the dataset [19].

The second stage involves the identification and localization of the ROI [20]. Traditional methods such as thresholding, edge detection, and K-means clustering [21,22], which are used to generate preliminary boundary information for isolating target areas, thereby preparing for precise segmentation in subsequent steps.

Finally, fine-grained image segmentation is performed to assign each pixel or voxel to a specific anatomical class or structure [23]. These segmentation maps are then used for quantitative analysis, morphological measurement, and feature extraction [24]. In medical imaging, particularly in Echo, accurate segmentation is essential for a wide range of clinical tasks, including the detection of pathological abnormalities, localization of structural defects [25], volumetric assessments, and real-time procedural guidance [26]. As such, segmentation plays a foundational role in enabling intelligent diagnostic tools and facilitating clinical decision-making across various cardiovascular disease settings.

Segmentation for Valvular Heart Disease

Valvular Heart Disease (VHD) is a common category of cardiovascular disorders and represents a leading cause of cardiac morbidity and mortality worldwide [27]. VHD refers to structural or functional abnormalities affecting one or more cardiac valves, either due to primary etiologies such as degenerative, congenital, or rheumatic causes or secondary factors, for example ventricular remodeling, resulting in stenosis or regurgitation [28]. These valvular lesions alter the heart's preload or afterload, leading to compensatory ventricular remodeling and, ultimately, heart failure. Based on anatomical location, VHD can be classified into four major types: mitral valve (MV) disease, tricuspid valve disease, aortic valve disease, and pulmonary valve disease. These conditions may occur independently or coexist in combination [29].

The Echo is considered the gold standard for diagnosing VHD, offering real-time visualization of both valvular structure and function [30]. Precising segmentation of complex valvular components, such as leaflets and annuli, serves as a critical foundation for assessing disease severity and carries significant implications for

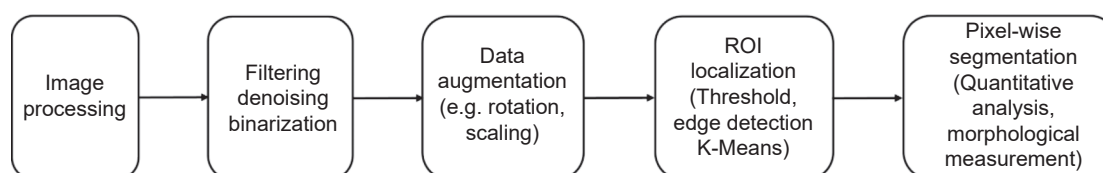


Figure 1 General workflow of ultrasound image segmentation, comprising preprocessing, ROI localization, and pixel-wise segmentation, which supports clinical tasks such as defect detection and quantitative analysis.

clinical decision-making and patient prognosis [31]. Nevertheless, the inherently thin and highly dynamic nature of valve structures, combined with their strong dependency on the orientation of the US transducer, presents substantial technical challenges for achieving accurate and reliable segmentation.

Traditional and machine learning methods in valvular segmentation

Early methods for valvular segmentation primarily relied on predefined mathematical rules and manually designed algorithms, including thresholding, edge detection, and region growing segmentation. These approaches are conceptually straightforward and computationally efficient, offering rapid segmentation. However, their accuracy deteriorates significantly in the presence of speckle noise or when valve boundaries are poorly defined.

To overcome the instability and operator dependence of traditional rule-based segmentation, researchers began to explore machine learning (ML) algorithms that could automatically learn discriminative features from image data rather than relying on rule-based segmentation strategies.

Benefiting from the continuous advancement of ML theories and algorithms, methods such as Support Vector Machines (SVMs) and Random Forests (RFs) have been applied to the automated segmentation of valvular structures. Çomak et al. proposed a wavelet feature-based diagnostic support system using SVM, which automatically identifies abnormalities in valvular color Doppler flow imaging (CDFI) and outperforms conventional neural network (CNN) models in terms of both diagnostic accuracy and efficiency [32]. Avci et al. developed a classification system for valvular CDFI signals in VHD using a genetic algorithm-enhanced SVM, achieving a classification accuracy of up to 95% [33]. Hernandez-Suarez et al. employed multiple ML algorithms, including logistic regression (LR), Naive Bayes (NB), and RF, to predict in-hospital mortality following transcatheter aortic valve replacement (TAVR) in a cohort of 10,883 patients [34]. In another study, Zhang et al. used an SVM classifier to automatically detect key hinge points of the mitral annulus, facilitating the quantitative assessment of mitral regurgitation (MR) [35].

While these ML methods have improved segmentation efficiency and stability to some extent, their performance remains constrained by several factors, such as dependence on handcrafted feature design, limited training data, and suboptimal generalizability across different patient populations. Specifically, handcrafted features such as texture, gradient, or shape descriptors are often tailored to static imaging assumptions and cannot adapt to the rapid valve motion, complex speckle pat-

terns, and variable acoustic quality that characterize Echo. As a result, feature-based models tend to lose discriminative power in low-contrast or dynamically changing regions.

Moreover, the limited size and diversity of training datasets restrict abilities of models to capture inter-patient and inter-device variability, further undermining robustness in multi-center clinical scenarios. These limitations explain why traditional ML approaches, although efficient in controlled experiments, often fail to achieve consistent and clinically reliable segmentation performance in real-world echocardiographic imaging, thereby motivating the shift toward deep learning frameworks capable of learning hierarchical spatiotemporal features directly from raw ultrasound data.

Deep learning methods in valvular segmentation

To address these limitations, researchers have increasingly turned to DL techniques, which offer powerful capabilities for automatic feature extraction and end-to-end model training, providing new solutions for high-precision segmentation of valvular structures.

An increasing number of studies have demonstrated that CNNs and their variants, particularly the U-Net architecture, have exhibited excellent performance in valvular segmentation tasks and have become one of the mainstream methodologies in the field. Vrudhula et al. developed and evaluated a DL-based automated model using 58,614 transthoracic echocardiographic (TTE) images to assess the severity of MR through CDFI in the apical four-chamber view (A4C) [36]. Aaron et al. focused on TTE as well, applying a DL model to process 61,689 complete echocardiographic studies. By incorporating multiple TTE views, they categorized MR severity into four grades, and demonstrated that the multi-view approach achieved a classification accuracy of 99.2%, outperforming the A4C view alone [37]. Wifstad et al. introduced an Attention U-Net with deep supervision and attention weight scheduling for automated mitral valve segmentation and motion tracking in transthoracic echocardiography [38]. The model architecture is illustrated in [figure 2](#).

More recently, the application of Transformer-based models has further enhanced the ability to capture long-range dependencies. An et al. proposed the CA-ISNet, which leverages attention mechanisms to significantly improve the accuracy and robustness of valvular segmentation in fetal and pediatric echocardiographic datasets [39].

These DL approaches have overcome the limitations of traditional and ML-based methods in terms of segmentation accuracy and generalizability and have gradually become the dominant direction in the research of valvular disease segmentation.

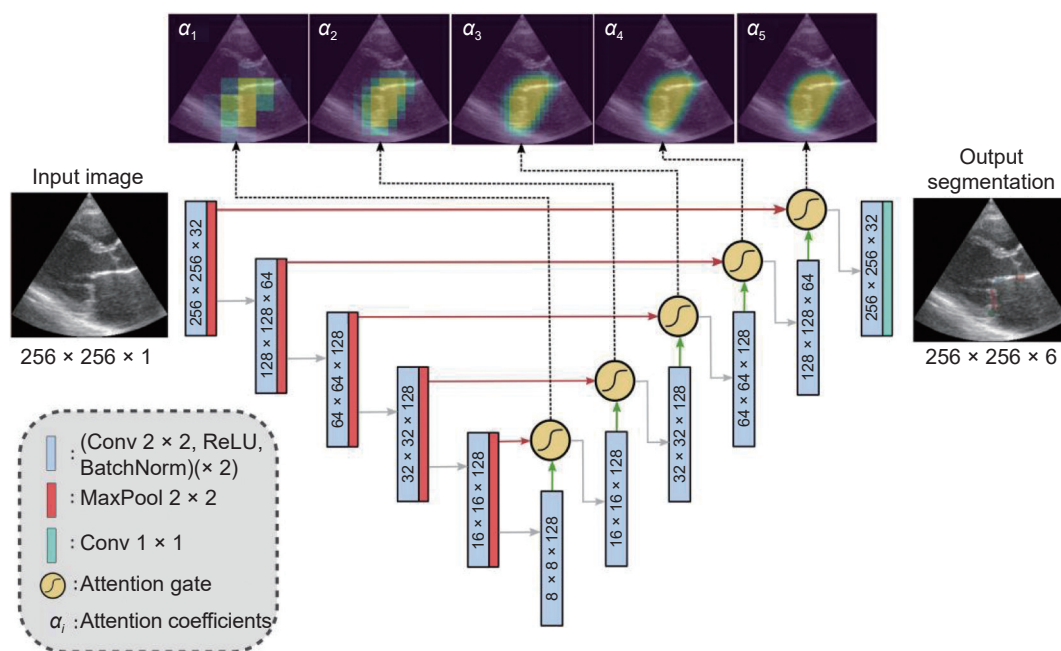


Figure 2 This picture shows the architecture of the Attention UNet with deep supervision. Attention gates generate spatial coefficient maps (α_1 – α_5) that highlight the thin, rapidly moving mitral valve leaflets while suppressing responses from surrounding myocardium and blood cavity. This selective focusing improves boundary delineation of low-contrast, fast-moving valve tissue. Deep supervision further stabilizes the attention weights across training epochs, ensuring consistent localization throughout the cardiac cycle. Together, these mechanisms address the major challenges of echocardiographic segmentation, poor contrast, motion blur, and leaflet thinness, thereby yielding more temporally consistent and anatomically reliable results. (Reprinted with permission from copyright Elsevier) [38]

Segmentation for Congenital Heart Disease

Congenital heart disease (CHD), encompassing a spectrum of cardiovascular malformations resulting from abnormal embryonic development of the heart and great vessels, remains a major cause of neonatal and pediatric mortality in China [40]. With an estimated prevalence of 0.5–0.7% among newborns [41], CHD poses significant clinical challenges, particularly due to the high incidence of specific subtypes such as ventricular septal defects (VSDs) and atrial septal defects (ASDs), which account for approximately 25% and 10–15% of cases, respectively [42,43]. The accurate segmentation and localization of these structural anomalies are critical for facilitating early diagnosis, optimizing treatment planning, and improving postoperative management.

However, despite ongoing progress in imaging technology and algorithmic development, the automatic and precise segmentation of CHD structures remain a difficult task. This is primarily due to the small anatomical scale of fetal and neonatal hearts, their structural complexity, and the inherently low quality of US imaging. These challenges highlight the urgent need for robust, automated segmentation techniques tailored to the specific demands of CHD assessment in pediatric Echo.

Traditional and machine learning methods in congenital heart diseases segmentation

In the early stage of research, the segmentation of US

images in cases of CHD primarily depended on manual delineation of anatomical contours or semi-automatic techniques based on classical image processing algorithms. As one representative approach, region growing methods, which expand seed regions based on grayscale intensity or textural similarity, achieved moderate success in segmenting structures such as ASDs and VSDs with well-defined boundaries. However, their effectiveness significantly declined in scenarios involving substantial image noise or pronounced anatomical variation.

Despite these encouraging results, the performance of ML-based segmentation remains far from ideal in practical echocardiographic applications. To improve segmentation efficiency and reduce operator dependence, researchers progressively introduced ML algorithms, including SVM, RF, and Extreme Gradient Boosting (XGBoost), to enable automatic identification and classification of abnormal cardiac structures. For instance, Yan et al. employed SVM, LR, RF, and XGBoost to predict the risk of postoperative arrhythmia in a cohort of 1,384 patients with peri membranous ventricular septal defect (PMVSD) [44]. Sun et al. constructed an artificial intelligence model incorporating natural language processing (NLP) and ML, trained on more than 29,000 PMVSD cases, which demonstrated excellent performance in predicting spontaneous closure at 1, 3, and 5 years, with an area under the curve (AUC) as high as 0.98 [45]. Du et al. developed an XGBoost-based predictive model using surgical risk indicators and preopera-

tive variables to estimate in-hospital mortality, achieving an AUC of 0.907, sensitivity of 0.785, and specificity of 0.824 [46].

Although these models incorporated multidimensional features such as intensity, texture, and shape, their segmentation performance remained constrained by strong dependence on handcrafted feature engineering and limited generalization across diverse patient populations and imaging conditions.

The variability in cardiac morphology, ranging from ASD and VSD to complex multi-chamber malformations, creates a highly heterogeneous imaging landscape in CHDs. Handcrafted features, often derived from static texture or boundary descriptors, lack the representational flexibility to capture such structural diversity, particularly when the defect geometry and acoustic window vary drastically among patients.

Moreover, handcrafted features are vulnerable to ultrasound-specific artifacts such as speckle noise, acoustic shadowing, and inconsistent probe orientation, all of which distort feature distributions and degrade model robustness.

Consequently, ML-based segmentation frameworks frequently fail to delineate malformed chambers or septal discontinuities with sufficient accuracy and continuity, especially in low-contrast or artifact-prone regions. These challenges underline the limitations of traditional

feature-driven approaches and highlight the necessity for deep learning models that can automatically learn hierarchical and spatially adaptive representations from heterogeneous echocardiographic data.

Deep learning methods in congenital heart diseases segmentation

An increasing number of researchers have sought to leverage end-to-end learning mechanisms and enhanced feature abstraction capabilities to achieve more accurate and robust segmentation of CHD anomalies. DL methods, particularly those that integrate CNN with attention mechanisms, have emerged as effective approaches for the automated and precise delineation of complex anatomical structures in neonatal echocardiographic images. A representative study by Nurmaini et al. developed a DL framework specifically tailored for fetal US that demonstrated superior capability in capturing intricate cardiac structures, significantly improving segmentation accuracy for critical regions such as the atrioventricular septum and heart valves [47]. Figure 3 shows the sample annotated images of defect position in ASD, VSD, and AVSD.

Furthermore, Transformer-based architectures, owing to their self-attention mechanisms, have shown great promise in modeling long-range spatial dependencies within cardiac anatomy, thereby contributing to greater

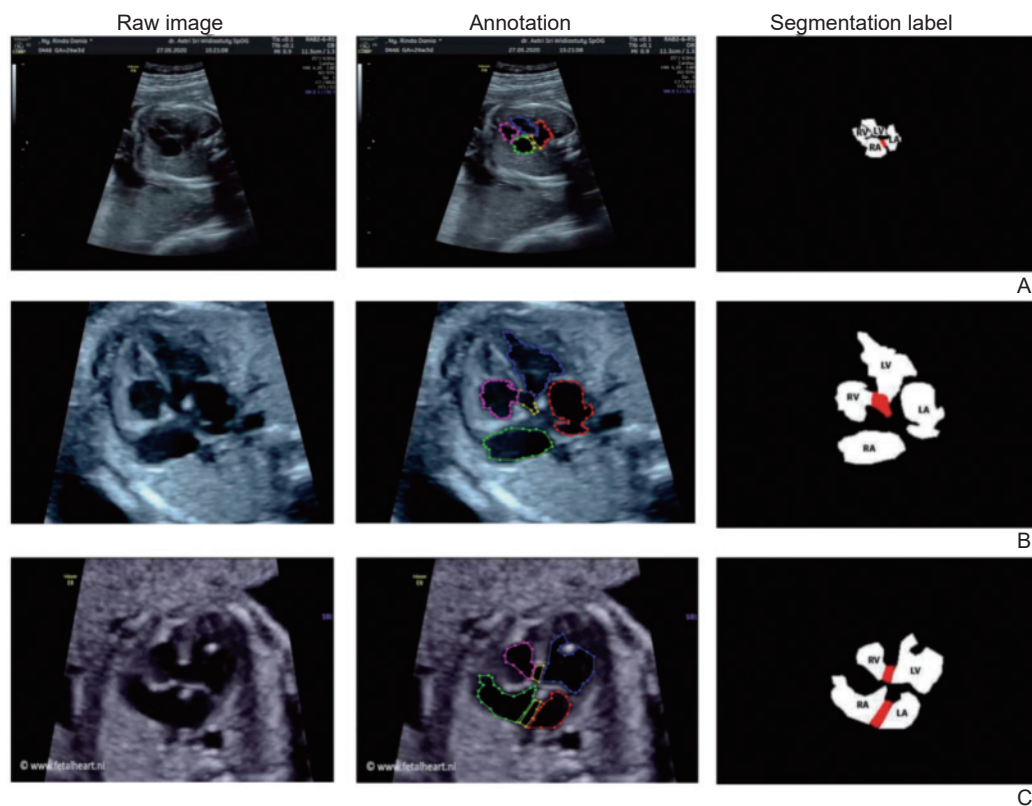


Figure 3 The sample of annotated image by maternal–fetal clinician for heart defect detection in case: (A) ASD; (B) VSD; and (C) AVSD. In the annotation, the green line is RA, the red line is LA, the purple line is RV, the blue line is LV, and the yellow line is defective. (Reprinted with permission from copyright MDPI) [47]

stability and accuracy in segmentation tasks [48,49]. Given the frequent scarcity of labeled clinical data, especially for rare or abnormal cases, few-shot learning strategies based on self-supervised or transfer learning paradigms have attracted growing interest. By pretraining model parameters or acquiring anatomical priorities through self-supervised tasks, such methods have successfully reduced reliance on large-scale annotated datasets and demonstrated encouraging translational potential in clinical applications.

Overall, echocardiographic segmentation for CHD has evolved from manual annotation and semi-automatic techniques such as region growing to increasingly sophisticated DL models driven by CNN, attention mechanisms, and Transformer architectures. This technological shift has significantly enhanced the level of automation and accuracy in segmentation tasks. Nonetheless, widespread clinical adoption remains hindered by persistent challenges, including annotation bottlenecks and limited model generalizability. Looking ahead, advancements in few-shot learning, self-supervised methods, and multimodal data fusion are expected to further accelerate the clinical translation and practical deployment of CHD ultrasound segmentation technologies.

Segmentation for Left Ventricular Dysfunction and Heart Failure

Heart failure (HF), a clinical syndrome representing the terminal stage of various cardiovascular diseases, is primarily characterized by impaired cardiac pumping function, with LV dysfunction being the most prevalent manifestation [50]. Among the key parameters for evaluating the severity and prognosis of HF, the left ventricular ejection fraction (LVEF) serves as a crucial quantitative indicator; however, its accurate assessment is critically dependent on the precise delineation of the endocardial and epicardial boundaries of the LV through image segmentation techniques [51,52]. Despite advances in imaging and algorithmic approaches, the task of automated and accurate segmentation remains highly challenging, owing to the substantial morphological variations of the LV throughout the cardiac cycle, the structural complexity of the myocardium, and the inherent limitations of echocardiographic images, including low contrast and the presence of noise.

Traditional and machine learning methods in left ventricular dysfunction and heart failure segmentation

In early clinical practice, segmentation of LV largely relied on manual contouring by experienced operators or

the use of semi-automatic annotation tools. This is exemplified by the work of Prashant et al. who developed an edge detection algorithm based on weighted radial filters, which effectively identified ventricular boundaries within cardiac chambers by leveraging spatially weighted gradients [53]. Similarly, Huang et al. proposed a motion segmentation method based on dynamic texture modeling, incorporating temporal constraints to improve boundary delineation performance across cardiac cycles [54]. While these techniques have been widely adopted in clinical workflows, their outcomes are inherently influenced by operator experience and interactive input, resulting in significant inter-observer variability and limited reproducibility.

To enhance segmentation efficiency and reduce subjective bias, traditional image processing techniques such as thresholding and region growing were introduced. Mühlenbruch et al. evaluated a novel semi-automated three-dimensional region-growing segmentation algorithm for functional LV analysis, reporting a 44.2% reduction in processing time and strong agreement with conventional methods [55]. Although these algorithms are computationally efficient and relatively simple to implement, their performance deteriorates in cases involving thin myocardial walls or blurred boundaries, rendering them inadequate for the precision demands of clinical diagnosis.

As traditional methods reached a plateau in segmentation accuracy, increasing attention has been directed toward learning-based approaches. ML algorithms, by extracting discriminative features from annotated datasets, have significantly enhanced the capacity for automated boundary recognition in LV segmentation tasks. Classifiers such as RF, SVM, and others have been deployed to automatically delineate LV contours. To illustrate, Khellaf et al. integrated a structured RF with an active shape model to segment the LV in both end-systolic and end-diastolic frames of echocardiographic sequences [56]. These methods have demonstrated notable improvements in the consistency and robustness of myocardial structure segmentation. As ML-based segmentation techniques continue to evolve, the level of automation in LV analysis has markedly improved, substantially reducing reliance on manual intervention and laying a solid foundation for standardized and time-efficient clinical assessment workflows.

Deep learning methods in left ventricular dysfunction and heart failure segmentation

As segmentation tasks increasingly demand finer delineation of structural boundaries and greater adaptability to morphological variability, the limitations of traditional ML models, particularly in terms of expres-

sive power and spatial modeling capacity, have become progressively more apparent. In contrast, DL approaches, with their end-to-end training paradigm and hierarchical feature abstraction capabilities, are better suited to address the challenges of segmenting dynamic structures such as the LV and have gradually become the central focus of research in this domain.

Among various DL architectures, U-Net has gained widespread adoption for precise delineation of both endocardial and epicardial boundaries in LV segmentation. One such effort is the CNN model proposed by Yu et al. who investigated LV contour segmentation in infant echocardiograms and proposed a CNN incorporating multi-scale feature extraction, which produced segmentation results closely aligned with clinical ground truth values [57]. Moreover, advanced U-Net variants, such as nnUNet and Attention U-Net, have demonstrated superior performance and robust generalization in echocardiographic segmentation tasks. Amer et al. intro-

duced ResDUNet, an enhanced model that integrates residual blocks and semantic features encoding into the U-Net framework, achieving a 1.2% improvement in Dice score over the original architecture [58]. Similarly, Moradi et al. proposed the MFP-Unet model, which addresses the semantic deficiency of conventional U-Net through a novel architecture and achieved an average Dice coefficient of 0.953 [59]. Carneiro et al. developed a novel supervised learning model capable of delivering reliable segmentation performance using only 20 annotated echocardiographic images [14]. Huang et al. designed a deep learning network VDS-UNET to achieve the classification of mitral regurgitation [60]. In [figure 4](#), the red contour is marked by professionals, and the green contour is the result of network segmentation. It shows that the LA can be segmented well in all three sections of echocardiography.

In parallel, attention-based architectures, most notably Transformer models, have begun to be applied

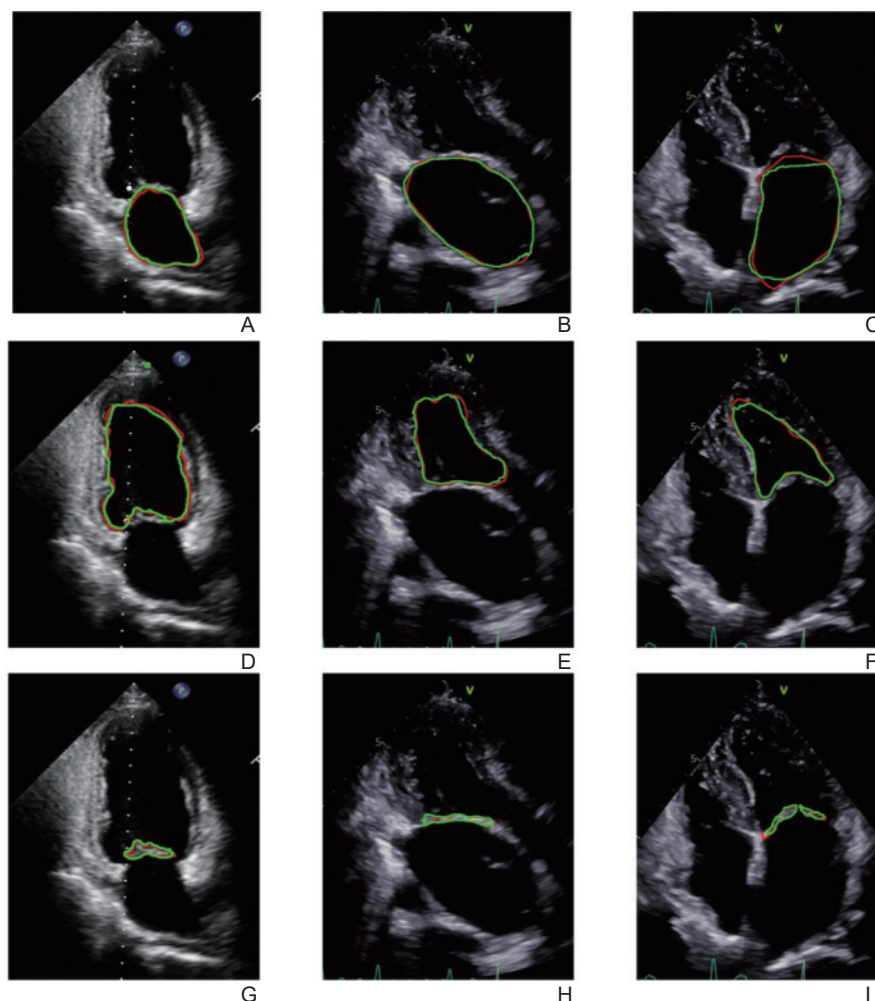


Figure 4 These nine images illustrate segmentation outcomes across three echocardiographic views. Images (A, D, G) correspond to apical two-chamber views with normal mitral valves, (B, E, H) represent apical three-chamber views with functional MR, and (C, F, I) show apical four-chamber views with mitral valve prolapse. The Dice scores for LA, LV and MV were 0.935, 0.915 and 0.757, respectively, indicating strong performance in normal and functional MR cases but reduced accuracy in prolapse, where leaflet thickening, asymmetric deformation and blurred coaptation boundaries make segmentation substantially more challenging. (Reprinted with permission from MDPI) [60]

to LV segmentation, significantly improving the accuracy and stability of segmentation across different echocardiographic views and cardiac cycle phases. Liao et al. proposed and validated two Transformer-based LV segmentation models utilizing Swin Transformer and Segformer backbones to overcome the limited receptive field of CNNs in echocardiographic imaging, ultimately achieving a Dice score of 92.79% [61].

With continued advancements, segmentation technologies for LV dysfunction and HF are evolving from static structural recognition toward dynamic functional modeling, from two-dimensional contour extraction to the integration of temporal features, and from auxiliary assessment tools toward alignment with clinically quantifiable metrics. In this context, DL models, particularly those augmented with attention mechanisms and temporal modeling capabilities, are gradually supplanting traditional approaches that rely on handcrafted features and

static images, becoming the dominant paradigm in both research and application. Future efforts should focus on enhancing model generalizability across devices and patient populations, reducing dependence on large-scale annotated datasets, and improving the real-time performance and interpretability of segmentation outcomes to accelerate clinical translation and deployment.

Challenges and Prospects

Although echocardiographic image segmentation has achieved notable progress in clinical contexts such as VHD, CHD, and LV dysfunction and HF (table 1), several fundamental challenges continue to hinder its broader clinical translation.

First, publicly available datasets remain scarce, and access to high-quality, expertly annotated clinical US data is limited. The demanding nature of manual annota-

Table 1 Representative literature on heart segmentation

Category	Author	Model	Content	Dataset	Evaluation
VHD	Çomak et al. [32]	LS-SVM (<i>ML</i>)	Identify abnormalities in valvular CDFI	Patients: 215	AUC: 0.950
	Avci [33]	GSVM (<i>ML</i>)	Classify the CDFI of heart valve diseases	Patients: 215	Accuracy: 0.950
	Hernandez et al. [34]	LR, ANN, NB, RF (<i>ML</i>)	Predict in-hospital mortality after TAVR	Patients: 10883	AUC: 0.920
	Vrudhula et al. [36]	3D CNN (<i>DL</i>)	Detect mitral regurgitation severity	Videos: 80833	AUC: 0.998
	Holste et al. [62]	CNN (<i>DL</i>)	Detect severe aortic stenosis	Videos: 17570	AUROC: 0.978
	Yazan et al. [63]	CNN, LSTM (<i>DL</i>)	Classify different heart valvular diseases	Videos: 4541	Accuracy: 0.979
CHD	Yan et al. [44]	SVM, LR, RF, XGBoost (<i>ML</i>)	Predict postoperative arrhythmia	Patients: 1384	AUC: 0.863
	Sun et al. [45]	NLP, RSF (<i>ML</i>)	Predict spontaneous closure of PMVSD	Patients: 29142	AUC: 0.950
	Nurmaini et al. [47]	Mask-RCNN (<i>DL</i>)	Segment and detect congenital heart defects	Images: 1149	Dice: 0.983
	Nova et al. [48]	U-Net, V-Net (<i>DL</i>)	Segment cardiac septal defects from 2D Echo images	Images: 2609	Accuracy: 0.991
	Tan et al. [64]	CNN (<i>DL</i>)	Detect CHD from fetal US images	Patients: 100	Precision: 0.890 Recall: 0.820
	Jiang et al. [65]	ResNet50 (<i>DL</i>)	Classify seven CHD views	Patients: 1411 Images: 14838	Accuracy: 0.923
	Nemri et al. [66]	Swin Transformer, U-Net (<i>DL</i>)	Segment echocardiographic images	Videos: 10030	Dice: 0.870
Ventricle	Huang et al. [54]	DMM-ACM (<i>Traditional Method</i>)	Localize and segment the heart	Images: 1280	FNF: 0.033 FPF: 0.057 TPF: 0.967
LV	Khellaf et al. [56]	SRF (<i>ML</i>)	Segment the LV in 3D echocardiograms	Videos: 45	Dice (ED): 0.899 Dice (ES): 0.864
	Carneiro et al. [14]	DBN, PCA (<i>DL</i>)	Segment LV in small dataset	Images: 480	TPR: 0.950 FPR: 0.008
	Jafari et al. [67]	3D U-Net (<i>DL</i>)	Estimate LV ejection fraction from Echo	Patients: 45	Dice (ED): 0.933 Dice (ES): 0.909
LVEF	Liao et al. [60]	Transformer, K-Net (<i>DL</i>)	Segment LV correctly and assess LVEF	Videos: 10030	Dice: 0.929
	Jafari et al. [68]	Deep Bayesian (<i>ML</i>)	Estimate ejection fraction and detect foreshortening	Patients: 2680	Dice: 0.920

VHD, valvular heart disease; SVM, support vector machines; CDFI, color doppler flow imaging; AV, aortic valve; MV, mitral valve; LR, logistics regression; ANN, artificial neural network; NB, naive bayes; RF, random forest; STS, the Society of Thoracic Surgeons; ACC, The American College of Cardiology; TAVR, transcatheter valve therapy; TAVR, transcatheter aortic valve replacement; AUC, area under curve; DL, deep learning; CNN, convolutional neural network; LA, left atrium; AUROC, area under receiver operating characteristic curve; LSTM, long short-term memory; AS, aortic stenosis; MS, mitral stenosis; MVP, mitral valve prolapse; MR, mitral regurgitation; CHD, congenital heart disease; XGBoost, eXtreme Gradient Boosting; NLP, natural language processing; RSF, random survival forest; PMVSD, perimembranous ventricular septal defects; ML, machine learning; IoU, Intersection over Union; Echo, echocardiography; DICOM, digital imaging and communications in medicine; DMM, dynamic metric map; ACM, active contour model; FNF: false negative fraction; FPF: false positive fraction; TPF: the positive fraction; LV, left ventricular; ED, end diastolic; ES, end systole; DBN, deep belief network; PCA, principal component analysis; EF, ejection fractions; LVEF, left ventricular ejection fractions.

tion, combined with institutional and ethical constraints, results in datasets that are often insufficient in scale, diversity, and anatomical complexity. These limitations significantly affect model training and hinder generalizability to real-world scenarios.

Second, pronounced heterogeneity across imaging equipment, acquisition protocols, and operators leads to performance degradation when models are applied in multi-center or cross-device environments. Furthermore, segmentation performance remains suboptimal in challenging cases, such as those involving indistinct anatomical boundaries or motion-related artifacts.

At the clinical level, concerns persist regarding the interpretability of AI model outputs, limited clinician trust, and the technical complexity of integrating segmentation tools into existing diagnostic workflows.

To address these challenges, future research should advance along several strategic and transformative directions. At the data level, establishing harmonized data ecosystems spanning institutions and populations will be essential to advance echocardiographic AI. Federated learning can enable collaborative model training across institutions without sharing patient data, promoting data diversity while preserving privacy.

From a methodological standpoint, weakly supervised and self-supervised learning can mitigate the reliance on exhaustive manual annotations, whereas domain adaptation and transfer learning enhance model generalization across centers and devices. Architecturally, integrating temporal attention or transformer-based modules can improve the capture of spatiotemporal cardiac dynamics. Furthermore, incorporating uncertainty quantification and explainable AI frameworks can increase transparency and clinician confidence.

In summary, future echocardiographic segmentation research should establish a synergistic pathway across data sharing, algorithmic innovation, and clinical translation, ultimately enabling deployable, interpretable, and trustworthy AI systems that can sustainably support precision cardiovascular diagnostics.

Acknowledgment

Not applicable.

Authors' Contributions

Li Zuoyan drafted the manuscript. Li Fan gave guidance to the manuscript, polished the language. Du Lianfang reviewed and approved the final version. All authors contributed to the revision of the manuscript.

Funding

This work was supported by the *Shanghai Chest Hospital Science and*

Technology Development Fund.

Data Availability

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Conflict of interest

The authors have no conflict of interest to declare.

Consent for publication

Not applicable.

References

- [1] Zhou J, Du M, Chang S, Chen ZY. Artificial intelligence in echocardiography: detection, functional evaluation, and disease diagnosis. *Cardiovasc Ultrasound* 2021;19:11.
- [2] Mielnicki W, Dyla A, Malyszek-Tumidajewicz J. A systematic approach to transoesophageal echocardiography in the intensive care unit - a practical guide for intensivists. *Anaesth Intensive Ther* 2021;53:329-335.
- [3] Li Y. Computer-aided segmentation and quantification in contrast-enhanced echocardiography. Imperial College London 2018.
- [4] Leclerc S, Smistad E, Pedrosa J, Ostvik A, Cervenansky F, Espinosa F, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. *IEEE Trans Med Imaging* 2019;38:2198-2210.
- [5] Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, et al. Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy. *Circulation* 2018;138:1623-1635.
- [6] Dreyfus J, Taramasso M, Kresoja K-P, Omran H, Iliadis C, Russo G, et al. Prognostic implications of residual tricuspid regurgitation grading after transcatheter tricuspid valve repair. *JACC Cardiovasc Interv* 2024;17:1485-1495.
- [7] Lancellotti P, Pibarot P, Chambers J, La Canna G, Pepi M, Dulgheru R, et al. Multi-modality imaging assessment of native valvular regurgitation: an EACVI and ESC council of valvular heart disease position paper. *Eur Heart J Cardiovasc Imaging* 2022;23:E171-E232.
- [8] Morris DA, Belyayskiy E, Aravind-Kumar R, Kropf M, Frydas A, Braunauer K, et al. Potential usefulness and clinical relevance of adding left atrial strain to left atrial volume index in the detection of left ventricular diastolic dysfunction. *JACC Cardiovasc Imaging* 2018;11:1405-1415.
- [9] Thavendiranathan P, Poulin F, Lim K-D, Plana JC, Woo A, Marwick TH. Use of myocardial strain imaging by echocardiography for the early detection of cardiotoxicity in patients during and after cancer chemotherapy: a systematic review. *J Am Coll Cardiol* 2014;63:2751-2768.
- [10] Satou GM, Rheuban K, Alverson D, Lewin M, Mahnke C, Marcini J, et al. Telemedicine in pediatric cardiology: a scientific statement from the American Heart Association. *Circulation* 2017;135:e648-e678.
- [11] Trippi JA, Lee KS, Kopp G, Nelson D, Kovacs R. Emergency

- echocardiography telemedicine: An efficient method to provide 24-hour consultative echocardiography. *J Am Coll Cardiol* 1996;27:1748-1752.
- [12] Garcia-Canadilla P, Sanchez-Martinez S, Crispi F, Bijmens B. Machine learning in fetal cardiology: what to expect. *Fetal Diagn Ther* 2020;47:363-372.
 - [13] Kumar N, Verma R, Anand D, Zhou Y, Onder OF, Tsougenis E, et al. A multi-organ nucleus segmentation challenge. *IEEE Trans Med Imaging* 2020;39:1380-1391.
 - [14] Carneiro G, Nascimento JC, Freitas A. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Trans Image Process* 2012;21:968-982.
 - [15] Zhang X, Dai L. Image enhancement based on rough set and fractional order differentiator. *Fractal and Fractional* 2022;6:214.
 - [16] Liu XB, Song LP, Liu S, Zhang YD. A review of deep-learning-based medical image segmentation methods. *Sustainability* 2021;13:1224.
 - [17] Kotaridis I, Lazaridou M. Remote sensing image segmentation advances: A meta-analysis. *ISPRS-J Photogramm Remote Sens* 2021;173:309-322.
 - [18] Zhang W, Zhuang P, Sun H-H, Li G, Kwong S, Li C. Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. *IEEE Trans Image Process* 2022;31:3997-4010.
 - [19] Li Y, Zhang Y, Cui W, Lei B, Kuang X, Zhang T. Dual encoder-based dynamic-channel graph convolutional network with edge enhancement for retinal vessel segmentation. *IEEE Trans Med Imaging* 2022;41:1975-1989.
 - [20] Pachitariu M, Stringer C. Cellpose 2.0: how to train your own model. *Nat Methods* 2022;19:1634-1641.
 - [21] Liu Y, Li H, Hu C, Luo S, Luo Y, Chen CW. Learning to aggregate multi-scale context for instance segmentation in remote sensing images. *IEEE Trans Neural Netw Learn Syst* 2025;36:595-609.
 - [22] Ragab DA, Sharkas M, Marshall S, Ren J. Breast cancer detection using deep convolutional neural networks and support vector machines. *Peerj* 2019;7:e6201.
 - [23] Sinha A, Dolz J. Multi-scale self-guided attention for medical image segmentation. *IEEE J Biomed Health Inform* 2021;25:121-130.
 - [24] Mou L, Zhao Y, Fu H, Liu Y, Cheng J, Zheng Y, et al. CS²-Net: Deep learning segmentation of curvilinear structures in medical imaging. *Med Image Anal* 2021;67:101874.
 - [25] Nieminen MS, Heikkilä J, Karjalainen J. Echocardiography in acute infectious myocarditis: relation to clinical and electrocardiographic findings. *Am J Cardiol* 1984;53:1331-1337.
 - [26] Jin Q, Meng Z, Tuan DP, Chen Q, Wei L, Su R. DUNet: A deformable network for retinal vessel segmentation. *Knowledge-Based Systems* 2019;178:149-162.
 - [27] Timmis A, Aboyans V, Vardas P, Townsend N, Torbica A, Kavousi M, et al. European society of cardiology: the 2023 atlas of cardiovascular disease statistics. *Eur Heart J* 2024;45:4019-4062.
 - [28] Chan V, Malas T, Lapierre H, Boodhwani M, Lam B-K, Rubens FD, et al. Reoperation of left heart valve bioprostheses according to age at implantation. *Circulation* 2011;124:S75-S80.
 - [29] Timmis A, Townsend N, Gale C, Grobbee R, Maniadakis N, Flather M, et al. European society of cardiology: cardiovascular disease statistics 2017. *Eur Heart J* 2018;39:508-579.
 - [30] Edler I, Lindström K. The history of echocardiography. *Ultrasound in medicine & biology* 2004;30:1565-1644.
 - [31] Hirasawa K, Izumo M, Akashi YJ. Stress echocardiography in valvular heart disease. *Front Cardiovasc Med* 2023;10:1233924.
 - [32] Çomak E, Arslan A, Türkoğlu İ. A decision support system based on support vector machines for diagnosis of the heart valve diseases. *Comput Biol Med* 2007;37:21-27.
 - [33] Avci E. A new intelligent diagnosis system for the heart valve diseases by using genetic-SVM classifier. *Expert Systems with Applications* 2009;36:10618-10626.
 - [34] Hernandez-Suarez Dagmar F, Kim Y, Villablanca P, Gupta T, Wiley J, Nieves-Rodriguez Brenda G, et al. Machine learning prediction models for in-hospital mortality after transcatheter aortic valve replacement. *JACC Cardiovasc Interv* 2019;12:1328-1338.
 - [35] Zhang J, Liu Y, Xu W. Identification of mitral annulus hinge point based on local context feature and additive SVM classifier. *Comput Math Methods Med* 2015;2015:419826.
 - [36] Vrudhula A, Duffy G, Vukadinovic M, Liang D, Cheng S, Ouyang D. High-throughput deep learning detection of mitral regurgitation. *Circulation* 2024;150:923-933.
 - [37] Long A, Haggerty CM, Finer J, Hartzel D, Jing L, Keivani A, et al. Deep learning for echo analysis, tracking, and evaluation of mitral regurgitation (DELINEATE-MR). *Circulation* 2024;150:911-922.
 - [38] Wifstad SV, Kildahl HA, Grenne B, Holte E, Hauge SW, Saebo S, et al. Mitral valve segmentation and tracking from transthoracic echocardiography using deep learning. *Ultrasound Med Biol* 2024;50:661-670.
 - [39] An S, Zhu H, Wang Y, Zhou F, Zhou X, Yang X, et al. A category attention instance segmentation network for four cardiac chambers segmentation in fetal echocardiography. *Comput Med Imaging Graph* 2021;93:101983.
 - [40] Van der Linde D, Konings EE, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJ, et al. Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis. *J Am Coll Cardiol* 2011;58:2241-2247.
 - [41] Zhao LJ, Chen LZ, Yang TB, Wang TT, Zhang SM, Chen LT, et al. Birth prevalence of congenital heart disease in China, 1980-2019: a systematic review and meta-analysis of 617 studies. *Eur J Epidemiol* 2020;35:631-642.
 - [42] De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;24:1342-1350.
 - [43] Esteve A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-118.
 - [44] Yan L, Meng Y, Sun H, Liu X, Han B. Application of machine learning in predicting postoperative arrhythmia following transcatheter closure of perimembranous ventricular septal defects. *Kardiologi Pol* 2025;83:295-304.
 - [45] Sun J, Feng T, Wang B, Li F, Han B, Chu M, et al. Leveraging artificial intelligence for predicting spontaneous closure of perimembranous ventricular septal defect in children: a multicentre, retrospective study in China. *Lancet Digit Health* 2025;7:e44-e53.
 - [46] Du X, Wang H, Wang S, He Y, Zheng J, Zhang H, et al. Machine learning model for predicting risk of in-hospital mortality after surgery in congenital heart disease patients. *Rev Cardiovasc Med* 2022;23:376.
 - [47] Nurmaini S, Rachmatullah MN, Sapitri AI, Darmawahyuni A, Tutuko B, Firdaus F, et al. Deep learning-based computer-aided fetal echocardiography: application to heart standard view segmentation for congenital heart defects detection. *Sensors (Basel)* 2021;21:8007.
 - [48] Nova R, Nurmaini S, Partan RU, Putra ST. Automated image segmentation for cardiac septal defects based on contour region with convolutional neural networks: A preliminary study. *Informatics in Medicine Unlocked* 2021;24:100601.

- [49] Nurmaini S, Rachmatullah MN, Sapitri AI, Darmawahyuni A, Jovandy A, Firdaus F, et al. Accurate detection of septal defects with fetal ultrasonography images using deep learning-based multiclass instance segmentation. *IEEE access* 2020;8:196160-196174.
- [50] Wang H, Chai K, Du M, Wang S, Cai J-P, Li Y, et al. Prevalence and incidence of heart failure among urban patients in china: a national population-based analysis. *Circ Heart Fail* 2021;14:e008406.
- [51] Dunlay SM, Roger VL, Redfield MM. Epidemiology of heart failure with preserved ejection fraction. *Nat Rev Cardiol* 2017;14:591-602.
- [52] MEMBERS WC, Hunt SA, Abraham WT, Chin MH, Feldman AM, Francis GS, et al. ACC/AHA 2005 Guideline update for the diagnosis and management of chronic heart failure in the adult. *Circulation* 2005;112:e154-e235.
- [53] Bansod P, B. DU, N. MS, and Burkule N. Segmentation of left ventricle in short-axis echocardiographic sequences by weighted radial edge filtering and adaptive recovery of dropout regions. *Comput Methods Biomech Biomed Engin* 2011;14:603-613.
- [54] Huang JZ, Huang XL, Metaxas D, Axel L. Dynamic texture based heart localization and segmentation in 4-D cardiac images. *I S Biomed Imaging* 2007:852-855.
- [55] Mühlenbruch G, Das M, Hohl C, Wildberger JE, Rinck D, Flohr TG, et al. Global left ventricular function in cardiac CT. Evaluation of an automated 3D region-growing segmentation algorithm. *Eur Radiol* 2006;16:1117-1123.
- [56] Khellaf F, Leclerc S, Voorneveld JD, Bandaru RS, Bosch JG, Bernard O, editors. Left ventricle segmentation in 3D ultrasound by combining structured random forests with active shape models. *Medical Imaging 2018: Image Processing. SPIE* 2018;10574:120-129.
- [57] Yu L, Guo Y, Wang Y, Yu J, Chen P. Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks. *IEEE Trans Biomed Eng* 2016;64:1886-1895.
- [58] Amer A, Ye XJ, Janan F. ResDUnet: A deep learning-based left ventricle segmentation method for echocardiography. *IEEE access* 2021;9:159755-159763.
- [59] Moradi S, Oghli MG, Alizadehasl A, Shiri I, Oveisi N, Oveisi M, et al. MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography. *Physica Medica* 2019;67:58-69.
- [60] Huang H, Ge Z, Wang H, Wu J, Hu C, Li N, et al. Segmentation of echocardiography based on deep learning model. *Electronics* 2022;11:1714.
- [61] Liao M, Lian Y, Yao Y, Chen L, Gao F, Xu L, et al. Left ventricle segmentation in echocardiography with transformer. *Diagnostics* 2023;13:2365.
- [62] Holste G, Oikonomou EK, Mortazavi BJ, Coppi A, Faridi KF, Miller EJ, et al. Severe aortic stenosis detection by deep learning applied to echocardiography. *Eur Heart J* 2023;44:4592-4604.
- [63] Al-Issa Y, Alqudah AM. A lightweight hybrid deep learning system for cardiac valvular disease classification. *Sci Rep* 2022;12:14297.
- [64] Tan J, Au A, Meng QJ, FinesilverSmith S, Simpson J, Rueckert D, et al. Automated detection of congenital heart disease in fetal ultrasound screening. *International Workshop on Advances in Simplifying Medical Ultrasound. Cham: Springer International Publishing* 2020: 243-252.
- [65] Jiang X, Yu J, Ye J, Jia W, Xu W, Shu Q. A deep learning-based method for pediatric congenital heart disease detection with seven standard views in echocardiography. *World J Pediatr Surg* 2023;6:e000580.
- [66] Nemri S, Duong L. Automatic segmentation of echocardiographic images using a shifted windows vision transformer architecture. *Biomed Phys Eng Express* 2024;10:065017.
- [67] Smistad E, Ostvik A, Salte IM, Melichova D, Nguyen TM, Haugaa K, et al. Real-time automatic ejection fraction and foreshortening detection using deep learning. *IEEE Trans Ultrason Ferroelectr Freq Control* 2020;67:2595-2604.
- [68] Jafari MH, Van Woudenberg N, Luong C, Abolmaesumi P, Tsang T. Deep bayesian image segmentation for a more robust ejection fraction estimation. *2021 IEEE 18th International Symposium on Biomedical Imaging (Isbi)* 2021:1264-1268.