Biomedical Signal Processing and Control 119 (2026) 109899

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Deep learning advancements for cardiovascular diseases (CVDs) diagnosis: Imaging modalities, challenges, and future perspectives

Inayatul Haq [a,b,c], Haomin Liang [a], Ke Zeng [a], Tao Wang [a], Izhar Uddin [d], Jielong Lin [e], Yan Kang [f,*], Bingding Huang [a,b,**]

[a] School of Artificial Intelligence, Shenzhen Technology University, Shenzhen 518118, China
[b] College of Applied Sciences, Shenzhen University, Shenzhen 518060, China
[c] Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen 518060, China
[d] School of Clinical Medicine, Ningxia Medical University, Yinchuan 750101, China
[e] Department of Radiology, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China
[f] College of Health Science and Environmental Engineering, Shenzhen Technology University, Shenzhen 518118, China

ARTICLE INFO

Keywords:
Deep learning
Cardiovascular diseases
Convolutional and transformer networks
Radiology
Cardiovascular imaging

ABSTRACT

Cardiovascular diseases (CVDs) remain the leading cause of global mortality and impose a substantial clinical and socioeconomic burden. Medical imaging plays a central role in assessing cardiac structure, function, and tissue characteristics. However, conventional image interpretation is limited by operator dependence, variable image quality, and time-consuming analysis. Deep learning (DL) enables automated, efficient, and reproducible analysis of cardiovascular images and has become an important tool in this domain. This review presents a comprehensive and modality-oriented synthesis of DL applications across five major cardiovascular imaging modalities: echocardiography, coronary CT angiography (CCTA), cardiac magnetic resonance imaging (CMRI), nuclear imaging (SPECT/PET), and X-ray angiography. We review the main DL architectures, including convolutional neural networks, recurrent models, transformer-based networks, graph-based models, hybrid frameworks, and generative models. Their performance is examined across core tasks such as segmentation, classification, detection, motion analysis, and functional assessment. The strengths and limitations of these approaches are summarized for each modality. Beyond modality-specific studies, evidence is integrated to identify cross-modality methodological trends, transferable design principles, and persistent technical gaps. Major barriers to clinical deployment are discussed, including dataset bias and imbalance, annotation variability, imaging artifacts, domain shift, limited multimodal integration, high computational cost, and limited interpretability. Future research directions are outlined, with emphasis on CNN–transformer hybrids, diffusion models, multimodal foundation models, federated learning, and explainable artificial intelligence. This review aims to support the development of robust and clinically reliable DL systems and to facilitate their translation into routine cardiovascular imaging practice.

1. Introduction

Cardiovascular diseases (CVDs) remain the leading cause of mortality worldwide. According to the World Health Organization, they account for approximately 17.9 million deaths annually, representing nearly one-third of global mortality [1]. Besides the high death rate, CVDs cause long-term disability, reduced quality of life, and significant economic losses due to healthcare costs and productivity decline [2]. Global mortality data from 1950 to 2022 illustrate persistent regional disparities and a sustained rise in the burden of CVDs (Fig. 1) [3]. Conditions such as heart failure, myocardial infarction, and stroke often lead to chronic complications that necessitate lifelong management [4]. In the United States alone, the combined direct and indirect costs of CVDs exceed 351 billion USD per year, reflecting substantial healthcare and social impact [5]. This burden is particularly severe in low- and middle-income countries, where access to early diagnosis and advanced

* Corresponding author.
** Corresponding author at: School of Artificial Intelligence, Shenzhen Technology University, Shenzhen 518118, China.
E-mail addresses: kangyan@sztu.edu.cn (Y. Kang), huangbingding@sztu.edu.cn (B. Huang).

https://doi.org/10.1016/j.bspc.2026.109899
Received 31 August 2025; Received in revised form 13 January 2026; Accepted 13 February 2026
Available online 20 February 2026

treatment remains inadequate [6].

CVDs include a wide range of conditions affecting the heart and vasculature, including coronary artery disease (CAD) [7], heart failure [8], arrhythmias [9], valvular heart disease [10], congenital heart disease [11], cardiomyopathy [12], peripheral artery disease [13], stroke [14], hypertension [15], and aortic aneurysm [16], as summarized in Fig. 2. Given the diversity and complexity of these conditions, accurate and timely diagnosis is essential for effective management and better patient outcomes.

Medical imaging plays a central role in diagnosing CVDs. Each modality offers distinct strengths and limitations. Echocardiography is widely available, noninvasive, and radiation-free, but its quality depends on patient factors and operator skill [17]. Coronary CT angiography (CCTA) provides high-resolution visualization of coronary anatomy but uses ionizing radiation and iodinated contrast agents, which may limit its suitability for some patients [18]. Cardiac Magnetic Resonance Imaging (CMRI) offers excellent soft-tissue contrast and comprehensive functional assessment without radiation exposure, while it remains costly and less accessible in many settings [19]. Nuclear imaging provides valuable information on myocardial perfusion and viability, but suffers from lower spatial resolution and the use of radioisotopes [20]. X-ray angiography remains the gold standard for coronary visualization and interventional planning, but it is invasive and requires contrast administration [21].

Artificial intelligence (AI), particularly deep learning (DL), is increasingly clinically used to overcome these challenges. DL models have shown strong performance in tasks such as left-ventricular ejection fraction estimation [22,23], anatomical segmentation, and disease detection across multiple modalities, including X-ray angiography, CCTA, CMRI, and ultrasound. They often match human performance in pattern recognition, as demonstrated in DL–based cardiac imaging tasks, such as the automated assessment of aortic stenosis from echocardiography [24] and the video-based assessment of systolic function from coronary angiograms [25], and MRI-based cardiac function analysis [26]. DL models in cardiovascular imaging provide significant diagnostic and prognostic value, enabling personalized treatment strategies. These models can analyze both structural and functional features from imaging data, supporting risk stratification and improving clinical decision-making [27–29]. Their ability to process complex imaging data consistently and rapidly reduces human error and improves workflow efficiency [30]. Integration with wearable devices and telemedicine platforms further extends DL applications to continuous cardiovascular monitoring and remote patient management [31]. Convolutional neural networks (CNN)-based DL models have shown strong accuracy in cardiovascular imaging, including interpreting echocardiograms [32], improving CCTA plaque quantification [33], and enhancing CMRI segmentation and functional measurement [34]. Recent advances show DL's growing value in automating cardiac function analysis, including ejection fraction calculation, strain estimation, right-ventricular assessment, and chamber measurement using routine echocardiography and CMRI. These models provide fast, reproducible results and help reduce operator dependency and image-quality variability [35–39].

However, despite these advances, existing studies vary widely in methodology, dataset quality, and clinical validation, creating uncertainty about the reliability and generalizability of DL systems in real-world cardiovascular practice. A structured, modality-oriented synthesis is therefore needed to clarify current capabilities, highlight persistent challenges, and guide future research.

Accordingly, this review provides a comprehensive analysis of DL applications across major cardiovascular imaging modalities, including echocardiography, CCTA, CMRI, nuclear imaging, and X-ray angiography. Specifically, it aims to: (i) summarize modality-specific applications of DL; (ii) compare the strengths and limitations of major architectures for segmentation, classification, and detection; (iii) identify key technical and clinical challenges affecting robustness and generalizability; (iv) discuss opportunities for multimodal data integration for personalized diagnosis and risk prediction; and (v) outline future research priorities, including explainable AI and federated learning, to support reliable clinical deployment.

1.1. Structure and logical flow of the paper

This review is organized as follows. Section 2 outlines the major challenges and limitations of DL in cardiovascular imaging. Section 3 reviews key DL architectures and methodological advances. Section 4 surveys modality-specific applications across echocardiography, CT,

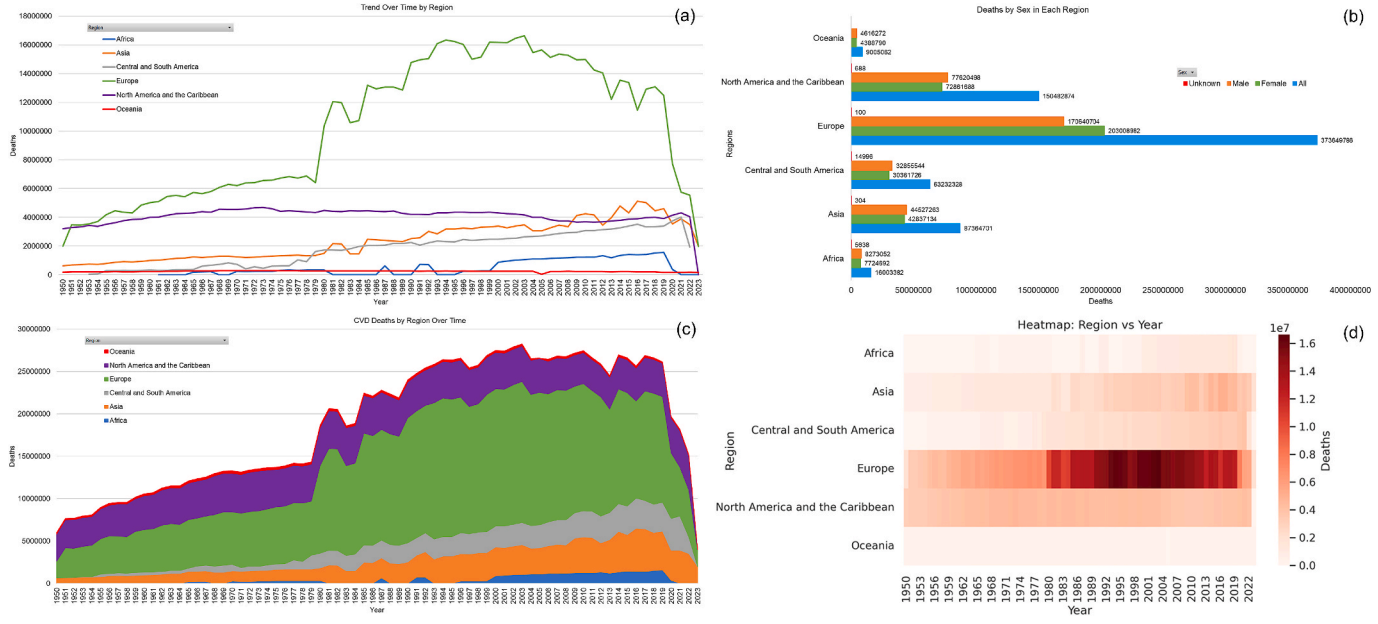

Fig. 1. The WHO record shows global trends in cardiovascular diseases (CVDs) mortality [304]. (a) shows the temporal trend of CVDs-related deaths from 1950 to 2022 across six global regions. (b) Distribution of CVDs mortality by sex (male, female, unknown) within each region. (c) A stacked area chart illustrates each region's cumulative contribution to global CVDs mortality over time. (d) The heat map shows the intensity of CVDs deaths by region and year, emphasizing regional and temporal disparities.

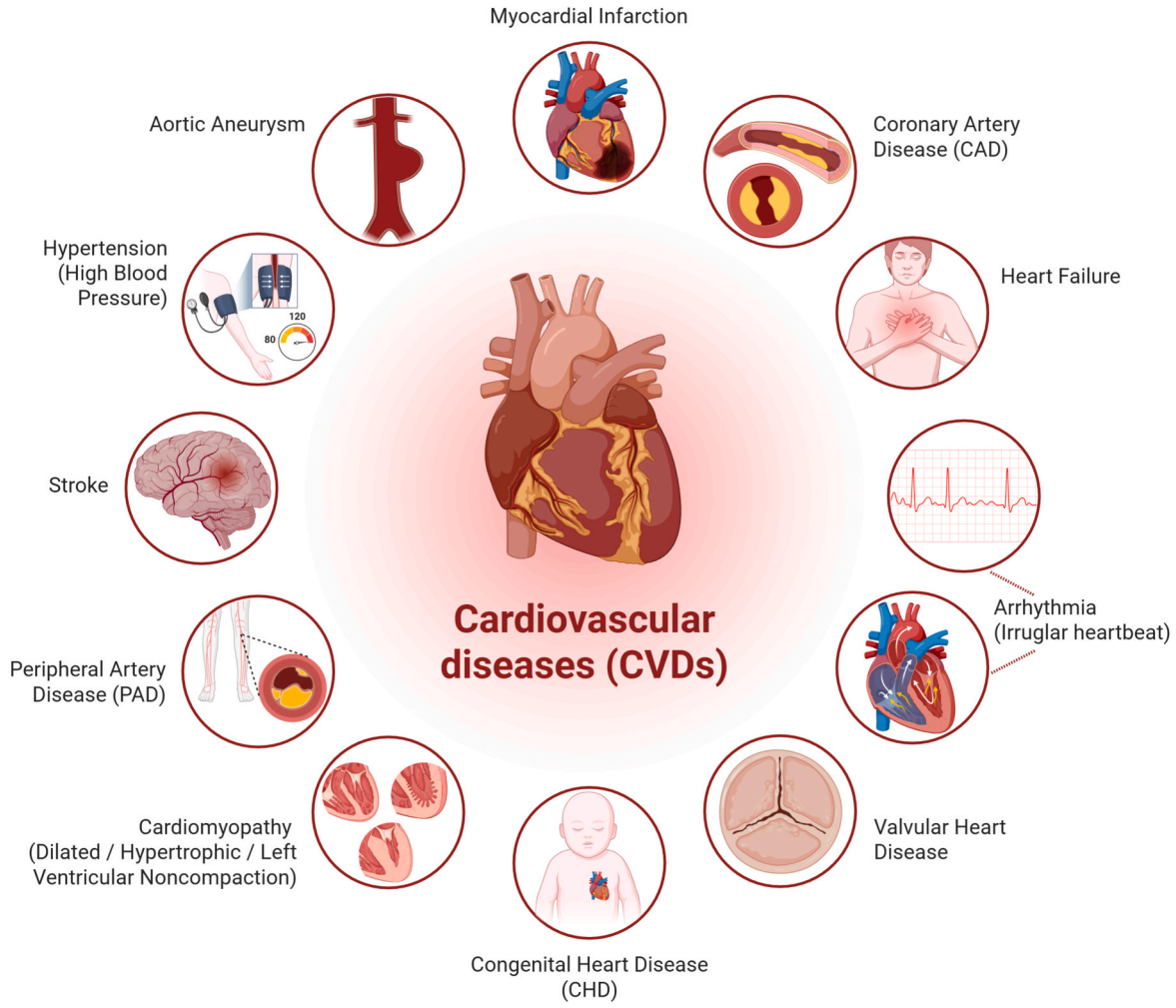

Fig. 2. Major categories of cardiovascular diseases. Created in BioRender. Huang, B. (2025) https://BioRender.com/zcylvgx.

MRI, nuclear imaging, and X-ray angiography, and summarizes their strengths and limitations. Section 5 compares major cardiovascular imaging modalities and summarizes DL tasks, performance characteristics, and limitations across them, and Section 6 discusses multimodal DL strategies. Section 7 highlights emerging trends and future directions. Finally, the paper concludes in Section 9.

2. Challenges and limitations of deep learning in cardiovascular imaging

DL models for CVDs detection, classification, segmentation, and prediction face persistent limitations across five modalities: echocardiography, CCTA, CMRI, nuclear imaging, and X-ray angiography. These challenges reduce robustness, generalizability, and clinical adoption.

2.1. Dataset limitations and bias

Across the five modalities mentioned above, dataset limitations remain a major barrier to robust model development. Many datasets are small and drawn from demographically skewed populations, often overrepresenting older or high-risk patients while underrepresenting younger or low-risk groups, leading to overfitting and poor generalizability [40–43]. Class imbalance is a recurring problem, such as the rare transthoracic echocardiographic views [44], small-caliber coronary branches in CCTA [45–47], thin myocardial walls in CMRI [48–50], subtle ischemic patterns in nuclear imaging [42,51,52], and small vessel pixels in X-ray angiography [53–58]. These imbalances bias training and reduce sensitivity to clinically significant but less frequent patterns.

2.2. Annotation quality and variability

Annotation inconsistencies across institutions and observers negatively impact model accuracy in all five modalities. Manual labeling often introduces observer bias, variability in ground truth definitions, and inconsistent annotation standards [45,47,59–65]. In echocardiography and CMRI, inter-observer variability in chamber and valve segmentation, partial chamber visibility, and misaligned slices reduces segmentation precision [66–70]. In nuclear imaging, labels are often derived from invasive coronary angiography rather than direct functional measurements, leading to misalignment between training labels and physiological reality [52,71–79]. In X-ray angiography, pixel-level vessel labeling, subjective frame selection, and inconsistent vessel boundary definitions hinder reproducibility [58,80–87].

2.3. Image quality, artifacts, and protocol variability

In echocardiography, view dependency, foreshortened views, shadowing, and acoustic noise degrade model performance for detection and segmentation tasks [44,88–91]. CCTA imaging faces challenges from motion blur, blooming artifacts caused by calcifications, partial volume effects, and variable contrast injection timing, which compromise vessel and plaque analysis [65,92–97]. CMRI suffers from low resolution in basal and apical slices, poor tissue boundary contrast, and motion artifacts, which reduce segmentation accuracy [98–102]. In nuclear

imaging, the use of polar map compression results in loss of spatial detail and reduces the ability to localize perfusion defects [41,78,103]. X-ray angiography is hindered by low vessel contrast, overlapping anatomical structures such as catheters and bones, and uneven contrast agent flow, leading to frequent misclassification [54,55,81,85,86,104–112].

2.4. Architectural and algorithmic limitations

Echocardiographic applications using 3D-CNNs for ejection fraction estimation are highly sensitive to label noise and limited temporal modeling, which reduces prediction stability [113–115]. In CCTA and CMRI, encoder–decoder architectures lose fine structural details during downsampling, resulting in incomplete segmentation of small vessels or cardiac cavities [50,116,117]. Nuclear imaging models that rely on compressed polar map inputs suffer from limited spatial resolution and reduced lesion localization accuracy [41,71,78,103]. In X-ray angiography, patch-wise CNNs and coarse localization methods often lack anatomical precision, producing false positives in stenosis detection [57,80,82,84,118,119].

2.5. Domain shift and generalization

All five cardiac modalities listed above experience performance degradation when models are applied to external datasets due to domain shifts arising from differences in scanner hardware, acquisition protocols, and population demographics [47,59,79,93,102,105,120–122]. Although domain adaptation techniques such as transfer learning and GAN-based synthesis have been explored, these can introduce unrealistic textures or fail to preserve anatomical fidelity, limiting their clinical usefulness [120,123,124].

2.6. Limited multimodal integration

In CCTA, CMRI, nuclear imaging, and echocardiography, combining imaging data with clinical metadata such as demographics and biomarkers could improve diagnostic performance. However, multimodal integration remains limited due to inconsistent data formats, missing variables, and the absence of standardized frameworks for fusion [41,63,64,125–127]. Aligning anatomical and functional datasets across modalities remains technically challenging, particularly when acquisition protocols differ.

2.7. Interpretability and clinical trust

The opaque "black-box" nature of many DL models in all five modalities hinders clinical adoption by reducing transparency in decision-making [47,71,73,108,128–130]. Clinicians may be reluctant to rely on automated outputs for high-stakes tasks such as treatment planning or risk stratification without interpretable outputs.

2.8. Computational demands and scalability

For CCTA and CMRI, processing high-resolution volumetric data requires significant computational resources, including high-memory GPUs and extended training times [69,93,117,131]. In echocardiography and X-ray angiography, real-time or near-real-time inference is often constrained by computational load, limiting deployment in time-sensitive clinical settings [108,129,130,132].

3. Deep learning architectures and methodological advances for CVDs diagnosis

DL encompasses various architectures and techniques deployed for different data types and diagnostic tasks. Below are the main types of DL models and their applications in diagnosing CVDs. Fig. 3 shows common DL architectures (CNN, U-Net, graph-based architectures, GAN, RNN, Transformer) and a typical AI workflow for cardiovascular imaging, from multi-modal data acquisition and preprocessing to feature extraction and clinical endpoints for diagnostic tasks.

3.1. Convolutional neural networks

CNNs have been extensively applied across cardiovascular imaging modalities to support the detection, characterization, and quantification of CVDs. In coronary CCTA, CNN-based architectures play an essential role in automated segmentation of coronary arteries, stenosis

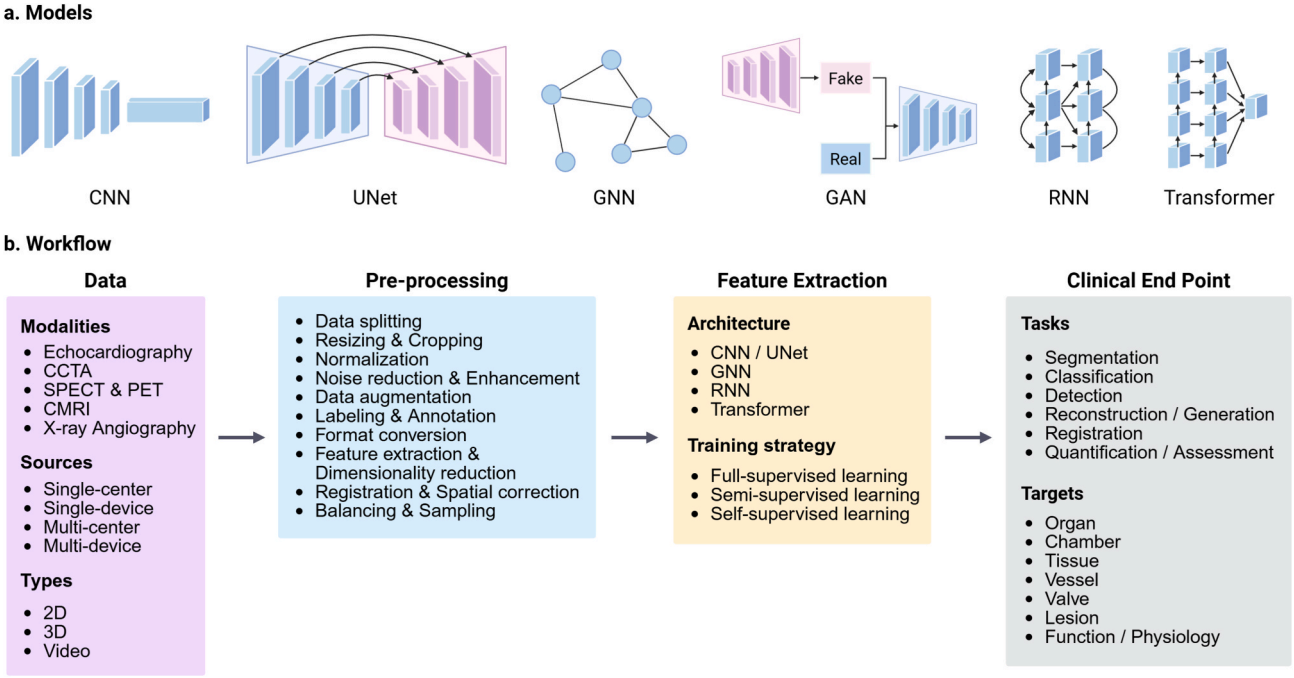

Fig. 3. Deep learning models and workflow for cardiovascular image analysis. Created in BioRender. Huang, B. (2025) https://BioRender.com/4c3pvm3.

assessment, and plaque characterization. Models such as 3D PSPNet and U-Net variants effectively capture complex vascular structures by leveraging multi-scale spatial features and mitigating class imbalance, which enhances stenosis localization and vessel boundary delineation [133,134]. Similar advances are reported with PSPNet-based coronary angiography segmentation and residual CNN frameworks, which enhance feature extraction for coronary morphology and disease assessment [135,136]. CNNs combined with recurrent or attention modules further improve cardiac image interpretation by capturing temporal continuity and refining feature saliency. For instance, Attention-Residual V-Net (ARVNet) and temporal feature-fused networks improve segmentation performance in fetal echocardiography and dynamic cardiac imaging scenarios by modelling spatial–temporal dependencies [137,138]. For preoperative planning and whole-heart analysis, U-Net-based GANs and Dense V-Net architectures provide robust segmentation of cardiac substructures across CT and multimodal datasets, supporting comprehensive anatomical assessment [139,140].

In CMRI, CNNs have shown high performance in segmenting myocardial tissue, measuring infarct size, and defining late gadolinium enhancement regions. DeepLabV3 + and transfer-learning-based CNN models enable accurate classification of myocardial infarction, myocarditis, and healthy myocardium, reducing observer variability and accelerating clinical workflows [141,142]. More advanced strategies, such as curriculum-learning-enhanced U-Net models, achieve reliable segmentation of myocardial edema and infarction across multi-sequence MRI acquisitions, supporting both diagnostic accuracy and computational efficiency [143]. For chest X-ray (CXR) imaging, ensemble CNN models integrating InceptionV3, MobileNet, and ResNet have improved the detection of cardiomegaly and pulmonary congestion, which are early indicators of cardiac dysfunction. These models use transfer learning to achieve rapid, consistent interpretation even in resource-limited environments [144]. Although CT is based on X-ray acquisition, it provides tomographic volumetric information rather than 2D projection imaging. CT-based DL has also been used to predict cardiovascular risk features from low-dose scans, illustrating the expanding role of CNNs in cross-modality CVDs risk stratification [145].

CNNs have driven significant advances in nuclear cardiac imaging, particularly in SPECT and PET. The DL models now facilitate CT-free attenuation correction, low-dose image denoising, and motion compensation, improving myocardial perfusion map quality and reducing radiation exposure [52,146–149]. Supplementary CNN frameworks enable automated ischemia classification and accurate myocardial blood-flow quantification from perfusion polar maps, strengthening prognostic evaluation in nuclear cardiology [150,151]. In echocardiography, CNNs are widely used to classify cardiomyopathy, valvular disorders, and regional wall-motion abnormalities. Architectures such as EfficientNetB3 achieve high diagnostic accuracy through optimized regularization and feature normalization strategies [152]. CNN-based models have also been applied for left atrial segmentation and chamber morphology analysis, improving structural assessment in both clinical and research settings [153]. In fetal cardiac imaging, attention-enhanced CNNs such as ARVNet are used for echocardiographic segmentation tasks, while deep residual networks developed for fetal cardiac CT support the classification of congenital abnormalities, enabling early-stage diagnosis and real-time screening [137,154]. Complementary CNN-based classification methods have also been applied to broader cardiovascular risk prediction tasks, further emphasising the versatility of DL in cardiac disease assessment [155].

3.2. Graph-based and capsule networks

Graph-based architectures and capsule networks enhance the ability of DL models to represent spatial hierarchies and anatomical relationships in cardiovascular imaging, improving structural understanding across modalities. Capsule networks improve upon conventional CNNs by encoding part–whole relationships and preserving pose-aware features, which enhances robustness under limited data availability or noisy imaging conditions. For instance, SegCaps, a capsule-based segmentation model, has been explored for left-ventricle segmentation in echocardiography, providing improved boundary delineation while using significantly fewer parameters than U-Net variants [156]. Capsule architectures have also been applied to intravascular ultrasound (IVUS), where they facilitate reliable delineation of vessel walls and identification of plaque components in small datasets by improving spatial feature retention [157].

Graph-based methods, particularly GCNs and graph attention networks (GATs), are increasingly used in tasks that require structured representations of cardiovascular anatomy. For example, GATs have been applied to model vessel topology in CCTA, enabling improved artery labeling through node-wise attention and structure-aware feature aggregation [158]. Similarly, graph-based post-processing applied to CNN-derived vessel masks enhances centerline extraction in X-ray angiography, improving segmentation robustness under low-contrast conditions [159]. In cine CMRI, spatiotemporal GNNs integrate graph-based structural encoding with temporal motion cues to detect myocardial infarction and characterize regional wall-motion abnormalities. These models use dynamically constructed graphs to represent changes in myocardial contraction patterns across the cardiac cycle, improving the localization of infarcted segments [160–162]. Graph-based feature fusion has also been incorporated into domain-adaptive frameworks that combine transfer learning with GCNs for enhanced cardiac disease prediction from CMRI-derived features [162].

Recent hybrid transformer–graph architectures further advance CVDs' image analysis. The Adaptive Attention-Enhanced Transformer with Modified Graph Cuts (AAET-MGC) integrates CNN-based local feature extraction with global transformer attention and graph-cut refinement, enabling precise myocardial segmentation and improved delineation of cardiac structures in CMRI while mitigating overfitting on limited datasets [163]. Similarly, the Adaptive Attention Multi-Graph Convolutional Network (AAM-GCN) introduces multi-relation graph learning and spatial–intensity attention mechanisms to enhance myocardial segmentation quality, particularly in structurally complex cardiac regions [164]. Although AAM-GCN was originally designed for ischemic stroke lesion detection on diffusion MRI, its underlying multi-graph attention framework is conceptually transferable to cardiac segmentation tasks. Graph-based learning has also been extended to risk stratification and disease subtype prediction, where GCNs use imaging-derived clinical features and topological representations of cardiovascular anatomy. For example, the potential of GCNs for CVDs risk prediction through improved multimodal feature integration through improved multimodal feature integration and class balancing [165]. Moreover, GCNs have been used to classify MPI polar maps, enabling enhanced localization of ischemic territories by modeling projected spatial adjacency and intensity gradients within polar maps [166].

3.3. Hybrid and specialized architectures

Hybrid and specialized DL architectures have strengthened CVDs diagnosis by combining spatial, temporal, and multimodal information across major cardiovascular imaging modalities. Hybrid CNN–GRU frameworks extract anatomical features and capture sequential trends in temporal or clinical time-series data, improving cardiovascular risk prediction and supporting stenosis assessment when temporal imaging sequences are available. These models have shown improved prediction of cardiovascular risk and stenosis severity by effectively learning both spatial patterns and longitudinal clinical dynamics [167,168]. CNN–LSTM-based hybrids have also been used for temporal cardiac analysis. Studies combining CNN encoders with LSTM layers for heart disease prediction show improved modeling of sequential clinical or imaging-derived time-series data [169]. In cardiac MRI, Siamese CNN architectures using frame-pair representations have been applied for short-axis cardiac motion tracking, thereby improving functional

assessment [170].

In CMRI and CCTA, several specialized hybrid models integrate attention mechanisms, residual connections, and deep supervision to improve segmentation and pathology localization. Multi-scale residual–dense hybrid networks capture detailed information and have shown robust performance in cardiac structure segmentation and automated diagnosis tasks [171]. Siamese U-shaped networks improve myocardial pathology segmentation on multimodal MR data by using paired feature representations that enhance contrast between normal and pathological regions [172]. Residual U-Net variants with deep supervision further improve left-ventricular segmentation in cardiac CT angiography, supporting accurate stenosis and plaque assessment [173,174]. Attention-augmented residual architectures also help to improve vessel and myocardial analysis. Three-dimensional squeeze-and-excitation residual networks boost stenosis detection in magnetic resonance angiography by emphasizing salient vascular features [175]. Moreover, attention-improved U-Net variants, such as multiscale attention residual U-Net and residual self-attention U-Net, provide improved delineation of ventricular and myocardial structures in CMRI by strengthening important spatial dependencies [176,177].

Multimodal fusion architectures have also advanced cardiac analysis by integrating anatomical and functional information. Recent PET–CT–based DL models use attenuation maps to capture metabolic and structural features jointly, improving chamber quantification and supporting comprehensive cardiac assessment [178]. Graph-based networks have also advanced coronary modeling by capturing the structural topology of vascular trees, allowing for more robust stenosis detection and vessel characterization in angiographic imaging. These methods use relational information among arterial branches, improving interpretability and performance in coronary analysis tasks [176,177].

3.4. Generative models

Generative models have become increasingly important in cardiovascular imaging by enabling realistic image synthesis, artifact removal, augmentation, and functional motion analysis [179]. Architectures such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), VAE-GAN hybrids, and diffusion models address challenges including limited labeled datasets, acquisition variability, and class imbalance, thereby improving generalization and diagnostic accuracy across CMRI, CCTA, and X-ray angiography [179]. Early applications in cardiac MRI demonstrated the utility of synthetic data to augment small datasets, where deep generative networks produced realistic CMRI images for congenital heart disease to support downstream segmentation tasks [123]. Subsequent work extended this capability by using paired and unpaired GAN frameworks to synthesize left-atrium MR images, improving annotation availability and segmentation performance [180]. Cycle-consistent adversarial models have also been applied to X-ray angiography for cross-domain style adaptation, enabling consistent vessel appearance and improving robustness in angiographic vessel analysis [181].

Generative models further enhance cardiac image quality by improving myocardial boundary clarity, reducing noise, and mitigating motion-related distortions [182,183]. Comparative studies show that GAN-based architectures yield sharper ventricular boundary delineation than U-Net in CMRI segmentation, demonstrating their ability to preserve fine anatomical details [181]. Apart from structural appearance, generative frameworks support functional and motion-based cardiac assessment. Convolutional autoencoders have been used to learn normal myocardial motion patterns in cine CMRI, enabling unsupervised detection of regional abnormalities associated with myocardial infarction [184]. GAN-based approaches have also been explored for improving myocardial perfusion, CMRI reconstruction, and motion modelling [179]. Recent hybrid architectures integrate semantic constraints into generative frameworks to produce anatomically coherent cardiac images that support the detection of functional abnormalities [185].

Generative frameworks also improve data organization and multimodal learning. Autoencoder-based feature extraction enables accurate CMRI view classification, reducing manual workload in cine-MRI repositories [186]. Cross-modal autoencoder models further support integrated cardiovascular state representation by learning holistic embeddings from multimodal cardiac data [187]. Semi-supervised CMRI segmentation also benefits from VAE-GAN hybrid models, reducing reliance on large annotated datasets across multi-vendor imaging sources [188].

3.5. Diffusion-based models

Diffusion-based generative models have recently emerged as a transformative class of techniques in cardiovascular imaging [189]. Unlike GANs, which may suffer from training instability and mode collapse, diffusion models progressively learn to denoise random noise distributions, producing highly stable and anatomically consistent reconstructions [190]. This makes them particularly suitable for cardiovascular imaging tasks requiring structural fidelity, temporal consistency, and robustness to noise [189].

In cardiac MRI, diffusion probabilistic models demonstrate superior performance for cine reconstruction, particularly in undersampled spiral acquisitions, producing more temporally coherent sequences and fewer artifacts than conventional reconstruction methods [189]. In echocardiography, diffusion-based denoising methods improve myocardial visibility, reduce speckle noise, and improve chamber boundary delineation in low-contrast ultrasound conditions [191]. Latent diffusion models extend these capabilities by enabling high-resolution cardiac image synthesis and data augmentation. Their ability to generate diverse, anatomically consistent cardiac images improves downstream segmentation, classification, and motion estimation performance, especially when training data are limited or heterogeneous [190].

3.6. Recurrent neural networks

Recurrent neural networks (RNNs), particularly LSTM and BiLSTM models, have shown strong performance in modeling the temporal progression of cardiac motion in dynamic cardiovascular imaging. Unlike CNNs, which primarily learn spatial patterns, RNNs capture sequential dependencies across frames in cine MRI and CMRI, allowing accurate reconstruction and interpretation of time-resolved cardiac dynamics. Convolutional RNN architectures, such as CRNNs, effectively integrate spatial encoding with temporal modeling and have shown improved reconstruction quality in dynamic MR sequences [192]. Hybrid CNN–RNN frameworks further improve performance in cardiac segmentation and function analysis. For example, residual CNN-RNN combinations improve cardiac image segmentation and disease classification by using both spatial features and temporal continuity [136]. Densely connected and dual-path recurrent models have also been applied to CMRI for left-ventricular segmentation, enabling more consistent tracking of myocardial wall motion across the cardiac cycle [193].

RNN-based architectures are very effective at capturing long-range temporal dependencies and subtle morphological changes across frames, thereby improving robustness to noise and generalization in clinically variable cardiac datasets. Bidirectional recurrent units have shown improved accuracy in identifying cardiac abnormalities by processing temporal sequences in both forward and backward directions [194]. Attention mechanisms are increasingly incorporated into recurrent models to focus on physiologically meaningful phases of the cardiac cycle. Densely gated recurrent networks have been used to detect cardiac phases in echocardiographic sequences, improving temporal localization and supporting automated functional assessment [195]. Attention-improved recurrent architectures for CMRI segmentation have also been reported, demonstrating improved tracking of motion-

salient regions across dynamic sequences [196].

3.7. Transformer-based and attention models

Transformer-based and attention-guided architectures have become increasingly important in cardiovascular imaging by modeling long-range dependencies, emphasizing clinically relevant anatomy, and improving robustness across modalities. In IVUS, transformer-driven segmentation models accurately define the lumen and media–adventitia boundaries, supporting coronary plaque assessment and stenosis quantification [197]. Vision Transformer (ViT) models have shown strong performance in CMRI, where global self-attention captures long-range spatial context, thereby improving the segmentation of myocardial structures. Feature-recombination and distillation strategies used in ViT-FRD improve boundary localization and segmentation accuracy in CMRI [198]. Hybrid ViT–CNN frameworks further strengthen cardiac image analysis by combining global transformer attention with local convolutional priors, enabling improved delineation of ventricular cavities and myocardial tissues [198].

Transformers also contribute to cross-modality adaptation in CMRI. For example, ST-GAN employs Swin Transformer blocks to support unsupervised domain adaptation and to improve segmentation consistency during transfer between MRI sequences [199]. In CT-based cardiac imaging, transformer-augmented architectures such as shape-aware contour attention networks and axial-attention fusion models provide more precise cardiac structure delineation, even under challenging acquisition conditions [200–202]. Fusion-attention Swin Transformer models similarly improve CMRI segmentation by integrating multi-scale contextual features [203]. Transformer-enhanced encoder–decoder networks, such as RotCAtt-TransUNet++ and MSF-TransUNet, improve segmentation accuracy in both cardiac CT and CMRI by modeling intra- and inter-slice dependencies and fusing hierarchical features [204,205]. Federated self-supervised transformer models further advance cine MRI analysis by learning temporal and spatial patterns from large amounts of unlabeled multi-center data, enabling accurate 4D cardiac segmentation while preserving data privacy [206].

In echocardiographic video analysis, advanced spatiotemporal DL models with attention mechanisms improve beat-to-beat functional assessment, enabling accurate estimation of ejection fraction and wall-motion abnormalities [207]. For chest X-ray–based cardiac disease detection, ViT models improve multi-label classification by capturing global contextual cues and outperforming CNNs in identifying cardiac abnormalities such as cardiomegaly [208]. Fig. 4 illustrates the comprehensive process of cardiovascular disease diagnosis, integrating traditional imaging techniques, radiological evaluation, and AI-driven analysis. It shows the collaborative role of medical professionals and advanced technology in enhancing diagnostic accuracy and treatment decisions.

3.8. Explainable and language-driven AI approaches

XAI is essential for safe and transparent integration of DL into cardiovascular imaging. Conventional models often function as opaque "black boxes," creating uncertainty for clinicians. XAI techniques such as Grad-CAM, SHAP, and saliency mapping help address this limitation by highlighting image regions that drive predictions, allowing physicians to verify whether model decisions align with cardiac pathology [209–211]. In myocardial perfusion imaging (MPI), XAI has been shown to improve interpretation by localizing perfusion defects and identifying areas contributing to risk estimation [212,213]. These visualization methods increase clinician confidence and support reliable deployment of AI systems in routine cardiac practice. Explainability has also been incorporated directly into clinical workflows. For example, XAI-enhanced SPECT analysis for CAD provides interpretable outputs during real-world assessments, improving diagnostic trust and user acceptance [210]. Integrated systems such as CardioNetFusion combine DL with structured explanation modules, offering patient-specific insights that clarify how cardiac imaging features contribute to disease detection [214]. Reviews of evaluation strategies emphasize that trustworthy explanations are essential for regulatory approval and responsible use of AI in cardiology [209,215].

Language-driven approaches enhance interpretability by linking visual cardiac features to semantic clinical descriptions. Vision–language foundation models trained on echocardiography can generate structured summaries, highlight abnormal regions, and translate image patterns into clinically relevant narratives [216]. Broader analyses of large

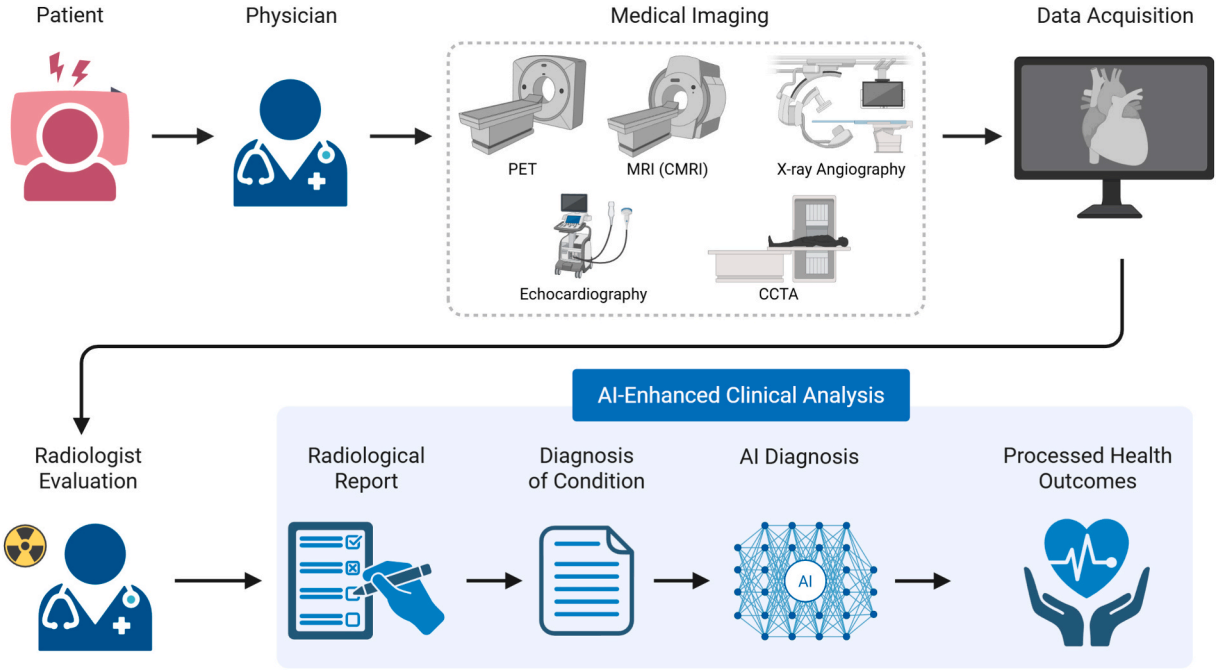

Fig. 4. Complete Workflow for cardiovascular diseases diagnosis: Integration of traditional imaging techniques, radiological evaluation, and AI-Enhanced clinical analysis. Created in BioRender. Huang, B. (2025) https://BioRender.com/1dvgh3g.

language models (LLMs) show their ability to support diagnostic reasoning by linking patient history, imaging findings, and clinical risk factors in cardiovascular contexts [217,218]. Early multimodal systems have also used language-guided representations to enhance interpretability in IVUS analysis [219]. Recent multimodal and contrastive learning frameworks offer additional interpretability. Medical foundation models that align cardiac images with clinical text improve semantic consistency and allow image features to be understood in relation to domain-specific clinical terms [220]. Meta-learning strategies further enhance robustness by adapting explanation mechanisms across patient subgroups and imaging domains, ensuring consistent interpretability in diverse cardiac populations [221,222]. Multimodal generative models can also provide intuitive, text-linked visual explanations for 3D cardiac images and videos, improving transparency in structural and functional assessments [223].

3.9. Comparative analysis of CNN and ViT architectures

A comparative analysis of CNNs and ViT architectures reveals complementary strengths in cardiovascular imaging. CNNs capture local spatial features efficiently and perform well on small or moderately sized clinical datasets. ViTs, through global self-attention, model long-range dependencies and complex spatial relationships, showing advantages when pretrained on large or multimodal data. The main distinctions between these architectures, including their data requirements, learning behavior, and clinical relevance, are summarized in Table 1 [216,224–228].

4. Deep learning applications for CVDs diagnosis across imaging modalities

This section reviews studies that use DL techniques to detect, classify, segment, and diagnose CVDs across various imaging modalities. Quantitative results summarized in later sections are compiled from published studies rather than reproduced under a unified experimental setup; therefore, the comparisons represent indicative performance trends rather than direct benchmarking.

Fig. 5 links major cardiovascular imaging devices, such as echocardiography, CCTA, nuclear imaging, CMRI, and X-ray Angiography, to their primary anatomical targets, including coronary arteries, aorta, cardiac chambers, and myocardium. Example outputs demonstrate the diversity of modalities and their complementary roles in diagnosis.

4.1. Deep learning for echocardiography-based diagnosis

DL has increasingly been applied to echocardiography for automated diagnosis of CVDs. These models address the challenges of subjectivity, operator dependence, and image quality variability. Early studies used established CNNs, such as VGG16 and EfficientNet, to classify conditions including angina, CAD, hypotension, and prosthetic valve identification. Techniques such as batch normalization, dropout, and transfer learning improved generalization, especially with limited data and class imbalance [67,229]. For smaller datasets, CNNs trained with cross-validation on multi-view echocardiography achieved high accuracy in distinguishing between apical and parasternal views, thereby addressing mislabeling across imaging systems [44,230]. Moreover, CNNs applied to 2D still images, Doppler sequences, and cine loops improved classification of normal versus abnormal cases, reducing reliance on manual interpretation [231]. In transthoracic echocardiography, including color Doppler views, models further enhanced cardiovascular assessment, with a video-based R(2 + 1)D network automating apical 4-chamber view classification and tricuspid regurgitation grading, reducing inter-observer variability [232]. Multi-task DL models have been developed for real-time view recognition and image-quality assessment, minimizing operator dependence and improving workflow [132]. AI-assisted systems for STEMI evaluation have shown expert-level performance, reduced clinician workload, and ensured accurate functional measurements [233].

3D-CNNs trained on cine-loop sequences from transthoracic echocardiography automatically classified ejection fraction categories without manual segmentation, effectively capturing temporal dynamics [113]. Hybrid models combining LSTMs with CNNs or VAEs extracted both spatial and temporal features, aiding the classification of valvular regurgitation [231]. Fusion networks that integrate spatial features with temporal optical flow cues enhanced view classification across challenging echocardiographic video clips [234].

A 3D multi-decoder residual U-Net with separate decoder paths for the mitral annulus and leaflets helps mitigate ultrasound noise, irregular leaflet geometry, and inconsistent manual contours. The decoders and post-processing pipeline improve automated mitral valve assessment by accurately reconstructing anatomical landmarks and surface geometry [235]. Encoder–decoder frameworks, especially U-Net and its variants, have been widely used for segmenting cardiac structures. These models have enabled the automated delineation of left ventricular endocardium, myocardium, and atrial structures, even in the presence of noise and contrast variability. Prominent innovations such as VDS-UNet (VGG16 + deep supervision), MFP-UNet (feature pyramid integration), and ResU-Net (residual encoder) have improved segmentation accuracy and robustness across various challenges, including poor-quality images and operator variability [66,68,88,236,237]. For left-ventricular analysis, a 3D nnU-Net framework uses CMR-derived labels to address low resolution, speckle patterns, and labeling inconsistencies in 3DE. This approach enables more reliable boundary recovery and reduces variability in functional measurements [238].

Table 1
Comparative characteristics of CNN and ViT architectures in cardiovascular imaging [216,224–228].

Aspect	CNNs	ViTs	Relevance to Cardiovascular Imaging
Feature Representation	Encode local spatial hierarchies using convolutional filters and pooling.	Model long-range spatial dependencies through global self-attention.	CNNs capture detailed vessel and tissue textures; ViTs integrate global cardiac geometry and motion context.
Data Requirement	Perform reliably with small or medium datasets and limited annotations.	Require larger datasets or pretrained weights (transfer learning) for stable convergence.	CNNs remain the preferred choice for limited clinical datasets; ViTs benefit from large or publicly pretrained models (e.g., ImageNet-21 K, Med-ViT).
Training Efficiency	Training is computationally efficient and converges faster on smaller datasets.	Training is more resource-intensive and sensitive to hyperparameter tuning.	CNNs are practical for routine experiments; ViTs demand higher GPU memory and longer training cycles.
Generalization Ability	Strong with appropriate augmentation but may struggle with cross-domain shifts.	Capture global context, improving cross-modality and domain adaptation when sufficiently trained.	Hybrid CNN–ViT models show enhanced robustness across imaging modalities (CT, MRI, Echo).
Explainability and Interpretability	Well-established visualization tools (Grad-CAM, LRP, SHAP) provide interpretable feature maps.	Attention maps offer coarse interpretability but lack standardized clinical validation.	CNNs currently provide clearer visual explanations for clinicians, while ViTs are advancing in explainable attention modeling.
Performance Trends	Maintain competitive or superior accuracy on small-scale datasets.	Match or exceed CNNs on large, balanced, or multimodal datasets when pretrained.	ViTs show promising results in CMRI segmentation, echocardiographic view classification, and coronary plaque analysis.

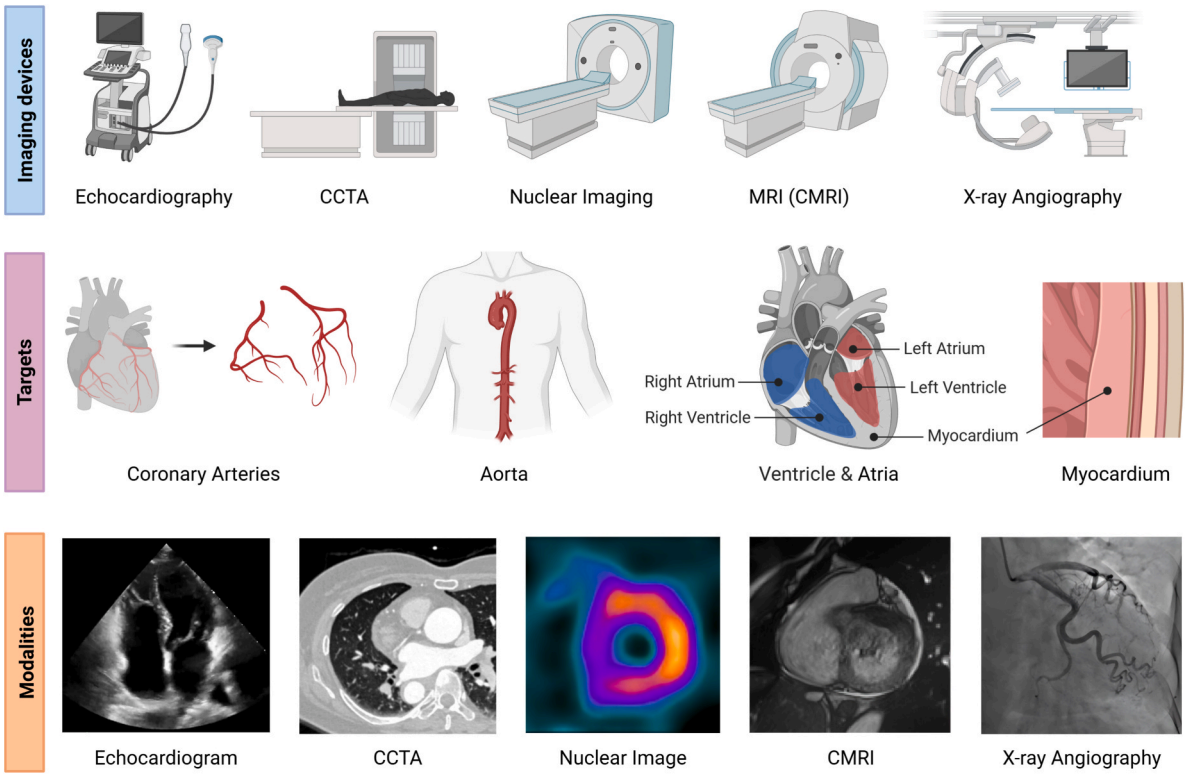

Fig. 5. Cardiovascular imaging devices, anatomical targets, and corresponding modalities: Echocardiography [207], CCTA [305], SPECT [71], nuclear image [306], X-ray angiography [129]. Created in BioRender. Huang, B. (2025) https://BioRender.com/3r6qw2a.

Combining detection and segmentation, such as using YOLOv7 for chamber localization with U-Net for detailed segmentation, further automates chamber quantification [66].

A Dual-Branch TransV-Net, integrating a V-Net backbone with an edge-enhancement branch and a transformer-based context module, strengthens boundary delineation and suppresses noise. This model also compensates for missing ventricular regions due to probe blind zones, thereby improving the geometric reconstruction of both ventricles [239]. Comparative studies of U-Net3D, ResU-Net3D, MultiResU-Net3D, and nnU-Net have shown that adaptive configurations like nnU-Net handle low contrast, noise, and complex ventricular anatomy more effectively. Automatic preprocessing and augmentation optimization further enhance stable ventricular geometry extraction from 3DE [240]. Attention-augmented residual V-Net (ARVNet) and hybrid ResNet–U-Net++ models have been applied to small lesion segmentation, addressing inter-class imbalance and inter-observer variability in hypertrophy and fetal cardiac tumor diagnoses [90,241]. AI-driven models for right-ventricular (RV) segmentation have addressed challenges such as incomplete RV capture, irregular chamber shape, and operator-dependent variation. By learning full-volume endocardial patterns, these models automate boundary detection and derive volumetric and functional metrics, offering more consistent RV evaluation [242].

Multi-task architectures that combine 3D segmentation with motion tracking utilize shared features and shape-consistency constraints. This approach effectively manages low SNR, weak boundaries, and out-of-plane motion, improving temporal coherence, myocardial strain estimation, and motion-based functional indices [243]. Besides segmentation, DL has been applied to functional analysis. Models like EchoNet, trained on millions of echocardiographic images, automate tasks such as chamber identification, volume estimation, and ejection fraction calculation. These models also predict systemic features like age and sex [244]. EchoPWC-Net, using optical flow-based motion estimation and ultrasound-specific augmentations, performs robust myocardial strain analysis, despite noise and vendor variability [245]. CNN– and ResNet-based classifiers outperform junior clinicians in identifying regional wall motion abnormalities, valve types, and hypertrophy, reducing inter-observer variability [90,115]. Applications of DL have extended to specialized domains. Zebrafish echocardiography validated segmentation frameworks with EfficientNet encoders and U-Net/U-Net++ decoders, achieving accurate ventricle boundary detection and automated ejection fraction estimation [246]. Clinical applications included fetal echocardiography for detecting cardiac rhabdomyomas using attention-based segmentation [241], transesophageal echocardiography during CPR for robust left ventricle segmentation [247], and pediatric echocardiography for congenital heart disease detection using ResNet50-based classifiers across multiple views [248].

Table 2 reviews DL models for diagnosing CVDs in the Echocardiography imaging modality. Only studies with clearly reported and comparable performance metrics are included. Supplementary models are discussed in the text for the broader context.

4.2. *Deep learning for CT-based cardiovascular diagnosis*

DL has been widely applied to the diagnosis of CVDs from CT images, particularly for segmentation and classification tasks. Early work using patch-based CNNs enabled accurate segmentation of cardiac chambers from chest CT scans, resolving issues of anatomical variability and eliminating the need for deformable atlas registration [45]. Extending this concept, U-Net variants, including 2D, 3D, and inception-enhanced designs, have been extensively used for coronary artery and aorta segmentation. The 3D Inception U-Net incorporated volumetric context and multi-scale feature extraction, addressing the discontinuity and fragmentation problems common in 2D approaches [249]. Similarly, modified 2D U-Net models with dropout and batch normalization provided efficient coronary artery segmentation with reduced overfitting, while maintaining fast computation suitable for clinical environments [46]. Enhanced architectures further incorporated squeeze-and-excitation blocks, deformable decoders, or UNet++ frameworks

Table 2
Summary of deep learning models for the diagnosis of cardiovascular diseases in echocardiography.

Study	Task	DL Model	Dataset	Model Performance
[229]	Classification	Modified VGG16	Kaggle echocardiogram dataset (2404 images annotated by Marian Cardiac Centre, Poland)	Accuracy: 94.92%, F1-score 94.50%
[44]	Classification	CNN	340 patients, 17,000 labelled images, five standard views	Accuracy: 98.1%
[113]	Classification	3D-CNN	5600 TTE exams (Apical 4-chamber view); 30 frames per exam;	Accuracy: 78%, F1 Score 54.6% – 72.3%
[286]	Classification	PRCNN	An open-source multivariate dataset contains 132 instances with 12 attributes from the UCI Machine Learning Repository.	Accuracy: 99.5%
[66]	Segmentation	YOLOv7 and U-Net	The CAMUS dataset of the University Hospital of St. Etienne, consisting of echocardiographic images from 450 patients	DSC: LVendo: 92.63%, LVepi: 85.59%, LA: 87.57%
[230]	Classification	CNN	267 TTE studies from UCSF clinical database, Total ~ 834,267 images (still and video frames)	Accuracy 91.7%, F1-score: 0.904 ± 0.058 for still images
[231]	Classification	LSTM and VAE-CNN	120 patient data from Hope Clinic, Shillong, India, including 2D echo images, 3D Doppler images, and videographic images	Accuracy of LSTM:100%, VAE-CNN: 98%
[239]	Segmentation	Dual-Branch TransV-Net (DBTV)	120 pediatric 3DE images from SCMC	DSC: 0.91 (LV), 0.880 (RV)
[238]	Segmentation	nnU-Net (self-configuring 3D U-Net)	MITEA (MR-Informed Three-dimensional Echocardiography Analysis) dataset; 536 annotated 3DE images	DSC for myocardium: 0.77, cavity: 0.87
[88]	Segmentation	VDS-UNet (combination of UNet and VGG16)	153 echocardiographic videos, 2183 images from 49 subjects	DSC: LA 0.935, LV 0.92, MV 0.757
[89]	Classification and segmentation	ConvNet	356 2D transthoracic echocardiographic sequences from 114 subjects	Accuracy of iPhCH 95%, ConvNet 94%. DSC: SDM 9.4%, ConvNet 12.8%
[67]	Recognition	EfficientNetB3 (A4C view) and EfficientNetB4 (PLA view)	2044 transthoracic echocardiographic studies (1597 natural mitral valves, 447 prosthetic valves)	AUC: 0.99
[236]	Segmentation	MFP-U-Net	Public dataset and a prepared dataset from Rajaie Cardiovascular Medical and Research Center and Intelligent Imaging Technology Research Center	DSC: 0.953, HD: 3.49, MAD: 1.12
[234]	Classification	Fused CNNs	432 video images of echocardiography collected from Tsinghua University Hospital and Fuzhou University Hospital, comprising 93 different patients	Precision: 92.1% for fused CNNs, 89.5% for spatial CNN network alone
[237]	Segmentation	U-Net 2	AMUS dataset – 500 patients with 2D echocardiography (apical 2-chamber and 4-chamber views)	LVEndo DSC: ~0.92 (ED), ~0.89 (ES). LVEF MAE: 5.6%. LVEDV/ESV MAE: ~9.5 ml
[114]	Detection	YOLOv5s	61 echocardiographic images from 41 patients. LOGIQ S8 ultrasound device	Precision 95%, Recall 100%, F1-score 0.98
[115]	Assessment of RWMA	Deep CNN using ResNet, DenseNet, Inception-ResNet, Inception, and Xception	300 patients with myocardial infarctions and 100 age-matched control patients with normal wall motion	AUC for RWMA detection: 0.99 and 0.97 for ResNet
[245]	Motion estimation	PWC-Net architecture	Simulated data from an access database, in vivo data from 30 patients	Average endpoint error of (0.06 ± 0.04) mm per frame
[241]	Detection and segmentation	ARVNet	AIIMS (All India Institute of Medical Sciences), New Delhi. Total 1,000 fetal echocardiographic images	Accuracy 0.99, Precision 0.98, Recall 0.97, DSC 0.99
[246]	Segmentation + prediction	EfficientNet-b4 (encoder), Unet++ (decoder)	B-mode echocardiography (sagittal plane), Adult zebrafish,	DSC 0.967, IoU 0.937,
[68]	Segmentation	ResDUnet	CAMUS from the University Hospital of St. Etienne, France. Patients 500, images 4000	DSC 95.1%, HD 4.21 ± 1.23 mm, MAD 1.31 ± 0.43 mm
[247]	Segmentation	Custom encoder-decoder based on U-Net	TEE images from the ICARUS project, conducted at Wonju Severance Christian Hospital, Korea. The count of patients is 9.	DSC 0.899, IoU 0.822, Recall 0.904, Precision 0.901
[248]	Detection	CNN	echocardiographic Videos of 1411 children from the Hospital of	AUC: 0.91, Accuracy: 92.3%

(continued on next page)

Table 2 (*continued*)

Study	Task	DL Model	Dataset	Model Performance
			Zhejiang University; the final number included: 1376	

optimized with metaheuristics, improving detection of aortic dissection and CAD under challenging imaging conditions [61,92]. These studies demonstrated the versatility of U-Net derivatives in handling vessel structures of varying size and complexity across different CT datasets.

Besides segmentation, CNN-based models have been applied for fibrosis detection, CAD identification, and coronary artery lumen analysis. Myocardial fibrosis, typically assessed by late gadolinium-enhanced MRI, was directly detected from early contrast-enhanced cardiac CT using CNNs, reducing patient exposure and broadening diagnostic availability [250]. In CAD diagnosis, DenseNet-based CNNs (PCNN) combined with feature selection algorithms reduced redundancy in high-dimensional CT data, enabling robust disease classification across multi-institutional datasets [47]. Other pipelines integrated YOLOv7 for feature extraction, with hyperparameter-tuned UNet++ for improved CAD detection, and supported fuzzy image enhancement for low-quality scans [92]. Meanwhile, 3D U-Net segmentation of the coronary artery lumen employed preprocessing strategies such as patch-based cropping, subject-specific normalization, and augmentation to overcome noise and small-vessel-size challenges, yielding superior results compared to traditional graph-cut or level-set methods [93]. In parallel, the CT-FFRAI framework applied DL to curved multiplanar reconstructions of CCTA, integrating stenosis and calcification cues to handle tortuous and calcified vessels and enabling rapid vessel-specific ischemia prediction without computational fluid dynamics, thereby markedly improving functional CAD assessment [251].

Several frameworks extended segmentation tasks to encompass multiple cardiovascular structures. U-Net architectures on multicenter CCTA datasets successfully delineated eight structures, including aorta, vena cavae, pulmonary artery, and coronary sinus, achieving clinically acceptable Dice scores and reducing manual effort [60]. Whole-heart segmentation pipelines integrated regression CNNs with Dense V-Net models to automate delineation of ten cardiovascular structures in transcatheter aortic valve implantation (TAVI) planning, offering reproducible accuracy within seconds [252]. Similarly, CNNs trained on dual-energy CT datasets exploited virtual non-contrast (VNC) and CCTA alignment to generate reliable whole-heart segmentations transferable to standard non-contrast-enhanced CT, expanding applicability across imaging protocols [94].

DL has also been utilized for plaque detection and stenosis quantification. Automated CCTA-AI systems reconstructed coronary branches, classified plaques, and detected significant stenosis, achieving higher diagnostic accuracy than conventional CCTA while reducing analysis time by up to 85% [62]. Complementary approaches integrated U-Net segmentation with 3DNet classification to reconstruct artery trees and classify CAD cases, combining anatomical and clinical data for improved diagnosis [95]. Coronary artery calcium (CAC) scoring systems based on U-Net enabled automated heart localization and quantification across multi-cohort datasets, overcoming the labor-intensive manual process and supporting scalable cardiovascular risk assessment [64]. Supplementary classification frameworks using ResNet-50, VGG, and Inception ResNet v2 achieved high performance in detecting small calcifications by preprocessing CT scans into cropped subregions, enhancing sensitivity to localized features [253].

Further applications include pulmonary vascular tree analysis and perfusion imaging. A 3D CNN with graph-cut refinement classified arteries and veins in chest CT scans, handling dense pulmonary vasculature with higher accuracy than random forests or 2D CNNs [254]. In myocardial CT perfusion (CTP), CNN-based models processed rest and stress datasets to detect functionally significant CAD, achieving high diagnostic accuracy and reducing interpretation time to under one minute compared to expert review [255]. Moreover, a hybrid CNN-LSTM model learned both spatial plaque morphology and temporal slice progression patterns from CT data, enabling early CVDs prediction and outperforming traditional approaches through enhanced spatio-temporal representation learning [256]. DL image reconstruction (DLIR) further improved image quality in CCTA by significantly reducing noise while preserving diagnostic fidelity, offering clearer visualization than iterative reconstruction methods [63]. Complementary to DLIR, a deep super-resolution network for ultra-low-dose CCTA significantly enhanced vessel sharpness and reduced noise, enabling reliable stenosis and plaque assessment at greatly reduced radiation exposure, thereby supporting safer high-quality CAD imaging [257]. Other models focused on quantifying epicardial adipose tissue volume (EATV), correlating with coronary disease risk. CNN-based vessel segmentation frameworks provided accurate EATV measurements with superior sensitivity and specificity compared to classical algorithms [96]. Similarly, fully convolutional networks (FCNs), such as FCN-all-at-once-VGG16, are efficient for left ventricular myocardial segmentation in large-scale CT datasets, significantly reducing time requirements compared to manual delineation [258].

A review of DL Models for diagnosing CVDs in the CCTA images is presented in Table 3. Only studies reporting clear, comparable performance metrics are included, while supplementary models are described in the text for a broader context.

4.3. Deep learning for cardiac MRI-based diagnosis

DL is now central to CMRI analysis, particularly for automated segmentation and functional assessment. FCNs, including VGG-16–based architectures, demonstrated strong performance on large population datasets such as UK Biobank, delivering fast, reproducible segmentation of ventricular and atrial structures [26]. Dual-stage FCN pipelines further improved segmentation accuracy by combining cavity localization with myocardium delineation and applying radial loss to reduce contour mismatch, which enhanced estimates of end-diastolic and stroke volumes [117]. Extensions of FCN models using residual and dilated blocks improved gradient flow and multi-scale feature extraction, which strengthened segmentation of small or irregular myocardial regions [49,116,124,125]. The nnU-Net has been successfully adapted from cine MRI to free-breathing real-time MRI. It showed good performance compared with commercial tools and reduced manual corrections under motion and stress-related artifacts [259]. For 4D flow MRI, a 3D U-Net achieved rapid, fully automated segmentation of cardiac chambers and great vessels, overcoming low contrast and cardiac motion and enabling reliable flow-based assessment within seconds [260].

Dense architectures also helped address data scarcity and class imbalance. Multi-scale residual DenseNet models, equipped with inception-style features and Dice–cross entropy loss, achieved state-of-the-art segmentation results on ACDC-2017, LV-2011, and the Kaggle Data Science Bowl datasets [171]. Region-of-interest cropping reduced background noise and improved segmentation of underrepresented classes. A Dense U-Net for phase-contrast MRI enabled automatic left atrial segmentation and flow quantification in atrial fibrillation, improving contrast handling and beat-to-beat reproducibility [261]. Hybrid approaches combined CNNs with deformable models or autoencoder-based strategies to improve robustness. Error-corrected CNNs reduced misclassification during myocardial infarction detection and right ventricular segmentation, while combined CNN-deformable model frameworks produced anatomically consistent left-ventricular contours in challenging datasets [50,98,131,171]. GANs expanded CMRI capabilities by augmenting datasets and refining segmentation. Progressive GANs synthesized CMRI images of congenital heart disease, addressing limited annotated data and privacy concerns [123].

Table 3
Summary of deep learning models for the diagnosis of cardiovascular diseases in CCTA images.

Study	Task	DL Model	Dataset	Model Performance
[45]	Segmentation	CNN	Chest CT images from 11 patients acquired at baseline prior to radiotherapy	Total accuracy of 87.2% ± 3.3%, individual chamber accuracy: LV: 87.8%, RV: 82.9%, LA: 88.6%, RA: 83.0%
[250]	Segmentation	3D Inception U-Net	51 annotated 3D CCTA volumes	DSC 0.917, HD 2.53
[46]	Segmentation	2D U-Net	CTCA scans of 69 subjects with chest pain from University College Hospital London and Barts Health NHS Trust	DSC: 91.20% for aorta and coronary arteries, 88.80% for coronary arteries alone
[251]	Segmentationand detection	Two-stage 3D U-Net, MFPN	Private dataset from Peking University First Hospital. 600 CCTA cases. Images Siemens SOMATOM Force dual-source CT scanner	Segmentation DSC 87.65%. Stenosis Detection accuracy: 92.00%, F1 Score 89.61%
[47]	Detection	CNN	Dataset 1 contains 500 patients,Dataset 2 has 200 patients	Accuracy: 99.2%, 98.73%; F1 score: 98.95, 98.82; AUC (ROC): 0.92, 0.91; AUC (PR): 0.96, 0.90
[92]	Detection	U-Net++ with YOLOv7	Dataset 1: 500 patients (2364 images). Dataset 2: 1000 patients (3D CCTA images)	Accuracy: Dataset 1 is 99.4%, Dataset 2 is 99.5%. F1 score: Dataset 1 - 98.60, Dataset 2 - 98.95
[93]	Segmentation	3D U-Net	Dataset 1: 34 subjects. Dataset 2: 18 subjects, segmented CTCA images	DSC of Dataset 1 is 0.7146 and Dataset 2 is 0.7406
[61]	Segmentationand detection	U-Net	20 contrast-enhanced CT datasets (10 with aortic dissection)	Sensitivity: 90.00%, Specificity: 80.00%, Accuracy: 85.00%
[62]	Detection	CNN	150 patients with confirmed coronary artery stenosis	Accuracy: 92%, Precision: 91%, Recall: 93%
[255]	Classification	3D CNN	COPDGene: 21 non-contrast CT scans. CTEPH: 33 contrast-enhanced CT scans (18 CTEPH patients, 15 controls).	Accuracy 93.6% on COPDGene, 89.1% on CT
[63]	Image restoration	DLIR	43 patients undergoing CCTA and ICA	sensitivity 92%, specificity 73%, accuracy 82%
[259]	Segmentation	FCN	1100 subjects with CAD, data obtained during end-diastolic phase	Sensitivity: 91.2%, Specificity: 99.7%, DSC: 88.3 ± 6.2%
[94]	Segmentation	3D CNN	18 patients scanned with a dual-layer detector CT scanner, 290 single-energy NCCT images from different vendors	DSC: 0.897 ± 0.034, ASSD: 1.42 ± 0.45 mm
[256]	Prediction	CNN	112 symptomatic patients undergoing CCTA, stress CTP, and ICA with FFR	AUC of CCTA alone is AUC 0.67,CCTA + stress CTP is 0.87, CCTA + CTP-DLrest is 0.96, CCTA + CTP-DL stress is 0.98
[95]	Segmentation and diagnosis	U-Net, 3DNet	443 patients with CCTA images	DSC: 0.771 ± 0.021, Accuracy: 0.750 ± 0.056, AUC: 0.737
[253]	Segmentation	CNN	71 patients undergoing TAVI, data split into training (55), validation (8), and test (8) sets	DSC: 0.920 (IQR: 0.906–0.925); computing time: 13.4s (range 11.9–14.9s)
[64]	Prediction	CNN	30,286 LDCTs from NLST for training; 2,085 subjects for validation; independent MGH dataset with 335 subjects	AUC: 0.871 on NLST test set, 0.924 on MGH dataset
[96]	Prediction	Deep CNN	150 patients (83 with vascular stenosis, 67 without)	Sensitivity: 0.8512, Specificity: 0.9899, Accuracy: 0.9623, AUC: 0.9813
[254]	Detection	Inception Resnet v2, VGG, ResNet-50	2400 CT images (1200 normal, 1200 with calcium)	Accuracy 98.52%

Conditional and dilation-based GAN variants improved ventricular boundary delineation, reconstructed corrupted slices, and stabilized performance under variable image quality [48,49,262].

DL-accelerated reconstruction methods have significantly improved 4D flow MRI. These approaches reduced scan time, enhanced spatial–temporal fidelity, and generated cleaner velocity fields, facilitating reliable assessment of valvular disease, diastolic dysfunction, shunt lesions, and pulmonary hypertension [263]. U-Net and its derivatives remain dominant in CMRI segmentation. Residual connections, generalized Dice or Jaccard losses, and batch normalization improved segmentation of the left ventricle and myocardium in the Sunnybrook and ACDC datasets, addressing intensity heterogeneity, papillary muscle interference, and motion artifacts [99,124,125]. Specialized extensions such as DeepWF performed water–fat separation and parametric mapping directly from multi-echo MRI, reducing artifacts produced by conventional reconstruction workflows [100]. Some frameworks

integrated U-Net with atlas-based localization or deformable constraints to enhance segmentation of anatomically variable chambers, including the left and right ventricles [50,131]. A late-fusion U-Net that combined magnitude and velocity information improved left-ventricle segmentation from 4D flow MRI, reduced dependence on cine MRI, and produced accurate volumetric and flow biomarkers [264].

DL also supports classification and regression tasks. CardioViewNet used transfer learning to perform cardiac view recognition and replaced inconsistent DICOM metadata with automated plane identification [265]. Multi-view CNN regression models estimated ventricular volumes and ejection fraction directly from slices, improving efficiency relative to segmentation-dependent workflows [266,267]. Indices-Net combined autoencoder-based features with multi-output regression to estimate wall thickness and cavity areas across the cardiac cycle [268]. Video-based Swin Transformers integrated cine and LGE sequences to classify multiple CVDs phenotypes and achieved diagnostic performance comparable to that of expert cardiologists [102].

Supplementary applications demonstrate the expanding role of DL in CMRI. Automated pipelines for 4D flow MRI performed static-tissue detection and phase-offset correction, improving pulmonary and systemic flow estimation and aligning closely with invasive measurements [269]. CNN-based detectors accurately identified end-diastolic and end-systolic frames in free-breathing cine MRI, reducing manual workload [120]. DL-ESPIRiT accelerated cine reconstruction up to 12-fold while preserving image fidelity [270]. DL classifiers of myocardial delayed enhancement detected fibrosis patterns with high diagnostic accuracy [70]. Mask R-CNN architectures further improved the segmentation of thin structures such as the right ventricle [271]. Multi-modal learning using a 3D U-Net on paired MRI/CT datasets enhanced the segmentation of chambers and pulmonary veins, supporting improved radiotherapy planning [101]. Besides segmentation and classification, DL has advanced CMRI-based functional assessment. Networks trained on tagged MRI estimated myocardial strain by jointly learning motion and deformation, achieving expert-level accuracy in large population datasets [272]. Cine-based models such as DeepStrain extended strain estimation to routine cine MRI and produced global and regional strain metrics that showed strong agreement with tagging-based references [36,273].

A summary of DL models for diagnosing CVDs in CMRI is provided in Table 4, focusing on studies with clearly reported and comparable metrics, whereas other relevant models are discussed in the text.

Table 4
Summary of deep learning models for diagnosing cardiovascular diseases in CMRI.

Study	Task	DL Model	Dataset	Model Performance
[26]	Segmentation + diagnosis	FCN	UK Biobank dataset consisting of 4,875 subjects with 93,500 pixel-wise annotated images	DSC of LV cavity 0.94, LV myocardium 0.88, RV cavity 0.90
[99]	Segmentation	FCN	Sunnybrook Database with 805 images	DSC: 0.92 (SGD), 0.90 (RMSProp); HD: 4.48 (SGD), 5.43 (RMSProp); Jaccard: 0.97 (SGD and RMSProp); Sensitivity: 0.92 (SGD), 0.90 (RMSProp); Specificity: 0.99 (SGD and RMSProp)
[262]	Segmentation	CCGAN	ACDC 2017 dataset (150 cases)	Dice index: 0.951 (LV), 0.921 (RV), 0.909 (MYO)
[265]	Classification	Fine-tuned CNN	Two cardiac studies (200 patients and 15 healthy volunteers)	Average F1 score: 97.66%; Precision: 97.82%; Recall: 97.62%
[266]	Prediction	Deep CNN	Data Science Bowl Cardiac Challenge Data (1140 subjects, 102,600 images)	RMS Error: EDV (10.6 ml), ESV (7.56 ml), EF (4.98%)
[124]	Segmentation	Residual Block-based DLN	ACDC training data (100 subjects) and free-breathing CMR data (12 subjects)	DSC: 0.919 (LV), 0.806 (myocardium), 0.818 (RV)
[125]	Segmentation	Modified U-Net	Sunnybrook Cardiac Dataset (45 cine-MRI images)	DSC: 0.9001, MAE: 0.0094, MSE: 0.0093
[120]	Detection	CNN	Free-breathing CMR data from 10 healthy subjects	Sensitivity: 90%, Specificity: 99%, Dice Index: 0.90
[287]	Detection and quantification	CNNEC	MRI images from MI patients and healthy controls	Accuracy: 95.3%, Sensitivity: 95.33%, Specificity: 98.70%
[116]	Segmentation	FR-Net	Sunnybrook public dataset (45 cases)	DSC: 0.93; APD: 1.41 mm
[117]	Segmentation and quantification	FCN	ACDC-2017 dataset and local dataset	DSC: 0.96; HD: 6.31 mm
[121]	Detection	CNN with YOLOv3	Multi-resource fetal echocardiogram dataset (Shenzhen Maternal and Child Health Hospital)	Accuracy: 94.84% and ADE: ES (1.25 frames)
[131]	Segmentation	CNN and deformable models	MICCAI 2009 LV segmentation challenge (45 datasets)	DSC: 0.94 and APD: 1.81 mm
[171]	Segmentation	FC-DenseNets	ACDC-2017, LV-2011, Kaggle Data Science Bowl	DSC: 0.96 (LV), 0.90 (RV), 0.91 (MYO); Accuracy: 100%

(continued on next page)

4.4. Deep learning for nuclear imaging-based diagnosis

DL has addressed multiple challenges in nuclear cardiology, particularly noise reduction, radiation exposure, and attenuation artifacts. A 3D convolutional autoencoder reconstructed standard-dose quality from low-dose SPECT-MPI images of 930 patients, enabling imaging at one-sixteenth of the clinical dose without compromising diagnostic fidelity [71]. Complementary approaches using DCNNs and GANs directly corrected attenuation and scatter artifacts from non-corrected SPECT, eliminating the need for CT-based correction, reducing misregistration errors, and producing clinically reliable reconstructions within seconds [42,52,74,76,122,274]. Recently, a 3D U-Net–based Feature Alignment Attenuation Correction Network (FA-ACNet) aligned SPECT features with CT guidance during training, so far operated CT-free at inference, reducing soft-tissue attenuation artifacts and improving perfusion defect detection without extra radiation exposure [146].

CNN-based architectures achieved strong diagnostic performance across different datasets for CAD detection and ischemia classification. VGG16 trained on 216 polar maps with transfer learning and augmentation improved accuracy compared with semi-quantitative polar map analysis [41]. Slice-based CNNs on 1,413 SPECT images retained anatomical and pathological features lost in polar maps, enhancing CAD detection [126]. Three-dimensional CNNs applied to 979 CZT SPECT scans combined with preprocessing pipelines (cropping, alignment, and

Table 4 (*continued*)

Study	Task	DL Model	Dataset	Model Performance
				on ACDC-2017 test set
[48]	Image imputation	cGAN	UK Biobank (4848 subjects), ACDC	SSIM: 0.872, PSNR: 26.88; Correlation: 0.991 (LV volume), 0.977 (LV mass), 0.961 (RV volume)
[49]	Image reconstruction	CNN, U-Net	90 cardiac MR exams (1200 acquisitions)	SSIM: 0.91 (water), 0.90 (fat); Pearson's correlation: R2 $\geq$ 0.97
[100]	Quantitative analysis	CNN	175 subjects (350 MRI scans)	DSC: 0.93 (bounding box), 0.80 (myocardial segmentation); ICC: 0.89
[50]	Segmentation	CNN and stacked autoencoder	MICCAI 2012 RV Segmentation Challenge (48 subjects)	DSC: 82.5%, HD: 7.85 mm
[101]	Segmentation	3D U-Net	32 left-sided whole-breast cancer patients (36 datasets)	DSC: 0.88 (chambers), 0.85 (great vessels), 0.77 (pulmonary veins); MDA: <2.0 mm
[270]	Image reconstruction	(2 + 1)D DL-ESPIRiT	2D cardiac cine MRI: 22 healthy volunteers, two pediatric patients	PSNR: Significantly higher ($P < 0.01$), SSIM: Significantly higher ($P < 0.01$)
[70]	Detection and classification	GoogLeNet, AlexNet, ResNet-152	1995 MDE images from 200 patients	Accuracy: 79.5% (GoogLeNet), 78.9% (AlexNet), 82.1% (ResNet-152); AUC: 0.938–0.948
[271]	Segmentation	Mask R-CNN with ResNet101 + FPN	16,200 NMR images from 128 patients	DSC: 0.92 (LV), 0.89 (RV); HD: 4.78 mm (LV), 7.03 mm (RV)
[102]	Diagnosis	VST	Multi-center CMR dataset (Beijing Fuwai Hospital and 7 others in China). Total subjects 9,719	AUC: 0.988 (screening), 0.991 (diagnosis); F1 score: 0.977 (screening), 0.906 (diagnosis)
[267]	Prediction	Multi-View Fusion CNN with VGG20-BN architecture	Data Science Bowl Cardiac Challenge Dataset (1140 subjects)	Volume prediction accuracy with an RMSE of 9.6 ml and correlation coefficient (R) of 0.974 for EDV, 7.1 ml RMSE and R of 0.976 for ESV, and 4.71% RMSE with a correlation of 0.828 for EF.
[268]	Prediction	DCAE and CNN	145 subjects (2900 images)	MAE: 1.44 $\pm$ 0.71 mm (wall thickness), 204 $\pm$ 133 mm^2 (areas);

Table 4 (*continued*)

Study	Task	DL Model	Dataset	Model Performance
				Correlation: 0.758 (wall thickness), 0.903 (areas)
[123]	Image generation	PG-GAN and U-Net	303 patients for PG-GAN training;	DSC: ~0.98 for PG-GAN, ~0.99 for actual data

sequence selection) improved classification and interpretability [103]. Lightweight models, such as RGB-CNNs trained on 224 SPECT MPI scans and CNN classifiers trained on 3,318 polar maps, reduced observer dependency and improved recognition of ischemia, infarction, and normal cases [43,51].

Hybrid and knowledge-based strategies were applied to small and imbalanced datasets. DL models incorporating VGG-16, VGG-19, and ResNet-18 with SVM achieved accuracies close to those of expert readers over 192 patient datasets, while heuristic methods based on color thresholding and segmentation complemented variability reduction [275]. On large-scale cohorts, CNNs trained on 37,243 stress-only SPECT scans learned directly from circumferential count profiles without reliance on population-specific normal databases [75]. Multicenter studies such as REFINE-SPECT (1,160 patients) demonstrated higher sensitivity and reproducibility than conventional perfusion deficit scoring, while ANN models trained on 1,001 Japanese multicenter SPECT studies improved ischemia detection by integrating perfusion, motion, and wall-thickening features [75,77].

DL has also been applied to auxiliary nuclear imaging tasks. CNN-based spatial transformer networks trained on 6,762 datasets automated reorientation, reducing operator dependence [276]. U-Net and DuRDN architectures generated attenuation μ-maps for dedicated and general-purpose cardiac SPECT, improving image reconstruction in scanners without CT capability [274]. Virtual attenuation correction models, such as DLAC polar maps trained on 11,532 studies, further enhanced diagnostic accuracy by learning spatial relationships between NAC and CTAC data [76].

In PET imaging, DL extended diagnostic capabilities. Modified ResNet-50 with multi-task learning applied to 944 PET polar maps improved impaired myocardial flow reserve prediction and stratification of cardiovascular risk traits [78]. Likewise, an nnU-Net V2 model trained on multi-tracer cardiac PET images performed robust heart-structure segmentation without CT, overcoming PET-CT misalignment and motion artifacts by learning directly from PET uptake patterns, thereby strengthening functional cardiac assessment [277]. Sequential hybrid imaging strategies combining CTA with selective PET MPI in 864 patients resolved CTA's limited predictive value by identifying functionally significant stenoses [72]. Moreover, convolutional models such as CAclassNet trained on [18F]-Florbetaben PET scans of 47 patients distinguished AL and ATTR amyloidosis subtypes from early acquisitions, addressing diagnostic challenges where visual interpretation was unreliable [79].

Table 5 reviews DL Models for diagnosing CVDs in nuclear imaging techniques, such as Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) modalities.

4.5. *Deep learning for X-ray angiography-based diagnosis*

DL has been widely applied to X-ray coronary angiography for vessel segmentation, stenosis detection, image enhancement, and motion correction. Vessel segmentation has been addressed with multiple architectures. Lightweight networks such as BRU-Net integrated bottleneck residual blocks, attention modules, and preprocessing (CLAHE, top-hat transforms) to handle noise, low contrast, and complex backgrounds [129]. DenseNet121-based U-Nets with penalty generalized

Table 5
Summary of deep learning models for diagnosing cardiovascular diseases in nuclear imaging.

Study	Task	DL Model	Dataset	Model Performance
[41]	Classification	VGG16	216 patient cases with AC and NAC polar maps	Accuracy: 74.53%, Sensitivity: 75.00%, Specificity: 73.43%
[126]	Diagnosing	VGG-based network	1413 SPECT images (772 CAD, 641 non-CAD)	Accuracy: 86.14% for slice images, 82.57% for polar map images
[122]	Image translation	DeepAC	4,886 training images from Yale University, 604 testing images from the University of Zurich, and the University of Calgary	AUC for CAD: 0.79 (DeepAC) vs. 0.70 (NC) and 0.81 (AC)
[42]	Image reconstruction	CNN	100 SPECT/CT datasets	Voxelwise correlations: 97.7% (SPECTDL), 92.2% (SPECTNC)
[103]	Classification	3D-CNN	979 subjects' images (70x70 pixels) from Kaohsiung Chang Gung Memorial Hospital, Taiwan	Accuracy: 87.64%, sensitivity: 81.58%, Specificity: 92.16%
[43]	Diagnosis	CNN	224 patients' SPECT MPI scans (stress and rest images)	Accuracy: 93.47%, AUC: 0.936
[51]	Classification	CNN	SPECT Myocardial Perfusion Imaging (MPI): 3318 images (67% women, 33% men)	Accuracy: 75.62%, AUC: 0.845, sensitivity: 78.56%, specificity: 74.34%
[275]	Classification	Transfer Learning (VGG-16)	192 patients' SPECT MPI images	Accuracy 86%, sensitivity 100%, specificity: 71%
[72]	Prediction	CNN	864 patients' data from Turku University Hospital	AE rate of normal perfusion is 0.5%, reduced perfusion is 2.5%
[71]	Image denoising	3D CAE	930 patients' SPECT MPI images	Improved image quality: p-value < 0.01)
[52]	Image reconstruction	ResNet, UNet	99 patients' MPI-SPECT images	ResNet: ME −6.99, SSI 0.99; UNet: ME −4.41, SSI 0.98
[73]	Diagnosis	ANN	1001 training images, 364 validation images	ANN AUC: 0.92 (stress defects), 0.90 (stress-induced ischemia)
[74]	Image generation	cGAN	65 patients' SPECT/CT scans	NMAE: 3.60% ± 0.85%, MSE: 1.89e-4 ± 0.89e-4
[75]	Diagnosis	Deep-CNN	37,243 patients' stress-only MPI scans	AUC: 0.872 ± 0.002, Accuracy: 83.0%, Specificity: 85.6%, Sensitivity: 75.0%
[274]	Image reconstruction	U-Net, DuRDN	270 clinical studies (dedicated), 400 clinical studies (general)	Dedicated SPECT: NMSE: 1.20 ± 0.72% (full μ-maps), 1.69 ± 0.82% (truncated μ-maps); General-purpose SPECT: NMSE: 1.37 ± 1.16% (indirect), 2.57 ± 1.06% (direct)
[76]	Image reconstruction	U-Net-based DLAC	11,532 patients' stress SPECT MPI exams	DLAC TPD correlation with CTAC TPD: R2 = 0.85, AUC: 0.827
[276]	Image transformation	CNN + STN	322 patients (augmented to 6762 samples), Center A. 52 patients (Center A), 23 patients (Center B)	R^2 up to 0.994, RD-all = 2.47% (Center A)
[77]	Prediction	Deep CNN	1,160 patients' stress MPI scans	AUC: per-patient 0.81 vs. TPD 0.78, per-vessel 0.77 vs. TPD 0.73
[78]	Multi-task learning	ResNet-50	1,185 patients' PET MPI scans	AUC for impaired MFR: 0.94; AUC for sex: 0.81; AUC for smoking: 0.71
[79]	Classification	CNN, CAclassNet	47 subjects' PET/CT scans	Accuracy: 95.49%, Sensitivity: AL-CA 1, ATTR-CA 0.935, CTRL 0.809

dice (pGD) loss further improved segmentation on large-scale datasets by reducing false positives/negatives, outperforming conventional dice loss in cases of vessel overlap and catheter interference [81]. Other CNN-based pipelines, including dual-channel and patch-based designs, enhanced vessel delineation in noisy images [58,130], while U-Net derivatives enabled specialized applications such as vessel geometry reconstruction in TAVI [56], 3D artery reconstruction using ResNet backbones [106], and hybrid frameworks integrating XA with CTA for multimodal registration and depth recovery [82]. Spatio-temporal models (ST-FCN, OSVOS) used angiographic sequences to improve vessel continuity across motion phases, reducing segmentation errors compared to U-Net or multiscale CNNs [82,278]. Attention-based and multi-channel frameworks such as ASCARIS, EfficientUNet++–based pipelines, and MSN-A further improved robustness under contrast variability, motion artifacts, and overlapping structures [84,109,111,279]. A double-head ResUNet-18 architecture further enhanced angiographic interpretation by pairing contrast and non-contrast frames to improve calcification visibility under low contrast and in the presence of motion artifacts, enabling more reliable identification of tiny calcified lesions for CAD assessment [280].

Automated stenosis detection has also progressed considerably. Architectures such as Inception-V3 with self-attention improved diagnostic reproducibility by automating frame selection, grading stenosis $\geq 50\%$, and lesion localization with Grad-CAM [80]. Object detectors such as YOLOv8 achieved high precision and recall on the Mendeley and ARCADE datasets, localizing stenotic regions even in poor-quality images [54,86], while EfficientDet D3, using BiFPN, reduced false negatives and enhanced multi-scale feature representation [118]. Lightweight alternatives such as LRSE-Net, with depthwise separable convolutions and squeeze–excitation modules, enabled efficient and explainable stenosis classification [57]. CNN–RNN frameworks and CNN–transformer

hybrids improved automation by eliminating manual frame selection and capturing both temporal and global vessel features, thereby enhancing stenosis classification and myocardial bridge detection [85,105]. Complementarily, DenseNet-169-based learning on angiography predicted physiologically significant stenosis by capturing subtle patterns in intermediate lesions, reducing dependency on subjective visual reading and invasive FFR evaluation [281]. Despite limitations in projection, traditional computer-aided frameworks that combine Hessian enhancement, skeletonization, and lumen measurement also contribute to stenosis grading [107].

Supplementary advances targeted motion correction, radiation reduction, and noise suppression. LSTM-based models predicted cardiorespiratory motion from simulated and real sequences with submillimeter accuracy, reducing artifacts during interventions [104]. Super-resolution methods such as ASRNet reconstructed intermediate frames in rotational angiography, reducing radiation exposure while preserving diagnostic quality [53]. Image denoising models like SAID applied spatially adaptive optimization and dual-domain filtering to suppress Gaussian noise while preserving small vessels [119].

Expanding beyond angiography, integrated frameworks have combined segmentation and classification pipelines. ASCARIS, which coupled preprocessing with nested U-Net segmentation and VGG-16 classification, improved vessel delineation and categorization [279]. DenseNet121–InceptionResNet-v2 backbones achieved accurate external validation for segmentation and stenosis identification [111]. A multimodal learning framework combining ResNet-50 features from XA images with C-arm angulation data enabled robust classification of coronary anatomy despite motion blur and low-contrast conditions, supporting automated view selection and safer angiographic workflows [282]. Cardiovascular border analysis on chest radiographs (CB_auto) automated measurements linked with valvular disease indices [110], while multimodal fusion of retinal fundus and DXA data offered complementary vascular and systemic biomarkers for CVDs risk stratification [87].

In addition to angiography, CXR imaging has increasingly been used with DL to detect cardiomegaly, elevated pulmonary artery wedge pressure (a surrogate of heart-failure burden), and signs of pulmonary congestion on chest radiographs, which offers a rapid, low-cost screening tool for early cardiovascular assessment [283–285].

A review of DL Models for diagnosing CVDs in X-ray angiography modalities is presented in Table 6. Studies with clearly defined, comparable performance metrics are presented here, whereas supplementary models are discussed in the main text to provide a broader context.

4.6. Strengths and limitations of deep learning across cardiovascular imaging modalities

The following are the strengths and limitations of DL reported in the reviewed literature across various cardiovascular imaging modalities, including echocardiography, coronary CCTA, CMRI, nuclear imaging, and X-ray angiography.

4.6.1. Strengths of deep learning in cardiovascular imaging

DL models have demonstrated considerable strengths in cardiovascular imaging, particularly in automating segmentation, classification, and disease detection. In echocardiography, CNNs and transfer learning approaches have achieved excellent accuracy in disease classification, effectively reducing observer variability and increasing diagnostic consistency [44,229]. Similarly, CCTA analysis has benefitted from U-Net-based architectures, which have provided robust segmentation of cardiac chambers and coronary vessels, achieving Dice scores above 0.8 and minimizing manual annotation [45,46,249]. In CMRI, DL models like fully FCNs have shown impressive performance, achieving Dice scores exceeding 0.9 for ventricular segmentation, which enables near-expert agreement and reduces processing times from minutes to seconds [26,116]. Nuclear imaging has also benefited from CNN-based

Table 6
Summary of deep learning models for diagnosis of cardiovascular diseases: X-ray angiography images.

Study	Task	DL Model	Dataset	Model Performance
[80]	Detection and localization	GoogleNet Inception-V3 with self-attention mechanism	452 right coronary artery angiography clips, internal and external validation datasets	Frame-wise AUC 0.971, accuracy 0.934; clip-wise accuracy 0.965; external validation AUC 0.925 (single) and 0.956 (ensemble).
[54]	Detection	YOLOv8	1934 images from Mendeley	Precision: 99.4%, Recall: 100%, mAP: 99.5%, F1 Score: 99.7%
[55]	Image reconstruction	Multi-view CNN, TDAEn Dense, En LSTM Dec	Synthetic 3D dataset of 10,000 models from real subjects	Chamfer Distance: MvFCNN 0.0099, TDAEn Dense 0.0328, En LSTM Dec 1.544. MSE for radius: MvFCNN 0.0000058, TDAEn Dense 0.000146, En LSTM Dec 0.291
[81]	Segmentation	DenseNet121 with pGD	Angiographic images of major vessels from 1980 patients	DSC: 91.9 ± 8.7%, Precision: 91.3 ± 8.8%, Recall: 92.6 ± 9.6%
[130]	Segmentation	Multiple CNNs	120 X-ray angiographic images were acquired from a collaborating hospital.	Precision: 83.03%, Sensitivity: 77.74%, Specificity: 99.34%, F1 Score: 0.8007
[118]	Detection	EfficientDet-D3, RetinaNet with ResNet-50 + FPN, Faster R-CNN with ResNet-101	Invasive coronary angiography sequences were collected from 438 patients from a hospital in Portugal.	EfficientDet: Precision 0.55, Recall 0.79, F1 0.64, mAP 0.67; RetinaNet: Precision 0.58, Recall 0.69, F1 0.63, mAP 0.45; Faster R-CNN: Precision 0.58, Recall 0.70, F1 0.63, mAP 0.57
[56]	Identification	U-Net	2827 frames from 50 angiographic sequences	D1: 0.91 ± 0.86, D2: 138.48 ± 75.76, Accuracy: 0.99, Precision: 0.73, Recall: 0.87, Dice: 0.15, Jaccard: 0.74, AUROC: 0.94
[82]	Segmentation	EfficientNet-based U-Net	394 XA series from 68 patients, CTA images from 38 patients	Precision: 0.7563, Recall: 0.6922, F1: 0.7176; ADD: 0.8705 mm, 1.06 mm (markers),

(continued on next page)

Table 6 (*continued*)

Study	Task	DL Model	Dataset	Model Performance
				1.5706 mm (bifurcation points)
[57]	Detection	LRSE-Net	Deep stenosis detection dataset, angiographic dataset	Accuracy: 0.9549/ 0.9543, Sensitivity: 0.6320/ 0.8792, Precision: 0.5991/ 0.8944, F1-score: 0.6103/ 0.8944, Specificity: 0.9620/0.9733
[58]	Segmentation	Deep CNN	44 X-ray angiograms (512x512)	Accuracy: 93.5%, Sensitivity: 90%, Specificity: 97%, PPV: 96.7%, NPV: 90.6%
[106]	Image reconstruction	U-Net with ResNet encoder	2,342 LAD, 1,907 LCX, 1,523 RCA images	Accuracy: LAD 97.7% start and 94.9% end, LCX97.5% start and 89.8% end, RCA 96.4% start and 94.6% end
[129]	Segmentation	BRU-Net	DCA1: 134 images, CCA: 150 images	SE: 0.8770, SP: 0.9789, Accuracy: 0.9729
[294]	Segmentation	ST-FCN	186 patients, 267 vessel segments	DSC: 0.90, AC: 0.92, SN: 0.89, SP: 0.92, F1: 0.87
[108]	Detection	VGG16, ResNet50, Inception-v3	Synthetic dataset of 10,000 images, real dataset of 250 images	Inception-v3: Accuracy 0.95, Precision 0.93, Sensitivity 0.98, Specificity 0.92, F1-score 0.95
[109]	Segmentation	Multichannel FCN	148 X-ray angiographic image sequences	Precision: 0.923, Sensitivity: 0.921, Specificity: 0.981, F1-score: 0.922
[279]	Segmentation	ASCARIS: Attention-based Nested U-Net, VGG-16	130 X-ray coronary angiograms (300x300 pixels)	Accuracy: 97%, Sensitivity: 95%, Specificity: 93%, DSC: 0.916, Jaccard: 0.89, MSE: 27.2, PSNR: 35.2
[110]	Segmentation	U-Net, Mask R-CNN	3 hospitals, public dataset; 1555 CXR images	Accuracy: 93.9%, ICC: >0.98
[84]	Segmentation	EfficientUNet++ with EfficientNet-B5	123 regions from 117 images	Accuracy: 99.9%, Sensitivity: 95.1%, DSC: 94.8%
[85]	Detection	CNNs, Transformer	80 patients, 262 stenoses	DSC: 0.917, ASD: 1.39,

Table 6 (*continued*)

Study	Task	DL Model	Dataset	Model Performance
				Accuracy: 92.7%
[86]	Diagnosis	U-Net, YOLOv8	ARCADE dataset	DSC: 0.75 (U-Net), Precision: 0.53 (YOLOv8)
[111]	Segmentation	U-Net, ResNet101, DenseNet121, InceptionResNet-v2	3302 images from 2042 patients, 181 external images	F1: 0.917, Recall: 0.921 (DenseNet121)
[87]	Diagnosis	Hybrid (Multi-modal) Model	500 participants (QBB dataset)	best model: Accuracy 78.3%, F1-score 76.7%, MCC 0.566
[278]	Segmentation	OSVOS-based CNN	50 training, 10 testing sequences	DSC: 0.88, Accuracy: 95.67%

models for myocardial perfusion analysis, demonstrating diagnostic accuracy comparable to that of human experts and improving reproducibility [41,43]. In X-ray angiography, DL models such as YOLOv8 have excelled at vessel segmentation and stenosis detection, achieving high precision and recall, thereby enhancing clinical utility for automated CAD diagnosis [54,118]. These models enable faster and more consistent diagnostics by reducing human error and variability. Moreover, hybrid models that combine CNNs with temporal models such as LSTMs and attention mechanisms have achieved strong results in capturing dynamic features, including myocardial motion, thereby enhancing performance in echocardiography and CMRI [113,231].

4.6.2. *Limitations of deep learning in cardiovascular imaging*

Despite the remarkable advances, DL in cardiovascular imaging still faces substantial challenges. Dataset limitations are a key issue, as many models rely on small, often single-center datasets, limiting generalizability across diverse patient populations and imaging conditions. In echocardiography, the lack of pediatric and fetal datasets restricts model generalization to these groups [241,247,248,286]. Similarly, CCTA studies often suffer from imbalanced datasets, particularly when analyzing small coronary branches or rare anatomical conditions, which hampers model performance [45,47]. CMRI is also limited by the scarcity of annotated data for rare diseases, affecting model performance in such cases [116,287]. Class imbalance remains a challenge across all modalities. In nuclear imaging, for instance, underrepresented conditions, such as mild ischemia, reduce sensitivity, thereby decreasing model performance in these scenarios [41,103]. This is also evident in echocardiography, where overfitting occurs in models trained on small, homogeneous datasets, diminishing model accuracy when applied to diverse conditions [231,246]. Overfitting and generalization issues are pervasive when training on imbalanced datasets, as seen in CMRI, where models trained on homogeneous datasets struggle with generalization across patient populations [26]. Image quality issues significantly affect DL model performance across all modalities. In CMRI, motion artifacts and low contrast in small structures such as the right ventricle degrade model accuracy [99,116]. In CCTA, motion blur, calcification artifacts, and partial-volume effects complicate the accurate segmentation of plaques and vessel morphology [61,92]. X-ray angiography faces similar challenges, with overlapping anatomical structures, such as catheters and bones, making accurate vessel segmentation difficult [58,130].

Computational demands of DL models, particularly 3D models and transformer-based architectures, remain a major limitation. These models require significant computational resources, including high-memory GPUs and long training times, making them less feasible for real-time use in clinical settings. Models such as 3D U-Net and transformer-based approaches face scalability issues in CCTA and CMRI,

requiring substantial resources to process volumetric or high-resolution data [249,262]. The complexity of models also limits their applicability in echocardiography, where real-time analysis may be constrained by the computational intensity of models like YOLOv7–U-Net and ensemble architectures [66,68]. Interpretability remains a critical barrier to the clinical deployment of DL models. Techniques like Grad-CAM and SHAP can provide post-hoc explanations, but these methods do not fully address the need for interpretable models that offer real-time feedback in clinical settings. In CMRI, for instance, GAN-generated data, while diverse, may introduce artifacts that are difficult for clinicians to recognize, undermining the model's reliability in clinical applications such as myocardial infarction detection [48,123]. Similarly, X-ray angiography models, despite their high precision, still struggle to provide clinically interpretable results for subtle lesions and small vessel abnormalities, which are crucial for accurate diagnosis [54,86].

4.7. Critical insights from reviewed studies

Recent studies show that greater model complexity does not necessarily yield higher diagnostic accuracy. Performance depends primarily on data quality, annotation precision, and task design rather than network depth or parameter count. Foundation and hybrid models provide benefits mainly through transfer learning, multimodal fusion, and domain adaptation.

Because the reported metrics originate from heterogeneous datasets and evaluation protocols, their numerical values should be interpreted as approximate indicators of progress rather than absolute comparisons across studies. This approach preserves fairness while highlighting the collective evolution of DL methods in cardiovascular imaging. Although comparing different architectures trained on diverse data sources and dataset sizes is not entirely equitable, such analysis remains valuable for identifying which models tend to perform better or converge faster under specific imaging modalities and data conditions. Besides quantitative performance, this review also synthesizes essential aspects of each study, including dataset characteristics, methodological frameworks, architectural innovations, identified challenges, and corresponding solutions.

Many investigations are limited by small, single-center datasets, where preprocessing quality, standardized labeling, and effective augmentation often outweigh architectural differences. Simpler CNNs sometimes outperform deeper or transformer-based networks under these conditions due to stable training and clean annotations. Limited external validation further restricts generalizability, emphasizing the need for multicenter datasets and standardized acquisition protocols.

For classification tasks, accuracy typically plateaus after several thousand well-labeled samples, whereas segmentation tasks require fewer but more precise annotations. Transfer learning consistently improves convergence and stability in small or imbalanced datasets, while scratch training performs best only with large, diverse, and reliable data. Models trained on balanced, high-quality, and multi-institutional datasets demonstrate superior robustness and clinical reliability.

Recent work also indicates that ViTs, especially when pretrained or combined with convolutional backbones, can match or exceed CNN performance in CMRI and echocardiography. Their strength lies in capturing long-range spatial dependencies and integrating multimodal information. Future cardiovascular AI systems are expected to favor hybrid CNN–ViT architectures optimized for data efficiency and interpretability.

Performance metrics such as Dice, accuracy, and AUC are useful, but their clinical value depends on how they improve diagnostic decisions, quantification accuracy, and patient management [65]. Segmentation accuracy matters only when it enhances estimation of parameters such as ejection fraction, plaque burden, or perfusion defects, which guide specific therapeutic actions [65,288]. Screening applications accept moderate sensitivity if the negative predictive value remains high, whereas confirmatory diagnostic tasks require near-expert precision to influence treatment [102]. Small numerical gains, such as a 2% Dice or 1% AUC improvement, are clinically meaningful only when they reduce misclassification, improve workflow efficiency, or support more reliable reporting [288,289].

5. Cross-Modality comparison of cardiovascular imaging and Deep learning applications

This section compares the major cardiovascular imaging modalities and reviews the application of DL in each. It first summarizes the clinical roles and technical characteristics of echocardiography, CCTA, CMRI, SPECT/PET, and X-ray angiography.

5.1. Clinical and technical comparison of cardiovascular imaging modalities

Cardiovascular imaging modalities differ in both the type of information they provide such as functional, anatomical, tissue, or physiological, and in their technical trade-offs. Transthoracic echocardiography is widely used as a first-line, noninvasive modality. It enables real-time assessment of cardiac structure and function. However, image quality and measurement reproducibility depend strongly on acquisition conditions and operator expertise [290]. CCTA provides noninvasive visualization of the coronary arteries. It is primarily used for the evaluation of coronary artery disease. However, its clinical use requires exposure to ionizing radiation and typically the administration of iodinated intravenous contrast [291]. CMRI provides multiparametric assessment of cardiac structure and function. It also enables detailed myocardial tissue characterization. It does not involve ionizing radiation. However, examinations are typically longer and technically more demanding than ultrasound-based assessments [292]. In contrast, SPECT and PET use radiotracers to assess myocardial perfusion and related physiological processes, such as metabolism and viability. These modalities primarily provide functional rather than anatomical information. However, they have lower spatial resolution than CT or MR and rely on radioactive tracers [293]. Invasive X-ray coronary angiography remains the reference standard for real-time assessment of the coronary lumen and for interventional guidance. This is due to its high temporal and spatial resolution. However, the procedure is invasive and involves exposure to ionizing radiation [278,279]. Table 7 shows the core clinical roles and technical characteristics of major cardiovascular imaging modalities.

5.2. Comparison of Deep learning applications across modalities

Based on the detailed modality-specific discussions in Section 4, this section provides a concise cross-modality synthesis of how DL is applied across cardiovascular imaging modalities, focusing on representative tasks, shared strengths, and recurring constraints rather than exhaustive method-level coverage. In echocardiography, DL supports view classification, structural assessment, and functional estimation. Models achieve accurate view recognition across clinical settings and help reduce operator variability during acquisition [44,230]. DL improves the identification of prosthetic valves [67] and supports the recognition of common cardiac abnormalities [231]. Segmentation networks such as MFP-UNet enhance left-ventricular definition despite acoustic noise and variable image quality [236]. Large-scale 2D echocardiography datasets improve generalization, although strong dependence on operator skill and heterogeneous acquisition protocols remains a limitation [237]. CCTA benefits from DL in vascular segmentation, plaque assessment, and structural analysis. 3D Inception U-Net architectures provide accurate segmentation of aortic and coronary structures [249]. DL improves the detection of coronary calcium, supporting early risk assessment [253]. DL techniques also assist in detecting myocardial fibrosis in contrast-enhanced CT [250] and contribute to myocardial perfusion analysis [255].

Table 7
Core clinical roles and technical characteristics of major cardiovascular imaging modalities, as reported in the cited studies.

Modality	Primary clinical information	Spatial resolution	Temporal resolution	Invasiveness
Echo	Cardiac structure, valve motion, function (real-time)	Moderate	High	Noninvasive
CCTA	Coronary anatomy and plaque/stenosis evaluation	High	Moderate–high	Noninvasive
CMRI	Function + tissue characterization (multiparametric)	High	Moderate	Noninvasive
SPECT/ PET	Perfusion / physiology; metabolism/viability with tracers	Low–moderate	Low–moderate	Noninvasive
X-ray angio	Coronary lumen (real-time), interventional guidance	High	High	Invasive

CMRI provides high tissue contrast and detailed structural and functional information. DL enhances myocardial segmentation, ventricular quantification, and infarction assessment through convolutional architectures and error-correcting networks [287]. DL models support the identification and functional mapping of myocardial fibrosis in integrated CT-MRI workflows [253,258], though long scan times and high computational demands remain significant challenges. DL-based reconstruction and segmentation approaches improve efficiency and enable more practical CMRI analysis for longitudinal studies [258,287]. Nuclear imaging benefits from DL in attenuation correction, low-dose reconstruction, and perfusion classification. DL-based attenuation correction improves SPECT image quality using both direct and indirect strategies [274]. Classification models contribute to the automated interpretation of perfusion abnormalities [275]. DL-based cardiac image reorientation increases consistency and reduces manual intervention [276]. X-ray angiography remains essential for real-time assessment of coronary anatomy during interventions. DL enables vessel segmentation, coronary-tree extraction, and stenosis identification in fluoroscopic sequences [278,279]. Spatio-temporal fully convolutional architectures improve robustness in dynamic angiographic imaging [294]. Table 8 presents a comparison of DL applications across cardiovascular imaging modalities: representative tasks, strengths, and constraints.

6. Multimodal Deep learning for cardiovascular imaging

Multimodal DL integrates heterogeneous data sources to improve predictive performance and is particularly relevant in cardiovascular medicine, where different imaging modalities capture complementary pathological information, i.e., anatomy, tissue characterization, and perfusion [295]. In this review, the scope is restricted to cardiovascular imaging modalities, such as echocardiography, CCTA, CMRI, SPECT/PET, and X-ray angiography. The following discussion summarizes representative DL strategies for integrating information across these imaging sources.

Table 8
Comparison of deep learning applications across cardiovascular imaging modalities: representative tasks, strengths, and constraints, as reported in the above cited studies.

Modality	Representative DL tasks	Practical strengths	Common constraints affecting DL
Echo	Automated functional assessment; echo interpretation/ phenotyping	Scalable; automates routine quantification	Image quality variability; acquisition dependence
CCTA	Coronary segmentation; calcium scoring; CT-based fibrosis analysis	High-detail anatomic modeling and quantification	Motion/artifacts; annotation burden; radiation/contrast constraints
CMRI	Automated CMR analysis; structure segmentation and quantification	Strong multi-structure segmentation; supports large cohorts	Long/complex workflows; compute and pipeline complexity
SPECT / PET	DL-based attenuation correction; image quality/quantification improvement	Reduces artifacts; improves consistency	Low spatial resolution; tracer workflows; multi-center variability
X-ray angio	Coronary segmentation with temporal info; stenosis analysis	Potential for real-time procedural support	Domain shift; invasive data limits diversity

.

A well-established multimodal setting is hybrid nuclear–CT imaging, where CT provides anatomical context (and attenuation information) while PET/SPECT provides physiological/perfusion information. DL has been applied to hybrid PET/CT myocardial perfusion imaging for ischemia detection and interpretability, demonstrating the feasibility of learning clinically meaningful patterns from combined perfusion PET and CT data [296]. Related work has also shown the utility of DL in quantitative PET myocardial perfusion analysis for coronary disease assessment, supporting the broader role of learned models in PET-based perfusion interpretation [297]. For SPECT workflows, DL has been applied to attenuation correction and image quality enhancement. This represents an important enabling step for SPECT/CT-style pipelines. It supports more consistent perfusion assessment, especially when CT-based correction is limited or variable [122].

A second major multimodal direction is multi-sequence CMRI fusion, where multiple MR sequences provide complementary contrasts for myocardial structure and pathology. The Multi-Sequence Cardiac MR (MS-CMR) Segmentation challenge and subsequent analyses highlight that jointly using auxiliary MR sequences alongside LGE can improve robustness in LGE-related segmentation tasks [298]. More recent multi-sequence pathology-focused networks explicitly combine multiple CMR sequences (e.g., bSSFP, LGE, T2-weighted, mapping sequences) within a unified architecture to segment myocardial pathology, demonstrating practical DL fusion across CMR sequences [299].

Beyond within-modality fusion, CT and CMR can be complementary in clinical risk assessment, motivating multimodal predictive modeling. For example, a Radiology study showed that combining CCTA and stress CMR using machine learning improved the prediction of major adverse cardiovascular events compared with traditional risk scores, highlighting the value of multimodal integration for patient-level prognosis [300]. At a broader level, current surveys note that most large-scale studies still rely on single-modality data, and truly paired multimodal datasets remain relatively limited [301].

Regarding X-ray angiography and CCTA, DL has achieved moderate-to-high diagnostic accuracy when these modalities are analyzed separately. However, despite the frequent combined use of these modalities in clinical workflows, end-to-end fusion models that jointly learn from angiography and CT remain relatively underexplored, as also noted in recent systematic reviews and *meta*-analyses [302,303].

Across multimodal settings, the main technical strategies include early, intermediate, and late fusion. Early fusion combines inputs when spatial or temporal alignment is reliable, intermediate fusion integrates modality-specific features through dedicated encoders and attention mechanisms, and late fusion combines predictions from each modality. In practice, multimodal cardiovascular imaging remains constrained by the general challenges of data availability, alignment, and inter-institutional variability discussed in Section 2. Addressing these issues through standardized acquisition, harmonized data curation, and multicenter validation will be essential for the clinical translation of multimodal DL systems [295].

7. Emerging trends and future research

The evolution of DL in cardiovascular imaging is moving toward integrated, context-aware, and patient-centric frameworks. Methodological innovation is driven by hybrid architectures that combine convolutional precision with transformer-based global reasoning and graph-based relational modeling. Cross-domain integration now links imaging with clinical, genomic, and physiological data to achieve personalized diagnosis and prognosis. Advances in real-time computing and FL are enabling continuous analysis through telemedicine and wearable devices while preserving data privacy. These converging trends emphasize explainability, interoperability, and clinical translation, positioning DL as a reliable assistant for cardiovascular decision support and individualized patient management. Fig. 6 illustrates the focal directions shaping the future of DL in CVDs diagnosis, as summarized below.

i. Multi-Modal Data Integration: DL is moving apart from single-modality analysis toward combining information from echocardiography, CT, MRI, nuclear imaging, and clinical or genomic data. Such integration enables a more holistic and personalized understanding of cardiovascular health.
ii. XAI: Developing transparent and interpretable models remains a priority to strengthen clinician trust, support regulatory compliance, and promote safe clinical adoption.
iii. Integration with Augmented and Virtual Reality (AR/VR): AR and VR technologies enhance 3-D visualization of cardiovascular structures for improved surgical planning, training, and image-guided interventions.
iv. Predictive Risk Modeling and Personalized Stratification: DL models are increasingly used to integrate imaging with longitudinal and real-time data to identify high-risk patients early, guide preventive care, and support individualized treatment planning.
v. Human–AI Collaboration: Combining human expertise with AI-driven analytics improves decision-making in complex cases and enhances workflow efficiency.
vi. Novel and Automated DL Architectures: Next-generation architectures use GANs for image enhancement, data augmentation, and synthetic data generation to overcome data scarcity. Hybrid networks merge CNN precision with transformer attention and graph-based relational reasoning. End-to-end learning frameworks can learn directly from raw imaging data, reducing the need for handcrafted features and revealing novel imaging biomarkers.
vii. Telemedicine and Wearable Integration: Continuous data streams from wearable sensors, combined with imaging and clinical information, enable real-time cardiovascular monitoring and remote diagnostics through federated and privacy-preserving learning.
viii. Point-of-Care (POC) Imaging Diagnostics: Portable, AI-driven imaging devices such as handheld ultrasound and low-dose mobile X-ray systems are advancing bedside diagnostics and improving accessibility in resource-limited settings.

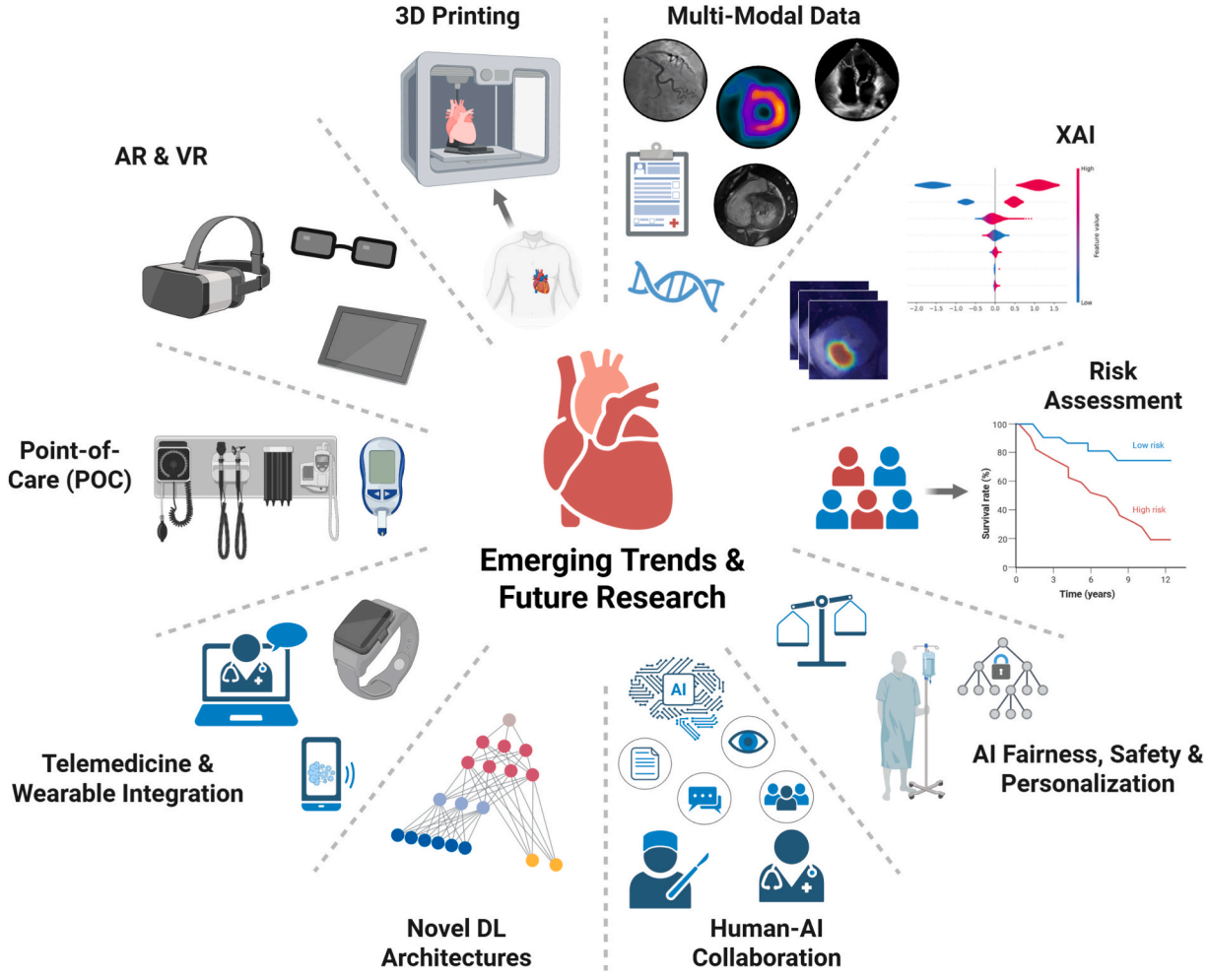

Fig. 6. Emerging trends and future research directions in deep learning for diagnosing cardiovascular diseases. Created in BioRender. Huang, B. (2025) https://BioRender.com/ktfaubh.

ix. Clinical and Commercial Integration: Major imaging vendors are incorporating DL algorithms into scanners and analysis software to automate acquisition, segmentation, and reporting, accelerating clinical translation and standardization.
x. Functional and Physiological Assessment: AI-based CT-fractional flow reserve (CT-FFR) and other physics-informed models extend DL applications apart from static image interpretation to noninvasive hemodynamic analysis and functional prediction. DL is also being applied to myocardial motion tracking, tissue characterization, and survival forecasting.
xi. 3D Printing: Advances in cardiac 3D printing enable DL models to generate anatomically accurate reconstructions that support preoperative planning, device sizing, and patient-specific simulations. Integrating imaging-driven DL predictions with physical models enhances visualization and personalized intervention planning, especially for complex congenital or structural heart disease.
xii. AI Fairness, Safety, and Personalization: Ensuring fairness, robustness, and transparency is essential as DL systems move closer to clinical deployment. Future research must address demographic bias, uncertainty quantification, and secure model adaptation to individual patients. These directions support equitable care, regulatory compliance, and trustworthy integration of AI into cardiovascular workflows.

8. List of acronyms and their full forms

Table 9 lists the acronyms and their full forms used in the preceding sections.

9. Conclusion

This review provides a comprehensive and modality-oriented synthesis of DL applications for CVDs diagnosis across echocardiography, CCTA, CMRI, nuclear imaging, and X-ray angiography. Across these modalities, DL has demonstrated substantial potential to improve segmentation, detection, classification, and functional assessment, thereby enhancing diagnostic accuracy, efficiency, and reproducibility. CNNs remain the dominant and most reliable backbone for spatial feature extraction in cardiovascular imaging, while transformer-based and graph-based architectures offer improved contextual modeling and structural reasoning, particularly for complex anatomy, multi-view data, and dynamic imaging sequences. However, their performance is strongly influenced by image quality, motion artifacts, acquisition protocols, and dataset diversity, and no single architecture consistently outperforms others across all modalities and tasks. Hybrid models that integrate convolutional, attention-based, and graph-based components currently represent the most promising direction, as they balance local precision with global contextual understanding and improved adaptability.

Despite these methodological advances, several critical barriers continue to limit clinical translation. These include limited and

Table 9
List of acronyms and their full form.

Acronym	Meaning	Acronym	Meaning
3D-CAE	3D convolutional autoencoder	HR	Hazard ratio
AB	Adaptive Boosting	IE	Interpolation error
ADE	Average detection error	IWNN	Iterative weighted nuclear norm
APD	Average perpendicular distance	LR	Logistic regression
ARVNet	Attention-residual V-Net	LRSE-Net	Low-rank and sparse encoding network
ASRNet	Angular super-resolution network	LSTM	Long Short-Term Memory
BiLSTM	Bidirectional Long Short-Term Memory	LV	Left ventricular
BRU-Net	Bottleneck Residual U-Net	LVH	Left ventricular hypertrophy
CA	Classification accuracy	MAD	Mean absolute distance
CAE	Convolutional autoencoder	MAE	Mean absolute error
CAMUS	Cardiac acquisitions for multi-structure ultrasound segmentation	MFPN	Multiscale feature pyramid network
CBAM	Convolutional block attention module	MFP-U-Net	Multi-feature pyramid U-Net
CCGAN	Conditional convolution generative adversarial network	MPE	Mean prediction error
CCTA	Cardiac computed tomography angiography	MPI	Myocardial perfusion imaging
cGAN	Conditional GAN	NB	Naive bayes
CHD	Coronary heart disease	NN	Neural network
CMRI	Cardiac magnetic resonance imaging	PCA	Principal Component Analysis
CNNEC	CNN system with error correction capabilities	PET	Positron emission tomography
CNNs	Convolutional neural networks	PG-GAN	Progressive generative adversarial networks
COPD	Chronic obstructive pulmonary disease	PRCNN	Pre-trained recurrent convolutional neural networks
CVDs	Cardiovascular diseases	PWC-Net	Pyramid, Warping, and Cost Volume Network
DCAE	Deep convolutional autoencoder	R(2 + 1)D CNN	Residual (2 + 1)D convolutional neural network
DL	Deep learning	RF	Random forest
DLAC	Deep learning attenuation correction	R-FCN	Region-based fully convolutional network
DLIR	Deep-learning image reconstruction	RNNs	Recurrent neural networks
DLN	Deep learning network	RWMA	Regional wall motion abnormality
DNN	Deep neural network	SAID	Spatially adaptive image denoising
DSC	Dice similarity coefficient	SCGAN	Semi-centralized generative adversarial network
DT	Decision trees	SPECT	Single-photon emission computed tomography
DuRDN	Dual squeeze-and-excitation residual dense network	ST-FCN	Spatio-temporal fully convolutional network
EATV	Epicardial adipose tissue volume	STN	Spatial transformer network
ESPIRiT	Eigenvector-based self-consistent parallel imaging reconstruction	SVM	Support vector machine
FC-DenseNets	Fully convolutional multi-scale residual DenseNets	TDA	Topological data analysis
FCN	Fully convolutional network	TDAEn	Topological data analysis encoder
FL	Federated learning	TL	Transfer learning
FR-Net	Focal residual neural network	VAE	Variational autoencoder
GB	Gradient boosting	VAE-CNN	Variational autoencoder with a convolutional neural network
GCNs	Graph convolutional networks	VDS-U-Net	VGG16-based deep supervision U-Net
GNNs	Graph neural networks	VST	Video-based swin transformer
HD	Hausdorff distance	XAI	Explainable AI
HF	Heart failure	XGB	Extreme gradient boosting

imbalanced annotated datasets, inter-institutional variability, domain shift, and the persistent lack of interpretability and standardized evaluation protocols. Moreover, the integration of multimodal imaging with clinical data remains underdeveloped, despite its clear potential to support more comprehensive and personalized diagnosis.

The major future research directions and translational priorities for cardiovascular AI are summarized in Section 7.

Declaration of Generative AI and AI-assisted technologies in the writing process

The authors used Grammarly for language refinement and BioRender for diagram creation, with all content subsequently reviewed, edited, and approved under their full responsibility.

Ethical approval.

Not applicable.

Data availability.

Not applicable.

CRediT authorship contribution statement

Inayatul Haq: Writing – review & editing, Writing – original draft, Validation, Resources, Methodology, Data curation, Conceptualization. **Haomin Liang:** Validation, Data curation. **Ke Zeng:** Validation. **Tao Wang:** Formal analysis, Investigation. **Izhar Uddin:** Validation, Data curation. **Jielong Lin:** Investigation, Resources. **Yan Kang:** Supervision, Resources, Investigation. **Bingding Huang:** Supervision, Investigation, Funding acquisition, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was supported by the Shenzhen Science and Technology Program (KJZD20240903095605007, JCYJ20250604145033042) and the Shenzhen Medical Research Fund (D250402003).

Data availability

No data was used for the research described in the article.

References:

[1] WHO (2021) Cardiovascular diseases (CVDs) fact sheet. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). Accessed April 2024.
[2] CDC. (2024) Heart Disease Facts. Centers for Disease Control and Prevention. https://www.cdc.gov/heart-disease/data-research/facts-stats/?CDC_AAref_Val=https://www.cdc.gov/heartdisease/facts.htm. Accessed April, 2024.
[3] WHO (2020) Death rate from cardiovascular diseases, 2019. Our World in Data. https://ourworldindata.org/grapher/death-rate-from-cardiovascular-disease-age-standardized-ghe?time=2019. Accessed April, 2024.
[4] Benjamin EJ, et al. (2018) Heart disease and stroke statistics—2018 update: a report from the American Heart Association. circulation 137(12):e67-e492. Doi: 10.1161/CIR.0000000000000558.
[5] Calafiero E, Jané-Llopis E (Year) The Global Economic Burden of Noncommunicable diseases. In: London: World Ecomnomic Forum, pp 1-48. https://ideas.repec.org/p/gdm/wpaper/8712.html.
[6] T.A. Gaziano, et al., The global cost of nonoptimal blood pressure, J. Hypertens. 27 (7) (2009) 1472–1477, https://doi.org/10.1097/HJH.0b013e32832a9ba3.
[7] A.V. Khera, S. Kathiresan, Genetics of coronary artery disease: discovery, biology and clinical translation, Nat. Rev. Genet. 18 (6) (2017) 331–344.
[8] G. Savarese, L.H. Lund, Global public health burden of heart failure, Card. Fail. Rev. 3 (1) (2017) 7.
[9] C.T. January, et al., 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice guidelines and the Heart Rhythm Society in collaboration with the Society of Thoracic Surgeons, Circulation 140 (2) (2019).
[10] C.M. Otto, B. Prendergast, Aortic-valve stenosis—from patients at risk to severe valve obstruction, N. Engl. J. Med. 371 (8) (2014) 744–756.
[11] D. Van Der Linde, et al., Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis, J. Am. Coll. Cardiol. 58 (21) (2011) 2241–2247.
[12] P. Elliott, et al., Classification of the cardiomyopathies: a position statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases, Eur. Heart J. 29 (2) (2008) 270–276.
[13] F.G.R. Fowkes, et al., Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis, Lancet 382 (9901) (2013) 1329–1340.
[14] E.S. Donkor, Stroke in the 21st century: a snapshot of the burden, epidemiology, and quality of life, Stroke Research and Treatment 2018 (1) (2018) 3238165.
[15] P.K. Whelton, et al., 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice guidelines, J. Am. Coll. Cardiol. 71 (19) (2018) e127–e248.
[16] E.M. Isselbacher, Thoracic and abdominal aortic aneurysms, Circulation 111 (6) (2005) 816–828.
[17] R.M. Lang, et al., Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging, European Heart Journal-Cardiovascular Imaging 16 (3) (2015) 233–271.
[18] S. Abbara, et al., SCCT guidelines for the performance and acquisition of coronary computed tomographic angiography: a report of the society of Cardiovascular Computed Tomography guidelines Committee: Endorsed by the north American Society for Cardiovascular Imaging (NASCI), J. Cardiovasc. Comput. Tomogr. 10 (6) (2016) 435–449, https://doi.org/10.1016/j.jcct.2016.10.002.
[19] J. Schulz-Menger, et al., Standardized image interpretation and post-processing in cardiovascular magnetic resonance-2020 update: Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing, J. Cardiovasc. Magn. Reson. 22 (1) (2020) 19.
[20] S. Dorbala, et al., Single photon emission computed tomography (SPECT) myocardial perfusion imaging guidelines: instrumentation, acquisition, processing, and interpretation, J. Nucl. Cardiol. 25 (5) (2018) 1784–1846.
[21] A. Gurav, et al., Coronary angiography: a review of the state of the art and the evolution of angiography in cardio therapeutics. Frontiers in Cardiovascular Medicine 11(1468888, 2024.
[22] F. Fortuni, et al., Advancements and applications of artificial intelligence in cardiovascular imaging: a comprehensive review, European Heart Journal-Imaging Methods and Practice (2024), 2(4):qyae136.
[23] B. He, et al., Blinded, randomized trial of sonographer versus AI cardiac function assessment, Nature 616 (7957) (2023) 520–524.
[24] H. Krishna, et al., Fully automated artificial intelligence assessment of aortic stenosis by echocardiography, J. Am. Soc. Echocardiogr. 36 (7) (2023) 769–777.
[25] R. Avram, et al., Automated assessment of cardiac systolic function from coronary angiograms with video-based artificial intelligence algorithms, JAMA Cardiol. 8 (6) (2023) 586–594, https://doi.org/10.1001/jamacardio.2023.0968.
[26] W. Bai, et al., Automated cardiovascular magnetic resonance image analysis with fully convolutional networks, J. Cardiovasc. Magn. Reson. 20 (1) (2018) 65, https://doi.org/10.1186/s12968-018-0471-x.
[27] A. Lin, et al., Deep learning-enabled coronary CT angiography for plaque and stenosis quantification and cardiac risk prediction: an international multicentre study, The Lancet Digital Health 4 (4) (2022).
[28] J. Narula, et al., Prospective deep learning–based quantitative assessment of coronary plaque by computed tomography angiography compared with intravascular ultrasound: the REVEALPLAQUE study, European Heart Journal-Cardiovascular Imaging 25 (9) (2024) 1287–1295.
[29] H.T. Soe, H. Iwata, Enhancing Myocardial Infarction Diagnosis: LSTM-based Deep Learning Approach Integrating Echocardiographic Wall Motion Analysis, Journal of Medical and Biological Engineering 44 (5) (2024) 704–710.
[30] V. Gulshan, et al., Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, JAMA 316 (22) (2016) 2402–2410.
[31] G.H. Tison, et al., Passive detection of atrial fibrillation using a commercially available smartwatch, JAMA Cardiol. 3 (5) (2018) 409–416.
[32] Y. Cao, et al., Utilization of artificial intelligence in echocardiographic imaging for congenital heart disease: a narrative review, Pediatric Medicine 8(9 (2025).
[33] P. Shrivastava, et al., A systematic review on deep learning-enabled coronary CT angiography for plaque and stenosis quantification and cardiac risk prediction, European Journal of Radiology Open 14(100652 (2025).
[34] M. Alsharqi, E.R. Edelman, Artificial Intelligence in Cardiovascular Imaging and Interventional Cardiology: Emerging Trends and Clinical Implications, Journal of the Society for Cardiovascular Angiography & Interventions 4 (3) (2025) 102558, https://doi.org/10.1016/j.jscai.2024.102558.
[35] M. Lachmann, et al., Deep Learning–Enabled Assessment of right Ventricular Function Improves Prognostication after Transcatheter Edge-to-Edge Repair for Mitral Regurgitation, Circ. Cardiovasc. Imaging 18 (1) (2025) e017005.
[36] Y. Wang, et al., StrainNet: improved myocardial strain analysis of cine MRI by deep learning from DENSE, Radiol. Cardiothorac. Imaging 5 (3) (2023) e220196.
[37] A.C. Kwan, et al., Deep learning-derived myocardial strain, Cardiovascular Imaging 17 (7) (2024) 715–725.
[38] S. Li, et al., Monitoring blood pressure and cardiac function without positioning via a deep learning–assisted strain sensor array, Sci. Adv. (2023), 9(32): eadh0615.

[39] O.A. Smiseth, et al., Myocardial strain imaging: theory, current practice, and the future, Cardiovascular Imaging 18 (3) (2025) 340–381.
[40] N. Yuan, et al., Prediction of coronary artery calcium using deep learning of echocardiograms, J. Am. Soc. Echocardiogr. 36 (5) (2023).
[41] N.D. Papathanasiou, et al., Automatic characterization of myocardial perfusion imaging polar maps employing deep learning and data augmentation, Hell. J. Nucl. Med. (2020).
[42] J. Yang, et al., Direct attenuation correction using deep learning for cardiac SPECT: a feasibility study, J. Nucl. Med. 62 (11) (2021) 1645–1652.
[43] N. Papandrianos, E. Papageorgiou, Automatic diagnosis of coronary artery disease in SPECT myocardial perfusion imaging employing deep learning, Appl. Sci. 11 (14) (2021) 6362.
[44] K. Kusunose, et al., Clinically feasible and accurate view classification of echocardiographic images using deep learning, Biomolecules 10 (5) (2020) 665.
[45] Dormer JD, et al. (Year) Heart chamber segmentation from CT using convolutional neural networks. In: Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging, pp 659-64. Doi: 10.1117/12.2293554.
[46] Cheung WK, et al. (2021) A computationally efficient approach to segmentation of the aorta and coronary arteries using deep learning. Ieee Access 9(108873-88.
[47] A.F. AlOthman, et al., Detecting coronary artery disease from computed tomography images using a deep learning technique, Diagnostics 12 (9) (2022) 2073, https://doi.org/10.3390/diagnostics12092073.
[48] Y. Xia, et al., Recovering from missing data in population imaging–Cardiac MR image imputation via conditional generative adversarial nets, Med. Image Anal. 67(101812 (2021).
[49] X. Yang, et al., DBAN: Adversarial network with multi-scale features for cardiac MRI segmentation, IEEE J. Biomed. Health Inform. 25 (6) (2020) 2018–2028.
[50] M.R. Avendi, et al., Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach, Magn. Reson. Med. 78 (6) (2017) 2439–2448, https://doi.org/10.1002/mrm.26631.
[51] Zahiri N, et al. (2021) Deep learning analysis of polar maps from SPECT myocardial perfusion imaging for prediction of coronary artery disease.
[52] S. Mostafapour, et al., Deep learning-guided attenuation correction in the image domain for myocardial perfusion SPECT imaging, J. Comput. Des. Eng. 9 (2) (2022) 434–447.
[53] T.Y. Siow, et al., Angular super-resolution in X-ray projection radiography using deep neural network: Implementation on rotational angiography, Biomedical Journal 46 (1) (2023) 154–162.
[54] Osama M, et al. (2024) Blood Vessels Disease Detection of Coronary Angiography Images using Deep learning Model.
[55] Atlı İ, Gedik OS (2022) 3D reconstruction of coronary arteries using deep networks from synthetic X-ray angiogram data. Communications Faculty of Sciences University of Ankara Series A2-A3 physical sciences and engineering 64 (1):1-20. Doi: 10.33769/aupse.1020175.
[56] L. Busto, et al., Automatic Identification of Bioprostheses on X-ray Angiographic Sequences of Transcatheter Aortic Valve Implantation Procedures using Deep Learning, Diagnostics 12 (2) (2022) 334.
[57] E. Ovalle-Magallanes, et al., LRSE-Net: Lightweight residual squeeze-and-excitation network for stenosis detection in X-ray coronary angiography, Electronics 11 (21) (2022) 3570.
[58] Nasr-Esfahani E, et al. (Year) Vessel extraction in X-ray angiograms using deep learning. In: 2016 38th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 643-6. Doi: 10.1109/EMBC.2016.7590784.
[59] Wang C, Smedby Ö (Year) Automatic whole heart segmentation using deep learning and shape context. In: Statistical Atlases and Computational Models of the Heart ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8, pp 242-9. Doi: 10.1007/978-3-319-75541-0_26.
[60] L. Baskaran, et al., Automatic segmentation of multiple cardiovascular structures from cardiac computed tomography angiography images using deep learning, PLoS One 15 (5) (2020), https://doi.org/10.1371/journal.pone.0232573.
[61] J. Cheng, et al., A deep learning algorithm using contrast-enhanced computed tomography (CT) images for segmentation and rapid automatic detection of aortic dissection. Biomedical Signal Processing and Control 62(102145, 2020.
[62] D. Han, et al., Deep learning analysis in coronary computed tomographic angiography imaging for the assessment of patients with coronary artery stenosis. Computer Methods and Programs in Biomedicine 196(105651, 2020.
[63] D.C. Benz, et al., Validation of deep-learning image reconstruction for coronary computed tomography angiography: impact on noise, image quality and diagnostic accuracy, J. Cardiovasc. Comput. Tomogr. 14 (5) (2020) 444–451, https://doi.org/10.1016/j.jcct.2020.01.002.
[64] R. Zeleznik, et al., Deep convolutional neural networks to predict cardiovascular risk from computed tomography, Nat. Commun. 12 (1) (2021) 715.
[65] D. De Santis, et al., Deep learning image reconstruction algorithm: impact on image quality in coronary computed tomography angiography, Radiol. Med. 128 (4) (2023) 434–444.
[66] M.J. Mortada, et al., Segmentation of anatomical structures of the left heart from echocardiographic images using Deep Learning, Diagnostics 13 (10) (2023) 1683.
[67] M. Vafaeezadeh, et al., A deep learning approach for the automatic recognition of prosthetic mitral valve in echocardiographic images. Computers in Biology and Medicine 133(104388, 2021.
[68] Y. Ali, et al., Echocardiographic image segmentation using deep Res-U network, Biomed. Signal Process. Control 64(102248 (2021), https://doi.org/10.1016/j.bspc.2020.102248.
[69] Wolterink JM, et al. (Year) Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease. In: International Workshop on Reconstruction and Analysis of Moving Body Organs, pp 95-102. Doi: 10.1007/978-3-319-52280-7_9.
[70] Y. Ohta, et al., Detection and classification of myocardial delayed enhancement patterns on MR images with deep neural networks: a feasibility study. Radiology, Artif. Intell. 1 (3) (2019).
[71] Ramon AJ, et al. (Year) Initial investigation of low-dose SPECT-MPI via deep learning. In: 2018 IEEE nuclear science symposium and medical imaging conference proceedings (NSS/MIC), pp 1-3. Doi: 10.1109/NSSMIC.2018.8824548.
[72] T. Maaniitty, et al., Prognostic value of coronary CT angiography with selective PET perfusion imaging in coronary artery disease, J. Am. Coll. Cardiol. Img. 10 (11) (2017) 1361–1370.
[73] K. Nakajima, et al., Diagnostic accuracy of an artificial neural network compared with statistical quantitation of myocardial perfusion images: a Japanese multicenter study. European Journal of Nuclear Medicine and Molecular Imaging 44(2280-9, 2017.
[74] L. Shi, et al., Deep learning-based attenuation map generation for myocardial perfusion SPECT, Eur. J. Nucl. Med. Mol. Imaging (2020).
[75] H. Liu, et al., Diagnostic accuracy of stress-only myocardial perfusion SPECT improved by deep learning, Eur. J. Nucl. Med. Mol. Imaging (2021).
[76] T. Hagio, et al., "Virtual" attenuation correction: improving stress myocardial perfusion SPECT imaging using deep learning, Eur. J. Nucl. Med. Mol. Imaging 49 (9) (2022) 3140–3149.
[77] J. Betancur, et al., Deep learning analysis of upright-supine high-efficiency SPECT myocardial perfusion imaging for prediction of obstructive coronary artery disease: a multicenter study, J. Nucl. Med. 60 (5) (2019) 664–670.
[78] M.W. Yeung, et al., Multi-task deep learning of myocardial blood flow and cardiovascular risk traits from PET myocardial perfusion imaging, J. Nucl. Cardiol. 29 (6) (2022) 3300–3310.
[79] M.F. Santarelli, et al., Deep-learning-based cardiac amyloidosis classification from early acquired pet images, Int. J. Cardiovasc. Imaging 37 (7) (2021) 2327–2335.
[80] J.H. Moon, et al., Automatic stenosis recognition from coronary angiography using convolutional neural networks. Computer Methods and Programs in Biomedicine 198(105819, 2021.
[81] Yang S, et al. (2019) Major vessel segmentation on x-ray coronary angiography using deep networks with a novel penalty loss function.
[82] T. Park, et al., Deep learning segmentation in 2d x-ray images and non-rigid registration in multi-modality images of coronary arteries, Diagnostics 12 (4) (2022) 778.
[83] E. Ovalle-Magallanes, et al., Attention Mechanisms Evaluated on Stenosis Detection using X-ray Angiography Images. Journal of advances, Applied & Computational Mathematics (2022).
[84] M. Nobre Menezes, et al., Coronary X-ray angiography segmentation using Artificial Intelligence: a multicentric validation study of a deep learning model, Int. J. Cardiovasc. Imaging 39 (7) (2023) 1385–1396.
[85] P. Zhou, et al., A framework of myocardial bridge detection with x-ray angiography sequence, Biomed. Eng. Online 22 (1) (2023) 101.
[86] M. Popov, et al., Dataset for Automatic Region-based Coronary Artery Disease Diagnostics using X-Ray Angiography Images, Sci. Data 11 (1) (2024) 20.
[87] H.R. Al-Absi, et al., Cardiovascular disease diagnosis from DXA scan and retinal images using deep learning, Sensors 22 (12) (2022) 4310, https://doi.org/10.3390/s22124310.
[88] H. Huang, et al., Segmentation of Echocardiography based on Deep Learning Model, Electronics 11 (11) (2022) 1714.
[89] C. Raynaud, et al., (Year) Handcrafted features vs ConvNets in 2D echocardiographic images, in: In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI, 2017, pp. 1116–1119, https://doi.org/10.1109/ISBI.2017.7950712.
[90] X. Yu, et al., Using deep learning method to identify left ventricular hypertrophy on echocardiography, Int. J. Cardiovasc. Imaging (2021) 1–11.
[91] J. Zhang, et al., Ensemble machine learning approach for screening of coronary heart disease based on echocardiography and risk factors, BMC Med. Inf. Decis. Making 21 (1) (2021) 187.
[92] A.R. Wahab Sait, A.K. Dutta, Developing a Deep-Learning-based coronary artery disease detection technique using computer tomography images, Diagnostics 13 (7) (2023) 1312.
[93] Huang W, et al. (Year) Coronary artery segmentation by deep learning neural networks on computed tomographic coronary angiographic images. In: 2018 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 608-11.
[94] S. Bruns, et al., Deep learning from dual-energy information for whole-heart segmentation in dual-energy and single-energy non-contrast-enhanced cardiac CT, Med. Phys. 47 (10) (2020) 5048–5060.
[95] Y. Li, et al., Automatic coronary artery segmentation and diagnosis of stenosis by deep learning based on computed tomographic coronary angiography, Eur. Radiol. 32 (9) (2022) 6037–6045.
[96] H. Wang, et al., (2021) Deep Learning-Based Computed Tomography Images for Quantitative Measurement of the Correlation between Epicardial Adipose Tissue volume and Coronary Heart Disease, Sci. Program. 1 (2021).

[97] M. Chen, et al., Diagnostic performance of deep learning-based vascular extraction and stenosis detection technique for coronary artery disease, Br. J. Radiol. 93 (1113) (2020).

[98] A. Bhan, et al., An assessment of machine learning algorithms in diagnosing cardiovascular disease from right ventricle segmentation of cardiac magnetic resonance images, Healthcare Anal. 3(100162 (2023).

[99] Romaguera LV, et al. (Year) Left ventricle segmentation in cardiac MRI images using fully convolutional neural networks. In: Medical Imaging 2017: Computer-Aided Diagnosis, pp 760-70. Doi: 10.1117/12.2253901.

[100] J.W. Goldfarb, et al., Water–fat separation and parameter mapping in cardiac MRI via deep learning with a convolutional neural network, J. Magn. Reson. Imaging 50 (2) (2019) 655–665.

[101] E.D. Morris, et al., Cardiac substructure segmentation with deep learning for improved cardiac sparing, Med. Phys. 47 (2) (2020) 576–586.

[102] Y.-R. Wang, et al., Screening and diagnosis of cardiovascular disease using artificial intelligence-enabled cardiac magnetic resonance imaging, Nat. Med. (2024) 1–10.

[103] J.-J. Chen, et al., Convolutional neural network in the evaluation of myocardial ischemia from CZT SPECT myocardial perfusion imaging: comparison to automated quantification, Appl. Sci. 11 (2) (2021) 514.

[104] F. Azizmohammadi, et al., Patient-specific cardio-respiratory motion prediction in x-ray angiography using LSTM networks, Phys. Med. Biol. 68 (2) (2023) 025010, https://doi.org/10.1088/1361-6560/acaba8.

[105] Cong C, et al. (Year) Automated stenosis detection and classification in x-ray angiography using deep neural network. In: 2019 IEEE international conference on bioinformatics and biomedicine (BIBM), pp 1301-8. Doi: 10.1109/BIBM47256.2019.8983033.

[106] M. Hwang, et al., A simple method for automatic 3D reconstruction of coronary arteries from X-ray angiography, Front. Physiol. 12(724216 (2021).

[107] T. Wan, et al., Automated identification and grading of coronary artery stenoses with X-ray angiography. Computer Methods and Programs in Biomedicine 167 (13-22, 2018.

[108] E. Ovalle-Magallanes, et al., Transfer learning for stenosis detection in X-ray coronary angiography, Mathematics 8 (9) (2020) 1510.

[109] J. Fan, et al., Multichannel fully convolutional network for coronary artery segmentation in X-ray angiograms, IEEE Access (2018).

[110] C. Kim, et al., A deep learning–based automatic analysis of cardiovascular borders on chest radiographs of valvular heart disease: development/external validation, Eur. Radiol. (2022) 1–12.

[111] S. Yang, et al., Deep learning segmentation of major vessels in X-ray coronary angiography, Sci. Rep. 9 (1) (2019) 16897.

[112] Z. Huang, et al., Iterative weighted nuclear norm for X-ray cardiovascular angiogram image denoising, SIViP (2017).

[113] Silva JF, et al. (Year) Ejection fraction classification in transthoracic echocardiography using a deep learning approach. In: 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), pp 123-8. Doi: 10.1109/CBMS.2018.00029.

[114] J. Chen, Gao Y (2021) the Role of Deep Learning-Based Echocardiography in the Diagnosis and Evaluation of the Effects of Routine Anti-Heart-Failure Western Medicines in elderly patients with Acute Left Heart failure, Journal of Healthcare Engineering 1 (2021).

[115] K. Kusunose, et al., A deep learning approach for assessment of regional wall motion abnormality from echocardiographic images, Cardiovascular Imaging (2020), 13(2_Part_1):374–81.

[116] M. Chen, et al., (Year) FR-NET: Focal loss constrained deep residual networks for segmentation of cardiac MRI, in: In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI, 2019, pp. 764–767, https://doi.org/10.1109/ISBI.2019.8759556.

[117] H. Abdeltawab, et al., A deep learning-based approach for automatic segmentation and quantification of the left ventricle from cardiac cine MR images, in: Computerized Medical Imaging and Graphics 81(101717, 2020, https://doi.org/10.1016/j.compmedimag.2020.101717.

[118] Stralen PV, et al. (Year) Stenosis detection in X-ray coronary angiography with deep neural networks leveraged by attention mechanisms. In: Proceedings of the 9th International Conference on Bioinformatics Research and Applications, pp 123-8. Doi: 10.1145/3569192.3569212.

[119] Z. Huang, et al., Spatially adaptive denoising for X-ray cardiovascular angiogram images, Biomed. Signal Process. Control (2018).

[120] F. Yang, et al., (2017) Convolutional Neural Network for the Detection of End-Diastole and End-Systole Frames in Free-Breathing Cardiac magnetic Resonance Imaging, Comput. Math. Methods Med. 1 (2017).

[121] B. Pu, et al., Fetal cardiac cycle detection in multi-resource echocardiograms using hybrid classification framework, Futur. Gener. Comput. Syst. (2021).

[122] A.D. Shanbhag, et al., Deep learning–based attenuation correction improves diagnostic accuracy of cardiac SPECT, J. Nucl. Med. 64 (3) (2023) 472–478.

[123] Diller G-P, et al. (2020) Utility of deep learning networks for the generation of artificial cardiac magnetic resonance images in congenital heart disease. BMC medical imaging 20(1-8.

[124] F. Yang, et al., (2019) a deep learning segmentation approach in free-breathing real-time cardiac magnetic resonance imaging, Biomed Res. Int. 1 (2019).

[125] Curiale AH, et al. (Year) Automatic myocardial segmentation by using a deep learning network in cardiac MRI. In: 2017 XLIII Latin American computer conference (CLEI), pp 1-6. Doi: 10.1109/CLEI.2017.8226420.

[126] N.T. Trung, et al., A deeplearning method for diagnosing coronary artery disease using SPECT images of heart, J. Sci. Technol. (2020).

[127] E.M. de Souza Filho, et al., Machine learning algorithms to distinguish myocardial perfusion SPECT polar maps. Frontiers in Cardiovascular Medicine 8(741667, 2021.

[128] K. Yasutomi, et al., A submillimeter range resolution time-of-flight range imager with column-wise skew calibration, IEEE Trans. Electron Devices 63 (1) (2015) 182–188.

[129] X. Tao, et al., A lightweight network for accurate coronary artery segmentation using x-ray angiograms. Frontiers in Public Health 10(892418, 2022.

[130] Yang S, et al. (Year) Automatic coronary artery segmentation in X-ray angiograms by multiple convolutional neural networks. In: Proceedings of the 3rd international conference on multimedia and image processing, pp 31-5. Doi: 10.1145/3195588.3195592.

[131] Avendi MR, et al. (2016) A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Medical image analysis 30(108-19. Doi: 10.1016/j.media.2016.01.005.

[132] X. Li, et al., A multi-task deep learning approach for real-time view classification and quality assessment of echocardiographic images, Sci. Rep. 14 (1) (2024) 20484.

[133] K. Chachadi, et al., Automated Coronary Artery Segmentation with 3D PSPNET using Global Processing and Patch based Methods on CCTA Images, Cardiovasc. Eng. Technol. (2025) 1–15.

[134] L.-S. Pan, et al., Coronary artery segmentation under class imbalance using a U-Net based architecture on computed tomography angiography images, Sci. Rep. 11 (1) (2021) 14493.

[135] X. Zhu, et al., Coronary angiography image segmentation based on PSPNet. Computer Methods and Programs in Biomedicine 200(105897, 2021.

[136] T. Liu, et al., Residual convolutional neural network for cardiac image segmentation and heart disease diagnosis, IEEE Access (2020).

[137] Jamuna V, et al. (Year) ARVNet: Attention-Residual V-Net for Fetal Cardiac Rhabdomyoma Segmentation in Echocardiography. In: 2025 8th International Conference on Trends in Electronics and Informatics (ICOEI), pp 1023-8. Doi: 10.1109/ICOEI65986.2025.11012982.

[138] Z. Lin, et al., (2025) Temporal features-fused vision retentive network for echocardiography image segmentation, Sensors 25 (6) (1909).

[139] Kanakatte A, et al. (Year) Heart region segmentation using dense VNet from multimodality images. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp 3255-8. Doi: 10.1109/EMBC46164.2021.9630303.

[140] Lou Z, et al. (Year) Whole heart auto segmentation of cardiac CT images using U-Net based GAN. In: 2020 13th international congress on image and signal processing, BioMedical engineering and informatics (CISP-BMEI), pp 192-6. Doi: 10.1109/CISP-BMEI51763.2020.9263532.

[141] Awang Damit DS, et al. (2024) Automated DeepLabV3+ based model for left ventricle segmentation on short-axis late gadolinium enhancementmagnetic cardiac resonance imaging images. International Journal of Electrical & Computer Engineering (2088-8708) 14(3):Doi: 10.11591/ijece.v14i3.pp3362-3371.

[142] A. Ben Khalifa, et al., Deep transfer Learning for Classification of late Gadolinium Enhancement Cardiac MRI Images into Myocardial Infarction, Myocarditis, and healthy classes: Comparison with Subjective Visual Evaluation, Diagnostics 15 (2) (2025) 207, https://doi.org/10.3390/diagnostics15020207.

[143] Cui H, et al. (2022) Deep U-Net architecture with curriculum learning for myocardial pathology segmentation in multi-sequence cardiac magnetic resonance images. Knowledge-based systems 249(108942.

[144] Atyam H, et al. (Year) Screening of Cardiomegaly using Ensemble Model of InceptionV3 and MobileNet. In: 2024 10th International Conference on Communication and Signal Processing (ICCSP), pp 1426-31. Doi: 10.1109/ICCSP60870.2024.10543609.

[145] H. Chao, et al., Deep learning predicts cardiovascular disease risks from lung cancer screening low dose computed tomography, Nat. Commun. 12 (1) (2021) 2963.

[146] P. Yang, et al., Deep learning-based CT-free attenuation correction for cardiac SPECT: a new approach, BMC Med. Imaging 25 (1) (2025) 38.

[147] M.A. Prieto Canalejo, et al., Synthetic attenuation correction maps for SPECT imaging using deep learning: a study on myocardial perfusion imaging, Diagnostics 13 (13) (2023) 2214.

[148] Y. Du, et al., Deep learning-based multi-frequency denoising for myocardial perfusion SPECT, EJNMMI Phys. 11 (1) (2024) 80.

[149] I.D. Apostolopoulos, et al., Deep learning-enhanced nuclear medicine SPECT imaging applied to cardiac studies, EJNMMI Phys. 10 (1) (2023) 6, https://doi.org/10.1186/s40658-022-00522-7.

[150] E. Alskaf, et al., Deep learning applications in myocardial perfusion imaging, a systematic review and meta-analysis, in: Informatics in Medicine Unlocked 32 (101055, 2022, https://doi.org/10.1016/j.imu.2022.101055.

[151] J. Teuho, et al., Classification of ischemia from myocardial polar maps in 15O–H2O cardiac perfusion imaging using a convolutional neural network, Sci. Rep. 12 (1) (2022) 2839.

[152] M. Mahtab, et al., Enhancing Heart Disease Detection in Echocardiogram Images using Optimized EfficientNetB3 Architecture, Journal of Computing & Biomedical Informatics 7 (02) (2024).

[153] Kausar A, et al. (Year) An improved dense v-network for fast and precise segmentation of left atrium. In: 2021 international joint conference on neural networks (IJCNN), pp 1-8. 10.1109/IJCNN52387.2021.9534418.

[154] Lei L, et al. (Year) A deep residual networks classification algorithm of fetal heart CT images. In: 2018 IEEE international conference on imaging systems and techniques (IST), pp 1-4. Doi: 10.1109/IST.2018.8577179.

[155] K. Mohite, et al., Amalgam based cardiovascular disease prediction using Xception with XGBoost model. Front Health, Inform (2024).
[156] Naghne R, et al. (Year) An efficient capsule-based network for 2D left ventricle segmentation in echocardiography images. In: 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp 1-4. Doi: 10.1109/EMBC40787.2023.10340175.
[157] L. Bargsten, et al., Capsule networks for segmentation of small intravascular ultrasound image datasets, Int. J. Comput. Assist. Radiol. Surg. 16 (8) (2021) 1243–1254, https://doi.org/10.1007/s11548-021-02417-x.
[158] Hampe N, et al. (Year) Graph attention networks for segment labeling in coronary artery trees. In: Medical Imaging 2021: Image Processing, pp 410-6. Doi: 10.1117/12.2581219.
[159] M. Lalinia, A. Sahafi, Coronary vessel segmentation in x-ray angiography images using edge-based tracking method, Sens. Imaging 25 (1) (2024) 32.
[160] Lu P, et al. (Year) Dynamic spatio-temporal graph convolutional networks for cardiac motion analysis. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), pp 122-5. Doi: 10.1109/ISBI48211.2021.9433890.
[161] Wang Q, et al. (Year) Automatic severity classification of coronary artery disease via recurrent capsule network. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp 1587-94. Doi: 10.1109/BIBM.2018.8621136.
[162] H. Lin, et al., Coronary heart disease prediction method fusing domain-adaptive transfer learning with graph convolutional networks (GCN), Sci. Rep. 13 (1) (2023) 14276.
[163] Vaanathi S, et al. (Year) Adaptive Attention-Enhanced Transformer With Modified Graph Cuts (AAET-MGC) Algorithm for Cardiovascular Disease Diagnosis. In: 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), pp 944-53. Doi: 10.1109/ICSSAS64001.2024.10760431.
[164] S. Aouthu, et al., Automated Diagnosis of Acute Cerebral Ischemic Stroke Lesions using Capsule Graph Neural Networks from Diffusion-weighted MRI Scans, Journal of Electrical Engineering & Technology 20 (4) (2025) 2631–2650, https://doi.org/10.1007/s42835-024-02120-2.
[165] Breen WA, et al. (Year) Predicting Cardiovascular Disease Risk and Outcomes Using a Graph Convolutional Network Algorithm. In: 2025 3rd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), pp 1-6. Doi: 10.1109/ICAECA63854.2025.11012177.
[166] N. Spier, et al., Classification of polar maps from cardiac perfusion imaging with graph-convolutional neural networks, Sci. Rep. 9 (1) (2019) 7569.
[167] Soni T, et al. (Year) Ensembled Deep Learning Technique for Cardiovascular Disease Detection: A CNN-GRU Approach. In: 2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC), pp 484-8. Doi: 10.1109/AIC61668.2024.10730954.
[168] Prasanna V, Divyaprabha P (Year) Optimized Hybrid CNN-GRU Approach for Accurate Prediction of Cardiovascular Diseases: Integrating Medical Imaging and Clinical Data. In: 2024 International Conference on Communication, Control, and Intelligent Systems (CCIS), pp 1-6. Doi: 10.1109/CCIS63231.2024.10931927.
[169] V. Sudha, D. Kumar, Hybrid CNN and LSTM network for heart disease prediction, SN Comput. Sci. 4 (2) (2023) 172.
[170] Tang J, et al. (Year) Cardiac motion tracking in short-axis MRI using Siamese convolution network. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp 865-70. Doi: 10.1109/BIBM47256.2019.8982995.
[171] M. Khened, et al., Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Medical Image Analysis 51(21-45, 2019.
[172] W. Li, et al., Myocardial pathology segmentation of multi-modal cardiac MR images with a simple but efficient Siamese U-shaped network. Biomedical Signal Processing and Control 71(103174, 2022.
[173] C. Li, et al., An 8-layer residual U-Net with deep supervision for segmentation of the left ventricle in cardiac CT angiography. Computer Methods and Programs in Biomedicine 200(105876, 2021.
[174] B. Zhao, et al., Deep Learning-based Segmentation and Localization in CT Angiography for Coronary Heart Disease Diagnosis, IEEE Access (2025).
[175] H. Chung, et al., Stenosis detection from time-of-flight magnetic resonance angiography via deep learning 3d squeeze and excitation residual networks, IEEE Access (2020).
[176] K.R. Singh, et al., MADRU-Net: Multiscale attention-based cardiac MRI segmentation using deep residual U-Net. IEEE Transactions on Instrumentation and Measurement 73(1-13, 2023.
[177] Y.-Z. Li, et al., RSU-Net: U-net based on residual and self-attention mechanism in the segmentation of cardiac magnetic resonance images. Computer Methods and Programs in Biomedicine 231(107437, 2023.
[178] W. Hijazi, et al., Deep learning-derived cardiac chamber volumes and mass from PET/CT attenuation scans: associations with myocardial flow reserve and heart failure, Circ. Cardiovasc. Imaging 18 (7) (2025).
[179] Y. Skandarani, et al., Generative adversarial networks in cardiology, Can. J. Cardiol. 38 (2) (2022) 196–203.
[180] Widodo RSS, et al. (Year) GAN Synergy: Transfer Learning-based Annotated Left Atrium MR Imaging Generation with Paired and Unpaired Models. In: 2024 International Seminar on Intelligent Technology and Its Applications (ISITIA), pp 627-32. Doi: 10.1109/ISITIA63062.2024.10668068.
[181] O. Tmenova, et al., CycleGAN for style transfer in X-ray angiography. International Journal of Computer Assisted Radiology and Surgery 14(1785-94, 2019.
[182] S. Karimi-Bidhendi, et al., Fully-automated deep-learning segmentation of pediatric cardiovascular magnetic resonance of patients with complex congenital heart diseases, J. Cardiovasc. Magn. Reson. 22 (1) (2020) 80.
[183] Y. Zhang, et al., Comparative analysis of U-Net and TLMDB GAN for the cardiovascular segmentation of the ventricles in the heart. Computer Methods and Programs in Biomedicine 215(106614, 2022.
[184] K. Sugimoto, et al., Detection and localization of myocardial infarction based on a convolutional autoencoder, Knowl.-Based Syst. (2019).
[185] S. Gurusubramani, B. Latha, Enhancing cardiac diagnostics through semantic-driven image synthesis: a hybrid GAN approach, Neural Comput. & Applic. 36 (14) (2024) 8181–8197.
[186] Shaker MS, et al. (Year) Cardiac MRI view classification using autoencoder. In: 2014 Cairo International Biomedical Engineering Conference (CIBEC), pp 125-8. Doi: 10.1109/CIBEC.2014.7020935.
[187] A. Radhakrishnan, et al., Cross-modal autoencoder framework learns holistic representations of cardiovascular state, Nat. Commun. 14 (1) (2023) 2436.
[188] Li S, et al. (Year) Semi-supervised cardiac MRI segmentation based on generative adversarial network and variational auto-encoder. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp 1402-5. Doi: 10.1109/BIBM52615.2021.9669685.
[189] O. Schad, et al., Real-time cardiac cine MRI: a comparison of a diffusion probabilistic model with alternative state-of-the-art image reconstruction techniques for undersampled spiral acquisitions. Magnetic Resonance in Medicine, 2025.
[190] Rombach R, et al. (Year) High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10684-95. 10.1109/CVPR52688.2022.01042.
[191] Sivaanpu A, et al. (Year) Denoising Echocardiography with an Improved Diffusion Model. In: 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp 1-4. Doi: 10.1109/EMBC53108.2024.10782561.
[192] C. Qin, et al., Convolutional recurrent neural networks for dynamic MR image reconstruction, IEEE Trans. Med. Imaging 38 (1) (2018) 280–290.
[193] Y. Wang, W. Zhang, A dense RNN for sequential four-chamber view left ventricle wall segmentation and cardiac state estimation. Frontiers in Bioengineering and Biotechnology 9(696227, 2021.
[194] A. Darmawahyuni, et al., Unidirectional-bidirectional recurrent networks for cardiac disorders classification, TELKOMNIKA (telecommunication Computing Electronics and Control) 19 (3) (2021) 902–910.
[195] F.T. Dezaki, et al., Cardiac phase detection in echocardiograms with densely gated recurrent neural networks and global extrema loss, IEEE Trans. Med. Imaging 38 (8) (2018) 1821–1832.
[196] C. Chen, et al., Deep learning for cardiac image segmentation: a review. Frontiers in Cardiovascular Medicine 7(25, 2020.
[197] V. Kumari, et al., Transformer and Attention-based Architectures for Segmentation of Coronary Arterial Walls in Intravascular Ultrasound: a Narrative Review, Diagnostics 15 (7) (2025) 848.
[198] C. Fan, et al., ViT-FRD: a vision transformer model for cardiac MRI image segmentation based on feature recombination distillation, IEEE Access (2023), 11 (129763–72.
[199] Y. Zhang, et al., ST-GAN: a Swin Transformer-based Generative Adversarial Network for Unsupervised Domain Adaptation of Cross-Modality Cardiac Segmentation, IEEE J. Biomed. Health Inform. 28 (2) (2023) 893–904.
[200] S. Park, M. Chung, Cardiac segmentation on CT Images through shape-aware contour attentions. Computers in Biology and Medicine 147(105782, 2022.
[201] A.E. Addo, et al., Transformer-based heart organ segmentation using a novel axial attention and fusion mechanism, The Imaging Science Journal 72 (1) (2024) 121–139, https://doi.org/10.1080/13682199.2023.2198394.
[202] Y. Xu, et al., Hybrid U-Net and Swin-transformer network for limited-angle cardiac computed tomography, Phys. Med. Biol. 69 (10) (2024) 105012.
[203] R. Yang, et al., A fusion-attention swin transformer for cardiac MRI image segmentation, IET Image Proc. 18 (1) (2024) 105–115.
[204] Nguyen-Le Q-B, et al. (Year) RotCAtt-TransUNet++: Novel Deep Neural Network for Sophisticated Cardiac Segmentation. In: 2024 International Conference on Multimedia Analysis and Pattern Recognition (MAPR), pp 1-6. Doi: 10.1109/MAPR63514.2024.10660759.
[205] Zhang N, Li J (Year) MSF-TransUNet: A Multi-Scale Fusion Approach for Precise Cardiac Image Segmentation. In: Proceedings of the 2024 4th International Conference on Artificial Intelligence, Big Data and Algorithms, pp 1139-46. Doi: 10.1145/3690407.3690596.
[206] M. Mazher, et al., Self-supervised spatial–temporal transformer fusion based federated framework for 4D cardiovascular image segmentation, Inf. Fusion 106 (102256 (2024).
[207] D. Ouyang, et al., Video-based AI for beat-to-beat assessment of cardiac function, Nature 580 (7802) (2020) 252–256.
[208] Raghu M, et al. (2021) Do vision transformers see like convolutional neural networks? Advances in neural information processing systems 34(12116-28.
[209] A.M. Salih, et al., A review of evaluation approaches for explainable AI with applications in cardiology, Artif. Intell. Rev. 57 (9) (2024) 240.
[210] Y. Otaki, et al., Clinical deployment of explainable artificial intelligence of SPECT for diagnosis of coronary artery disease, Cardiovascular Imaging 15 (6) (2022) 1091–1102.
[211] A. Salih, et al., Explainable artificial intelligence and cardiac imaging: toward more interpretable models, Circ. Cardiovasc. Imaging 16 (4) (2023).
[212] R.J. Miller, et al., Explainable deep learning improves physician interpretation of myocardial perfusion imaging, J. Nucl. Med. 63 (11) (2022) 1768–1774.

[213] A. Singh, et al., Direct risk assessment from myocardial perfusion imaging using explainable deep learning, Cardiovascular Imaging 16 (2) (2023) 209–220.
[214] Raja G, et al. (Year) CardioNetFusion: A Deep Learning and Explainable AI Integrated System for Cardiovascular Disease Detection. In: 2024 IEEE International Conference on Data Mining Workshops (ICDMW), pp 388-95. Doi: 10.1109/ICDMW65004.2024.00056.
[215] L. Szabo, et al., Clinician's guide to trustworthy and responsible artificial intelligence in cardiovascular imaging. Frontiers in Cardiovascular Medicine 9 (1016032, 2022.
[216] M. Christensen, et al., Vision–language foundation model for echocardiogram interpretation, Nat. Med. 30 (5) (2024) 1481–1488.
[217] Zhou S, et al. (2025) Large language models for disease diagnosis: A scoping review. npj Artificial Intelligence 1(1):9.
[218] T.F. Lüscher, et al., Artificial intelligence in cardiovascular medicine: clinical applications, Eur. Heart J. 45 (40) (2024) 4291–4304.
[219] Li J, et al. (Year) Language-driven multimodal model for intravascular ultrasound image segmentation. In: IET Conference Proceedings CP862, pp 225-30. Doi: 10.1049/icp.2023.3315.
[220] F. Liu, et al., A multimodal multidomain multilingual medical foundation model for zero shot clinical diagnosis. Npj Digital Medicine 8(1):86, 2025.
[221] A.M. Alsaleh, et al., Few-shot learning for medical image segmentation using 3D U-Net and model-agnostic meta-learning (MAML), Diagnostics 14 (12) (2024) 1213, https://doi.org/10.3390/diagnostics14121213.
[222] M. Haupt, et al., Explainable Artificial Intelligence in Radiological Cardiovascular Imaging—A Systematic Review, Diagnostics 15 (11) (2025) 1399.
[223] J.-O. Lee, et al., Multimodal generative AI for interpreting 3D medical images and videos. NPJ Digital Medicine 8(1):273, 2025.
[224] J. Alvén, et al., PlaqueViT: a vision transformer model for fully automatic vessel and plaque segmentation in coronary computed tomography angiography, Eur. Radiol. 1–11 (2025), https://doi.org/10.1007/s00330-025-11410-w.
[225] Le A, et al. (Year) ViTAL-CT: Vision Transformers for High-Risk Plaque Classification in Coronary CTA. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp 681-90. Doi: 10.1007/978-3-032-04978-0_65.
[226] H. Aghapanah, et al., CardSegNet: an adaptive hybrid CNN-vision transformer model for heart region segmentation in cardiac MRI, Comput. Med. Imaging Graph. 115(102382 (2024), https://doi.org/10.1016/j.compmedimag.2024.102382.
[227] Granizo S, et al. (Year) A comparative analysis of vision transformers and convolutional neural networks in cardiac image segmentation. In: 2024 12th International Symposium on Digital Forensics and Security (ISDFS), pp 1-7. Doi: 10.1109/ISDFS60797.2024.10527254.
[228] D.J. Chung, et al., Echocardiogram vector embeddings via R3D transformer for the advancement of automated echocardiography. JACC, Advances 3 (9_Part_2) (2024) 101196.
[229] M. Raoof, et al., Heart Disease Classification from Echocardiogram Images using Deep Learning, IEEE Access (2024).
[230] A. Madani, et al., Fast and accurate view classification of echocardiograms using deep learning, npj Digital Med. 1 (1) (2018) 6.
[231] I. Wahlang, et al., Deep Learning methods for classification of certain abnormalities in Echocardiography, Electronics 10 (4) (2021) 495.
[232] A. Vrudhula, et al., Automated Deep Learning Phenotyping of Tricuspid Regurgitation in Echocardiography, JAMA Cardiol. 10 (6) (2025) 595–602.
[233] Y. Jang, et al., An artificial intelligence-based automated echocardiographic analysis: enhancing efficiency and prognostic evaluation in patients with revascularized STEMI, Korean Circulation Journal 54 (11) (2024) 743–756.
[234] Gao X, et al. (2017) A fused deep learning architecture for viewpoint classification of echocardiography. Information Fusion 36(103-13.
[235] Munafo R, et al. (2024) A deep learning-based fully automated pipeline for regurgitant mitral valve anatomy analysis from 3D echocardiography. IEEE Access 12(5295-308.
[236] S. Moradi, et al., MFP-Unet: a novel deep learning based approach for left ventricle segmentation in echocardiography, Phys. Med. (2019).
[237] S. Leclerc, et al., Deep learning for segmentation using an open large-scale dataset in 2D echocardiography, IEEE Trans. Med. Imaging 38 (9) (2019) 2198–2210.
[238] D. Zhao, et al., MITEA: a dataset for machine learning segmentation of the left ventricle in 3D echocardiography using subject-specific labels from cardiac magnetic resonance imaging. Frontiers in Cardiovascular Medicine 9(1016703, 2023.
[239] J. Zhang, et al., Dual-branch transV-net for 3-D echocardiography segmentation, IEEE Trans. Ind. Inf. 19 (12) (2023) 11675–11686.
[240] Papangelopoulou K, et al. (Year) Efficient left ventricle segmentation in 3D echocardiography using deep nnU-Net. In: 2023 IEEE International Ultrasonics Symposium (IUS), pp 1-4. Doi: 10.1109/IUS51837.2023.10306804.
[241] S. Sengan, et al., Echocardiographic image segmentation for diagnosing fetal cardiac rhabdomyoma during pregnancy using deep learning, IEEE Access (2022), 10(114077–91.
[242] M. Randazzo, et al., Added value of 3D echocardiography in the diagnosis and prognostication of patients with right ventricular dysfunction. Frontiers in Cardiovascular Medicine 10(1263864, 2023.
[243] K. Ta, et al., Multi-task learning for motion analysis and segmentation in 3D echocardiography, IEEE Trans. Med. Imaging 43 (5) (2024) 2010–2020.
[244] A. Ghorbani, et al., Deep learning interpretation of echocardiograms, npj Digital Med. 3 (1) (2020) 10.
[245] Østvik A, et al. (2021) Myocardial function imaging in echocardiography using deep learning. ieee transactions on medical imaging 40(5):1340-51.
[246] M.-H. Huang, et al., Assessing cardiac functions of zebrafish from echocardiography using deep learning, Information 14 (6) (2023) 341.
[247] S. Kang, et al., Left ventricle segmentation in transesophageal echocardiography images using a deep neural network, PLoS One 18 (1) (2023).
[248] X. Jiang, et al., A deep learning-based method for pediatric congenital heart disease detection with seven standard views in echocardiography, World Journal of Pediatric Surgery 6 (3) (2023).
[249] Ravichandran SR, et al. (Year) 3D inception U-Net for aorta segmentation using computed tomography cardiac angiography. In: 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), pp 1-4. Doi: 10.1109/BHI.2019.8834582.
[250] M. Penso, et al., A deep-learning approach for myocardial fibrosis detection in early contrast-enhanced cardiac CT images. Frontiers in Cardiovascular Medicine 10(1151705, 2023.
[251] B. Peters, et al., Diagnostic performance of a coronary CT angiography-based deep learning model for the prediction of vessel-specific ischemia, Eur. Radiol. (2025) 1–12.
[252] S. Sharobeem, et al., Validation of a whole heart segmentation from computed tomography imaging using a deep-learning approach, J. Cardiovasc. Transl. Res. 15 (2) (2022) 427–437.
[253] S. Lee, et al., Deep-learning-based coronary artery calcium detection from CT image, Sensors 21 (21) (2021) 7059.
[254] P. Nardelli, et al., Pulmonary artery–vein classification in CT images using deep learning, IEEE Trans. Med. Imaging 37 (11) (2018) 2428–2440.
[255] G. Muscogiuri, et al., Diagnostic performance of deep learning algorithm for analysis of computed tomography myocardial perfusion, Eur. J. Nucl. Med. Mol. Imaging 49 (9) (2022) 3119–3128.
[256] Lu J (Year) CT Image Data Analysis Framework Based on Computer Deep Learning: Cardiovascular Disease Prediction. In: Proceedings of the 2025 International Conference on Health Big Data, pp 23-7. Doi: 10.1145/3733006.3733010.
[257] L.-M. Zou, et al., Ultra-low-dose coronary CT angiography via super-resolution deep learning reconstruction: impact on image quality, coronary plaque, and stenosis analysis, Eur. Radiol. (2025) 1–11.
[258] H.J. Koo, et al., Automated segmentation of left ventricular myocardium on cardiac computed tomography using deep learning, Korean J. Radiol. 21 (6) (2020) 660.
[259] M. Schilling, et al., Assessment of deep learning segmentation for real-time free-breathing cardiac magnetic resonance imaging at rest and under exercise stress, Sci. Rep. 14 (1) (2024) 3754.
[260] M. Bustamante, et al., Automatic time-resolved cardiovascular segmentation of 4D flow MRI using deep learning, J. Magn. Reson. Imaging 57 (1) (2023) 191–203.
[261] J. Baraboo, et al., Deep learning based automated left atrial segmentation and flow quantification of real time phase contrast MRI in patients with atrial fibrillation, Int. J. Cardiovasc. Imaging 41 (6) (2025) 1197–1208, https://doi.org/10.1007/s10554-025-03407-9.
[262] Zhang H, et al. (Year) Conditional convolution generative adversarial network for Bi-ventricle segmentation in cardiac MR images. In: Proceedings of the third international symposium on image computing and digital medicine, pp 118-22. Doi: 10.1145/3364836.3364860.
[263] G. Reiter, et al., Four-dimensional flow MRI for a dynamic perspective on the heart and adjacent great vessels, Radiology 316 (2) (2025).
[264] X. Sun, et al., Deep learning based automated left ventricle segmentation and flow quantification in 4D flow cardiac MRI, J. Cardiovasc. Magn. Reson. 26 (1) (2024) 100003.
[265] J. Margeta, et al., Fine-tuned convolutional neural nets for cardiac MRI acquisition plane recognition, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 5 (5) (2017) 339–349.
[266] Luo G, et al. (Year) Cardiac left ventricular volumes prediction method based on atlas location and deep learning. In: 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp 1604-10. Doi: 10.1109/BIBM.2016.7822759.
[267] G. Luo, et al., Multi-views fusion CNN for left ventricular volumes estimation on cardiac MR images, IEEE Trans. Biomed. Eng. 65 (9) (2017) 1924–1934.
[268] W. Xue, et al., Direct multitype cardiac indices estimation via joint representation and regression learning, IEEE Trans. Med. Imaging 36 (10) (2017) 2057–2067.
[269] E. Valdeolmillos, et al., 4D flow cardiac MRI to assess pulmonary blood flow in patients with pulmonary arterial hypertension associated with congenital heart disease, Diagn. Interv. Imaging 105 (7–8) (2024) 266–272.
[270] C.M. Sandino, et al., Accelerating cardiac cine MRI using a deep learning-based ESPIRiT reconstruction, Magn. Reson. Med. 85 (1) (2021) 152–167.
[271] B. Yong, et al., Automatic ventricular nuclear magnetic resonance image processing with deep learning, Multimed. Tools Appl. (2021).
[272] E. Ferdian, et al., Fully automated myocardial strain estimation from cardiovascular MRI–tagged images using a deep learning framework in the UK Biobank, Radiol. Cardiothorac. Imaging 2 (1) (2020).
[273] M.A. Morales, et al., DeepStrain: a deep learning workflow for the automated characterization of cardiac mechanics. Frontiers in Cardiovascular Medicine 8 (730316, 2021.
[274] X. Chen, et al., Direct and indirect strategies of deep-learning-based attenuation correction for general purpose and dedicated cardiac SPECT, Eur. J. Nucl. Med. Mol. Imaging 49 (9) (2022) 3046–3060.
[275] S.K. Berkaya, et al., Classification models for SPECT myocardial perfusion imaging, in: Computers in Biology and Medicine 123(103893, 2020, https://doi.org/10.1016/j.compbiomed.2020.103893.

[276] D. Zhang, et al., A novel deep-learning–based approach for automatic reorientation of 3D cardiac SPECT images, Eur. J. Nucl. Med. Mol. Imaging 48 (11) (2021) 3457–3468.
[277] Y. Salimi, et al., Deep Learning-based CT-Less Cardiac Segmentation of PET Images: a Robust Methodology for Multi-Tracer Nuclear Cardiovascular Imaging, Journal of Imaging Informatics in Medicine (2025) 1–15.
[278] Zhao L, et al. (Year) Automated coronary tree segmentation for x-ray angiography sequences using fully-convolutional neural networks. In: 2018 IEEE Visual Communications and Image Processing (VCIP), pp 1-4. Doi: 10.1109/VCIP.2018.8698707.
[279] M. Algarni, et al., Multi-constraints based deep learning model for automated segmentation and diagnosis of coronary artery disease in X-ray angiographic images, PeerJ Comput. Sci. 8(e993 (2022), https://doi.org/10.7717/peerj-cs.993.
[280] E. De Rose, et al., A multi-model deep learning approach for the identification of coronary artery calcifications within 2D coronary angiography images, Int. J. Comput. Assist. Radiol. Surg. (2025) 1–9.
[281] F. Arefinia, et al., Non-invasive fractional flow reserve estimation using deep learning on intermediate left anterior descending coronary artery lesion angiography images, Sci. Rep. 14 (1) (2024) 1818, https://doi.org/10.1038/s41598-024-52360-5.
[282] A. Ravi, et al., Deep learning-based classification of coronary arteries and left ventricle using multimodal data for autonomous protocol selection or adjustment in angiography, Sci. Rep. 15 (1) (2025) 15186.
[283] K. Kusunose, et al., Deep learning approach for analyzing chest x-rays to predict cardiac events in heart failure. Frontiers in Cardiovascular Medicine 10(1081628, 2023.
[284] S.E. Sorour, et al., (2024) a deep learning system for detecting cardiomegaly disease based on cxr image, Int. J. Intell. Syst. 1 (2024).
[285] P. Thiam, et al., Unsupervised domain adaptation for the detection of cardiomegaly in cross-domain chest X-ray images. Frontiers in Artificial Intelligence 6(1056422, 2023.
[286] C. Balakrishnan, V. Ambeth Kumar, IoT-enabled classification of echocardiogram images for cardiovascular disease risk prediction with pre-trained recurrent convolutional neural networks, Diagnostics 13 (4) (2023) 775, https://doi.org/10.3390/diagnostics13040775.
[287] Wang S-H, et al. (Year) Myocardial infarction detection and quantification based on a convolution neural network with online error correction capabilities. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp 1-8. Doi: 10.1109/IJCNN48605.2020.9207090.
[288] E.K. Oikonomou, et al., Artificial intelligence in medical imaging: a radiomic guide to precision phenotyping of cardiovascular disease, Cardiovasc. Res. 116 (13) (2020) 2040–2054.
[289] K. Seetharam, et al., The role of artificial intelligence in cardiovascular imaging: state of the art review. Frontiers in Cardiovascular Medicine 7(618849, 2020.
[290] M.D. Grant, et al., Transthoracic echocardiography: beginner's guide with emphasis on blind spots as identified with CT and MRI, Radiographics 41 (4) (2021) 1022–1042.
[291] P.I. Ngam, et al., Computed tomography coronary angiography–past, present and future, Singapore Med. J. 61 (3) (2020) 109.
[292] J.I. Tudela Martínez, et al., Cardiac magnetic Resonance in adults: an Updated Review of the Diagnostic Approach to Major Heart Diseases, J. Clin. Med. 14 (20) (2025) 7323.
[293] R.S. Driessen, et al., Myocardial perfusion imaging with PET, Int. J. Cardiovasc. Imaging 33 (7) (2017) 1021–1031.
[294] T. Wan, et al., Automatic vessel segmentation in x-ray angiogram using spatio-temporal fully-convolutional neural network. Biomedical Signal Processing and Control 68(102646, 2021.
[295] H.-H.-V. Tran, et al., AI-Driven Multimodality Fusion in Cardiac Imaging: Integrating CT, MRI, and Echocardiography for Precision, Cardiol. Rev. 10 (2025) 1097.
[296] J. Teuho, et al., Explainable deep-learning-based ischemia detection using hybrid O-15 H2O perfusion positron emission tomography and computed tomography imaging with clinical data, J. Nucl. Cardiol. 38(101889 (2024).
[297] L.E. Juarez-Orozco, et al., Deep learning in quantitative PET myocardial perfusion imaging: a study on cardiovascular event prediction, Cardiovascular Imaging (2020), 13(1_Part_1):180–2.
[298] X. Zhuang, et al., Cardiac segmentation on late gadolinium enhancement MRI: a benchmark study from multi-sequence cardiac MR segmentation challenge, Med. Image Anal. 81(102528 (2022).
[299] J. Qiu, et al., MyoPS-Net: Myocardial pathology segmentation with flexible combination of multi-sequence CMR images. Medical Image Analysis 84(102694, 2023.
[300] T. Pezel, et al., A Machine Learning Model using Cardiac CT and MRI Data Predicts Cardiovascular events in Obstructive Coronary Artery Disease, Radiology 314 (1) (2025).
[301] X.-y. Yang, et al., Utilizing multimodal artificial intelligence to advance cardiovascular diseases, Precis. Clin. Med. (2025), 8(3):pbaf016.
[302] L. Tu, et al., Accuracy of deep learning in the differential diagnosis of coronary artery stenosis: a systematic review and meta-analysis, BMC Med. Imaging 24 (1) (2024) 243.
[303] Amal S, et al. (2022) Use of multi-modal data and machine learning to improve cardiovascular disease care. Frontiers in cardiovascular medicine 9(840262. Doi: 10.3389/fcvm.2022.840262.
[304] WHO (2022) WHO Mortality Database: Interactive platform visualizing mortality data. World Health Organization. https://platform.who.int/mortality/themes/theme-details/topics/topic-details/MDB/cardiovascular-diseases. Accessed December, 2025.
[305] A. Zeng, et al., ImageCAS: a large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images, Comput. Med. Imaging Graph. 109(102287 (2023).
[306] O. Bernard, et al., Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Trans. Med. Imaging 37 (11) (2018) 2514–2525, https://doi.org/10.1109/TMI.2018.2837502.