scientific reports

Check for updates

OPEN

A multimodal multipath AI system for assessing PAH after VSD correction on echocardiography and chest radiography images

Gang Luo[1,3], Huashu Liu[2], Zhixin Li[1], Zhixian Ji[3] & Silin Pan[3]✉

Developing a novel artificial intelligence (AI) system that can automatically detect pulmonary arterial hypertension (PAH) after correcting the ventricular septal defect (VSD) and to help clinicians make reasonable treatment plans. We analyzed data from 1,316 patients under 1 year old who underwent VSD surgery at Women and Children's Hospital, Qingdao University and Qingdao Central Hospital from January 2017 to December 2023. Pediatric patients were classified into two groups based on postoperative echocardiography and cardiac catheterization results: a normal pulmonary artery pressure group (NG) and PAH after correcting VSD group (CD). We trained and validated a multimodal multipath AI system (MMAI) using echocardiography (Echo) and chest digital radiography (DR) dataset. Dice similarity coefficient (DSC) is used to measure the effectiveness of the model in automatic contour segmentation of images. We assessed the recognition performance of MMAI using the area under the receiver operating characteristic curve (AUC), accuracy, and F1-score through internal and external test sets. The ResNet-50 model demonstrates good performance in automatic cardiac contour segmentation, with DSC values of 0.950 ± 0.017 (Echo) and 0.946 ± 0.020 (DR). Compared to single image types, the multimodal model based on the ResNet-50 model performs better in the binary classification task in the training and validation sets, with both AUC and accuracy exceeding 90%. In multipath detection, the MMAI system performs well in the NG and CD by combining internal and external test set detection, with AUC, accuracy, and F1-score all exceeding 0.9. Our preliminary study developed an MMAI system using Echo and chest DR images, showing potential for assisting in the detection of CD in VSD patients under 1 year of age prior to surgery. Further validation is needed to confirm clinical applicability.

Keywords Pulmonary arterial hypertension, Ventricular septal defect, Echocardiography, Chest digital radiography, Artificial intelligence

Pulmonary arterial hypertension related to congenital heart disease (PAH-CHD), a crucial type of pulmonary hypertension (PH), is the leading cause of PH in children[1]. PAH progression involves pathological changes such as elevated pulmonary blood flow and pressure, damage to endothelial cells, intimal hyperplasia, and thickening of elastic fibers. Ventricular septal defect (VSD) is the most common form of CHD and is strongly associated with PAH[2].

Hemodynamically mediated PAH, characterized by increased pulmonary blood flow, may normalize after surgery. Following VSD repair, reduced pulmonary blood flow and vascular resistance often allow the recovery of the associated PAH. An abnormal increase in pulmonary flow in PAH after a surgically corrected VSD (CD) can cause endothelial cell damage under shear stress, leading to vascular reconstruction and transformation into obstructive PH[3,4]. Once PAH advances to Eisenmenger syndrome (a surgical contraindication), correcting the VSD can worsen the condition and cause right heart failure[5]. Nonetheless, patients with VSD undergoing defect correction face significant risks such as extended periods of mechanical ventilation, prolonged intensive care unit stays and hospitalizations, high costs, and substantial resource demands. From an etiological perspective, patients with CD may be more inclined to have primary PAH, and their prognosis may be similar to or even

[1]Qindao Medical College of Qingdao University, Qingdao 266071, China. [2]The Department of Pediatrics, Qingdao Central Hospital, University of Health and Rehabilitation Sciences (Qingdao Central Hospital), Qingdao 266042, China. [3]Heart Center, Women and Children's Hospital, Qingdao University, Qingdao 266034, China. ✉email: silinpan@126.com

worse than that of patients with primary PAH. Therefore, early identification and individualized PAH therapy are crucial for patients with CD.

Clinical decision-making is challenging and requires doctors to efficiently translate medical knowledge into swift, accurate, and personalized treatment plans. For the PAH-VSD subtype CD, preoperative diagnosis and prognosis assessment based on experience alone are challenging. Furthermore, precise drug regimen control remains elusive. Artificial intelligence (AI), specifically machine learning, has the potential to improve patient care, from detection and diagnosis to treatment planning, streamlining, and personalizing clinical decisions. Owing to significant progress in computer vision, AI has been at the forefront of the development of information extraction from medical images. Convolutional neural networks (CNNs) have made remarkable strides in the identification of medical images, including object detection, disease classification, and prognosis prediction[6].

Right heart catheterization remains the gold standard for the diagnosis and evaluation of PAH. However, it has some drawbacks: its invasiveness and cost limit its accessibility in resource-limited settings, and it cannot predict prognosis. Echocardiography (Echo) has emerged as the primary screening and evaluation tool for PAH, offering insights into cardiac anatomy and hemodynamics. Echo is noninvasive, cost-effective, operator-friendly, accurate, and has high specificity and sensitivity. Moreover, Echo is the preferred method globally for diagnosing PAH, with a mean pulmonary arterial pressure (mPAP) > 20 mmHg as an established criterion[3,7]. Echo can reveal right atrial and ventricular dilation, as well as the main pulmonary artery. Compared with traditional radiography, chest digital radiography (DR) offers superior image quality and resolution, providing more detailed information. Chest DR typically shows right heart enlargement, dilated pulmonary arteries, and prominent hilar shadows[8]. Our study developed a multimodal multipath artificial intelligence system (MMAI) by integrating Echo and DR, achieving noninvasive medical image recognition and segmentation through a novel multipath architecture. The system demonstrated breakthrough performance in CD diagnosis, providing a robust, multicenter-validated AI solution for early screening.

Methods

Study overview

This study was divided into the following key steps: First, based on clinical requirements, we established the inclusion and exclusion criteria for patient selection and divided the dataset into training and test groups. Subsequently, medical imaging data were systematically collected, and standardized preprocessing was performed. During the model development phase, the optimal base network architecture was selected by evaluating the image segmentation performance on the training set using cross-validation methods. Next, the model was further optimized by integrating the Transformer attention mechanism and the bidirectional short-term memory (Bi-LSTM) temporal modeling module. The best hyperparameter combinations were determined, and the appropriate experimental hardware environment was configured. Finally, after achieving satisfactory performance in the training set validation, the model was systematically evaluated on an independent test set to comprehensively assess its clinical applicability and generalization capability.

Study dataset

This retrospective study included 1,316 VSD patients who were under1 year of age who underwent surgery at Women and Children's Hospital, Qingdao University and Qingdao Central Hospital from January 2017 to December 2023. PAH was diagnosed per 2015 European Society of Cardiology (ESC) Guidelineswhen mean pulmonary artery pressure (mPAP) ≥ 20 mmHg[3], initially assessed intraoperatively or by. Echo. Patients with abnormal preoperative pulmonary pressures underwent confirmatory cardiac catheterization.

Patients with VSD who underwent surgical correction and achieved normal pulmonary artery pressure were included in the normal group (NG). The inclusion criteria for the CD group: Patients who have undergone surgical repair and no longer have shunt in the heart, but still have residual PAH for more than 1 year or pulmonary artery pressure that has increased again several months after surgery[1,3]. During clinical practice, this investigation monitored each patient's progress at least once every 3 months for a period of 1 year or longer. During the follow-up period, stable patients underwent Echo to assess the tricuspid regurgitation velocity (peak > 3.4 m/s) as a PAH indicator[2]. Cardiac catheterization was repeated if PAH symptoms worsened or treatment adjustments were required. Final grouping was based on ≥ 1-year postoperative data from Echo and catheterization. NG required two consecutive Echo-confirmed PASP ≤ 40 mmHg after discontinuing PAH therapy.

Exclusion criteria: Preoperative complications include primary PAH according to the guidelines of ESC; Other types of CHD and VSD who have not undergone surgery; Severe myocarditis, cardiomyopathy, and other organic heart diseases; Hypertension and hyperthyroidism.

Image acquisition

Echo examination: Ultrasound diagnostic instrument (Philips, iE33), probe frequency is 2.5–3.5 MHz. The patient is placed in a supine or left lateral position, and the maximum tricuspid regurgitation velocity (V) is measured three times in the apical four chamber section. The average value is taken, with pulmonary artery systolic pressure (PASP) = $4 \times V^2$ + right atrial pressure. Chest DR examination: Chest DR were obtained by using DRX-Evolution Plus (Reichert), with fully automatic exposure. The parameter is set to 110 kv, 8 mAs.

All Echo and Chest DR images were taken from the last examination before VSD surgery. The hospital medical imaging system exports all Echo and Chest DR images in JPEG format and they are then reviewed by two experienced ultrasound or imaging physicians with 5 years of expertise. This study excluded unqualified images based on the following criteria: ①Images with poor quality, such as blurring or noise, were not considered; ②Any images depicting unrelated lesions, such as pneumonia, were disregarded; ③Duplicate images were also excluded.

Ethical and information governance approvals

This retrospective study has received approval from the Hospital Ethics Committee (No. QFELL-KY-2016-22) and follows the ethical guidelines of medical research with human participants, as outlined in the Helsinki Declaration. The need for consent was waived owing to the retrospective nature of the study. Consent forms for the surgery in this study have been signed by the parents of all selected patients. Identify and censor any confidential patient information from image data, including name, date of birth, year of admission, admission number, and image number. All authors had access to the research data and have reviewed and approved the final manuscript.

Construction of AI model

We developed an MMAI to identify CD in patients with VSDs using Echo and/or chest DR images. This study employed Transformer's global learning to construct a semantic segmentation network that blended the concepts of Transformer and ResNet. The network is structurally similar to the U-Net model[9], comprising an encoder, decoder, and skip connections. The encoder is designed using ResNet and Vision Transformer. A multilayer ResNet feature extraction structure is used to segment the input medical images into image blocks. The segmented image blocks are merged to form a series of image blocks, subsequently fed into a Transformer-based encoder to extract global features. It then enters a CNN-based decoding layer to upsample the extracted multilevel features and fuses them with multiscale features during the downsampling process through skip connections to restore the feature map to the approximate resolution size of the input image. Finally, pixel-level segmentation predictions are performed.

The optimization training difficulties faced by deep neural networks can be effectively addressed by the concept of residual blocks in a residual network. We designed and employed five single-modal residual convolutional networks (ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152) to evaluate the effectiveness of the multimodal model. To effectively classify CD using Echo or chest DR images, we utilized the DCNN-LSTM model to incorporate the interpretation of multimodal images. The interaction between the Echo or chest DR images was modeled using an attention-based Bi-LSTM network, as shown in Fig. 1.

The proposed MMAI system comprises three components: region of interest (RoI) extraction, image segmentation, and automatic diagnostic assessment. Four resident physicians with expertise in medical imaging were recruited and instructed to delineate the cardiac contours of the Echo and chest DR images. In the final quality-control phase, all segmentation labels were comprehensively evaluated by the same senior radiologist. To ensure diagnostic validity, this verification process incorporated correlations with clinical imaging interpretations and relevant clinical outcomes when available. These delineations were used as RoI labels for training the model and as criteria for evaluating the effectiveness of the automated delineation. The self-attention mechanism of the Transformer enables global context modeling across image regions, allowing it to capture long-range spatial dependencies that improve contour precision in segmentation tasks. A training set is necessary for transformer-based models because they must be equipped with appropriate positional embeddings. However, acquiring precise annotations of medical images is not only expensive but also arduous.

Experimental environment

Patients were randomly divided into training and validation sets (80%) and test set (20%) using a fixed random seed (1, 2, 3). The training data (DATA I; n = 1,052) comprised patients with VSD from the Women and Children's Hospital, Qingdao University. The test set (DATA Ⅱ) included an internal cohort (n = 130, same hospital) and an external cohort (n = 134, Qingdao Central Hospital) in a 1:1 ratio.

Our study utilized a 12-core Intel (R) Xeon (R) Platinum 8255C CPU and a GeForce RTX 3090 GPU environment with PyTorch (Python-based) as the framework (random seed = 1024). Image upscaling was performed using the PyTorch transformation module. During the ResNet model training process, a weighted random sampling approach and a weighted cross-entropy loss function were used to address the effect of data imbalance on performance degradation. Through systematic ablation experiments, we identified the optimal weight constraint range as [0.5, 5.0] from seven candidate intervals ([0.1, 2.0], [0.1, 5.0], [0.1, 10.0], [0.5, 5.0], [0.5, 10.0], [1.0, 5.0], and [1.0, 10.0]). This range was validated over 100 training epochs on the internal dataset, demonstrating the best balance between model stability and segmentation performance. Its overall performance significantly surpassed that of the focal loss approach. We set the training step to 400 epochs, batch size to 4, and initial learning rate to 0.001. The MMAI was trained using the training and validation sets (DATA Ⅰ), covering RoI extraction, image segmentation, and rule-based diagnostic evaluation (Fig. 2). Ten-fold cross-validation and the Adam optimizer with a grid search were employed. The final model was selected based on the lowest validation loss over 100 epochs. The cross-validated average loss decreased from an initial 1.25 to 0.38, with the optimal model achieving a validation loss of 0.32 ± 0.05.

Clinical labeling

The diagnosis and evaluation criteria for PAH were established considering Echo reports, chest DR reports, surgical records, follow-up data, and expert reviews. Four resident physicians collected Echo reports, chest DR reports, surgical records, and medical follow-up data. To avoid adverse effects, each resident physician was only responsible for gathering and extracting one type of information.

The Echo and chest DR images were resized and cropped to 256×256 pixels to standardize the input image size. Because 256 (2^8) is a power-of-two dimension, it perfectly aligns with the pooling and downsampling operations in CNN architectures, ensuring precise feature map size matching across all network layers. The 256×256 pixels represent a well-validated technical balance in medical image analysis, allowing the simultaneous optimization of the computational efficiency, feature preservation[10], and domain adaptability[11,12]. Moreover, we augmented the images using random horizontal flipping, center cropping, and random vertical flipping

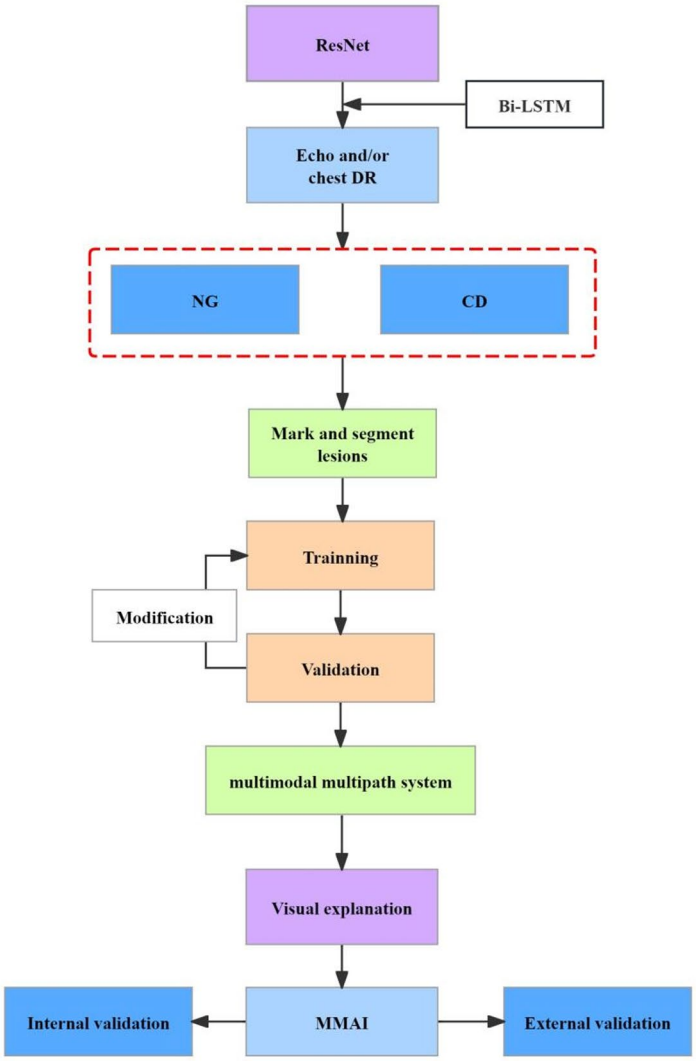

Fig. 1. Flowchart of the study design. The construction of multimodal, multipath artificial intelligence (MMAI) procedure included images for training, validation and testing. CD: correcting the defect; Echo, echocardiography; DR, digital radiography.

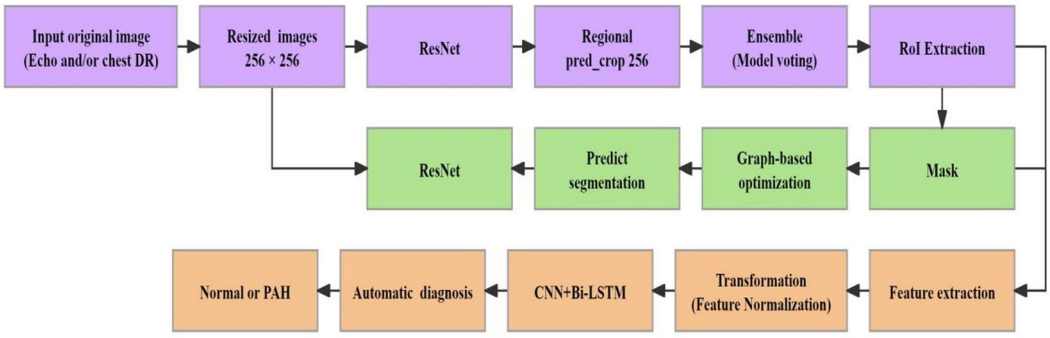

Fig. 2. The MMAI system incorporates region of interest (RoI) extraction, advanced image segmentation, and rule-based diagnostic evaluation capabilities.

	Male (n)	Age (months)	Weight (kg)	VSD size (mm)	Preoperative PAPS (mmHg)
NG (n = 652)	324	8.48 ± 1.46	8.44 ± 0.69	10.60 ± 1.37	60.95 ± 8.54
CD (n = 400)	184	8.51 ± 1.56	8.69 ± 0.79	10.78 ± 1.46	63.53 ± 7.44
t/χ^2	1.499	−0.332	3.268	−2.059	−4.995
P	0.221	0.740	0.001	0.040	0.000

Table 1. Baseline characteristics of NG and CD groups in Dataset I.

	Age (months)	Weight (kg)	VSD size (mm)	PAPS (mmHg)	NG (n,%)	CD (n,%)
Internal data (n = 130)	7.97 ± 2.53	9.33 ± 0.42	11.57 ± 4.01	62.21 ± 10.38	80 (61.5%)	50 (38.50%)
External data (n = 134)	8.15 ± 2.08	8.98 ± 0.52	11.66 ± 3.69	62.23 ± 9.21	96 (71.6%)	38 (28.4%)
t/χ^2	−1.003	−1.105	−0.297	−0.023	4.000	
P	0.302	0.271	0.767	0.982	0.261	

Table 2. Comparison of internal and external basic information in the test set (DATA Ⅱ).

operations to mitigate overfitting. Compared to the 512 × 512 pixels, the 256 × 256 pixels achieves a 3.2 × training acceleration on the RTX 3090 GPU, with a 73% reduction in GPU memory usage ($P < 0.01$). In sensitivity analysis, the Kullback–Leibler divergence of the grayscale value distribution in the RoI region between the 256 × 256 pixels and the original image shows no significant difference compared to the 512 × 512 pixels (0.037 vs. 0.029, $P > 0.05$).

General information

Doctors were considered the gold standard for manual contour segmentation, and the effectiveness of automatic network contour segmentation was measured using the Dice similarity coefficient (DSC). With a DSC of 0–1, the closer the value is to 1, the closer the automated segmentation aligns with the manual segmentation, suggesting a greater level of predictive precision[13]. The formula for calculating DSC is as follows:

$$DSC(A,B) = 2 \times |A \cap B| \div (|A| + |B|), 0 \le DSC \le 1$$

where A is the area manually segmented by the doctor, and B is the automatically segmented area. The Wilcoxon signed-rank test was used to identify significant differences in the data using paired samples and a nonparametric approach. Statistical significance was set at $P < 0.05$.

We integrated fusion gradient-weighted class activation mapping (GradCAM) into the multimodal model to analyze attention[14]. This improved the interpretability of the MMAI and bolstered the trust of the medical experts in their prognostications. Global average pooling was performed on the last convolutional layer of the trained AI model using classification activation maps. The weights allotted to each output in the global average pooling layer during training represent the relevance of the feature maps derived from the preceding convolutional layers. These weights are subsequently utilized to modulate the corresponding feature maps, thereby producing saliency maps. The test set (DATA Ⅱ) was used to evaluate the performance of the MMAI system.

Results

Dataset consistency verification results

The study enrolled 1,316 patients with VSD (47.94% male; mean age 7.90 ± 2.42 months, weight 9.13 ± 0.55 kg). Baseline measurements included VSD size (11.52 ± 3.89 mm) and PASP (62.56 ± 10.42 mmHg), with a follow-up of 40.19 ± 17.56 months. The clinical classifications were as follows: 828 cases (62.9%) in the NG group and 488 cases (37.1%) in the CD group. The training set (DATA I) included 1,052 patients (NG: n = 652, CD: n = 400) from the Women and Children's Hospital, Qingdao University. No age/gender differences existed between the NG and CD groups in DATA I ($P > 0.05$); however, weight, VSD size, and preoperative PASP differed significantly ($P < 0.05$; Table 1).

In the test dataset (DATA Ⅱ), 264 patients from two medical centers were included, with 176 in the NG group and 88 in the CD group. There were no significant differences in the overall clinical characteristics between patients in the external and internal test sets, indicating good consistency (Table 2).

Differences in automatic image segmentation between different ResNet network models

In DATA Ⅰ, the DSC for cardiac segmentation accuracy of Echo and chest DR images in ResNet-50 were 0.950 ± 0.017 and 0.946 ± 0.020, respectively, significantly higher than those of ResNet-18, ResNet-152, and ResNet-101 (Table 3). The ResNet-50 segmentation contours exhibited high agreement with the manual annotations (Fig. 3), prompting its selection for MMAI development.

Performance of MMAI in distinguishing CD

Bi-LSTM was utilized to integrate a hybrid model, which could predict CD in patients with VSD using either Echo or chest DR images in our study. In DATA I, the hybrid model integrating Echo and chest DR data predicted

	ResNet18	ResNet152	ResNet101	ResNet50
Echo	0.926 ± 0.030*	0.939 ± 0.023*	0.941 ± 0.026*	0.950 ± 0.017
Chest DR	0.912 ± 0.030*	0.918 ± 0.025*	0.940 ± 0.019*	0.946 ± 0.020

Table 3. DSC of multimodal heart segmentation accuracy in the DATA Ⅰ. DSC, dice similarity coefficient; * represents comparison with ResNet50, $P < 0.05$.

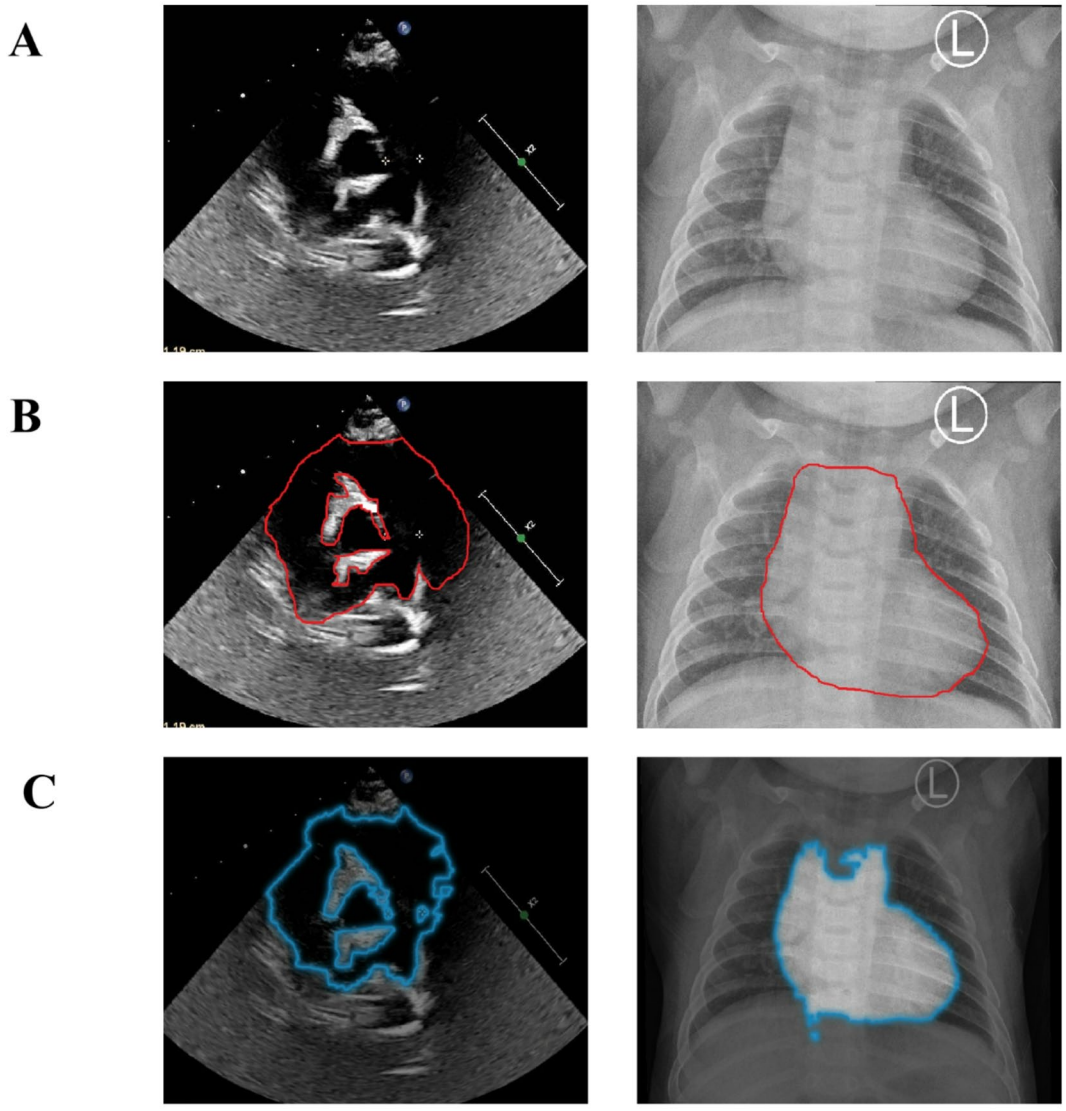

Fig. 3. Automatic and manual delineation of cardiovascular regions in Echo and Chest DR images. The segmentation contours generated by ResNet50 showed high morphological congruence with manual segmentations. (**A**) the original image, (**B**) the manual sketching effect, (**C**) the ResNet50 automatic sketching effect.

CD with area under the curve (AUC) = 0.94 and accuracy = 0.94, outperforming unimodal approaches (Table 4, Fig. 4). In the analysis of the Echo and chest DR images, the hybrid model prioritized the pulmonary artery region using heatmap visualization (Fig. 5).

MMAI performance in test sets

MMAI maintained robust performance across internal (AUC = 0.916, accuracy = 0.91) and external (AUC = 0.905, accuracy = 0.90) test sets, with balanced sensitivity/specificity (0.90–0.91) and F1-score (0.90; Table 5, Fig. 6).

	AUC (%)	95% CI	Accuracy (%)
The Echo-path model	0.86	0.83–0.88	0.85
The chest DR-path model	0.84	0.81–0.86	0.84
The hybrid model	0.94	0.90–0.96	0.94

Table 4. Comparison of CD recognition performance using multipath system in DATA Ⅰ.

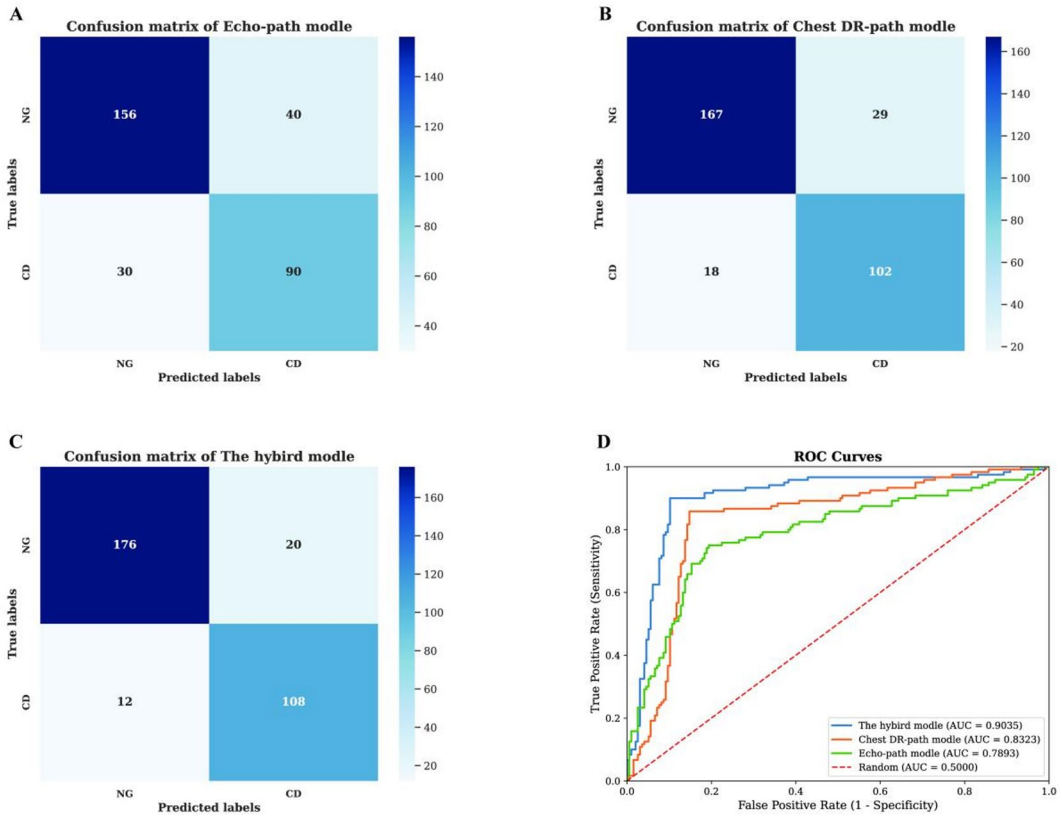

Fig. 4. Performance of multipath system to distinguish CD in the DATA Ⅰ. (**A**) Confusion matrix of Echo-path model. (**B**) Confusion matrix of Chest DR-path model. (**C**) Confusion matrix of the hybrid model. (**D**) Receiver Operating Characteristic of the hybrid model.

Discussion

In this paper, AI technology was applied to identify CD from Echo and/or chest DR images. The proposed MMAI model demonstrated effective performance in identifying CD in patients with VSD under one year of age, though further validation is warranted.

PAH is a frequently encountered disease that ultimately leads to chronic heart failure and significantly reduces life expectancy. Often asymptomatic in early stages, PAH is frequently misdiagnosed, making early detection crucial for timely intervention, disease progression delay, and mortality reduction. Children with VSD typically exhibit significant left-to-right cardiac shunts with systemic pulmonary circulation imbalance, often accompanied by PAH. Infants are at a disadvantage because their organs are not yet fully developed, making them more susceptible to congestive heart failure and frequent respiratory infections, which can lead to a critical state[7,15]. PAH-VSD includes various clinical subtypes with distinct symptoms, cardiac functions, hemodynamics, and medium- to long-term survival rates. Patients with CD still require anti-PAH medication after VSD surgery, and early intervention may explain their better prognosis.

PAH diagnosis currently involves both invasive and noninvasive methods. Although cardiac catheterization remains the gold standard, its prognostic limitations, high cost, complexity, and the risk of complications restrict its widespread use. Noninvasive techniques include Doppler Echo, computed tomography, and chest radiography. Echo measures the maximum tricuspid regurgitation velocity for PAH assessment[16]. Chest DR reveals increased pulmonary blood flow, pulmonary artery elevation, right lower pulmonary artery thickening, and right ventricular hypertrophy[7,8].

AI applications in Echo and DR image analyses enable automatic PAH-VSD detection, offering improved sensitivity (reducing missed diagnoses) and specificity (avoiding unnecessary referrals), thereby optimizing

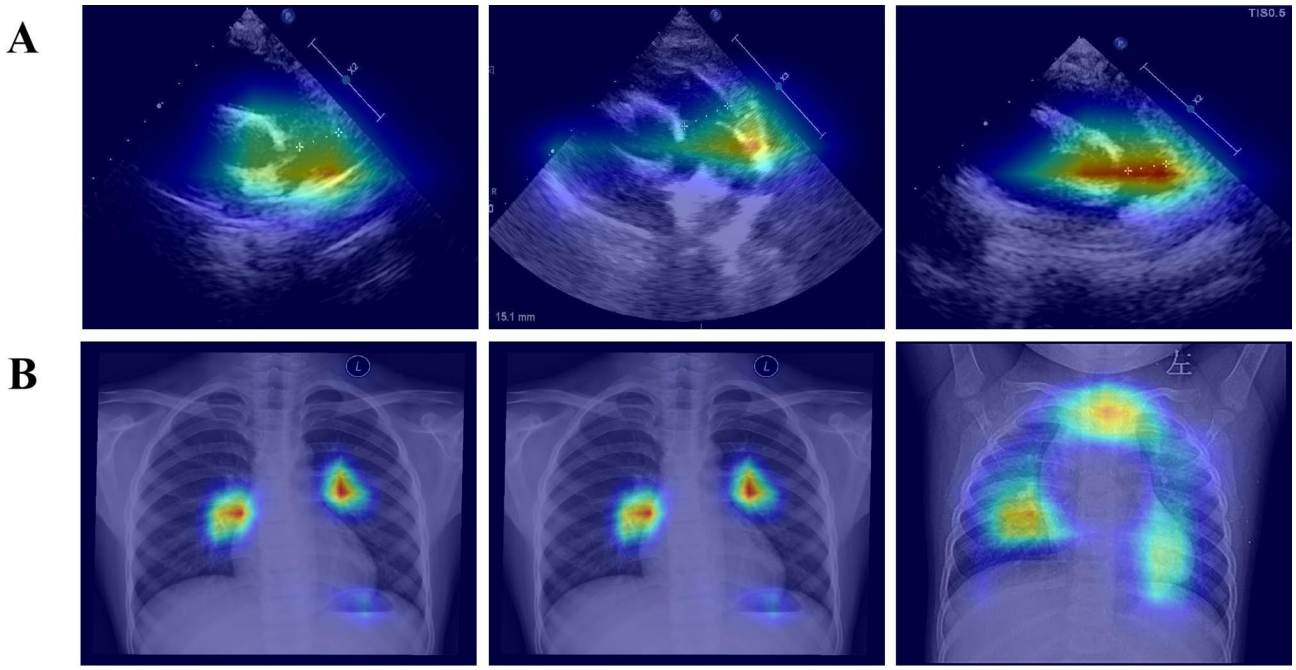

Fig. 5. The hybrid path model prioritizes the pulmonary artery region by utilizing a heatmap visualization in Echo (**A**) and Chest DR (**B**) images.

	AUC (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score
The internal test set(n = 130)	0.916	0.91	0.90	0.90	0.90
The external test set(n = 134)	0.905	0.90	0.90	0.91	0.90

Table 5. Performance of MMAI in different test sets.

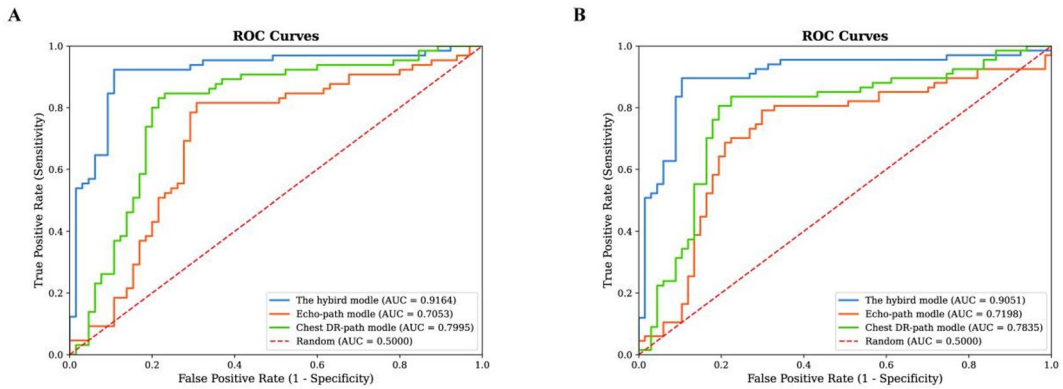

Fig. 6. Binary classification performance using the proposed Multimodal Multipath Artificial Intelligence (MMAI) system on Internal test set (**A**) and External test set (**B**).

resource allocation and minimizing errors. Previous study utilized AI models pretrained on ImageNet databases for congenital heart disease diagnosis[17]. ResNet-18 exhibited the highest sensitivity (0.970), specificity (0.982), and AUC (0.948) in this study, significantly improving the diagnostic accuracy of radiologists. In our previous study, the ResNet-50 model was trained on chest DR images from 831 patients with VSD, which allowed us to retrospectively predict the development of postoperative PAH. In the independent test set, ResNet-50 significantly outperformed the physicians (AUC = 0.81 vs. 0.65)[15]. The performance of the MMAI system developed by our institute has been further improved, with an AUC > 0.9 in the validation and test sets. GradCAM enhances model interpretability by generating heatmaps (e.g., highlighting pulmonary artery segments in ResNet-50). However, these heatmaps only indicate correlation, not causality, and may deviate

from clinicians' diagnostic criteria, necessitating physician validation. GradCAM has become a standard tool in visual interpretability research because of its lightweight design and compatibility with existing models. Diller et al. analyzed 405 individuals diagnosed with PAH, 308 patients with right ventricular enlargement but without PAH, and 67 healthy subjects. They trained a deep learning network to segment the heart cavity in Echo images and extracted geometric information from the cardiac cycle with a sensitivity of 100% and specificity of 95% for detecting PAH[18]. GradCAM demonstrated that the ResNet-50 model focused on the pulmonary artery segment, demonstrating good interpretability[15]. The heatmaps generated using GradCAM only indicate regions of correlation that the model focuses on, but cannot establish a causal relationship between these regions and pathological outcomes. There may be discrepancies between the saliency in heatmaps and clinicians' visual diagnostic criteria. Therefore, the heatmap results require cross-validation by physicians, in conjunction with the medical imaging characteristics.

The evaluation using Echo has limitations, and the PAH evaluation is based on the degree of tricuspid regurgitation. In the absence of significant or severe tricuspid regurgitation, an accurate PASP cannot be obtained by Echo[16]. Multiple data modalities often exhibit a certain degree of complementarity, which can provide more comprehensive information and help solve complex problems. This study utilized MMAI technology to efficiently identify and evaluate the PAH-VSD subtypes by drawing on multimodal image data from Echo and chest DR.

Limitations: While demonstrating promising results, our study has several important limitations that warrant discussion. ①Computational and Clinical Constraints: The deep learning framework employed carries inherent limitations in generalizability. Although sensitivity analysis is essential for clinical applicability evaluation, comprehensive assessments were precluded by: (a) GPU memory constraints impacting computational capacity, and (b) logistical challenges in multicenter clinical data acquisition. ②Although our study determined the hyperparameter combination through cross-validation and grid search, the hyperparameter optimization was solely based on conventional methods (e.g., validation loss minimization) without further robustness verification via sensitivity analysis. The grid search range might not fully cover potentially superior parameter spaces. Future work could enhance model reliability through systematic hyperparameter validation. ③Several critical validation gaps remain in the clinical verification requirements: Given the heterogeneity of CHD types associated with CHD-PAH and the limited data size of this study, external validation with larger cohorts is necessary. Although the proposed MMAI system showed a promising classification performance, rigorous clinical validation is required before its deployment to ensure safety and efficacy. While the annotations by senior radiologists help ensure diagnostic quality, there is a lack of quantitative inter-rater reliability assessments. We emphasize that all AI-based diagnostics must be integrated with the following: comprehensive patient symptoms, physical examination findings, and relevant auxiliary test results. The current MMAI model is strictly an auxiliary decision-support tool that requires clinician supervision.

Conclusions

Our preliminary study demonstrates that the MMAI system, developed based on preoperative Echo and chest DR images from VSD patients under 1 year of age, shows potential for assisting in preoperative classification of CD prior to surgery. These findings warrant further validation through larger-scale clinical studies.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper. Should any raw data files be needed in another format they are available from the corresponding author upon reasonable request.

Received: 23 January 2025; Accepted: 9 September 2025
Published online: 13 October 2025

References

1. Ruopp, N. F. & Cockrill, B. A. Diagnosis and treatment of pulmonary arterial hypertension: a review. *JAMA* **14**, 1379–1391 (2022).
2. Rosenzweig, E. B. et al. Paediatric pulmonary arterial hypertension: updates on definition, classification, diagnostics and management. *Eur. Respir. J.* **1**, 1801916 (2019).
3. Galiè, N. et al. ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The joint task force for the diagnosis and treatment of pulmonary hypertension of the European society of cardiology (ESC) and the European respiratory society (ERS) endorsed by: Association for European paediatric and congenital cardiology (AEPC), international society for heart and lung transplantation (ISHLT). *Eur. Respir. J.* **6**(2015), 903–975 (2015).
4. Vijarnsorn, C. et al. Contemporary survival of patients with pulmonary arterial hypertension and congenital systemic to pulmonary shunts. *PLoS ONE* **13**, e0195092 (2018).
5. Kempny, A. et al. Predictors of death in contemporary adult patients with eisenmenger syndrome: A multicenter study. *Circulation* **15**, 1432–1440 (2017).
6. Winter, P. & Carusi, A. Professional expectations and patient expectations concerning the development of artificial intelligence (AI) for the early diagnosis of Pulmonary hypertension (PH). *J. Responsib. Technol.* **12**, 100052 (2022).
7. Duffels, M. G. et al. Pulmonary arterial hypertension in congenital heart disease: An epidemiologic perspective from a Dutch registry. *Int. J. Cardiol.* **2**, 198–204 (2007).
8. Luo, G. et al. Residual networks models detection of atrial septal defect from chest radiographs. *Radiol. med.* **1**, 48–55 (2023).
9. Tran, D., Bourdev, L. D., & Fergus, R., et al. Deep end2end voxel2voxel prediction. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015.
10. He, D. et al. De-Noising of photoacoustic microscopy images by attentive generative adversarial network. *IEEE Trans. Med. Imaging.* **5**, 1349–1362 (2023).
11. Kott, O. et al. Development of a deep learning algorithm for the histopathologic diagnosis and Gleason grading of prostate cancer biopsies: A pilot study. *Eur. Urol. Focus.* **2**, 347–351 (2021).

12. Rasmussen, S. A., Taylor, V. J., Surette, A. P., Barnes, P. J. & Bethune, G. C. Using deep learning to predict final HER2 status in invasive breast cancers that are equivocal (2+) by immunohistochemistry. *Appl. Immunohistochem. Mol. Morphol.* **10**, 668–673 (2022).
13. Crum, W. R., Camara, O. & Hill, D. L. G. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Trans. Med. Imaging.* **11**, 1451–1461 (2006).
14. Mamalakis, M. et al. A transparent artificial intelligence framework to assess lung disease in pulmonary hypertension. *Sci. Rep.* **1**, 3812 (2023).
15. Li, Z., Luo, G., Ji, Z., Wang, S. & Pan, S. Explanatory deep learning to predict elevated pulmonary artery pressure in children with ventricular septal defects using standard chest x-rays: a novel approach. *Front. Cardiovasc. Med.* **11**, 1330685 (2024).
16. Janda, S., Shahidi, N., Gin, K. & Swiston, J. Diagnostic accuracy of echocardiography for pulmonary hypertension: A systematic review and meta-analysis. *Heart* **8**, 612–622 (2011).
17. Han, P. L. et al. Artificial intelligence-assisted diagnosis of congenital heart disease and associated pulmonary arterial hypertension from chest radiographs: A multi-reader multi-case study. *Eur. J. Radiol.* **171**, 111277 (2024).
18. Diller, G. P. et al. A framework of deep learning networks provides expert-level accuracy for the detection and prognostication of pulmonary arterial hypertension. *Eur. Heart J. Cardiovasc. Imaging.* **11**, 1447–1456 (2022).

Author contributions

Gang Luo: Data curation, Writing- Original draft preparation. Huashu Liu: Conceptualization, Methodology, Software. Zhixin Li: Visualization, Investigation. Zhixian Ji: Supervision, Software, Validation. Silin Pan: Writing- Reviewing and Editing.

Funding

This study was supported by the National Key Research and Development Program of China (2023YFC2705700) and the National Natural Science Foundation of China (82271725).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.