

RESEARCH

Open Access



# Differential diagnosis of congenital ventricular septal defect and atrial septal defect in children using deep learning–based analysis of chest radiographs

Huihui Jia<sup>1</sup>, Songqiao Tang<sup>2</sup>, Wanliang Guo<sup>1</sup>, Peng Pan<sup>1</sup>, Yufeng Qian<sup>1</sup>, Dongliang Hu<sup>1</sup>, Yakang Dai<sup>3</sup>, Yang Yang<sup>1</sup>, Chen Geng<sup>3,4\*†</sup> and Haitao Lv<sup>5\*†</sup>

## Abstract

**Background** Children with atrial septal defect (ASD) and ventricular septal defect (VSD) are frequently examined for respiratory symptoms, even when the underlying disease is not found. Chest radiographs often serve as the primary imaging modality. It is crucial to differentiate between ASD and VSD due to their distinct treatment.

**Purpose** To assess whether deep learning analysis of chest radiographs can more effectively differentiate between ASD and VSD in children.

**Methods** In this retrospective study, chest radiographs and corresponding radiology reports from 1,194 patients were analyzed. The cases were categorized into a training set and a validation set, comprising 480 cases of ASD and 480 cases of VSD, and a test set with 115 cases of ASD and 119 cases of VSD. Four deep learning network models—ResNet-CBAM, InceptionV3, EfficientNet, and ViT—were developed for training, and a fivefold cross-validation method was employed to optimize the models. Receiver operating characteristic (ROC) curve analyses were conducted to assess the performance of each model. The most effective algorithm was compared with the interpretations provided by two radiologists on 234 images from the test group.

**Results** The average accuracy, sensitivity, and specificity of the four deep learning models in the differential diagnosis of VSD and ASD were higher than 70%. The AUC values of ResNet-CBAM, InceptionV3, EfficientNet, and ViT were 0.87, 0.91, 0.90, and 0.66, respectively. Statistical analysis showed that the differential diagnosis efficiency of InceptionV3 was the highest, reaching 87% classification accuracy. The accuracy of InceptionV3 in the differential diagnosis of VSD and ASD was higher than that of the radiologists.

<sup>†</sup>Chen Geng and Haitao Lv contributed equally to this work.

\*Correspondence:

Chen Geng  
gengc@sibet.ac.cn  
Haitao Lv  
haitaosz@163.com

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Conclusions** Deep learning methods such as InceptionV3 based on chest radiographs in the study showed good performance for differential diagnosis of congenital VSD and ASD, which may be able to assist radiologists in diagnosis, education, and training, and reduce missed diagnosis and misdiagnosis.

**Keywords** Congenital heart disease, Chest radiography, Deep learning, Differential diagnosis

## Introduction

Congenital heart disease (CHD) is the most common congenital birth defect. Ventricular septal defect (VSD) and atrial septal defect (ASD) are the most common non-cyanotic CHDs, and they account for about 25–35% of all CHDs [1]. Currently, echocardiography is the predominant screening method employed for CHD. However, the necessity for sedation or even anesthesia in some children, due to a lack of cooperation, poses significant challenges. Additionally, the requirement for professional echocardiographers means that many primary hospitals may not have the necessary expertise, thereby increasing the risk of missed diagnoses.

Children with ASD and VSD are often medically examined for respiratory symptoms. Notably, chest radiography has the advantage of X-ray photography—which displays both the heart and the lungs, i.e., the shape of the heart is shown, in addition to a reflection of the state of pulmonary circulation [2, 3]—over ultrasound. Therefore, the potential ability to use chest radiography for differential diagnosis between VSD and ASD would be very important. However, considering that making a diagnosis of CHD requires highly specialized knowledge, imaging physicians in nonspecialist hospitals have a high rate of misdiagnosis and missed diagnosis. Currently, the combination of artificial intelligence and imaging data is increasingly being applied to various medical image analysis tasks such as lesion segmentation, disease detection, and assisted diagnosis [3–7]. This approach has demonstrated considerable application value in the detection and diagnosis of systemic diseases across imaging modalities, including the brain [8], heart and chest [9], and abdomen [10]. Deep learning based on chest radiographs has been used to diagnose diseases such as pneumonia and pneumothorax [11–14], as well as predict the long-term prognosis of asymptomatic individuals.

The purpose of this study was to investigate the application of artificial intelligence methods in the automatic differentiation of ASD and VSD using chest radiographs. We proposed a deep learning approach that utilizes chest radiographs to distinguish between VSD and ASD, with the aim of enabling imaging physicians in nonspecialized hospitals to achieve rapid and accurate differential diagnoses. This research seeks to advance the goal of artificial intelligence-assisted diagnosis of simple CHD.

## Methods

This retrospective study was approved by the institutional review board (NO.2022-02-006-H01), and the need for patients' informed consent was waived.

## Subjects

We conducted a retrospective analysis of the chest radiographs data of 1194 patients who had undergone digital radiographic examination (GE Healthcare, Discovery, USA) between June 2017 and May 2023, including 489 males and 705 females (with an average age of  $5.56 \pm 2.67$  years). The patients were divided into the ASD group and the VSD group.

The inclusion criteria were as follows: (1) no treatment for CHD received before chest radiography; and (2) gold standard angiocardiography and/or surgical results available for comparison. The exclusion criteria were as follows: (1) history of CHD surgery or other related treatments; (2) history of other heart diseases; (3) incomplete patient information; and (4) image quality of chest plain film insufficient to meet the diagnostic requirements.

## Image collection

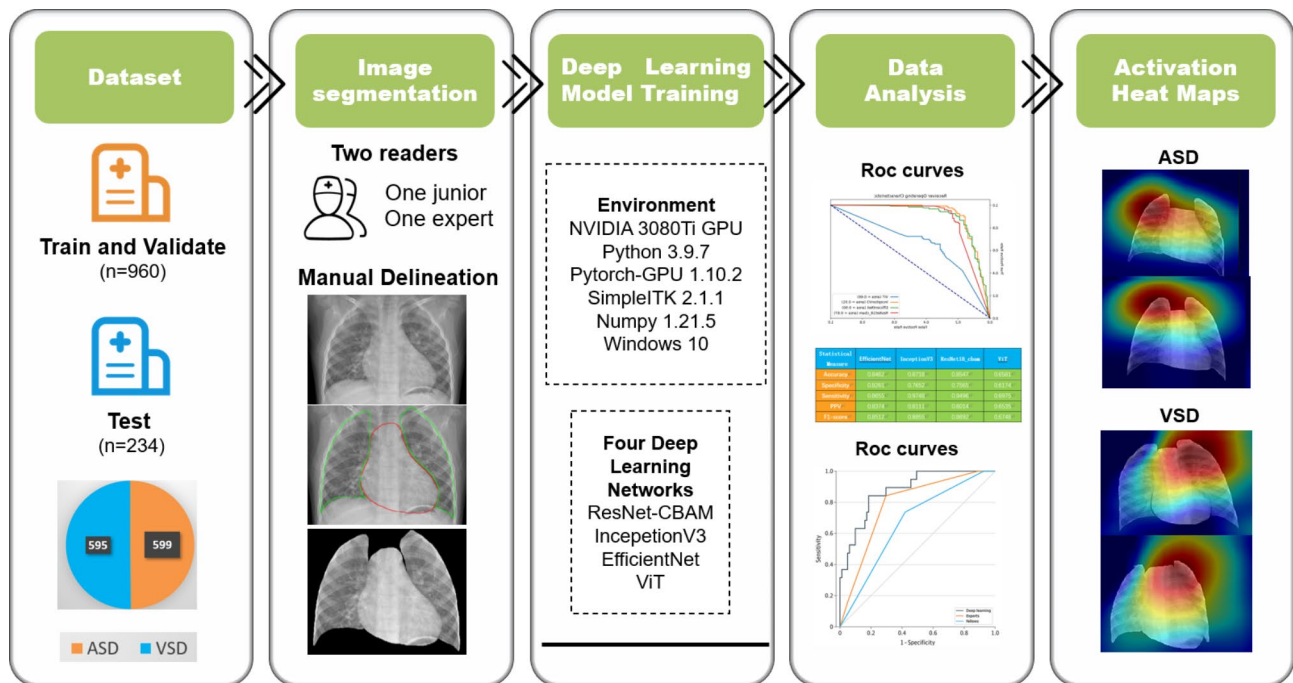
The children were placed in the standing position, instructed to inhale calmly, and the digital X-ray images were taken with the X-ray machines (GE Healthcare, Discovery, USA). The posterior and anterior chest films were collected, with tube current and tube voltage of 5–8 mA and 75–85 kV, respectively, and the filming distance of 180 cm.

## Training, validation, and test sets

The entire dataset contained 1194 images, including 595 of ASD and 599 of VSD. The dataset was divided into the training and validation set and the test set. A total of 960 samples, including 480 cases of ASD and 480 cases of VSD, were allocated to the training and validation set, while the remaining 115 cases of ASD and 119 cases of VSD were reserved for testing.

## Image pretreatment and deep learning model training

As shown in the flowchart (Fig. 1), our environment was NVIDIA 3080Ti GPU with 12 GB memory, Python 3.9.7, Pytorch-GPU 1.10.2, SimpleITK 2.1.1, Numpy 1.21.5, Windows 10. Due to variations in image sizes, all images were resampled to a size of  $1024 \times 1024$  pixels. The batch size was set at 4; the optimization function was Adam; the weights were initialized using the default initializer



**Fig. 1** Flowchart of the study. ASD, atrial septal defect. VSD, ventricular septal defect

(Standard normal distribution) of Pytorch; and the initial learning rate was set at 0.0001. We used a dynamic learning rate adjustment strategy, and the learning rate was adjusted to one-tenth of the original every 50 epochs. After 200 epochs, the training procedure stopped.

With the same training environment, four deep learning networks were trained and validated in this study, they were ResNet-CBAM, InceptionV3, EfficientNet, and ViT. Among them, ResNet-CBAM was the integration of the ResNet network, which was the most commonly used network for medical image classification, with the CBAM attention module. InceptionV3 was one of the most commonly used classification networks and achieved good performance on a variety of image classification tasks, EfficientNet was a network that achieved excellent classification performance with a small number of parameters, and ViT was a typical application of transformer in the image field. After performance comparison, InceptionV3 [15] was finally selected as the differential diagnostic network. This network is characterized by its unique architectural feature of using inception modules, which are composed of multiple parallel convolutional and pooling layers at different kernel sizes. These modules enable the network to capture a wide range of features, from fine details to larger-scale patterns, thereby effectively improving its ability to represent complex visual information. Additionally, InceptionV3 incorporates factorization of convolutions, and this architecture balances model depth and computational efficiency.

### Comparison between the models and diagnosis of radiologists

In order to assess the diagnostic performance of the deep learning models with that of the radiologists, 234 images in the validation group were visually diagnosed by the two radiologists mentioned above who independently categorized the cases as normal adenoids or adenoid hypertrophy. The differences in diagnostic ability between deep learning models and radiologists were compared using the Delong test. The Cohen Kappa test was used to assess the agreement between the deep learning models and radiologists, and between different radiologists.

### Evaluation and data analysis

Statistical analyses were performed using SPSS Statistics for Windows, version 25.0 (IBM Corp.) We performed a normality test on the obtained data, and normally distributed data were presented by mean (standard deviation) and skewed data were presented by median (upper and lower quartiles). We calculated the accuracy, sensitivity, specificity, and other indicators of defect detection; obtained the ROC curve of the model based on sensitivity and specificity; and calculated the area under the ROC curve (AUC). Kappa values were used to determine inter-observer agreement. Kappa values of 0.81–1.00 indicated perfect consistency, 0.61–0.80 indicated high consistency, 0.41–0.60 indicated moderate consistency, 0.21–0.40 indicated fair consistency, and 0–0.20 indicated

**Table 1** Clinical and demographic characteristics of the study

Characteristic	Training group (n = 762)	validation group (n = 198)	Test group (n = 234)
Age (year)	5.50 ± 2.68	5.56 ± 2.64	5.57 ± 2.67
<b>Sex, n (%)</b>			
Male	319(41.9)	81(40.9)	89(36.3)
Female	443(58.1)	117(59.1)	145(63.7)
<b>Clinical evaluation, n (%)</b>			
ASD	443 (58.2)	69(34.8)	83 (35.5)
VSD	319 (41.8)	129 (65.2)	151 (64.5)

very low consistency.  $P < 0.05$  was considered as a level with significant difference.

## Results

### General data

The present study collected posterior and anterior chest films images from 1194 children, including 489 males (41.0%) and 705 females (59.0%), with a mean age of  $5.53 \pm 2.66$  years (from 6 months to 12 years; Table 1). There was no statistically significant difference in the gender and age of the children in the three groups ( $P > 0.05$ ).

### Image segmentation results

Figure 2 presents representative X-ray images of ASD and VSD in children. Figure 3 illustrates the comparison between the artificial segmentation and the results of

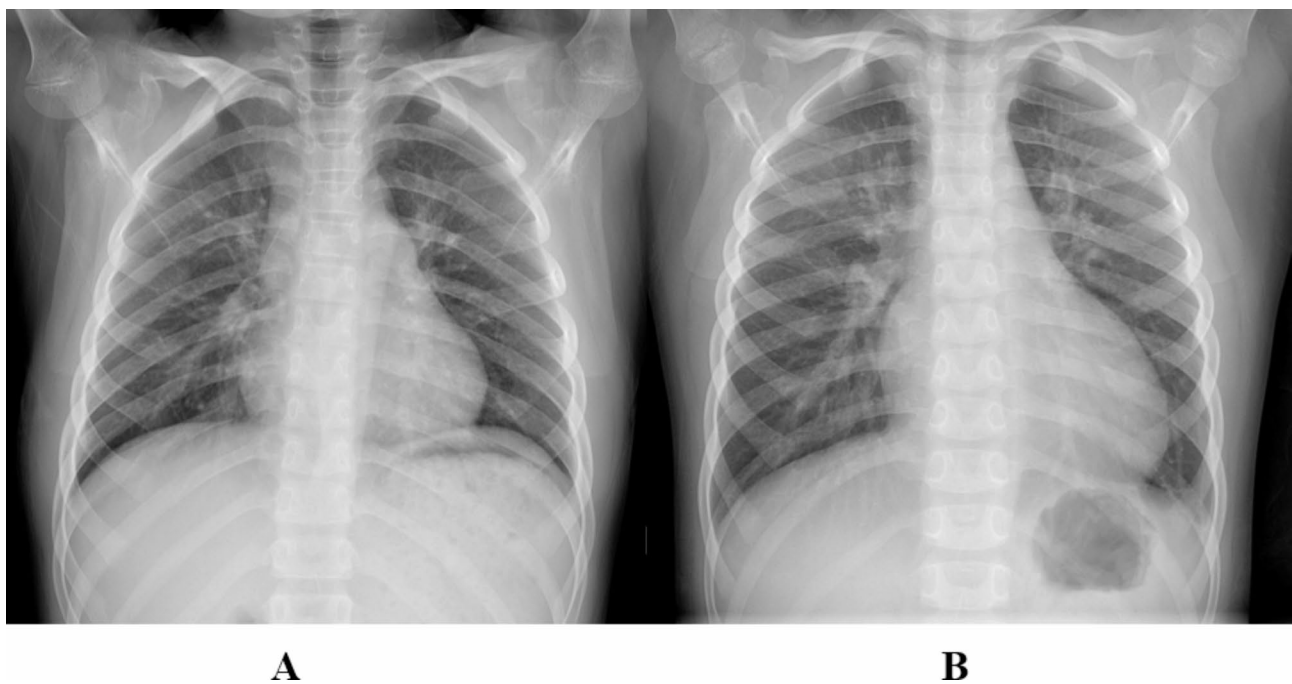
model predictions. The results demonstrate near-human performance in segmentation, with the predicted lung domain exhibiting a high degree of accuracy, thereby effectively fulfilling the requirements for image classification tasks.

### Detection efficiency of the Deep Learning Model

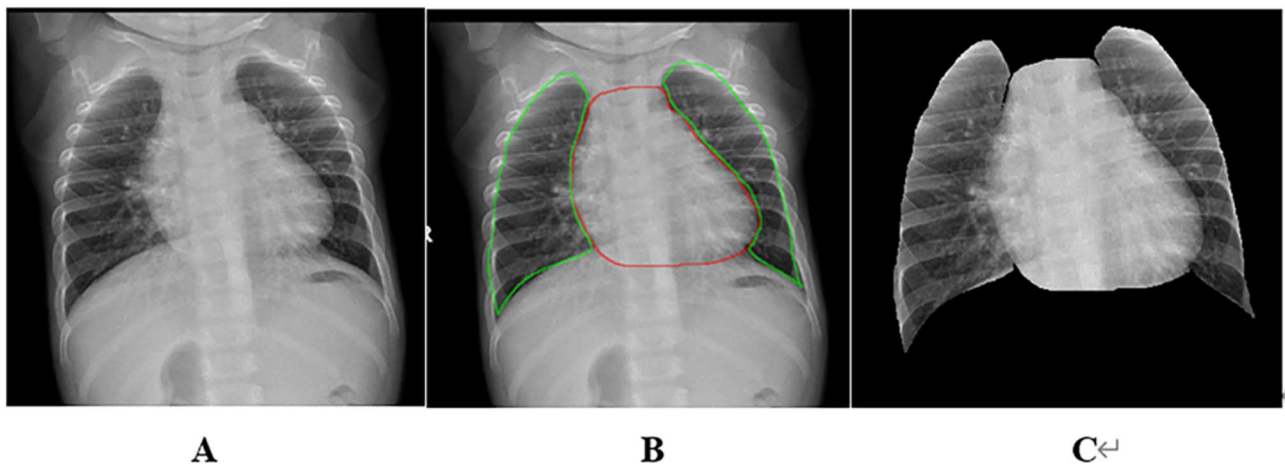
We used four networks in total, including ResNet-CBAM and others such as Inceptionv3 and ViT. We calculated the accuracy, specificity, sensitivity, PPV, TPR, TNR, and F1 scores of all networks in the test set. Four networks were trained for image classification and the reliability of the models in the differential diagnosis of ASD and VSD was verified using the activated regions in their classification. The results are shown in Table 2. The accuracy rate of InceptionV3 was the highest, reaching 0.872, and the sensitivity, specificity, positive predictive value, and F1 score were 0.975, 0.765, 0.811 and 0.886, respectively. The ROC curves of the four networks were plotted. The results showed that InceptionV3 had the largest area under the ROC curve, indicating its strongest ability to distinguish between ASD and VSD (Fig. 4).

### Comparison between models and radiologists

In this study, the deep learning model showing the best performance, InceptionV3, was selected for comparison with the radiologists' diagnosis (Table 3). Although Delong's test showed no significant difference between InceptionV3 and each radiologist in the test group



**Fig. 2** The representative X-ray images of ASD and VSD in children. **A** shows the chest X-ray preprocessing image of the representative ASD case; **B** is the representative X-ray image of VSD



**Fig. 3** Comparing the results of manual segmentation and model segmentation. **A** shows the chest X-ray preprocessing image of an ASD case; **B** is the lung field contour manually segmented by the physician; **C** is the image of the input model

**Table 2** Classified network test datas

Statistical Measure	EfficientNet	InceptionV3	ResNet18_cbam	ViT
ACC	0.8461	0.8718	0.8547	0.6582
SPE	0.8261	0.7652	0.7565	0.6174
SEN	0.8656	0.9748	0.9496	0.6975
PPV	0.8374	0.8112	0.8014	0.6535
NPV	0.8731	0.9003	0.8896	0.7352
AUC	0.901	0.9142	0.8715	0.6641
F1-score	0.8512	0.8855	0.8692	0.6748

ACC accuracy, SPE specificity, SEN sensitivity; PPV positive predictive value, NPV negative predictive value, AUC area under curve

( $P=0.241$ ,  $P=0.368$ ), InceptionV3 demonstrated higher AUC and sensitivity than that of the expert and the fellow (Fig. 5).

The Cohen Kappa test was used to determine the consistency among radiologists. The result showed that the Kappa value between the expert and the fellow was approximately 0.81, indicating excellent agreement.

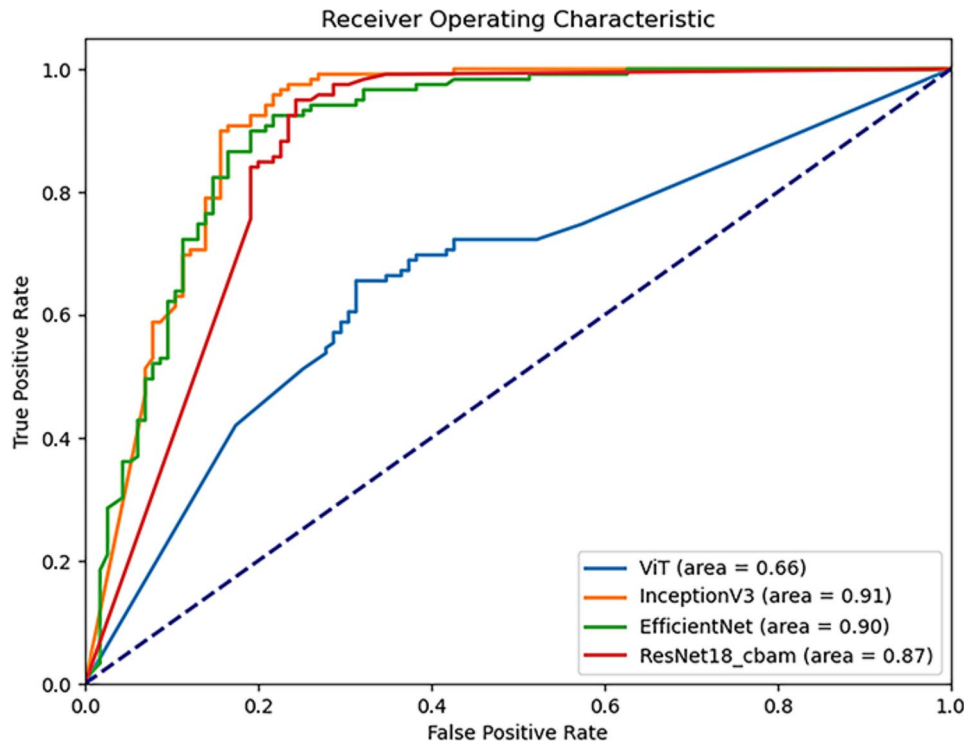
#### Activation heat maps from chest radiographs

For the network with the best results, Grad-CAM was used to draw class activation mapping (CAM) to visualize typical representatives of correct classification and misclassification and consider visualizing internal and external test data. Color indicated the amount of attention the model paid to a particular region, with red indicating more attention and blue indicating less attention. As shown in Fig. 6, both in the test and in the validation set, the red area of the true positive case of ASD was located in the right upper lung, while the red area of the true positive case of VSD was located in the left upper lung.

#### Discussion

In the present study, we employed four network models—EfficientNet, InceptionV3, ResNet18, CBAM, and ViT—utilizing the original chest X-ray images as input and classifying them into VSD or ASD categories as output. To objectively assess the performance of the deep learning models in detecting VSD or ASD, we implemented a fivefold cross-validation method. This involved randomly dividing the defect dataset into five segments, using four segments for training and reserving one for testing, while supplementing each test with an additional segment of normal data. This procedure was repeated five times, ensuring that the validation data varied with each iteration. This approach helped mitigate discrepancies in results and reduce the risk of overfitting due to the imbalance in VSD and ASD cases. The model that demonstrated the best performance during the fivefold cross-validation was subsequently employed for external testing.

The results of this study demonstrated that the average accuracy, sensitivity, and specificity of the four deep learning models for the differential diagnosis of VSD and ASD across five different test datasets were high, exceeding 0.70, with an average AUC value greater than 0.80. Among these models, InceptionV3 achieved the highest accuracy rate of 0.872, with sensitivity, specificity, positive predictive value, and F1 score recorded at 0.975, 0.765, 0.811, and 0.886, respectively. In addition, the Delong test and Kappa test of the test group showed no significant difference between InceptionV3 and each radiologist in the differential diagnosis ability of imaging specialists for VSD and ASD, and the diagnostic accuracy of the deep learning model was higher than that of the imaging physicians. These findings suggest that our model could assist physicians in refining their differential diagnoses, thereby enabling them to diagnose patients



**Fig. 4** Diagnostic performance of four models in differentiating ASD and VSD. Data are area under the receiver operating characteristic curve. The AUCs of the four classification networks are 0.87 (ResNet18\_cbam), 0.90 (EfficientNet), 0.91 (InceptionV3), and 0.66 (ViT)

**Table 3** Diagnostic performance of InceptionV3, the expert, and the fellow in differentiating ASD and VSD

Statistical Measure	The expert	InceptionV3	The fellow
ACC	0.8041	0.8617	0.7845
SPE	0.9520	0.9543	0.9437
SEN	0.1642	0.4749	0.1145
PPV	0.4337	0.6911	0.2904
NPV	0.8312	0.8946	0.8215
AUC	0.7823	0.8821	0.6749
P value	0.2416	Reference	0.3681

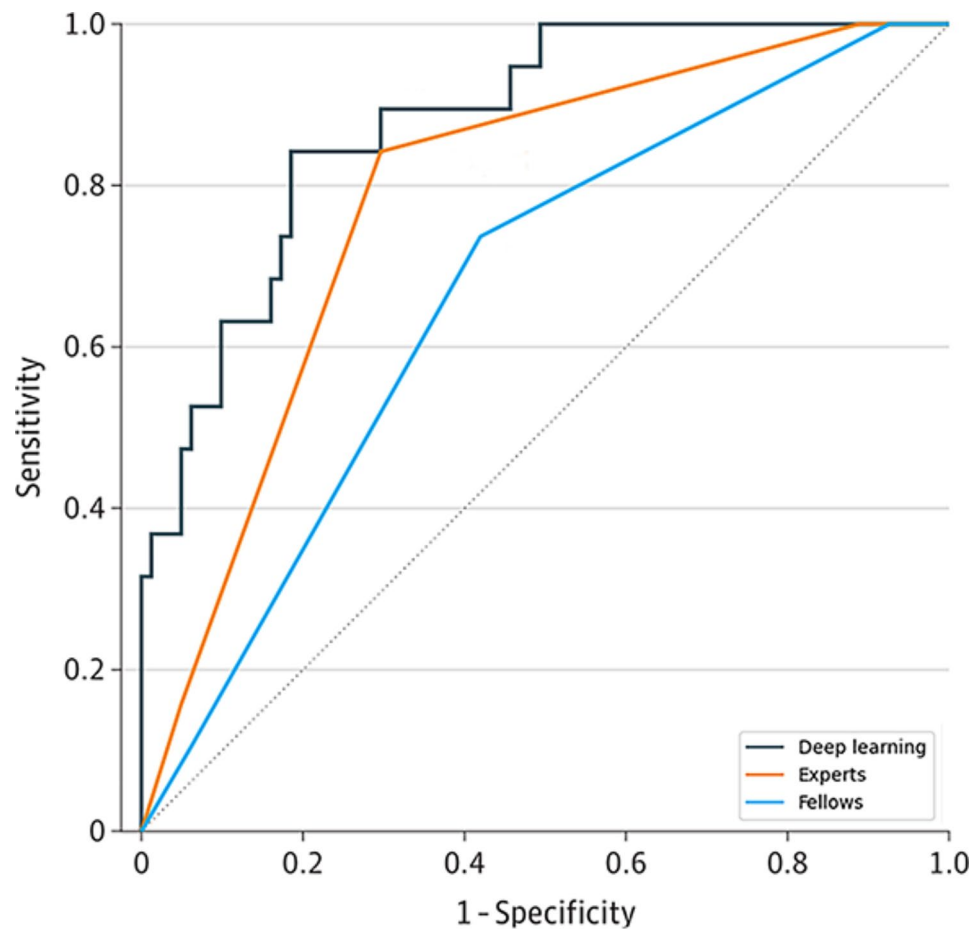
ACC accuracy, SEN sensitivity, SPE specificity; PPV positive predictive value, NPV negative predictive value, AUC area under curve

more effectively with the support of artificial intelligence. The model exhibited high diagnostic sensitivity and a low probability of missed diagnoses, which could significantly alleviate the workload of imaging physicians and enhance their diagnostic efficiency.

The deep learning model demonstrated greater sensitivity to VSD cases compared to ASD, as evidenced by a higher false detection rate for ASD. This discrepancy may be attributed to two primary factors. Firstly, from a hemodynamic perspective, the blood volume in the left atrium is smaller than that in the left ventricle, and the pressure difference between the atria is significantly less than that of the left ventricle. Additionally, the shunt in ASD is generally smaller, leading to milder cardiac structural changes compared to VSD. Secondly, the

increase in pulmonary circulation blood volume associated with ASD occurs later and to a lesser degree, resulting in relatively mild signs of increased vascular texture and thickening on chest radiography. Consequently, the deep learning model may have missed detecting some instances of ASD.

Compared with the imaging physicians, the deep learning model showed a higher false-positive rate. In the test and the validation set, 33 cases of ASD were on average falsely detected as VSD in the four deep learning models. We found that most of the false-positive cases were primary-foramen ASD or ASD with longer lesion time and larger defects. The location of the primary-foramen ASD was close to the mitral and tricuspid valve annulus, so tricuspid and mitral regurgitations may occur, resulting in the enlargement of the left heart circulation. In addition, if a larger ASD was not treated for a long time, pulmonary vascular resistance gradually increased, leading to pulmonary hypertension. When the pressure of the right atrium was greater than that of the left atrium, the right-to-left shunt could occur, and the left heart circulation of the patient also increased. Thus, in these two cases, patients could have left atrial and left ventricular enlargement, and the deep learning model was prone to mistaking ASD for VSD. Therefore, in the future, it is necessary to add training data on multiple defect types and constantly iteratively update the model to improve



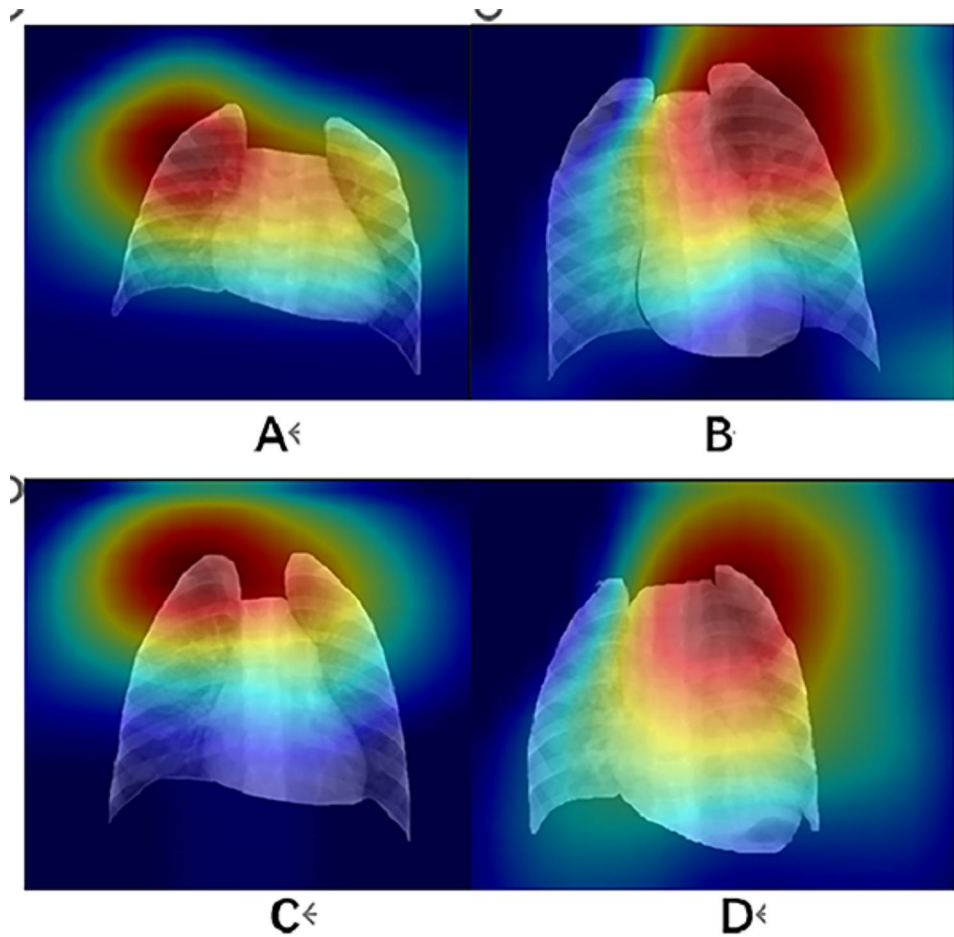
**Fig. 5** The ROC curves of the deep learning model, the expert, and the fellow for differentiating ASD and VSD. The AUCs of the three classification networks are 0.86 (the deep learning model), 0.80 (the experts), and 0.78 (the fellows)

the sensitivity and accuracy of the model for defect detection, especially for VSD.

Previous studies have shown that the use of artificial intelligence technology to assist in the detection of VSD or ASD has certain advantages. Gharehbaghi et al. [16] used a machine learning method, a time growth neural network, to distinguish between VSD heart sound, atrioventricular valve regurgitation heart sound, and normal heart sound in children, and achieved accuracy and sensitivity of 86.7% and 83.3%, respectively. Liu et al. [17] constructed The residual convolution recurrent neural network classification model based on deep learning to analyze children's heart sounds, which was able to preliminarily determine the type of left-to-right shunt CHD. Indeed, the diagnostic result of the model was better than expert auscultation, with an accuracy value of 0.940–0.994, which could improve the efficiency of CHD diagnosis. Toba et al. [18] predicted the pulmonary-to-systemic circulation flow ratio of patients with CHD based on deep learning chest radiograph analysis, quantitatively and objectively, which may confer an opportunity to quantify otherwise qualitative and subjective findings

of pulmonary vascularity in the clinical setting. Kim et al. [19] provided highly reliable cardiovascular boundary measurement according to the deep learning algorithm to diagnose and quantitatively evaluate valvular heart disease. Based on chest radiographs, this study applied the deep learning method to the differential diagnosis of VSD and ASD, and the detection effect was better than that in the previous studies. The average accuracy and sensitivity for the total test cases reached more than 80%, indicating improved diagnostic efficiency. The fivefold cross-validation method showed that the four models had good differential diagnostic ability for VSD and ASD and good robustness.

Importantly, the deep learning model showed good discrimination ability as well as displayed the location and range of specific identification points in the form of a heat map. The main feature of the class activation graph was to use the characteristic graph of the last convolution layer to find the corresponding weight of each channel through back-propagation. The larger the weight, the more important the corresponding characteristic graph. Then, the corresponding weights and feature maps



**Fig. 6** Activation Heat Maps from Chest Radiographs. **A** and **B** are from the internal test set, and **C** and **D** are from the external validation set. Chest radiograph with ASD (**A** and **C**) and VSD (**B** and **D**) were visualized using grad-CAM, which represents the area (yellow and red) that the deep learning model considered important for predicting an increased pulmonary

were multiplied to obtain the final class activation map [20–23]. The process of obtaining the attention heat map was as follows: the image to be visualized was input into the network model, the category was determined, and the output feature map of the layer to be visualized was obtained; then, the weight was assigned through the category of the image; values were assigned to each channel of the obtained feature map and added to form a single channel feature map. Figure 5 is activation heat maps of a representative example. In this study, we utilized a color-coding scheme where red indicates a large value and blue signifies a small value; the greater the value, the more attention the model pays to the corresponding area. Our findings from both the test and validation sets revealed that the red area associated with true positive cases of ASD was located in the right upper lung, suggesting that alterations in circulation within this region significantly contributed to the diagnosis of ASD. Conversely, the red area for true positive cases of VSD was found in the left upper lung, indicating that changes in circulation in this area had the greatest impact on the diagnosis of VSD.

This phenomenon may be attributed to the fact that variations in pulmonary blood flow typically originate from the peripheral arteries before affecting the pulmonary hilum, leading to more pronounced peripheral changes in the early stages of the disease. Additionally, due to anatomical and gravitational effects, pulmonary vessels in the upper lobes are generally thinner than those in the lower lobes under normal conditions. Consequently, when pulmonary circulation is altered, these same vessels may exhibit increased thickness or thinning, making changes in the upper pulmonary vessels more detectable by the deep learning model. Furthermore, in cases of ASD, the right atrium and right ventricle are enlarged, resulting in a relatively dense texture in the right lung, which enhances the visibility of increased blood flow in this region. Similarly, in VSD, the left atrium and left ventricle are enlarged, leading to a denser texture in the left lung, thereby making the manifestation of increased blood flow in the left lung more apparent. These speculations warrant further experimental validation.



This study had several limitations. Firstly, the number of cases in the test set was relatively small; however, we employed fivefold cross-validation to mitigate the impact of a potentially unbalanced test set on the results. Future research should aim to increase the number of defect cases in external testing to achieve a more objective evaluation of the deep learning model. Secondly, while our model demonstrated effectiveness in detecting VSD and ASD, it is important to note that most of the VSD cases tested were perimembranous VSD, and the majority of ASD cases were secundum ASD. These represent relatively simple and singular defect types, which may not encompass the full spectrum of defects encountered in clinical practice. In fact, apart from straightforward VSD and ASD cases, there exist other CHD malformations that are more challenging to detect. Currently, VSD and ASD are frequently diagnosed through cardiac computed tomography and echocardiography in clinical settings. Future efforts should involve incorporating chest radiograph data for various types of VSD and ASD into the training process, thereby developing a deep learning model that is better suited for clinical application. Furthermore, the training of deep learning classification models relies on parameter adjustments of annotated training sets to achieve discriminative performance that closely aligns with actual labels; however, this process is susceptible to overfitting and may introduce potential biases. Moving forward, we will concentrate on adaptive deep learning methods for image classification based on risk analysis to enhance the model's accuracy.

## Conclusions

In this study, deep learning using different convolutional network models was applied to differentiate and diagnose ASD and VSD using children's chest posterior anterior X-ray images. The results showed that despite different variations, all models exhibited good performance, with the Inception3 model being the best and having diagnostic efficacy no worse than that of radiology experts. This study indicates that deep learning models have clinical value in distinguishing between ASD and VSD, which can effectively improve the diagnostic efficiency of ASD and VSD in primary hospitals.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12887-024-05141-y>.

Supplementary Material 1

## Acknowledgements

The authors thank all the staff in Department of radiology, Children's Hospital of Soochow University, Suzhou, Jiangsu Province, People's Republic of China.

## Author contributions

HTL and WLG designed the study. HHJ, PP and YFQ collected data. HHJ, SQT and YKD wrote the draft. HTL, WLG, and CG contributed to the elaboration of the ideas developed in the manuscript and made critical amendments. GC, PP and YFQ contributed to the data collection and interpretation. DLH and YY provided the statistical analysis. The authors read and approved the final manuscript.

## Funding

This work was partially supported by Scientific Research Project of the Suzhou Health Commission (M2021029).

## Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

## Declarations

### Ethics approval and consent to participate

The authors ensured that questions related to the accuracy or integrity of any part of this work were appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (2013 revision), and the study protocol was approved by the Ethics Committee of the Children's Hospital of Soochow University (NO.2022-02-006-H01). Written informed consent was provided by the parents or legal guardians of the children.

### Consent for publication

Not Applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Department of Radiology, Children ' s Hospital of Soochow University, 215025 Suzhou, P. R. China

<sup>2</sup>School of Electronic & Information Engineering, Suzhou University of Science and Technology, 215009 Suzhou, China

<sup>3</sup>Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, 215163 Suzhou, China

<sup>4</sup>Jinan Guoke Medical Technology Development Co., Ltd, 250102 Shandong, China

<sup>5</sup>Department of Pediatric Cardiology, Children ' s Hospital of Soochow University, 215025 Suzhou, P. R. China

Received: 24 October 2023 / Accepted: 9 October 2024

Published online: 15 October 2024

## References

- Zhao QM, Liu F, Wu L, et al. Prevalence of congenital heart disease at live birth in China [J]. *Pediatr*. 2019;204:53–8.
- Liang H, Tsui BY, Ni H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence[J]. *Nat Med*. 2019;25:433–8.
- Diller GP, Babu-Narayan S, Li W, et al. Utility of machine learning algorithms in assessing patients with a systemic right ventricle[J]. *Eur Heart J Cardiovasc Imaging*. 2019;20:925–31.
- Egger J, Gsaxner C, Pepe A, et al. Medical deep learning—a systematic meta-review [J]. *Compute Meth Prog Bio*. 2022;221:106874.
- CHEN Y H, MAI Y C, FENG, R, et al. An adaptive threshold mechanism for accurate and efficient deep spiking convolutional neural networks[J]. *Neurocomputing*. 2022;469:189–97.
- FEI Z X, YANG E F, YU L, J, et al. A novel deep neural network-based emotion analysis system for automatic detection of mild cognitive impairment in the elderly[J]. *Neurocomputing*. 2022;468:306–16.
- HUANG K K, LI S, DENG W F et al. Structure inference of networked system with the synergy of deep residual network and fully connected layer network[J]. *Neural networks*, 2022;145:288–99.
- Nadeem MW, Ghamdi MAA, Hussain M et al. Brain Tumor Analysis Empowered with Deep Learning: A Review, Taxonomy, and Future Challenges[J]. *Brain Sciences*. 2020;10(2):1–33.

9. Halder A, Dey D, Sadhu AK. Lung nodule detection from Feature Engineering to Deep Learning in thoracic CT images: a comprehensive Review[J]. *J Digit Imaging*. 2020;33(4):655–77.
10. Rehman A, Khan FG. A deep learning based review on abdominal images[J]. *Multimedia Tools Appl*. 2020;80(3):30321–52.
11. Niehues SM, Adams LC, Gaudin RA, et al. Deep-learning-based diagnosis of bedside chest X-ray in Intensive Care and Emergency Medicine[J]. *Invest Radiol*. 2021;56(8):525–34.
12. Hwang EJ, Nam JG, Lim WH, et al. Deep learning for Chest Radiograph Diagnosis in the Emergency Department[J]. *Radiology*. 2019;293(3):573–80.
13. Kim JH, Han SG, Cho A, Shin HJ, Baek SE. Effect of deep learning-based assistive technology use on chest radiograph interpretation by emergency department physicians: a prospective interventional simulation-based study[J]. *BMC Med Inf Decis Mak*. 2021;21(1):311.
14. Lu MT, Ivanov A, Mayrhofer T, Hosny A, Aerts HJWL, Hoffmann U. Deep learning to assess long-term mortality from chest Radiographs[J]. *JAMA Netw Open*. 2019;2(7):e197416.
15. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision[J]. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.
16. Gharehbaghi A, Sepehri AA, Linden M, et al. Intelligent phonocardiography for screening ventricular septal defect using time growing neural network[J]. *Stud Health Technol Inf*. 2017;238:108–11.
17. Liu J, Wang H, Yang Z, et al. Deep learning-based computer-aided heart sound analysis in children with left-to-right shunt congenital heart disease[J]. *Int J Cardiol*. 2022;348:58–64.
18. Toba S, Mitani Y, Yodoya N, et al. Prediction of pulmonary to systemic Flow ratio in patients with congenital heart Disease using deep learning-based analysis of chest radiographs [J]. *JAMA Cardiol*. 2020;5(4):1–8.
19. Kim C, Lee G, Oh H et al. A deep learning-based automatic analysis of cardiovascular borders on chest radiographs of valvular heart disease: development/external validation[J]. *Eur Radiol*. 2022(3):32–44.
20. Rudie JD, Duda J, Duong MT, et al. Brain MRI deep learning and bayesian inference system augments radiology resident performance[J]. *J Digit Imaging*. 2021;34(4):1049–58.
21. Gao J, Jiang Q, Zhou B, et al. Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: an overview [J]. *Math Biosci Eng*. 2019;16(6):6536–61.
22. Wong PK, Yan T, Wang H, et al. Automatic detection of multiple types of pneumonia: open dataset and a multi-scale attention network [J]. *Biomed Signal Proces*. 2022;73:103415.
23. Angayarkanni SP. Hybrid convolution neural network in classification of cancer in histopathology images[J]. *J Digit Imaging*. 2022;35(2):248–57.

### Publisher's note

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