TOPICAL REVIEW

Application Status and Prospect of Deep Learning in Echocardiography

QI QI[1], XIAOXIANG HAN[2], (Graduate Student Member, IEEE), AND YIMAN LIU[3,4], (Member, IEEE)
[1]Department of Biomedical Statistics, Wisdom Lake Academy of Pharmacy, Xi'an Jiaotong-Liverpool University, Suzhou 215123, China
[2]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China
[3]Department of Pediatric Cardiology, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China
[4]Shanghai Engineering Research Center of Intelligence Pediatrics (SERCIP), Shanghai 200127, China
Corresponding author: Yiman Liu (LiuyimanSCMC@163.com)

ABSTRACT Echocardiography is essential for the diagnosis and treatment of cardiovascular diseases, especially congenital heart disease. However, the interpretation of echocardiography requires the accumulation of abundant professional experience for cardiologists, which may lead to missed diagnoses and misdiagnoses due to differences between operators. The development of artificial intelligence and its subset, deep learning, has altered this potential crisis in recent years with their rapid, accurate, and consistent nature. While the application of deep learning in echocardiography is still in its infancy, recent studies demonstrate that deep learning models can quickly obtain information by extracting samples from large databases. Moreover, growing evidence suggests that deep learning can be used for standard section recognition in echocardiography and auxiliary diagnosis of heart disease. In this review, we begin by outlining the principles of deep learning. Then, we investigate the current application of deep learning in echocardiography, underscoring its significance. Furthermore, we discuss its limitations and finally highlight future development prospects.

INDEX TERMS Artificial intelligence, deep learning, echocardiography.

I. INTRODUCTION

Accurate quantitative evaluation of cardiac morphology and functionality underpins the formulation of clinical diagnoses and effective treatment strategies. The intricacies of human perception play a crucial role, as echocardiographers traditionally rely on their visual faculties to interpret static and dynamic cardiac imagery. This conventional diagnostic process involves the conversion of light signals into electrochemical impulses by retinal photoreceptors, followed by their transmission through an intricate neuronal network to the brain, where they culminate in a blend of conscious and unconscious interpretations [1]. Echocardiography is paramount in cardiovascular diagnostics and therapy, noted for its unique ability to visualize the heart's dynamics and

detect anomalies in real-time [2]. Clinicians depend on this modality to discern intricate cardiac motions and patterns, translating them into diagnostic insights and therapeutic directives. Yet, the intrinsic subjectivity of this interpretation, especially in the quantification of two-dimensional echocardiographic images, is fraught with potential inaccuracies. Variances in diagnostic techniques and interpretations among physicians, exacerbated by suboptimal image quality, present a significant challenge [3].

Conversely, Deep Learning (DL), a subset of Artificial Intelligence (AI), heralds a paradigm shift, endowing machines with the capability to discern intricate features within a myriad of echocardiographic images. These systems can then apply their learned knowledge to new images, empowering them to pinpoint regions of interest and identify pathologies with increased precision [4]. As detailed in Fig. 1, while echocardiographic practice has progressed from

The associate editor coordinating the review of this manuscript and approving it for publication was Hasan S. Mir.

mere two-dimensional interpretation to semi-quantitative analyses, it remains hindered by human variability and the labor-intensive nature of manual evaluation. Traditionally, echocardiographers undergo extensive training over the years to hone their diagnostic acumen. However, the advent of AI, particularly DL, has sparked a renaissance in medical imaging research. DL's profound capabilities for rapid data analysis and pattern recognition offer a robust solution to these challenges, augmenting the echocardiographer's role by assimilating insights from an expansive compilation of echocardiographic studies.
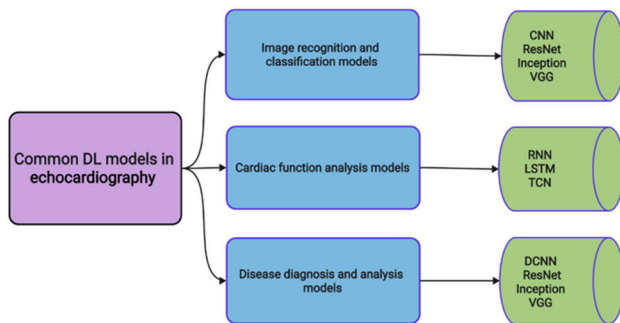

FIGURE 1. Common DL models in echocardiography.

This manuscript provides a succinct exposition of deep learning's foundational principles and delineates its burgeoning applications in echocardiography. These applications encompass the recognition of canonical imaging sections, the automated quantification of cardiac functions, the systematic evaluation of valvular pathologies, and the nuanced diagnosis and differentiation of cardiac diseases.

II. BACKGROUND

In recent years, the expansive utility of Artificial Intelligence (AI) has penetrated diverse sectors, notably in medical imaging analysis. This section unfolds the essential principles of Deep Learning (DL) and its pivotal component, the Convolutional Neural Network (CNN), which tackles key challenges in medical imagery interpretation. Furthermore, this segment offers an all-encompassing review of DL's multifaceted applications and delves into the inherent limitations that mark its current state of evolution.

A. BASIC PRINCIPLE OF DL

DL stands at the forefront of AI, endowed with the profound ability to emulate and potentially exceed human cognitive processes. It allows computational systems to perceive and analyze their environment, solving complex problems with heightened efficiency and goal-orientation, akin to human reasoning [1]. Within the broader AI spectrum, Machine Learning (ML) serves as an integral subset whereby computers derive knowledge from data, not through explicit programming, but via adaptive learning from labeled, partially labeled, or even unlabeled datasets. This subset is further stratified into supervised, unsupervised, and

reinforcement learning paradigms, each distinguished by the nature of feedback provided during the training process. Fig. 2 illustrates the intricate interplay between AI, ML, and DL, mapping out their symbiotic relationship.
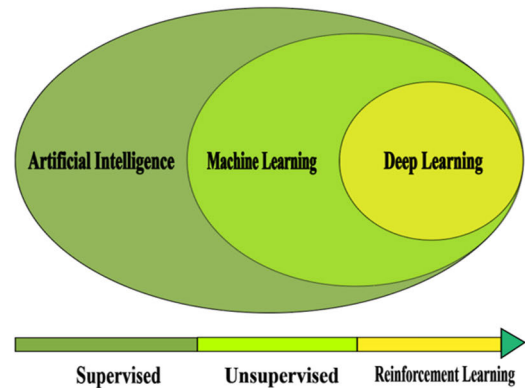

FIGURE 2. Relationship of artificial intelligence, machine learning, and deep learning.

B. CNN DYNAMICS

DL artfully replicates the workings of human neural networks by orchestrating a symphony of multi-layered neuronal cascades. This sophisticated approach distills complex data into a hierarchy of abstract features through successive nonlinear processing layers, each contributing to the system's capacity for making predictive classifications. The CNN, a preeminent algorithm within the DL paradigm, comprises a series of interconnected neuronal layers, each mirroring the intricate interplay of synapses in the human brain, modulated by variable weights.

Within these networks, each artificial neuron assimilates input signals weighted by their relative importance. Summation of these inputs leads to activation when a certain threshold is surpassed, thus propelling a signal forward that modulates the activity of subsequent neurons. Such intricate layering escalates the network's complexity and, consequently, the need for extensive training data. Traditionally, achieving a leaner neural network necessitated laborious manual extraction of features from raw data. Yet, DL has ushered in a new era where such limitations are transcended by automated learning algorithms, enabling refined pattern recognition with minimal human intervention [5].

C. DL-DRIVEN FIELD

It is demonstrated that DL has revolutionized the domain of feature analysis by supplanting the need for manual feature delineation with algorithms capable of unsupervised and semi-supervised learning, alongside hierarchical feature extraction [4], [6], [7]. This shift towards automation facilitates the application of these sophisticated techniques in practical, real-world scenarios. Presently, DL's capabilities are not only on par with, but in some instances, exceed human proficiency in sectors such as speech and image recognition,

and in prognosticating the efficacy of pharmaceutical compounds [8], [9]. Specifically, within drug discovery, AI acts as an accelerant, parsing through extensive chemical datasets to predict potential efficacy and toxicity, thereby expediting the journey from conception to clinical application of new medicinal drugs.

D. RESTRICTION OF DL

While AI harbors the transformative potential across numerous facets of human life, its translational capacity in clinical contexts demands rigorous scrutiny. DL models, proficient in parsing extensive datasets, are contingent upon the volume and integrity of data for their training efficacy. The healthcare sphere, in particular, mandates stringent validation of data to confirm its sufficiency, integrity, and precision before its employment in model training. Paramount to this is the assurance of dataset representativeness to circumvent biases and data omissions that could undermine a model's predictive accuracy [10]. AI's prospective utility in addressing these fundamental challenges is reflected in the advent of various echocardiography software tools, leveraging AI to enhance the analysis of cardiac imagery and facilitate structural and functional assessments [11], [12].

III. APPLICATION OF DL IN ECHOCARDIOGRAPHY

Clinical protocols advocate for the utilization of echocardiography in guiding therapeutic decisions via quantitative evaluation of the heart's chambers and valvular structures [13]. Yet, this quantitative scrutiny is intricate and time-intensive, often encumbered by manual delineation, rendering it impractical amidst the exigencies of a dynamic clinical setting [14]. Consequently, qualitative visual assessments, which hinge on substantial expertise in image procurement and interpretative acumen, prevail in practice—a level of proficiency not always possessed by less experienced sonographers. The complexity escalates for internal medicine practitioners who encounter ultrasonography more frequently in emergent care yet lack specialized training [15]. DL stands as an instrumental breakthrough, potentially elevating diagnostic precision. It emulates the analytical insights of a seasoned sonographer, discerning and extracting subtle diagnostic cues imperceptible to the clinician's eye [16]. Fig. 3 graphically elucidates how DL algorithms streamline and demystify the echocardiographic evaluation process.

Although in its incipient phases within the echocardiographic field, deep learning has been progressively adopted across a spectrum of applications. These include the automated recognition of standard echocardiographic sections, precise quantification of cardiac functions, systematic evaluation of valvular pathologies, and the intricate diagnosis and differentiation of cardiac diseases.

A. IMAGE RECOGNITION

In the realm of echocardiography, DL is now heralded as a robust and comprehensive instrument, essential for
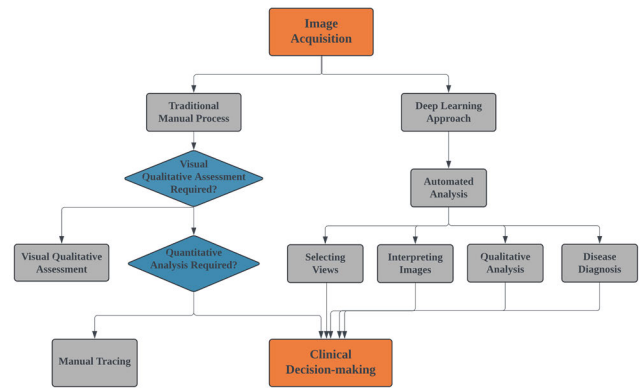

FIGURE 3. Comparison between traditional manual process and DL approach.

formulating clinical verdicts. Central to this advancement is the capacity for image recognition—critical for fully automated assessment of cardiac function. The initial classification of standard imagery during preprocessing is fundamental, laying the groundwork for the ensuing automated analysis. This field is characterized by a variety of scholarly methods developed to address the complexities inherent in both static and dynamic images.

1) STILL FRAME APPROACHES

Employing spatiotemporal feature extraction and supervised learning, a segmentation algorithm has been developed to autonomously discern and categorize key echocardiographic views—apical two-chamber and four-chamber, as well as parasternal long-axis—with impressive accuracy rates of 97%, 91%, and 97% respectively, culminating in an average recognition rate of 95% [17].

To anchor the training, validation, and refinement of deep learning models in a robust standard, a qualitative assessment framework for the apical four-chamber view was instituted. An expert-derived scale from 0 to 5 gauges image quality (0 indicating minimal detail, up to 5 indicating comprehensive chamber visualization). The model's proficiency is underscored by a minimal average absolute error margin of 0.71 ± 0.58 when juxtaposed with expert evaluations [18].

Further enhancing this domain, Vaseli et al. [19] introduced a streamlined classification model for echocardiographic analysis, utilizing sophisticated architectures like VGG-16, DenseNet, and ResNet via knowledge distillation techniques. The model showcased an 88.1% accuracy in identifying 12 canonical views and boasts compatibility with mobile technology, offering the potential for immediate diagnostic application, a crucial feature for urgent care settings.

The integration of Convolutional Neural Network (CNN) models has notably advanced the automated quality assessment of echocardiographic views. Illustrated in Fig. 4 is a schematic representation of a CNN model tailored for the classification of echocardiographic images. Wu et al. [20] and colleagues pioneered a knowledge distillation-based CNN,

specifically aimed at recognizing standard echocardiographic views to evaluate the precision and practicality of diagnosing congenital heart disease in pediatric patients. The model's training and validation capitalized on an extensive set of 350,654 echocardiographic images from 3,505 individuals, identifying 23 standard views frequently employed in the diagnosis of pediatric congenital heart conditions. The model achieved F1 scores exceeding 0.90 for the majority of these views, which include crucial perspectives such as the subcostal coronal/sagittal views of the atrial septum and the parasternal short-axis views at the level of the mitral valve.
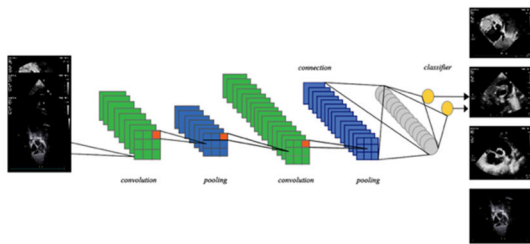

FIGURE 4. CNN model for classification.

In a related development, Madani et al. [21] proficiently trained a CNN to categorize 15 quintessential echocardiographic views, encompassing both video and still images, with the model attaining accuracies of 97.8% for video and 91.7% for still images. Validation exercises, including masking of the cardiac regions in images, underscored the model's dependency on the visualization of the cardiac structure for accurate classification, with performance diminishing substantially when the heart was obscured, underscoring the model's reliance on cardiac features for precise identification.

2) DYNAMIC FRAME APPROACHES

Dynamic frame methodologies in deep learning confront the challenge of lacking cohesive temporal and spatial data across image sequences, which is essential for tasks like detecting motion abnormalities. Pioneering strides in this realm involve the creation of algorithms tailored to compensate for these informational gaps.

Gao et al. [22] innovatively deployed a pair of 2D CNNs to harness the dynamism of video data, markedly achieving a classification correctness of 92.1%, thus outperforming traditional manual techniques. Fig. 5 illustrates this achievement. Extending these advancements, Shahin and Almotairi [23] developed an intricate classification mechanism, leveraging the combined spatial and temporal analytics afforded by ResNet and LSTM algorithms. This synergy not only bolstered the accuracy but also streamlined the training process. The constructed system exhibited formidable precision, automatically discerning standard cardiac ultrasound views with an accuracy of 96.3% and a sensitivity of 95.75%. Notably, it demonstrated an exceptional 99.1% accuracy rate in identifying standard views from various physiological

positions, including the apical region, parasternal long axis, and parasternal short axis, showcasing the profound impact of integrating spatial-temporal features in deep learning models.
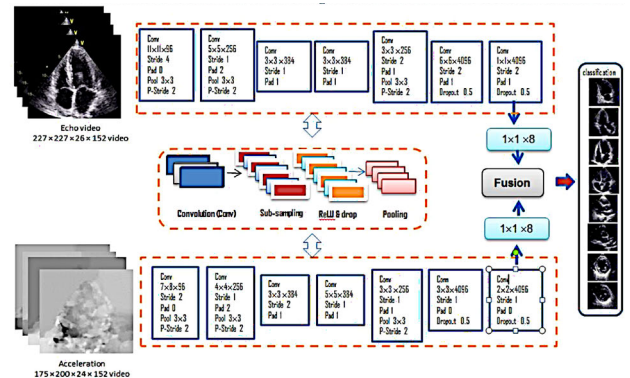

FIGURE 5. Architecture for fusing spatial and temporal information (Gao et al. [23], 2017).

In an innovative step, Huang et al. [24] crafted a 3D convolutional algorithm that synergizes temporal and spatial data, initially deployed to discern and curate cardiac ultrasound videos across four established views: the parasternal long-axis, short-axis at the papillary muscle level, and the apical 4-chamber and 2-chamber views. The algorithm's capabilities were further harnessed to segment and annotate the left ventricular wall, culminating in the nuanced assessment of ventricular wall motion abnormalities through pre- and post-segmentation video analysis.

Building on these advancements, Kumar et al. [25] unveiled a rank-based feature selection strategy within the R-DCNN framework, attaining a remarkable 96.7% accuracy in classifying cardiac ultrasound views, thereby surpassing traditional methods like Support Vector Machines (SVM) and Machine Learning (ML) boosting. In parallel, the CNN-derived CVC methodology [26] demonstrated exceptional prowess in classifying seven cardiac ultrasound views with overall accuracies of (98.3 ± 0.6) % for single-frame and (98.9±0.6) % for sequential imaging modalities.

Moreover, introducing a trailblazing self-supervised technique, Echo-SyncNet [27] was developed to ensure precise video synchronization between apical two-chamber and apical four-chamber views, a breakthrough depicted in Fig. 6.

Building on contemporary research, Ye et al. [28] introduced the innovative ECHO-Attention model, which integrates an attention mechanism to enhance the recognition of standard echocardiographic views. This model was trained and validated on a comprehensive dataset comprising 2,693 ultrasound videos from 267 patients, achieving an impressive overall recognition accuracy of 94.81%. Specifically, the model distinguished between the closely related views of the parasternal short-axis apical and the parasternal short-axis
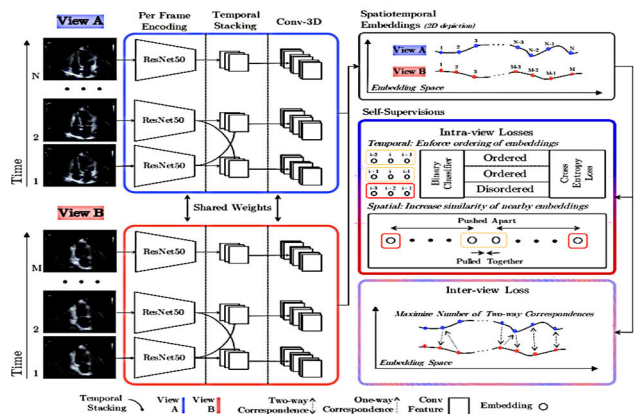
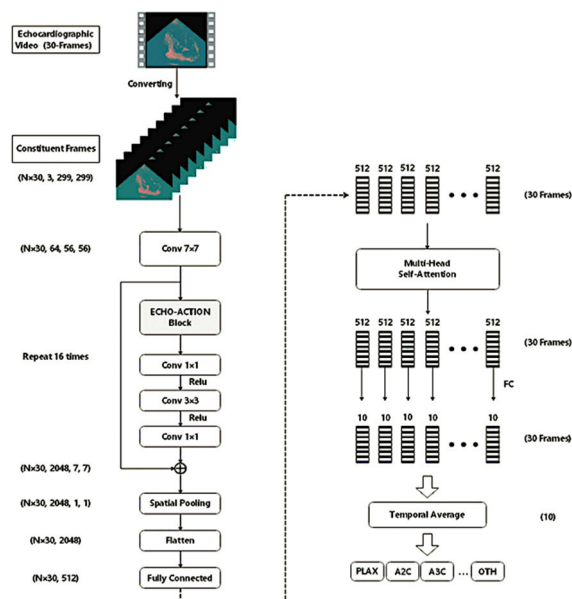**FIGURE 6.** General framework of Echo-SyncNet (Dezaki et al. [27], 2021).

FIGURE 7. ECHO-Attention for ResNet-50 architecture (Ye et al. [28], 2022).

nipple with accuracies of 88.65% and 81.70%, respectively, as illustrated in Fig. 7.

Zhu et al. [29] innovated with EchoV-Net, a profound deep residual network designed to autonomously discern various echocardiogram views, both contrast-enhanced and without contrast. This includes the apical two-chamber, three-chamber, and four-chamber views, alongside the parasternal short axis of the left ventricle (LV). The model exhibited stellar performance, attaining commendable metrics across the board—achieving an average overall accuracy, recall, precision, specificity, and F1 score of 97.0%, 96.9%, 96.9%, 100.0%, and 96.9%, respectively.

The challenge of identifying fetal cardiac sections is amplified due to the diminutive size of the fetal heart and indistinct anatomical structures on ultrasound imagery. A particular study [30] addressed this issue by employing the

FCN-16stride model, which realized a fetal cardiac ultrasound classification error rate of 23.48%.

Moreover, deep learning software has demonstrated efficacy in supporting nursing staff to capture high-quality echocardiographic images. These images span a range of cardiac views, such as the parasternal long-axis, short-axis, and a suite of apical chamber views, as well as subcostal four-chamber and the Inferior Vena Cava (IVC) views, instrumental in assessing left ventricular dimensions. Significantly, the software enabled measurements pertinent to cardiac functionality, pericardial effusion, and right ventricular size, achieving a level of precision on par with that of expert sonographers [31].

B. AUTOMATED QUANTIFICATION OF CARDIAC FUNCTION

The meticulous quantification of left ventricular size and functional assessment is a critical goal in echocardiography. Prior to the integration of ML and DL, deformation modeling provided promising avenues for edge detection, segmentation, and tracking of cardiac motions [32]. Contemporary research confirms that DL models have streamlined the measurement process, bolstering repeatability, closing gaps in expert knowledge, and enhancing overall efficiency [33]. Advancements in DL-based automated software now offer rapid and reliable metrics for left ventricular (LV) volumes, ejection fractions, and biplane longitudinal strains, significantly diminishing subjective variability in interpretation [34], [35], [36], [37], [38].

1) LV VOLUME QUANTIFICATION

Precision in segmentation is pivotal for an accurate functional evaluation. Leclerc et al. [39] conducted a comparative study, revealing that DL algorithms, particularly U-Net and its variants, outperform traditional ML methods in 2D echocardiogram segmentation, demonstrating remarkable accuracy even under challenging imaging conditions. Building on this, the LU-Net [40] emerged, incorporating an attention network to reinforce the precision of LV segmentation, paralleling expert evaluations with a high average similarity score and minimal error margin.

2) EJECTION FRACTION CALCULATION

Ouyang et al. [41] deployed a DL model that utilized a spatio-temporal CNN structure to predict ejection fractions with minimal error and high reliability in identifying cases of heart failure with decreased ejection fractions, as validated by external datasets. Further innovating in this field, Tian et al. [42] applied the Combined Channel and Spatial Attention Mechanism (CBAM) alongside U-Net for cardiac segmentation, enhancing the model's diagnostic focus. Moreover, Sarkar and Chandra [43] employed the ResNet50 model for patient classification based on ejection fraction, achieving high accuracy and F1 scores, showcasing the model's capacity to distinguish between normal and reduced ejection fraction statuses.

3) BIPLANE LONGITUDINAL STRAIN ANALYSIS

Salte et al. [38] engineered a sophisticated multi-network system encompassing a classification network to identify standard echocardiographic views, a temporal network to demarcate systolic and diastolic phases, a segmentation network dedicated to isolating the LV, and an optic flow network designed to map out the LV optic flow field. This integrative strategy has enhanced the accuracy of longitudinal strain calculations, delivering performance on par with commonly employed semi-automated systems. Building upon these innovations, Østvik et al. [44] introduced a groundbreaking fully automated myocardial functional imaging system derived from the traditional PWC-Net framework, as demonstrated in Fig. 8. This system adeptly performs automatic estimations of the LV's longitudinal strain, showcasing the potential of deep learning applications to refine and automate complex cardiac imaging analyses.
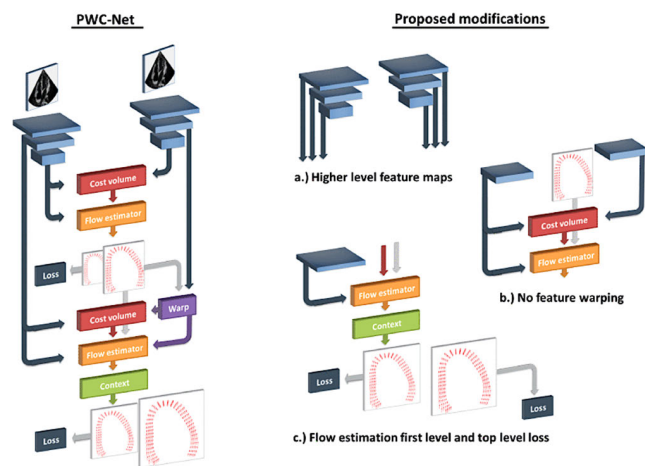
FIGURE 8. Traditional PWC-Net architecture and EchoPWC-Net (Østvik et al. [44],2021).

Romero-Pacheco et al. [45] have put forward a novel methodology for calculating the overall longitudinal strain of the left ventricle (LV) in echocardiographic analysis, leveraging the capabilities of the GMA algorithm. Complementing this, Tromp et al. [46] harnessed the ATTRACT platform to develop an advanced CNN-based intelligent recognition system. This system is adept at classifying, segmenting, and quantifying echocardiographic video data, enabling the autonomous annotation of 2D cardiac echocardiography and Doppler modalities. It proficiently quantifies cardiac chamber volumes, assesses LV systolic and diastolic functions, and precisely annotates the endocardium and chambers of the LV and left atrium across apical 2-chamber and 4-chamber views. Importantly, the system's reliability has been corroborated through validation on datasets spanning Canada, Taiwan, China, and the United States, underscoring the dependability of automated diagnostic methods.

In terms of 3D echocardiography, pioneering efforts have been made to segment the left ventricle utilizing a depth

mapping network—marking the first foray into high-dimensional LV segmentation with minimal annotated data [47]. In a significant leap forward, Laumer et al. [48] unveiled an innovative technique that transforms single-view 2D echocardiographic videos into a personalized 4D cardiac mesh, using the novel 4DHM architecture depicted in Fig. 9. This method, using solely 2D echocardiograms as input, establishes a mapping function between two distinct visual domains: echocardiographic video and cardiac mesh video. This self-supervised approach generates a dynamic 4D cardiac mesh capable of autonomously extracting key echocardiographic variables, including ejection fraction, myocardial mass, and volumetric alterations in ventricular volume, all with remarkable temporal precision.
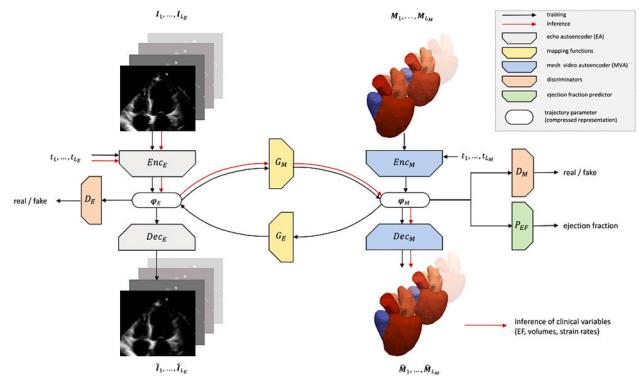
FIGURE 9. System architecture of 4DHM (Laumer et al. [48],2023).

C. AUTOMATED ASSESSMENT OF VALVULAR DISEASE

Contemporary echocardiographic guidelines advocate for a comprehensive evaluation of valvular regurgitation, integrating both semiquantitative and quantitative methods to precisely ascertain regurgitation severity [49]. Prominently, the clinical adoption of the proximal iso-velocity surface area (PISA) method is encouraged. While traditional 2D echocardiographic PISA approaches are constrained by assumptions of hemispheric symmetry, 3D echocardiographic PISA has demonstrated superior correlation with cardiac magnetic resonance (CMR), enhancing the accuracy in determining regurgitant severity [50].

De Agustín et al. [51] showcased the effectiveness of automated software for quantifying mitral PISA within 3D echocardiography. Additionally, using color Doppler echocardiography to measure aortic regurgitation volume revealed a strong positive correlation with CMR findings, providing a more detailed representation of aortic regurgitation severity compared to 2D echocardiographic PISA. Moreover, automated transesophageal 3D echocardiographic mitral valve analysis software has proven to be more reliable and consistent for the quantitative evaluation of the mitral annulus, surpassing manual assessment in both size and morphological precision [53].

The emergence of AI technology in valvular disease assessment presents a promising avenue, minimizing procedural complexities and significantly enhancing the precision and reliability of quantitative evaluations. The advancements and their implications are vividly illustrated in Fig. 10.
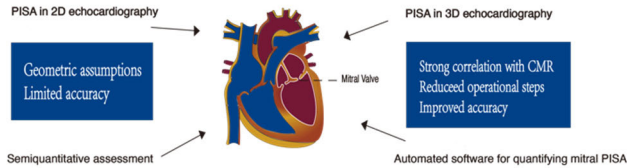

FIGURE 10. Comparison of PISA between 2D and 3D.

In a notable application of deep learning to echocardiography, one study [54] transformed transesophageal 3D echocardiograms into 2D planar representations to enable automatic mitral annulus segmentation. The developed method demonstrated potential for assisting mitral valve treatments, evidencing an average error of 2.0 mm with a standard deviation of 1.9 mm. Expanding the scope of computational cardiology, Herz et al. [55] implemented the V-Net architecture [56], employing a fully convolutional network to delineate left heart dysplasia and evaluate the tricuspid valve, demonstrating the versatility of deep learning in complex cardiac segmentation tasks.

Yu et al. [57] designed a dynamic CNN that capitalizes on multi-scale data to precisely track the mitral valve, refining fetal left ventricle segmentation with an average Dice similarity coefficient of 0.945 in testing scenarios. In a similar vein, Chandra et al. [58] integrated the Yolo3 detection system to monitor the mitral valve through echocardiograms, achieving remarkable detection accuracies of 98% for the mitral valve and 90% for the tricuspid valve.

Further illustrating the power of deep learning in cardiology, Vafaeezadeh et al. [59] developed and trained a suite of 13 models specializing in the identification of prosthetic valves. Their models achieved an impressive AUC of 98%, showcasing performance on par with seasoned cardiologists. The efficacy of their deep learning framework is visually represented in Fig. 11, providing a compelling testament to the advancements in automated cardiac diagnostics.

Guided by the 2017 American Society of Echocardiography guidelines, which delineate mitral regurgitation (MR) into four distinct grades based on severity, Zhang et al. [60] innovated the labeling process of color Doppler echocardiographic images using LabelMe software. Following the labeling, they harnessed the capabilities of the Mask R-CNN model for the automated evaluation of MR severity. The accuracy of their model in predicting the correct grade of MR was impressive, achieving precision scores of 0.90, 0.87, 0.81, and 0.91 for grades I through IV, respectively. These performance metrics are depicted in Fig. 12, demonstrating the precision and reliability of their deep learning solution in clinical diagnostics.
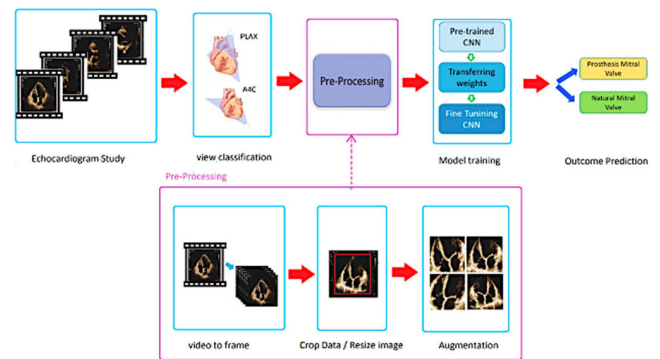

FIGURE 11. DL framework for prosthetic mitral valve recognition (Yu et al. [59], 2017).
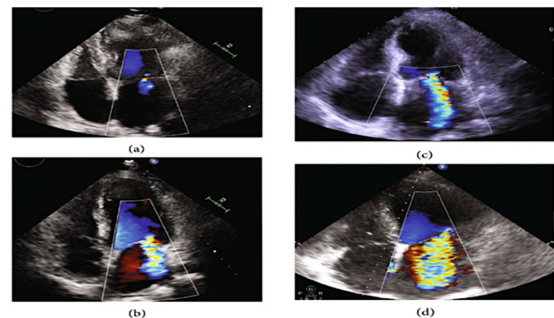

FIGURE 12. The original image of MR of grade I to grade IV (Zhang et al. [60], 2021).

Yang et al. [61] have made strides in echocardiographic analysis by devising a self-supervised learning algorithm, utilizing an apical four-chamber view dataset for the evaluation of mitral regurgitation severity on 2D color Doppler echocardiograms. To reduce the labor-intensive task of manual annotation, they pre-trained and introduced the CD-SSL model, which not only surpassed the traditional ResNet-UNet model with an accuracy of 95.9% for identifying the MR jet region but also offered six quantitative metrics. These metrics—proportions and areas such as MR jet length relative to left atrium (LA) length, MR jet length, LA width, LA region, MR jet area, and the MR jet region relative to LA region—provide a robust framework for aiding clinicians in diagnostic processes.

Complementing these technological advances, Edwards et al. [62] focused on pediatric echocardiography, crafting a deep learning model aimed at detecting mitral regurgitation for rheumatic heart disease screening among children. This model demonstrated an impressive 98% accuracy in classifying standard echocardiographic views and an 86% success rate in pinpointing mitral regurgitation, underscoring the potential of DL models in enhancing diagnostic precision in rheumatic cardiology.

D. DIAGNOSIS AND DIFFERENTIATION OF DISEASE

Hong et al. [63] have crafted a CNN tailored to color Doppler echocardiography images, automating the detection of

secundum atrial septal defects in pediatric patients. Their system executes a tripartite strategy: initially, it identifies four crucial echocardiographic views that include the subcostal view targeting the atrial septum, the apical and low paraspinal 4-chamber views, and the parasternal short-axis view. The subsequent phase involves meticulous segmentation of pertinent cardiac structures and identification of probable atrial septal defect sites. The final phase integrates the segmentation and detection data to conclude the diagnostic process. Through rigorous development and validation phases, the system showcased commendable image-level performance, with averages of 0.8545 for recall, 0.8577 for precision, 0.9136 for specificity, and 0.8546 for the F1 score, respectively. These findings illuminate the vast potential of machine learning to revolutionize the intelligent diagnosis of congenital heart conditions.

1) MYOCARDIAL HYPERTROPHY

Differentiating between various pathological conditions that share echocardiographic characteristics is often challenging, especially for less experienced practitioners. Cardiac amyloidosis and hypertrophic cardiomyopathy, for instance, both manifest as myocardial hypertrophy in imaging studies [64]. In a significant research effort analyzing 23,745 cardiac ultrasound examinations, a sophisticated end-to-end model was designed to evaluate left ventricular (LV) dimensions and wall thickness, with the specific intent to diagnose left ventricular hypertrophy (LVH). The implementation of a 3-dimensional CNN with residual learning pathways enabled this model to accurately differentiate cardiac amyloidosis (with an AUC of 0.83) from hypertrophic cardiomyopathy (with an AUC of 0.98) and other LVH etiologies [65].

Complementing these advancements, Yu et al. [66] developed a semi-automatic system that applies ResNet for classification tasks and U-net++ for segmentation tasks. The resultant integrated system, which is visually detailed in Fig. 13, proficiently sorts echocardiographic data into distinct categories, effectively discerning normal cardiac anatomy from conditions like hypertrophic cardiomyopathy (HCM), cardiac amyloidosis (CA), and hypertensive heart disease (HHD).

Xu et al. [67] leveraged transthoracic echocardiography (TTE) video data to advance the differential diagnosis of pathological left ventricular hypertrophy, distinguishing between hypertrophic cardiomyopathy (HCM), hypertensive heart disease (HHD), and uremic cardiomyopathy (UCM) with a nuanced approach. Furthering this field, Zhang et al. [68] engineered an automated system that revolutionizes cardiac ultrasound analysis. This versatile model proficiently performs multiple tasks: it recognizes 23 different cardiac ultrasound image sections, achieves segmentation of common cardiac chambers with approximately 85% accuracy, accurately computes cardiac chamber volumes and ejection fraction, and adeptly detects complex conditions
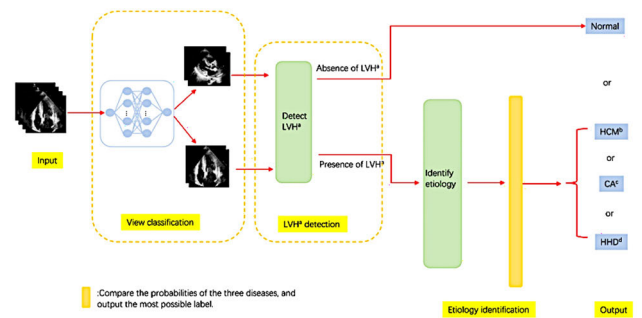

FIGURE 13. Identification of different disease frameworks for LV hypertrophy (Yu et al. [66], 2022).

such as hypertrophic cardiomyopathy, cardiac amyloidosis, and pulmonary hypertension with an overall view classification accuracy of 84%.

These pioneering studies showcase the remarkable capacity of deep learning systems to not only automate intricate diagnostic procedures but also provide accurate and comprehensive cardiac assessments, representing significant progress in the use of AI for cardiac diagnostics.

2) MYOCARDIAL INFARCTION

Kusunose et al. [69] carried out a pioneering study using short-axis echocardiographic views from 300 patients with a history of myocardial infarction alongside 100 patients exhibiting normal ventricular motion. Their model, feeding on a triad of images from distinct cardiac phases, demonstrated remarkable concordance with expert human assessment in detecting regional wall motion abnormalities, with a ResNet model achieving an AUC of 0.97.

Expanding upon these findings, Muraki et al. [70] developed a sophisticated model integrating CNNs with LSTM networks to differentiate Acute Myocardial Infarction (AMI) from normal myocardial conditions. The model attained overall accuracies of 83.2% and 85.1% for classifying myocardial states in short-axis and left ventricular long-axis views, respectively, underscoring the profound impact of DL in improving AMI detection and categorization.

Moreover, Hamlia et al. [71] engineered a fully automated CNN-based model capable of real-time early MI prediction via echocardiography. This model utilized a 2D CNN to segment the left ventricle with a stellar 97.18% accuracy from the apical four-chamber view and deployed a 3D CNN for myocardial infarction classification, attaining an accuracy of 90.9%. These figures notably exceed those achieved by prior active polynomial methods, which score.d 87.94% accuracy [72].

3) OTHER HEART DISEASES

Martins et al. [73] leveraged the capabilities of 3D CNNs along with supervised integration techniques to

autonomously identify rheumatic heart disease (RHD), achieving a diagnostic accuracy of 72.77%. Building on these technological advances, Han et al. [74] introduced an innovative dual-network model that integrates a spatial attention module. This model aids in the precise segmentation of abnormal cardiac structures, thereby streamlining the early screening process for congenital heart disease (CHD), a critical step in pediatric cardiology.

Complementing this progress, Nova et al. [75] employed a U-Net-based model specializing in the segmentation of infant cardiac septal anomalies. Their model demonstrated exceptional proficiency, distinguishing atrial septal defects (ASDs), ventricular septal defects (VSDs), atrioventricular septal defects (AVSDs), and normal cardiac anatomy with accuracies of 99.05%, 98.62%, 99.39%, and 98.97%, respectively.

IV. CONCLUSION

Echocardiography, a crucial noninvasive tool for cardiac assessment, traditionally depends on the expertise of the interpreting physicians. Models such as ResaNet and U-Net have risen to prominence, revolutionizing both segmentation and classification within this field. With a wave of innovation, these models are at the heart of automated systems, some enhanced by attention mechanisms for superior performance. Yet, there remains a noticeable research gap in the segmentation and classification of the right heart and related pathologies.

A recent landmark study [76] has validated DL's transformative power in echocardiogram interpretation, benchmarking DL-generated cardiac parameters against those of human sonographers. The findings illustrate DL's ability to boost interpretive efficiency, producing more consistent and precise outcomes while curtailing the time required for analysis. Despite concerns over DL's role in the workforce, it emerges not as a replacement but as an invaluable support to clinicians, especially those at the early stages of their careers, and extends its utility to noncardiac ultrasound experts.

However, the journey of DL in echocardiography is just beginning. The promise of unsupervised learning looms on the horizon, hinting at untapped potentials in autonomous echocardiogram analysis. Hearteningly, DL's scope is already expanding into the realm of fetal echocardiography, a testament to its clinical versatility [77].

In conclusion, despite the significant progress made, DL's path is one of ongoing discovery and enhancement. As it evolves, the expectation is that DL will redefine echocardiography, elevating diagnostic precision, operational efficiency, and broader access, all to the greater benefit of patient care and clinical outcomes.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[2] P. S. Douglas, M. J. Garcia, D. E. Haines, W. W. Lai, W. J. Manning, A. R. Patel, M. H. Picard, D. M. Polk, M. Ragosta, R. P. Ward, and R. B. Weiner, "ACCF/ASE/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 appropriate use criteria for echocardiography," *J. Amer. Soc. Echocardiography*, vol. 24, no. 3, pp. 229–267, Mar. 2011, doi: 10.1016/j.echo.2010.12.008.

[3] R. Hoffmann, H. Lethen, T. Marwick, M. Arnese, P. Fioretti, A. Pingitore, E. Picano, T. Buck, R. Erbel, F. A. Flachskampf, and P. Hanrath, "Analysis of interinstitutional observer agreement in interpretation of dobutamine stress echocardiograms," *J. Amer. College Cardiology*, vol. 27, no. 2, pp. 330–336, Feb. 1996, doi: 10.1016/0735-1097(95)00483-1.

[4] J.-G. Lee, S. Jun, Y.-W. Cho, H. Lee, G. B. Kim, J. B. Seo, and N. Kim, "Deep learning in medical imaging: General overview," *Korean J. Radiol.*, vol. 18, no. 4, p. 570, Jul. 2017, doi: 10.3348/kjr.2017.18.4.570.

[5] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[6] R. Dou, "Application of deep learning algorithms in medical imaging and research progress," *Chin. J. Med. Imag. Technol.*, vol. 24, no. 5, pp. 369–372, Oct. 2018, doi: 10.19627/j.cnki.cn31-1700/th.2018.05.001.

[7] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017, doi: 10.1038/nature21056.

[8] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. Mckeown, G. Yang, X. Wu, and F. Yan, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, Feb. 2018, doi: 10.1016/j.cell.2018.02.010.

[9] L. Xia, J. Shen, R. Zhang, S. Wang, and K. Chen, "Application of deep learning techniques in medical imaging," *Concord Med. J.*, vol. 9, no. 1, pp. 10–14, Feb. 2018, doi: 10.3969/j.issn.1674-9081.2018.01.003.

[10] F. Cabitza, R. Rasoini, and G. F. Gensini, "Unintended consequences of machine learning in medicine," *JAMA*, vol. 318, no. 6, p. 517, Aug. 2017, doi: 10.1001/jama.2017.7797.

[11] Q. Zhang, D. Zhao, and X. Chi, "A review of deep learning-based medical image diagnosis," *Comput. Sci.*, vol. 44, no. S2, pp. 1–7, Dec. 2017.

[12] P. Tao, Z. Fu, K. Zhu, and L. Wang, "Deep learning based echocardiographic cut surface recognition method," *Comput. Appl.*, vol. 37, no. 5, pp. 1434–1438, May 2017, doi: 10.11772/j.issn.1001-9081.2017.05.1434.

[13] R. M. Lang, L. P. Badano, V. Mor-Avi, J. Afilalo, A. Armstrong, L. Ernande, F. A. Flachskampf, E. Foster, S. A. Goldstein, T. Kuznetsova, P. Lancellotti, D. Muraru, M. H. Picard, E. R. Rietzschel, L. Rudski, K. T. Spencer, W. Tsang, and J.-U. Voigt, "Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American society of echocardiography and the European association of cardiovascular imaging," *Eur. Heart J. Cardiovascular Imag.*, vol. 16, no. 3, pp. 233–271, Feb. 2015, doi: 10.1093/ehjci/jev014.

[14] P. P. Sengupta and A. A. Adjeroh, "Will artificial intelligence replace the human echocardiographer? Clinical considerations," *Circulation*, vol. 138, no. 16, pp. 1639–1642, Oct. 2018, doi: 10.1161/circulationaha.118.037095.

[15] J. M. DeCara, R. M. Lang, R. Koch, R. Bala, J. Penzotti, and K. T. Spencer, "The use of small personal ultrasound devices by internists without formal training in echocardiography," *Eur. Heart J. Cardiovascular Imag.*, vol. 4, no. 2, pp. 141–147, Jun. 2003, doi: 10.1053/euje.4.2.141.

[16] A. Davis, K. Billick, K. Horton, M. Jankowski, P. Knoll, J. E. Marshall, A. Paloma, R. Palma, and D. B. Adams, "Artificial intelligence and echocardiography: A primer for cardiac sonographers," *J. Amer. Soc. Echocardiography*, vol. 33, no. 9, pp. 1061–1066, Sep. 2020, doi: 10.1016/j.echo.2020.04.025.

[17] H. Khamis, G. Zurakhov, V. Azar, A. Raz, Z. Friedman, and D. Adam, "Automatic apical view classification of echocardiograms using a discriminative learning dictionary," *Med. Image Anal.*, vol. 36, pp. 15–21, Feb. 2017, doi: 10.1016/j.media.2016.10.007.

[18] A. H. Abdi, C. Luong, T. Tsang, G. Allan, S. Nouranian, J. Jue, D. Hawley, S. Fleming, K. Gin, J. Swift, R. Rohling, and P. Abolmaesumi, "Automatic quality assessment of echocardiograms using convolutional neural networks: Feasibility on the apical four-chamber view," *IEEE Trans. Med. Imag.*, vol. 36, no. 6, pp. 1221–1230, Jun. 2017, doi: 10.1109/TMI.2017.2690836.

[19] H. Vaseli, Z. Liao, A. H. Abdi, H. Girgis, D. Behnami, C. Luong, F. Taheri Dezaki, N. Dhungel, R. Rohling, K. Gin, P. Abolmaesumi, and T. Tsang, "Designing lightweight deep learning models for echocardiography view classification," in *Proc. Med. Imag. : Image-Guided Procedures, Robotic Interventions, Model.* Bellingham, WA, USA: SPIE, Mar. 2019, pp. 1–14.

[20] L. Wu, B. Dong, X. Liu, W. Hong, L. Chen, K. Gao, Q. Sheng, Y. Yu, L. Zhao, and Y. Zhang, "Standard echocardiographic view recognition in diagnosis of congenital heart defects in children using deep learning based on knowledge distillation," *Frontiers Pediatrics*, vol. 9, Jan. 2022, Art. no. 770182, doi: 10.3389/fped.2021.770182.

[21] A. Madani, R. Arnaout, M. Mofrad, and R. Arnaout, "Fast and accurate view classification of cardiac echocardiograms using deep learning," *Circulation*, vol. 136, no. suppl_1, Mar. 2017, Art. no. A18074, doi: 10.1161/circ.136.suppl_1.18074.

[22] X. Gao, W. Li, M. Loomes, and L. Wang, "A fused deep learning architecture for viewpoint classification of echocardiography," *Inf. Fusion*, vol. 36, pp. 103–113, Jul. 2017, doi: 10.1016/j.inffus.2016.11.007.

[23] A. I. Shahin and S. Almotairi, "An accurate and fast cardio-views classification system based on fused deep features and LSTM," *IEEE Access*, vol. 8, pp. 135184–135194, 2020, doi: 10.1109/ACCESS.2020.3010326.

[24] M.-S. Huang, C.-S. Wang, J.-H. Chiang, P.-Y. Liu, and W.-C. Tsai, "Automated recognition of regional wall motion abnormalities through deep neural network interpretation of transthoracic echocardiography," *Circulation*, vol. 142, no. 16, pp. 1510–1520, Oct. 2020, doi: 10.1161/circulationaha.120.047530.

[25] B. P. Santosh Kumar, M. A. Haq, P. Sreenivasulu, D. Siva, M. B. Alazzam, F. Alassery, and S. Karupusamy, "Fine-tuned convolutional neural network for different cardiac view classification," *J. Supercomput.*, vol. 78, no. 16, pp. 18318–18335, Jun. 2022, doi: 10.1007/s11227-022-04587-0.

[26] A. Østvik, E. Smistad, S. A. Aase, B. O. Haugen, and L. Lovstakken, "Real-time standard view classification in transthoracic echocardiography using convolutional neural networks," *Ultrasound Med. Biol.*, vol. 45, no. 2, pp. 374–384, Feb. 2019, doi: 10.1016/j.ultrasmedbio.2018.07.024.

[27] F. T. Dezaki, C. Luong, T. Ginsberg, R. Rohling, K. Gin, P. Abolmaesumi, and T. Tsang, "Echo-SyncNet: Self-supervised cardiac view synchronization in echocardiography," *IEEE Trans. Med. Imag.*, vol. 40, no. 8, pp. 2092–2104, Aug. 2021, doi: 10.1109/TMI.2021.3071951.

[28] Z. Ye, Y. Jaya Kumar, G. Ong Sing, F. Song, and X. Ni, "Classification of echocardiographic standard views using a hybrid attention-based approach," *Intell. Autom. Soft Comput.*, vol. 33, no. 2, pp. 1197–1215, Jan. 2022, doi: 10.32604/iasc.2022.023555.

[29] Y. Zhu, J. Ma, Z. Zhang, Y. Zhang, S. Zhu, M. Liu, Z. Zhang, C. Wu, X. Yang and J. Cheng, "Automatic view classification of contrast and non-contrast echocardiography," *Frontiers Cardiovascular Med.*, vol. 9, Sep. 2022, Art. no. 989091, doi: 10.3389/fcvm.2022.989091.

[30] S. Vaanathi, P. C. P. Bridge, C. Ioannou and J. A. Noble, "Automated characterization of the fetal heart in ultrasound images using fully convolutional neural networks," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Piscataway, NJ, USA, 2017, pp. 671–674.

[31] A. Narang, R. Bae, H. Hong, Y. Thomas, S. Surette, C. Cadieu, A. Chaudhry, R. P. Martin, P. M. McCarthy, D. S. Rubenson, S. Goldstein, S. H. Little, R. M. Lang, N. J. Weissman, and J. D. Thomas, "Utility of a deep-learning algorithm to guide novices to acquire echocardiograms for limited diagnostic use," *JAMA Cardiology*, vol. 6, no. 6, p. 624, Jun. 2021, doi: 10.1001/jamacardio.2021.0185.

[32] S. K. Zhou, "Shape regression machine and efficient segmentation of left ventricle endocardium from 2D B-mode echocardiogram," *Med. Image Anal.*, vol. 14, no. 4, pp. 563–581, Aug. 2010, doi: 10.1016/j.media.2010.04.002.

[33] M. Cannesson, M. Tanabe, M. S. Suffoletto, D. M. McNamara, S. Madan, J. M. Lacomis, and J. Gorcsan, "A novel two-dimensional echocardiographic image analysis system using artificial intelligence-learned pattern recognition for rapid automated ejection fraction," *J. Amer. College Cardiology*, vol. 49, no. 2, pp. 217–226, Jan. 2007, doi: 10.1016/j.jacc.2006.08.045.

[34] F. Levy, E. Dan Schouver, L. Iacuzio, F. Civaia, S. Rusek, C. Dommerc, S. Marechaux, V. Dor, C. Tribouilloy, and G. Dreyfus, "Performance of new automated transthoracic three-dimensional echocardiographic software for left ventricular volumes and function assessment in routine clinical practice: Comparison with 3 Tesla cardiac magnetic resonance," *Arch. Cardiovascular Diseases*, vol. 110, no. 11, pp. 580–589, Nov. 2017, doi: 10.1016/j.acvd.2016.12.015.

[35] G. Tamborini, C. Piazzese, R. M. Lang, M. Muratori, E. Chiorino, M. Mapelli, L. Fusini, S. G. Ali, P. Gripari, G. Pontone, D. Andreini, and M. Pepi, "Feasibility and accuracy of automated software for transthoracic three-dimensional left ventricular volume and function analysis: Comparisons with two-dimensional echocardiography, three-dimensional transthoracic manual method, and cardiac magnetic resonance imaging," *J. Amer. Soc. Echocardiography*, vol. 30, no. 11, pp. 1049–1058, Nov. 2017, doi: 10.1016/j.echo.2017.06.026.

[36] W. Tsang, I. S. Salgo, D. Medvedofsky, M. Takeuchi, D. Prater, L. Weinert, M. Yamat, V. Mor-Avi, A. R. Patel, and R. M. Lang, "Transthoracic 3D echocardiographic left heart chamber quantification using an automated adaptive analytics algorithm," *JACC: Cardiovascular Imag.*, vol. 9, no. 7, pp. 769–782, Jul. 2016, doi: 10.1016/j.jcmg.2015.12.020.

[37] H. Kawakami, L. Wright, M. Nolan, E. L. Potter, H. Yang, and T. H. Marwick, "Feasibility, reproducibility, and clinical implications of the novel fully automated assessment for global longitudinal strain," *J. Amer. Soc. Echocardiography*, vol. 34, no. 2, pp. 136–145, Feb. 2021, doi: 10.1016/j.echo.2020.09.011.

[38] I. M. Salte, A. Østvik, E. Smistad, D. Melichova, T. M. Nguyen, S. Karlsen, H. Brunvand, K. H. Haugaa, T. Edvardsen, L. Lovstakken, and B. Grenne, "Artificial intelligence for automatic measurement of left ventricular strain in echocardiography," *JACC: Cardiovascular Imag.*, vol. 14, no. 10, pp. 1918–1928, Oct. 2021, doi: 10.1016/j.jcmg.2021.04.018.

[39] S. Leclerc, E. Smistad, J. Pedrosa, A. Østvik, F. Cervenansky, F. Espinosa, T. Espeland, E. A. R. Berg, P.-M. Jodoin, T. Grenier, C. Lartizien, J. D'hooge, L. Lovstakken, and O. Bernard, "Deep learning for segmentation using an open large-scale dataset in 2D echocardiography," *IEEE Trans. Med. Imag.*, vol. 38, no. 9, pp. 2198–2210, Sep. 2019, doi: 10.1109/TMI.2019.2900516.

[40] S. Leclerc, E. Smistad, A. Østvik, F. Cervenansky, F. Espinosa, T. Espeland, E. A. Rye Berg, M. Belhamissi, S. Israilov, T. Grenier, C. Lartizien, P.-M. Jodoin, L. Lovstakken, and O. Bernard, "LU-net: A multistage attention network to improve the robustness of segmentation of left ventricular structures in 2-D echocardiography," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 12, pp. 2519–2530, Dec. 2020, doi: 10.1109/TUFFC.2020.3003403.

[41] D. Ouyang, B. He, A. Ghorbani, N. Yuan, J. Ebinger, C. P. Langlotz, P. A. Heidenreich, R. A. Harrington, D. H. Liang, E. A. Ashley, and J. Y. Zou, "Video-based AI for beat-to-beat assessment of cardiac function," *Nature*, vol. 580, no. 7802, pp. 252–256, Apr. 2020, doi: 10.1038/s41586-020-2145-8.

[42] M. Tian and M. Zheng, "A heart segmentation algorithm based on dynamic ultrasound," *BioMed Res. Int.*, vol. 2022, pp. 1–8, Jun. 2022, doi: 10.1155/2022/1485584.

[43] P. G. Sarkar and V. Chandra, "A novel approach for detecting abnormality in ejection fraction using transthoracic echocardiography with deep learning," *Int. J. Online Biomed. Eng. (iJOE)*, vol. 16, no. 13, p. 99, Nov. 2020, doi: 10.3991/ijoe.v16i13.18483.

[44] A. Østvik, I. M. Salte, E. Smistad, T. M. Nguyen, D. Melichova, H. Brunvand, K. Haugaa, T. Edvardsen, B. Grenne, and L. Lovstakken, "Myocardial function imaging in echocardiography using deep learning," *IEEE Trans. Med. Imag.*, vol. 40, no. 5, pp. 1340–1351, May 2021, doi: 10.1109/TMI.2021.3054566.

[45] A. Romero-Pacheco, J. Perez-Gonzalez, and N. Hevia-Montiel, "Estimating echocardiographic myocardial strain of left ventricle with deep learning," in *Proc. 44th Annu. Int. Conf. IEEE EMBC*, Glasgow, Scotland, U.K., 2022, pp. 3891–3894.

[46] J. Tromp, P. J. Seekings, C.-L. Hung, M. B. Iversen, M. J. Frost, W. Ouwerkerk, Z. Jiang, F. Eisenhaber, R. S. M. Goh, H. Zhao, W. Huang, L.-H. Ling, D. Sim, P. Cozzone, A. M. Richards, H. K. Lee, S. D. Solomon, C. S. P. Lam, and J. A. Ezekowitz, "Automated interpretation of systolic and diastolic function on the echocardiogram: A multicohort study," *Lancet Digit. Health*, vol. 4, no. 1, pp. e46–e54, Jan. 2022, doi: 10.1016/s2589-7500(21)00235-1.

[47] S. Dong, G. Luo, C. Tam, W. Wang, K. Wang, S. Cao, B. Chen, H. Zhang, and S. Li, "Deep atlas network for efficient 3D left ventricle segmentation on echocardiography," *Med. Image Anal.*, vol. 61, Apr. 2020, Art. no. 101638, doi: 10.1016/j.media.2020.101638.

[48] F. Laumer, M. Amrani, L. Manduchi, A. Beuret, L. Rubi, A. Dubatovka, C. M. Matter, and J. M. Buhmann, "Weakly supervised inference of personalized heart meshes based on echocardiography videos," *Med. Image Anal.*, vol. 83, Jan. 2023, Art. no. 102653, doi: 10.1016/j.media.2022.102653.

[49] W. A. Zoghbi, D. Adams, R. O. Bonow, M. Enriquez-Sarano, E. Foster, P. A. Grayburn, R. T. Hahn, Y. Han, J. Hung, and R. M. Lang, "Recommendations for noninvasive evaluation of native valvular regurgitation: A report from the American Society of Echocardiography developed in collaboration with the Society for Cardiovascular Magnetic Resonance," *J. Amer. Soc. Echocardiography*, vol. 30, no. 4, pp. 303–371, Apr. 2017, doi: 10.1016/j.echo.2017.01.007.

[50] P. Thavendiranathan, S. Liu, S. Datta, S. Rajagopalan, T. Ryan, S. R. Igo, M. S. Jackson, S. H. Little, N. De Michelis, and M. A. Vannan, "Quantification of chronic functional mitral regurgitation by automated 3-Dimensional peak and integrated proximal isovelocity surface area and stroke volume techniques using real-time 3-Dimensional volume color Doppler echocardiography: In vitro and clinical validation," *Circulat., Cardiovascular Imag.*, vol. 6, no. 1, pp. 125–133, Jan. 2013, doi: 10.1161/circimaging.112.980383.

[51] J. A. de Agustín, P. Marcos-Alberca, C. Fernandez-Golfin, A. Gonçalves, G. Feltes, I. J. Nuñez-Gil, C. Almeria, J. L. Rodrigo, L. Perez de Isla, C. Macaya, and J. Zamorano, "Direct measurement of proximal isovelocity surface area by single-beat three-dimensional color Doppler echocardiography in mitral regurgitation: A validation study," *J. Amer. Soc. Echocardiography*, vol. 25, no. 8, pp. 815–823, Aug. 2012, doi: 10.1016/j.echo.2012.05.021.

[52] J. Choi, G.-R. Hong, M. Kim, I. J. Cho, C. Y. Shim, H.-J. Chang, J. Mancina, J.-W. Ha, and N. Chung, "Automatic quantification of aortic regurgitation using 3D full volume color Doppler echocardiography: A validation study with cardiac magnetic resonance imaging," *Int. J. Cardiovascular Imag.*, vol. 31, no. 7, pp. 1379–1389, Oct. 2015, doi: 10.1007/s10554-015-0707-x.

[53] F. Mahmood, J. Jeganathan, Z. Knio, Y. Amador, T. Hai, A. Khamooshian, R. Matyal, and K. Khabbaz, "Artificial intelligence in mitral valve analysis," *Ann. Cardiac Anaesthesia*, vol. 20, no. 2, p. 129, Apr. 2017, doi: 10.4103/aca.aca_243_16.

[54] B. S. Andreassen, F. Veronesi, O. Gerard, A. H. S. Solberg, and E. Samset, "Mitral annulus segmentation using deep learning in 3-D transesophageal echocardiography," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 4, pp. 994–1003, Apr. 2020, doi: 10.1109/JBHI.2019.2959430.

[55] C. Herz, D. F. Pace, H. H. Nam, A. Lasso, P. Dinh, M. Flynn, A. Cianciulli, P. Golland, and M. A. Jolley, "Segmentation of tricuspid valve leaflets from transthoracic 3D echocardiograms of children with hypoplastic left heart syndrome using deep learning," *Frontiers Cardiovascular Med.*, vol. 8, Dec. 2021, Art. no. 735587, doi: 10.3389/fcvm.2021.735587.

[56] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Stanford, CA, USA, Oct. 2016, pp. 565–571.

[57] L. Yu, Y. Guo, Y. Wang, J. Yu, and P. Chen, "Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 8, pp. 1886–1895, Aug. 2017, doi: 10.1109/TBME.2016.2628401.

[58] V. Chandra, P. G. Sarkar, and V. Singh, "Mitral valve leaflet tracking in echocardiography using custom Yolo3," *Procedia Comput. Sci.*, vol. 171, pp. 820–828, Jan. 2020, doi: 10.1016/j.procs.2020.04.089.

[59] M. Vafaeezadeh, H. Behnam, A. Hosseinsabet, and P. Gifani, "A deep learning approach for the automatic recognition of prosthetic mitral valve in echocardiographic images," *Comput. Biol. Med.*, vol. 133, Jun. 2021, Art. no. 104388, doi: 10.1016/j.compbiomed.2021.104388.

[60] Q. Zhang, Y. Liu, J. Mi, X. Wang, X. Liu, F. Zhao, C. Xie, P. Cui, Q. Zhang, and X. Zhu, "Automatic assessment of mitral regurgitation severity using the mask R-CNN algorithm with color Doppler echocardiography images," *Comput. Math. Methods Med.*, vol. 2021, pp. 1–10, Sep. 2021, doi: 10.1155/2021/2602688.

[61] F. Yang, J. Zhu, J. Wang, L. Zhang, W. Wang, X. Chen, X. Lin, Q. Wang, D. Burkhoff, S. K. Zhou, and K. He, "Self-supervised learning assisted diagnosis for mitral regurgitation severity classification based on color Doppler echocardiography," *Ann. Translational Med.*, vol. 10, no. 1, pp. 1–13, Jan. 2022, doi: 10.21037/atm-21-3449.

[62] L. A. Edwards, F. Feng, M. Iqbal, Y. Fu, A. Sanyahumbi, S. Hao, D. B. McElhinney, X. B. Ling, C. Sable, and J. Luo, "Machine learning for pediatric echocardiographic mitral regurgitation detection," *J. Amer. Soc. Echocardiography*, vol. 36, no. 1, pp. 96–104.e4, Jan. 2023, doi: 10.1016/j.echo.2022.09.017.

[63] W. Hong, Q. Sheng, B. Dong, L. Wu, L. Chen, L. Zhao, Y. Liu, J. Zhu, Y. Liu, Y. Xie, Y. Yu, H. Wang, J. Yuan, T. Ge, L. Zhao, X. Liu, and Y. Zhang, "Automatic detection of secundum atrial septal defect in children based on color Doppler echocardiographic images using convolutional neural networks," *Frontiers Cardiovascular Med.*, vol. 9, Apr. 2022, Art. no. 834285, doi: 10.3389/fcvm.2022.834285.

[64] J. Zhang, Y. Zhan, W. Ren, Z. Song, S. Liu, and F. Pan, "Comparison of subintimal and submembranous myocardial strain of left ventricle in cardiac amyloidosis and hypertrophic cardiomyopathy by two-dimensional spot tracking imaging," *Chin. J. Med. Imag.*, vol. 22, no. 1, pp. 36–40, Feb. 2014, doi: 10.3969/j.issn.1005-5185.2014.01.011.

[65] G. Duffy, P. P. Cheng, N. Yuan, B. He, A. C. Kwan, M. J. Shun-Shin, K. M. Alexander, J. Ebinger, M. P. Lungren, F. Rader, D. H. Liang, I. Schnittger, E. A. Ashley, J. Y. Zou, J. Patel, R. Witteles, S. Cheng, and D. Ouyang, "High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning," *JAMA Cardiology*, vol. 7, no. 4, p. 386, Apr. 2022, doi: 10.1001/jamacardio.2021.6059.

[66] X. Yu, X. Yao, B. Wu, H. Zhou, S. Xia, W. Su, Y. Wu, and X. Zheng, "Using deep learning method to identify left ventricular hypertrophy on echocardiography," *Int. J. Cardiovascular Imag.*, vol. 38, no. 4, pp. 759–769, Nov. 2021, doi: 10.1007/s10554-021-02461-3.

[67] Z. Xu, F. Yu, B. Zhang, and Q. Zhang, "Intelligent diagnosis of left ventricular hypertrophy using transthoracic echocardiography videos," *Comput. Methods Programs Biomed.*, vol. 226, Nov. 2022, Art. no. 107182, doi: 10.1016/j.cmpb.2022.107182.

[68] J. Zhang, S. Gajjala, P. Agrawal, G. H. Tison, L. A. Hallock, L. Beussink-Nelson, M. H. Lassen, E. Fan, M. A. Aras, C. Jordan, K. E. Fleischmann, M. Melisko, A. Qasim, S. J. Shah, R. Bajcsy, and R. C. Deo, "Fully automated echocardiogram interpretation in clinical practice: Feasibility and diagnostic accuracy," *Circulation*, vol. 138, no. 16, pp. 1623–1635, Oct. 2018, doi: 10.1161/circulationaha.118.034338.

[69] K. Kusunose, T. Abe, A. Haga, D. Fukuda, H. Yamada, M. Harada, and M. Sata, "A deep learning approach for assessment of regional wall motion abnormality from echocardiographic images," *JACC: Cardiovascular Imag.*, vol. 13, no. 2, pp. 374–381, Feb. 2020, doi: 10.1016/j.jcmg.2019.02.024.

[70] R. Muraki, A. Teramoto, K. Sugimoto, K. Sugimoto, A. Yamada, and E. Watanabe, "Automated detection scheme for acute myocardial infarction using convolutional neural network and long short-term memory," *PLoS ONE*, vol. 17, no. 2, Feb. 2022, Art. no. e0264002, doi: 10.1371/journal.pone.0264002.

[71] O. Hamila, S. Ramanna, C. J. Henry, S. Kiranyaz, R. Hamila, R. Mazhar, and T. Hamid, "Fully automated 2D and 3D convolutional neural networks pipeline for video segmentation and myocardial infarction detection in echocardiography," *Multimedia Tools Appl.*, vol. 81, no. 26, pp. 37417–37439, Jul. 2022, doi: 10.1007/s11042-021-11579-4.

[72] S. Kiranyaz, A. Degerli, T. Hamid, R. Mazhar, R. E. F. Ahmed, R. Abouhasera, M. Zabihi, J. Malik, R. Hamila, and M. Gabbouj, "Left ventricular wall motion estimation by active polynomials for acute myocardial infarction detection," *IEEE Access*, vol. 8, pp. 210301–210317, 2020, doi: 10.1109/ACCESS.2020.3038743.

[73] J. F. B. S. Martins, E. R. Nascimento, B. R. Nascimento, C. A. Sable, A. Z. Beaton, A. L. Ribeiro, W. Meira, and G. L. Pappa, "Towards automatic diagnosis of rheumatic heart disease on echocardiographic exams through video-based deep learning," *J. Amer. Med. Inform. Assoc.*, vol. 28, no. 9, pp. 1834–1842, Jul. 2021, doi: 10.1093/jamia/ocab061.

[74] G. Han, T. Jin, L. Zhang, C. Guo, H. Gui, R. Na, X. Wang, and H. Bai, "Adoption of compound echocardiography under artificial intelligence algorithm in fetal congenial heart disease screening during gestation," *Appl. Bionics Biomechanics*, vol. 2022, pp. 1–8, Jun. 2022, doi: 10.1155/2022/6410103.

[75] R. Nova, S. Nurmaini, R. U. Partan, and S. T. Putra, "Automated image segmentation for cardiac septal defects based on contour region with convolutional neural networks: A preliminary study," *Informat. Med. Unlocked*, vol. 24, Jan. 2021, Art. no. 100601, doi: 10.1016/j.imu.2021.100601.

[76] J. Tromp, D. Bauer, B. L. Claggett, M. Frost, M. B. Iversen, N. Prasad, M. C. Petrie, M. G. Larson, J. A. Ezekowitz, and S. D. Solomon, "A formal validation of a deep learning-based automated workflow for the interpretation of the echocardiogram," *Nature Commun.*, vol. 13, no. 1, Nov. 2022, Art. no. 6776, doi: 10.1038/s41467-022-34245-1.

[77] P. Sridar, A. Kumar, A. Quinton, R. Nanan, J. Kim, and R. Krishnakumar, "Decision fusion-based fetal ultrasound image plane classification using convolutional neural networks," *Ultrasound Med. Biol.*, vol. 45, no. 5, pp. 1259–1273, May 2019, doi: 10.1016/j.ultrasmedbio.2018.11.016.

QI QI was born in Jilin, in 2003. She is currently pursuing the bachelor's degree in biostatistics with Xi'an Jiaotong-Liverpool University, Suzhou, China. Her research interests include the intersection of technology and biology and the applications of deep learning and machine learning in the field of bioinformatics.

XIAOXIANG HAN (Graduate Student Member, IEEE) was born in Yancheng, Jiangsu, China, in 1998. He received the B.S. degree in computer science and technology from the Jinling Institute of Technology, in 2021. He is currently pursuing the master's degree in electronic information with the University of Shanghai for Science and Technology, Shanghai, China.

From 2021 to 2024, he was a Graduate Student with the School of Health Science and Engineering, University of Shanghai for Science and Technology, specializing in medical imaging and image processing. He has published multiple academic articles indexed by SCI or EI. His research interests include pattern recognition, medical image analysis, accelerated MRI reconstruction, intelligent ultrasound diagnosis, semi-/weakly-/self-supervised learning, and multimodal learning.

YIMAN LIU (Member, IEEE) received the B.E. degree in medical imaging engineering from the University of Shanghai for Science and Technology, in 2015, and the master's degree in medical artificial intelligence from East China Normal University, in 2024.

In 2015, he joined the Department of Pediatric Cardiology, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, mainly engaged in pediatric cardiac ultrasound diagnosis. He has published multiple academic articles indexed by SCI or EI. His research interests include medical image analysis, intelligent ultrasound diagnosis, self-supervised learning, and multimodal artificial intelligence.

● ● ●