

A Dense Cascaded Network Model for Outlier Prediction and Segmentation of Cardiac Images

Dr. Rasmi^{1*}, Dr. A. Jayanthiladevi² and Dr. Sunitha H. D.³

Submitted: 29/11/2023 Revised: 09/01/2024 Accepted: 19/01/2024

Abstract: Deep learning techniques have been effectively used for various applications to segment anatomical features in medical imaging. However, the images are split significantly and impact the outcome. Many clinical images with outliers caused by organ motion, patient movement, and image acquisition-related difficulties are frequently ignored topics in the medical image analysis community. This study compares various techniques for segmenting the heart cavity and compensating for outliers. We discuss how image motion outliers affect the segmentation of cardiac MR images. The method's foundation is a freshly developed integrated outlier detection and reconstruction method. Using a joint loss function and the enforcement of a data consistency term successfully transforms the outlier repair job into an under-sampled image reconstruction problem. This study proposes an end-to-end cascaded ResNet architecture with a segmentation network. The first two tasks improved by our instruction are image segmentation, outlier correction, and outlier identification. Using purposefully distorted and uncorrected cardiac images, the network reconstruction is done to fix motion-related aberrations automatically. We demonstrate that excellent segmentation accuracy and respectable image quality are attainable where MRI acquisitions are used as a test set. It was determined whether the fake motion outliers were present. In comparison to several image-correcting structures, this work demonstrates improved performance.

Keywords-an outlier, cardiac image, segmentation, image reconstruction, transforms.

1. Introduction

The three typical steps in medical image analysis are downstream tasks, such as image de-noising and reconstruction (including registration and segmentation). A cascade of mistakes from one task to the next can be introduced by applying medical image analysis algorithms to the unprocessed data serially, especially when the quality of the gathered data could be better. Although the use of medical images in clinical decision-making is growing, the assurance of image quality is a step that needs to be included in automated image analysis processes. It is a crucial step since good-quality medical images are essential for downstream activities like

segmentation [1].

Cardiac magnetic resonance (CMR) scans can be utilized to analyze local anomalies in myocardial wall motion, and measure measures of Volume, ejection fraction, and strain are examples of cardiac function [2]. These images are regularly taken to evaluate the condition of the heart. Patients who have the cardiovascular disease already frequently get CMR imaging. These people are more prone to get arrhythmias and may have issues holding their breath or remaining still when acquiring. Therefore, various image outliers may be present in the photographs [3], making it difficult to evaluate the level of detail in images produced by M.R. scanners. Clinicians may derive the wrong inferences from the imaging data due to deceptive segmentations [4]. Images are currently removed from further analysis and reacquired in clinical practice after being visually reviewed by one or more professionals and determined to be of insufficient quality. However, these data might be used for additional clinical analysis if effective image rectification and segmentation algorithms are used. We offer a deep learning-based method in this research to identify, correct, and segment joint motion outliers in cine short-axis CMR images. A novel end-to-end training setup is provided to recognize and repair motion outliers' and extract

*1*Post Doctoral Research Fellow, Srinivas University, Mangalore, Karnataka*

Associate Professor, Electronics & Communication Engineering, RR Institute of Technology, Bangalore, Karnataka

2Professor, Computer Science and Information Science, Srinivas University, Mangalore, Karnataka, India,

3Professor, Electronics & Communication Engineering, RR Institute of Technology, Bangalore, Karnataka

1rzwsa@yahoo.com,

2drjayanthila@srinivasuniversity.edu.in and

3suni742021@gmail.com

1https://orcid.org/0000-0002-0781-5574,

2https://orcid.org/0000-0002-6023-3899 and

3https://orcid.org/0000-0002-0609-914x

segmentations for the recovered images in a full, integrated framework. Also included is a study of several deep learning architectures and learning processes. In our previously published work [5], we suggested using it to find and correct motion outliers'; Convolutional Neural Network (CNN) architecture is used. This research furthers that idea.

- 1) Here, we expand on this concept to introduce a collaborative training strategy for segmenting, detecting and correcting image outliers'.
- 2) Then, a novel cascaded ResNet model is proposed for outlier prediction and segmentation process;
- 3) Various metrics like PSNR, MAE, SSIM and S.I. are evaluated and compared with other approaches.

The work is organized as follows: section 2 provides a detailed analysis of various prevailing approaches; section 3 is a methodology, and the numerical outcomes are provided in section 4. The work is summarized in section 5.

2. Related Works

Deep learning techniques have demonstrated outstanding success on benchmark data sets in the fascinating field of image quality assessment (IQA). According to [6], IQA is crucial in analyzing huge medical imaging data sets. Evaluating the image quality of M.R. images of the brain was the focus of early work in medical imaging. Deep learning's prominence has motivated medical image analysis to apply these techniques to various image quality assessment problems, utilizing 2D images and trained networks, such as echocardiography and fetal ultrasound. To spot motion outliers' in M.R. scans of the head and abdomen, Badrinarayanan et al. [7] developed spatially aware probability maps using a patch-based CNN architecture. More recently, the author recommended employing different attributes and instructions in a deep neural network for identifying outliers.

Problems with CMR image quality have primarily been researched due to their negative impacts on ventricular computing volume and, subsequently, ejection % in the presence of missing slices [8]. For the estimation of heart coverage, the cardiac region's inter-slice, motion detection, and image contrast, as stated by Castiglioni et al. [9], suggested using a decision forest approach. The automatic quality management of image

segmentation has also been connected to CMR image quality [10]. The author suggested using a random forest classifier in a complete segmentation framework to eliminate unsuccessful myocardial segmentation [11]. Using histogram, box, line, and texture data, Synthetic motion outliers' were studied by the authors of [12] and trained a random forest system to recognize various levels of outliers'. Sherubha et al. recent study [13] recommended using a curriculum-based approach to learning techniques to identify cardiac motion outliers' by utilizing various degrees of k-space corruption. These methods merely identify the outliers'; they do not fix the data they have affected. As a result, they can be utilized in clinical settings as a rejection mechanism but cannot permit the use of damaged data [14]. Our work is distinct from conventional approaches since it emphasizes under-sampling to account for speed [15]. We attempt to focus on outliers' for better segmentation and image quality. As far as we know, our research is the first to examine combined in a single image; image outlier identification, correction, and segmentation are performed. We use the cardiac image to illustrate how the framework may be applied, but it could also be useful for brain M.R. or other applications.

3. Methodology

This paper develops a novel imbalanced discriminant learning model for the first time synthesis of cardiac images from sMRI. The unbalanced model consists of three branches, each of which employs various numbers of unique ResNet-based substructures. Fig 1 shows the new unbalanced model's structural elements. Furthermore, the distinct ResNet-based substructure is composed of cascades of the newly formed Input block, intermediate block, and bottom block on each branch of the imbalanced model. An input block with three channels is constructed and shown in Fig. 4 to extract input latent and main features and enhance the structural diversity of the model. All three channels have identical topological topologies since each has 5×5 convolution. The architecture mentioned above was primarily influenced by utilizing the current craze for group convolution, which significantly lowers the convolutional parameters compared to many standard deep learning models' use of a single convolution per channel. Given the original input has the dimensions $192 \times 256 \times 64$, as shown in

Fig 1, According to [10], In this study, the z-axis is used to partition the input of InputBlock into three overlapping sub-blocks. Each subsection that has been separated is then supplied into various channels for convolution processes. As a result, each channel's learning of latent and primary characteristics focuses on a different input region. The above-mentioned overlapping design allows

for the implicit preservation of the spatial relationships between various input regions. Then, the outputs from three distinct channels will be combined, and B.N., Following a convolution, ReLU (or rectified linear unit) processing was used. Input block can determine the inputs' hidden and explicit attributes, which are then passed on to the block afterwards.

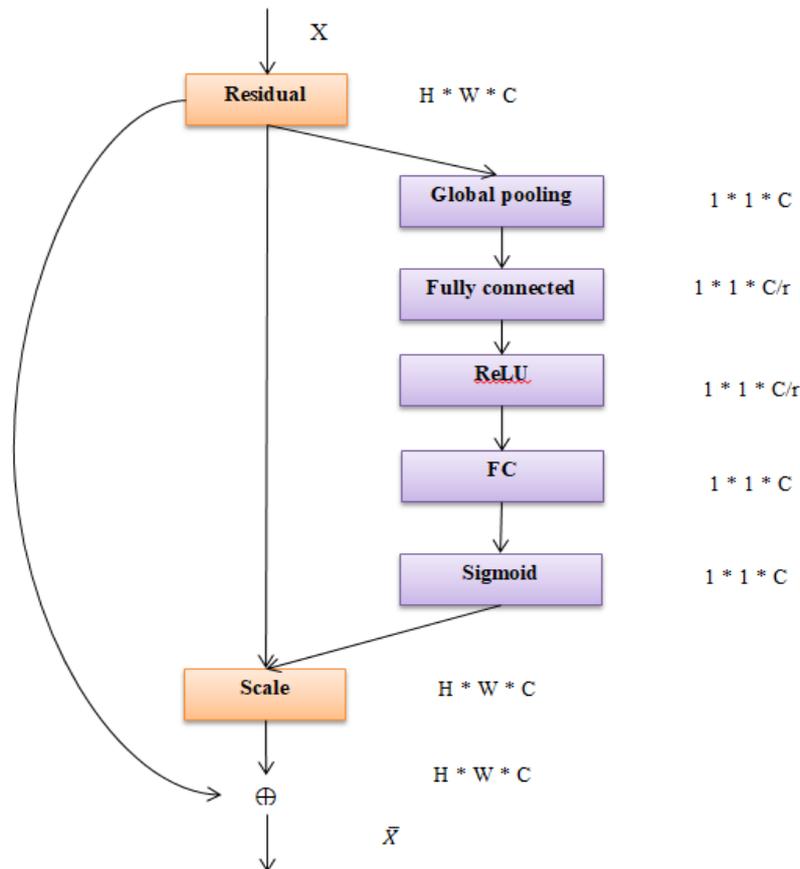


Fig 1 Cascaded ResNet model

Latent and coarse properties at various scales use basic blocks. Fig 1 illustrates structural elements. Here, several crucial issues need to be handled. First, to increase the model's structural multiplicity, the basic block uses two convolution channels in addition to the three used by the input block. Second, the two convolution channels use different sizes of convolutions (3×3 and 5×5) in contrast to InputBlock. Convolutions' receptive field will be expanded from 3×3 to 5×5 , enabling the enrichment of latent properties. Third, the blocks input sizes ($W \times H \times 192$ or $W \times H \times 512$) are different. (W and H represent an image's width and height, respectively.) As shown in Fig 1, each ResNet sub-structure consists of three blocks, with the first layer having a small input size ($W \times H \times 192$)

and the second and third blocks ($W \times H \times 512$) having huge input sizes. Input is divided and fed into the first layers with two channels where huge input will be divided unevenly between the second and third blocks, which are $W \times H \times 128$ for 3×3 convolutions and $W \times H \times 364$ for the 5×5 convolutions. The third and fourth blocks, which primarily highlight the sub-structure constructed on ResNet, have high-level latent features at the rear. These blocks may become unbalanced in two channels with unbalanced inputs and convolutions. Throughout blocks, latent properties of various scales can be enriched, discovered, and merged.

Learning latent and specific properties of various scales is designed to do. Fig 1 illustrates the block's

structural characteristics. Because block deploys four channels with various intricate architectures, it is comparable to block. However, it is clear from Fig. 6 that there is a large amount of imbalance in the block, which is useful for identifying the latent and specific characteristics of various scales. Another intriguing point is that each block channel contains a 1×1 convolution kernel. The reason is that the image channel will be smaller after the 1×1 convolution, going from $W_0 \times H_0 \times 192$ or 256 to $W_0 \times H_0 \times 128$, as seen in Fig 1. As a result, it is possible to lower both the number of feature maps and the number of parameters that must be calculated for the synthesis model. The synthesis model's training effectiveness can be increased in this way. It is also important to note that the well-known Inception V3 model also uses a similar concept. The three branches are made of imbalanced structure, and the primary realizations

$$MAE = \frac{1}{N_p} \sum_{p=1}^{N_p} |(I_x(p) - I_y(p))| \quad (1)$$

Here, N_p is the number of pixels in the image, I_x and I_y , and p stands for each pixel. Then, PSNR is depicted as in Eq. (2):

$$PSNR = 20 \log_{10}(\max(I)) - 10 \log_{10} \left(\frac{1}{N_p} \sum_{p=1}^{N_p} (I_x(p) - I_y(p))^2 \right) \quad (2)$$

Where $\max(I)$ stands for the ground truth image's maximum intensity value; for each image region x

$$SSIM(p) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3)$$

Here, μ_x and μ_y have average intensities of x and y , and regions σ_x and σ_y are the variances of those intensities, σ_{xy} represent the covariance and c_1 , and

$$SI(u) = \log_{10} \Phi \frac{\mu - TV(u)}{\sigma} \quad (4)$$

Where $\mu = \text{Var}[TV(I)]$ is the associated variance and $\mu = E[TV(I)]$ is the expectation of the overall fluctuation of the image. The dice overlap metric is

$$D(A, B) = \frac{2||A \cap B||}{||A|| \cup ||B||} \quad (5)$$

Two experiments were conducted to identify the best architectural architecture for correcting and segmenting CMR images, including visual outliers'. To select the ideal architecture, we considered several other layouts while maintaining the same total number of parameters for the

of the number of novel ResNet-based sub-structures changes in the new unbalanced model, which can be described as follows. Additionally, depth, wideness, and cardinality are indicated by the Input layer, blocks, and blocks inside the generalization power of the proposed unbalanced model can be further improved using ResNet-based sub-structures.

4. Numerical results and discussion

This section provides a detailed analysis of numerical outcomes. Mean Absolute Error (MAE), Peak Signal Noise Ratio (PSNR), and Structural Similarity Index are used to assess the quality of a photograph. Utilizing the Sharpness Index (S.I.), performance can be evaluated when no ground truth is available for analyzing future data. MAE is depicted as in Eq. (1):

and y , the following is how between two images is known as the SSIM:

c_2 represent the stabilizing coefficients for the denominator. The following sentence describes the S.I. between two images:

used to evaluate segmentations between areas A and B and has the following definition:

models: An example, a conventional U – net is a single U – net that manages segmentation and image correction. Cascaded ResNet with two output channels is additional that generates both the final segmentation and the corrected image. Third, the pair of serial Resnets output the

segmented image first and then the corrected image. We changed the number of filters in each design to make the total number of model parameters approximately identical, allowing a fair comparison. We employed a test set of 500 images, a training set of 3000 times from the subset feature, a validation set of 500 images and a validation set of 500 images to compare each design. We present

the segmentation findings from this research in Table I. We also provide the best outcome that can be attained, which uses one resnet model and does not train or test on any distorted photos (cascaded ResNet). Additionally, we segmented the low-quality images using the same model without additional training to obtain a baseline performance.

Table 1 Segmented outcome

Methods	LV	Myo	RV
U-net	98.5%	95%	94%
Baseline ResNet	93.2%	86%	87%
Single ResNet	94.8%	91%	91%
Cascaded U-net	90.8%	91%	90%
Cascaded 2-channel U-net	96.2%	91.5%	92%
Cascaded ResNet	98%	92%	93%

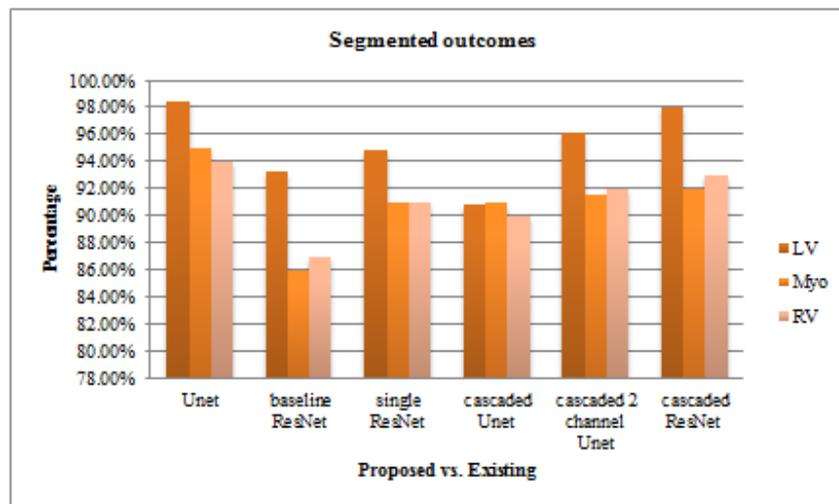


Fig 2 Segmented outcome analysis

Table 2 Performance evaluation

Methods	MAE	PSNR	SIM	SI
U-net	0.06	19	0.72	55.3
Baseline ResNet	0.07	24.3	0.74	64.3
Single ResNet	0.06	25.1	0.76	70.8
Cascaded U-net	0.05	26.2	0.78	70.9
Cascaded 2-channel U-net	0.04	27.3	0.80	71.8
Cascaded ResNet	0.02	29.5	0.82	72.5

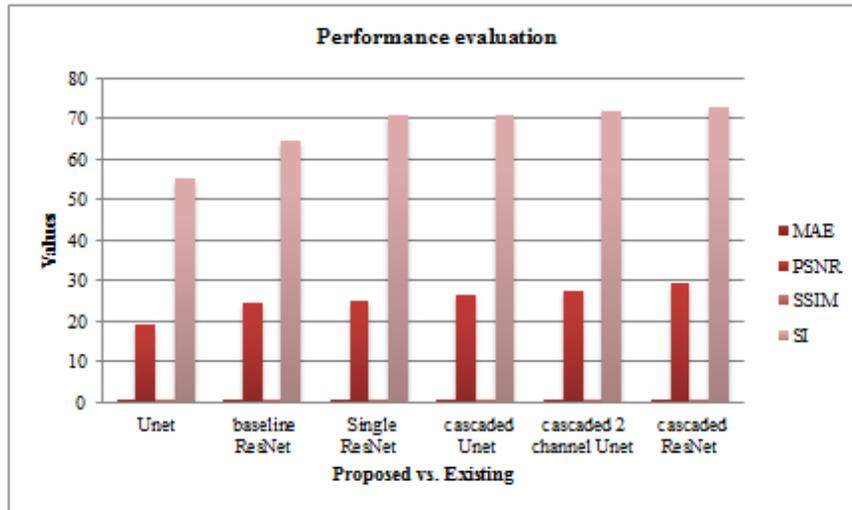


Fig 3 Outlier performance evaluation

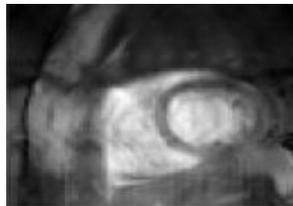


Fig 4a Outlier cardiac image

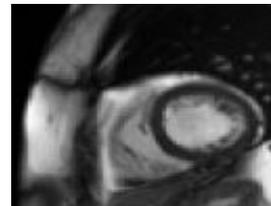


Fig 4b Proposed image

Based on our research findings, two serial networks are trained end-to-end to produce the best segmentation. This work decided on this approach for the remaining trials. We display the segmentation and image quality findings from this experiment in Table 1. The results demonstrate that the final segmentation outputs and image quality are enhanced by our unique technique (see Fig 2 and 3), in which we replace the corrective ResNet. Fig 4 represents the findings from this investigation for two sample scenarios. The suggested method can provide photos of good quality with little difference from the source images, with enhancements, particularly at the edges. Additionally, the enhanced image quality produces higher quality segmentations that accurately and clearly distinguish the myocardium from the LV and R.V. We compare the image quality outcomes obtained by various methods to the technique we suggest in Table 2 to address the S.I. and SSIM no-reference image quality metrics. The baseline of an outlier-filled image describes the acquisition without using a repair procedure. Fig 4 shows how the image quality has improved. The image obtained without mis-triggering is comparable to the excellent image created with our framework.

5. Conclusion

We have presented a comprehensive study on automatic cardiac motion outlier detection, correction, and segmentation in an end-to-end deep learning architecture that can be applied as a global image reconstructor. To test different network topologies for outlier rectification and downstream segmentation, we first created synthetic outliers' using high-quality data. In this paper, we collectively address the detection, rectification, and segmentation of image outliers, resulting in a network architecture capable of segmenting and reconstructing images with excellent quality. To fully assess the therapeutic usefulness of our framework, we also thought about how our algorithm might be applied in a future study.

The connection between the reconstruction and segmentation tasks can be taken advantage of thanks to the training of serial networks from beginning to end, resulting in higher-quality reconstructions optimized for the subsequent segmentation task. Additionally, we explored several techniques to fix image outliers' and put forth a strategy for directly correcting low-quality photos enhanced by segmentation loss with backward propagation. It is intriguing to note that

image quality and segmentation quality performance are improved by a mixed network using a combination of detection, rectification, and segmentation losses. The accuracy with which our technique accurately identifies cases devoid of motion outliers is further shown by the fact that, on high-quality data, the recommended method had no negative effects on image/segmentation performance. Depending on their needs, the user can select a high-quality image and a highly accurate segmentation by applying weighting to each loss to build the network to do any of the three tasks more efficiently.

References

- [1] Ide, T. Miyoshi, S. Katsuragi, R. Neki, K.-I. Kurosaki, I. Shiraishi, J. Yoshimatsu, and T. Ikeda, "Prediction of postnatal arrhythmia in fetuses with cardiac rhabdomyoma," *J. Maternal-Fetal Neonatal Med.*, vol. 32, no. 15, pp. 2463–2468, Aug. 2019.
- [2] Hong, Q. Sheng, B. Dong, L. Wu, L. Chen, L. Zhao, "Automatic detection of Secundum atrial septal defect in children based on colour Doppler echocardiographic images using convolutional neural networks," *Frontiers Cardiovascular Med.*, vol. 9, Apr. 2022.
- [3] Zhang, S. Gajjala, P. Agrawal, G. H. Tison, L. A. Hallock, "Fully automated echocardiogram interpretation in clinical practice," *Circulation*, vol. 138, no. 16, pp. 1623–1635, Oct. 2018.
- [4] Vullings, "Fetal electrocardiography and deep learning for prenatal detection of congenital heart disease," in *Proc. Comput. Cardiology Conf. (CinC)*, Singapore, Dec. 2019.
- [5] Vo, T. Le, A. M. Rahmani, N. Dutt, and H. Cao, "An efficient and robust deep learning method with 1-D octave convolution to extract fetal electrocardiogram," *Sensors*, vol. 20, no. 13, p. 3757, Jul. 2020.
- [6] Torrents-Barrena, G. Piella, N. Masoller, E. Gratacós, E. Eixarch, M. Ceresa, and M. Á. G. Ballester, "Segmentation and classification in MRI and U.S. fetal imaging: Recent trends and prospects," *Med. Image Anal.*, vol. 51, pp. 61–88, Jan. 2019.
- [7] Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
- [8] Kim, J. Yun, Y. Cho, K. Shin, R. Jang, H.-J. Bae, and N. Kim, "Deep learning in medical imaging," *Neurospine*, vol. 16, no. 4, pp. 657–668, 2019.
- [9] Castiglioni, L. Rundo, M. Codari, G. Di Leo, C. Salvatore, M. Interlenghi, F. Gallivanone, A. Cozzi, N. C. D'Amico, and F. Sardanelli, "A.I. applications to medical images: From machine learning to deep learning," *Phys. Medica*, vol. 83, pp. 9–24, Mar. 2021.
- [10] Chen, Z. Liu, M. Du, and Z. Wang, "Artificial intelligence in obstetric ultrasound: An update and future applications," *Frontiers Med.*, vol. 8, p. 1431, Aug. 2021.
- [11] Sherubha, "Graph-Based Event Measurement for Analyzing Distributed Anomalies in Sensor Networks", *Sādhanā*(Springer), 45:212, <https://doi.org/10.1007/s12046-020-01451-w>.
- [12] Sherubha, "An Efficient Network Threat Detection and Classification Method using ANP-MVPS Algorithm in Wireless Sensor Networks", *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, ISSN: 2278-3075, Volume-8 Issue-11, September 2019.
- [13] Sherubha, "An Efficient Intrusion Detection and Authentication Mechanism for Detecting Clone Attack in Wireless Sensor Networks", *Journal of Advanced Research in Dynamical and Control Systems (JARDCS)*, Volume 11, issue 5, Pg No. 55-68.
- [14] Arnaout, L. Curran, Y. Zhao, J. C. Levine, E. Chinn, and A. J. MoonGrady, "An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease," *Nature Med.*, vol. 27, no. 5, pp. 882–891, May 2021.
- [15] Chen, Z. Wang, M. Du, and Z. Liu, "Artificial intelligence in the assessment of female reproductive function using ultrasound: A review," *J. Ultrasound Med.*, vol. 41, no. 6, pp. 1343–1353, Jun. 2022.