Fetal congenital heart disease classification algorithm based on improved DenseNet

Jiesong Qi[a], Hua Fang*[b]
[a]School of Information Science and Engineering, University of Jinan, Jinan, China; [b]School of Information Science and Technology, Taishan University, Taian, China
* Corresponding author: fanghua_tsu@qq.com

ABSTRACT

Early prenatal ultrasound screening can significantly reduce neonatal mortality due to congenital heart disease(CHD). Due to the uniqueness of the fetal heart structures and the variety of fetal cases, the prenatal detection rate of fetal CHD is still quite low; therefore, an improved DenseNet is proposed to diagnose fetal CHD. Compared to the adult heart, the size of the fetal heart varies significantly and its location is not fixed, so a multi-scale feature fusion module is introduced into the network, which extracts multi-scale features in the fetal heart by combining convolutional kernels of various sizes. Secondly, there are complex structures and rich information in fetal ultrasound images, therefore the Efficient Channel Attention (ECA) mechanism is integrated into the network, which suppresses the expression of unimportant information and mentions the reliability of model classification. The experimental results demonstrate that the improved DenseNet achieves better results in the task of fetal CHD classification. Additionally, the improved DenseNet enhances the prenatal detection rate of fetal CHD by achieving the recall of 85% and the precision of 85.3% on the test set.

Keywords: Congenital heart disease, echocardiography, medical image classification, attention module, multi-scale feature

1. INTRODUCTION

Congenital heart disease (CHD) is a cardiovascular malformation caused by abnormal development of fetal heart and blood vessel tissue. In recent years, CHD has consistently had the highest incidence of all congenital disabilities. Prenatal diagnosis of CHD has been shown to improve survival and reduce long-term morbidity [1].

With the continuous development of modern ultrasound technology, fetal ultrasound screening has received more and more attention and emphasis. Echocardiography is a non-invasive, reproducible and convenient examination method, which can obtain the internal structure and function information of the heart through different views, and is the first choice for screening and preoperative diagnosis of CHD [2].

Despite the rapid development of fetal ultrasound imaging technology, the inconsistency rate between ultrasound reports and expert diagnostic opinions is still high, which may affect the treatment plan and treatment effect of fetal CHD. The reasons for this phenomenon are as follows: (1) Fetal echocardiographic images have low resolution and uneven quality, speckle noise and artifacts, which bring great obstacles to the diagnosis of CHD. (2) Due to the special anatomical structure of fetal heart compared with adult heart, fetal involuntary movement and other factors, the difficulty of fetal ultrasound examination is increased. (3) Echocardiography examination requires high professional knowledge and clinical experience of doctors, and doctors may miss or misdiagnose due to lack of clinical experience, fatigue and other factors.

To alleviate the above problems, this paper proposes a deep learning model with improved DenseNet to automatically screen and diagnose fetal CHD. The association pattern between echocardiography and fetal CHD was found automatically through computer-aided diagnostic analysis to improve the accuracy of classification.

2. RELATED WORK

2.1 CNN classification model

Convolutional Neural Network (CNN) is a deep learning model that often have different structures for different tasks. In 2012, a new CNN model called AlexNet [3] achieved great success in the ImageNet image recognition competition, which

International Conference on Future of Medicine and Biological Information Engineering
(MBIE 2024), edited by Yudong Yao, Xiaoou Li, Xia Yu, Proc. of SPIE Vol. 13270,
1327009 · © 2024 SPIE · 0277-786X · doi: 10.1117/12.3045858

owes its success to the deeper network structure as well as the development of GPU technology. Furthermore, deep CNNs are already widely used techniques for image classification applications thanks to the availability of the ImageNet dataset.

Currently, feature extractors for various computer vision tasks show excellent performance, such as VGGNet [4], ResNet [5], MobileNet [6], InceptionNet [7], and DenseNet [8].

Despite CNN's excellent performance in a variety of tasks, gradient disappearance or explosion problems still seriously hinder its performance improvement. ResNet introduces the concept of a residual block, the structure of which contains a shortcut connection across the layers, which allows the input information to be passed directly to deeper layers of the network, thus alleviating the gradient vanishing problem. 2016 Huang proposed DenseNet, a densely-connected convolutional neural network, which is designed with dense connections between each layer of the network to achieve feature map reuse and effectively solve the gradient vanishing problem.

2.2 Deep learning in fetal echocardiograms

Current research in the field of fetal congenital heart disease (CHD) focuses on normal and abnormal classification of the fetal heart. Gong et al. [9] initially cropped fetal heart regions from fetal four-chamber (FC) images using the Faster-RCNN model in order to identify fetal CHD. They then suggested a novel model, DGACNN, to categorize the extracted regions. Dong et al. [10] proposed a general deep learning framework for four-chamber planar automatic quality control of fetal heart. In order to successfully identify particular cardiac views, Arnaout et al. [11] investigated the use of neural network integration to develop a model through supervised learning. They also successfully employed a rule-based classifier to distinguish between structurally normal hearts and complex CHDs. Qiao et al. [12] proposed a straightforward and efficient residual network diagnostic system. This system greatly increases the diagnostic accuracy of fetal CHD by using a convolutional neural network and provides a global visual interpretation during the diagnostic process.

3. PROPOSED METHODOLOGY

3.1 DenseNet

DenseNet proposes a dense connection mechanism, is combining the outputs of all previous layers to get the input of the current layer. DenseNet has (L(L+1))⁄2 connections in total for a network with L layers. This dense connection mechanism shortens the distance between the layers of the network, further enhances the information flow between the layers, does feature reuse. DenseNet has shown higher performance than traditional networks in medical image processing, so the study chooses DenseNet as the base network for constructing the CHD classification model.

DenseNet network employs DenseBlock and TransitionLayer structure, where each DenseBlock contains n Denselayer layers. These layers output feature maps with the same dimensions, and adopt the above-mentioned dense connection mechanism between layers, and its basic structure consists of a series of BN, ReLU, and Conv operations, as shown in Figure 1. Two adjacent DensenBlocks are connected by the TransitionLayer structure, and the feature maps size is reduced by using convolution and pooling operations.

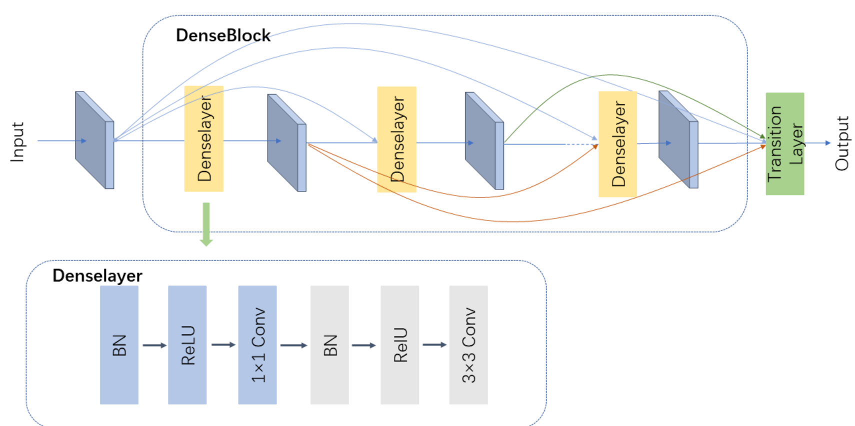

Figure 1. Structure of DenseBlock and TransitionLayer.

Since the diagnosis of CHD in echocardiography requires the observation of cardiac structures, such as the four chambers of the fetal heart, the interatrial septum, and the interventricular septum. These cardiac structures are usually of different scales and sizes, so we have enhanced the Denselayer by adding a multi-scale feature fusion module.

Inception Block [13] uses convolution kernels of various sizes to create receptive fields of various sizes, and finally performs fusion operations on the multi-scale features that are produced. The basic structure of the block consists of four parallel paths: one 1x1 convolution, one 1x1 convolution + 3x3 convolution, one 1x1 convolution + 5x5 convolution and one 3x3 maximum pooling + 1x1 convolution. In order to increase the computational efficiency of the model while keeping the same receptive field, two 3x3 convolutions are used in place of the 5 x 5 convolution in this instance. The module introduces a 1 x 1 convolution before each 3 x 3 convolution to adjust the number of channels in the feature map and add non-linear features to improve computational efficiency. The Inception Block is capable of extracting multi-scale features in a single module by using convolution kernel and pooling operations of various sizes in parallel. This makes the features richer, and also means that the final classification is more accurate.

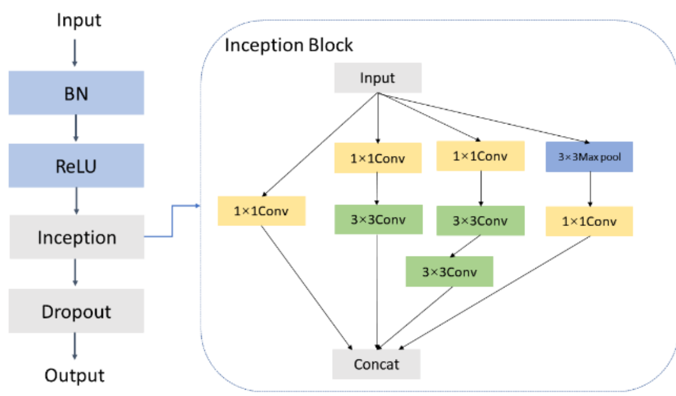

Figure 2. Improved Denselayer structure.

Figure 1 depicts the original network structure's Denselayer structure for deep feature extraction, which makes use of a single 1x1 and 3x3 convolution with a single receptive field. This structure is not conducive to multi-scale feature extraction. Consequently, the Inception Block is introduced into the above layer structure to increase the width of the network while obtaining a multi-scale receptive field, thus improving the representation capability of the network. Meanwhile, the combination of Dropout, can randomly select a part of neurons and set its output to 0 to reduce overfitting and enhances the generalisation ability of the model. The Improved Denselayer is shown in Figure 2.

3.2 ECA Block

In order to improve the model performance with less increase in model complexity, ECA attention mechanism is introduced in the network. It is an efficient channel attention module which determines the interactions between channels and extracts the dependencies between channels by fast 1D convolution.

The ECA Block [14] is comprised of an effective excitation module for modeling cross-channel interactions and a squeezing module for aggregating global spatial information. By considering only direct interactions between each channel and its k-nearest neighbors, this block controls model complexity. The ECA Block, after aggregating the convolutional features using the GAP without dimensionality reduction, first adapts itself to determine the size of the convolutional kernel k, then 1D convolution is performed, followed by a Sigmoid function to learn the weights of each channel.

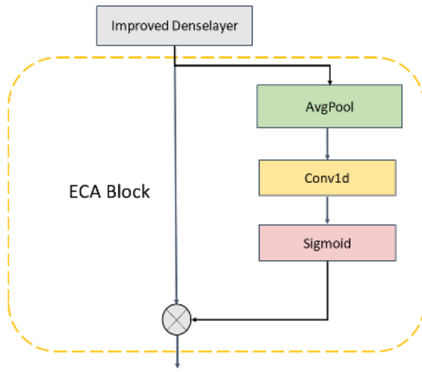

Figure 3. Structure of Improved Denselayer with the ECA Block.

In order to further strengthen the performance of the DenseNet model and improve the feature extraction ability of the network, the ECA attention block is added to the Improved Denselayer structure. The ECA Block can give different weights to the feature channels to help the model better capture the key information, thus improving the performance of the model. At the same time, with the characteristics of small parameter count and simple structure, the block can be conveniently integrated into the fetal congenital heart disease classification model. The improved layer structure is shown in Figure 3.

3.3 Improved DenseNet

The standard DenseNet121 consists of four DenseBlocks and three TransitionLayers, where each DenseBlock consists of 6, 12, 24, and 16 Denselayer layers, respectively. The layer structure in each DenseBlock was replaced with Improved Denselayer with multi-scale feature fusion and ECA Block, and the number of Improved Denselayer in each DenseBlock was adjusted to 2, 3, 3, 2 due to the high complexity of the improved model and the small amount of fetal ultrasound data, and the adjusted model is noted as Improved DenseNet, and Figure 4 depicts the network's general architecture.

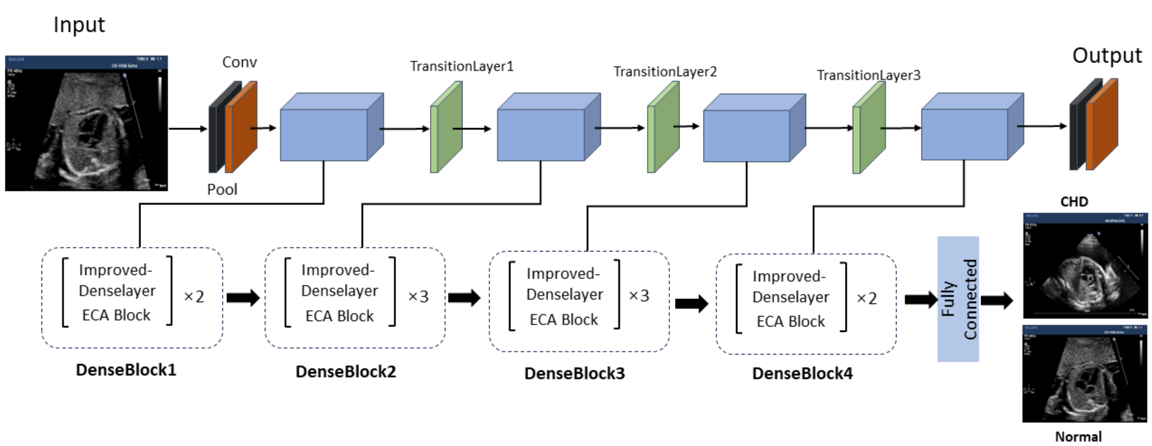

Figure 4. Network architecture.

4. EXPERIMENTAL RESULTS AND ANALYSES

4.1 Datasets and evaluation measures

The dataset used ultrasound data from the Shandong Provincial Hospital, which was mainly collected between 18 and 28 weeks of gestational age. The dataset includes 43 segments of echocardiography videos and 654 images of the fetal heart at the end-systole in fetal four-chamber（FC）view. End-systolic images of the heart are medically essential for assessing cardiac structure and function, which allows for a clearer view of the movement and thickness of the ventricular walls and assessment of whether the atria and ventricles are normal in size, especially in the presence of atrial septal defects or ventricular septal defects. Therefore, 81 images of the fetal heart at end-systole were obtained by processing 43 echocardiographic video clips.

The original FC views included some personal information about the patient and other biological tissues that were not relevant to the diagnosis of CHD, so image preprocessing was carried out on these FC views. In order to obtain the echocardiogram's region of interest, the fetal heart position was first identified and the region of interest in the images was extracted. Then, by clipping and resizing the echocardiogram to remove a large amount of invalid background, 256×256 image blocks containing the fetal heart are obtained. To avoid the network model overfitting, random vertical flipping, panning, and random horizontal flipping were also applied to the fetal heart dataset. The image preprocessing process for the fetal FC view is shown in Figure 5. Finally, the dataset was divided into a training set and a test set; 630 fetal FC views, including 420 images of a healthy fetal heart and 210 images of a CHD, made up the training set. To assess the model's performance, 80 FC views—40 of which are images of healthy fetal hearts and 40 of which are images of fetal hearts with CHD—that do not appear in the training set are utilized in the test set.

The deep learning framework used in this experiment is Pytorch. The batch size of each training session is set to 16, the training steps are 50 epochs, the optimization process employs the Adam optimizer, the dropout rate is set at 0.3, and the initial learning rate is 0.0001.

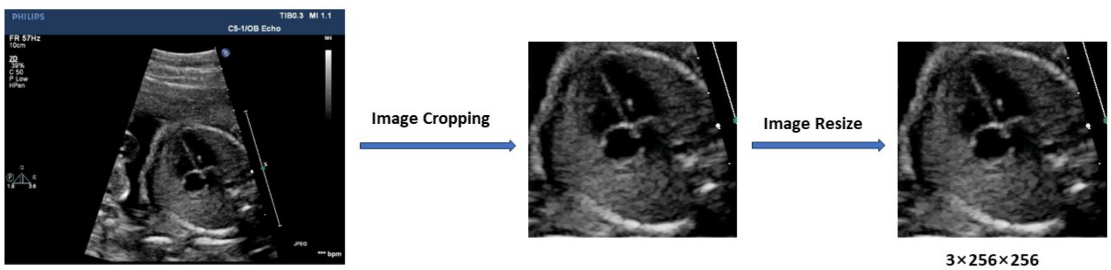

Figure 5. Image pre-processing process.

The confusion matrix yields three main metrics that we use to assess our proposed model, including precision, recall and F1-score as shown in Equation 1, Equation 2 and Equation 3. In the classification task, the F1-score is an evaluation metric that combines precision and recall to assess the performance of the model. Therefore, the model performs better in identifying fetal CHD when the precision values are greater, and the model is more sensitive to fetal CHD when the recall values are larger.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Reacll} \tag{3}$$

4.2 Comparison and analysis of experimental results

The experiments were conducted according to the experimental setup. In order to prove the rationality of the proposed network, VGG16, AlexNet, GoogLeNet, ResNet50, and DenseNet121 were used as comparison experiments, and the comparison results are shown in Table 1.

Firstly, the ultrasound heart images are split into two categories by the classification algorithm: normal images and cardiac diseased images. The classification results are shown in Table 1, and the table shows that the Improved DenseNet improves the classification precision, recall and F1-score over the standard DenseNet121, achieving the precision of 85.3%, the recall of 85% and F1-score of 85.1%. The standard DenseNet121 outperformed the other network models with a precision of 79.6%, a recall of 78.8% and an F1-score of 79.2%. The above data indicate that the improved model is more effective in the classification of fetal CHD.

Secondly, the effectiveness of the modules was verified by ablation experiments, and the effects of each module on the classification results are shown in Table 2. The table shows that after blending Inceptoin Block in the network to form

Improved Denselayer, the performance of the network is all higher than other traditional network structures. After fusion, multi-scale information is introduced into the network, which improves the model's ability to understand and represent images. The enhancement effect is obvious from the data, indicating that the fusion module is effective. Adding the ECA Block to Improved Denselayer improved the precision of the network model by 2.7%, demonstrating that the addition of the attention mechanism can improve the model's diagnosis of fetal CHD.

Table 1. Experimental comparison results.

Model	Precision	Recall	F1-score
VGG16	61.4%	61.2%	61.3%
AlexNet	68.6%	67.5%	68%
GoogLeNet	70.6%	68.4%	69.5%
ResNet50	75%	73.8%	74.4%
DenseNet121	79.6%	78.8%	79.2%
Ours	85.3%	85%	85.1%

Table 2. Results of ablation experiments.

Model	InceptionBlock	ECA Block	Precision	Recall	F1-score
	√		82.6%	82.5%	82.5%
		√	84.9%	81.2%	83%
Ours	√	√	85.3%	85%	85.1%

5. CONCLUSION

Despite advances in ultrasound imaging, the prenatal detection rate of fetal congenital heart disease remains meager. The complex anatomy of the fetal heart, as well as its small size and the diversity of fetal cardiac anomalies, add to the complexity of the examination. In this paper, we combine the multi-scale feature fusion module and the attention mechanism on the DenseNet model to enhance the extraction of multi-scale features and the capture of key information in fetal echocardiograms, thus improving the accuracy of fetal CHD classification. At the same time, in order to prevent overfitting of the network model, the number of Denselayer is adjusted. On the test dataset, the improved model had the precision of 85.3% and the recall of 85% in identifying fetal congenital heart disease, which effectively improved the prenatal detection rate of fetal CHD.

REFERENCES

[1] S. Pan, and G. Luo, "Application prospect of medical artificial intelligence in fetal echocardiography," 35(11), 850-853 (2020).

[2] V. T. Truong, B. P. Nguyen, T.-H. Nguyen-Vo et al., "Application of machine learning in screening for congenital heart diseases using fetal echocardiography," 38(5), 1007-1015 (2022).

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," 25(2012).

[4] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," (2014).

[5] K. He et al., "Deep residual learning for image recognition," Proceedings of the IEEE conference on computer vision and pattern recognition 770-778 (2016).

[6] A. G. Howard et al., "Mobilenets: Efficient convolutional neural networks for mobile vision applications," (2017).

[7] C. Szegedy et al., "Going deeper with convolutions," Proceedings of the IEEE conference on computer vision and pattern recognition 1-9 (2015).

[8] G. Huang et al., "Densely connected convolutional networks," Proceedings of the IEEE conference on computer vision and pattern recognition 4700-4708 (2017).
[9] Y. Gong et al., "Fetal congenital heart disease echocardiogram screening based on DGACNN: adversarial one-class classification combined with video transfer learning," 39(4), 1206-1222 (2019).
[10] J. Dong et al., "A generic quality control framework for fetal ultrasound cardiac four-chamber planes," 24(4), 931-942 (2019).
[11] R. Arnaout et al., "An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease," 27(5), 882-891 (2021).
[12] S. Qiao et al., "RLDS: An explainable residual learning diagnosis system for fetal congenital heart disease," 128(205-218 (2022).
[13] C. Szegedy et al., "Inception-v4, inception-resnet and the impact of residual connections on learning," Proceedings of the AAAI conference on artificial intelligence (2017).
[14] Q. Wang et al., "ECA-Net: Efficient channel attention for deep convolutional neural networks," Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 11534-11542 (2020).