Peer Review File

Article information: https://dx.doi.org/10.21037/tp-23-394

Reviewer A

The reviewed article is an interesting report on application of artificial intelligence to automatically detect landmarks that could be used to diagnose atrioventricular canal defect on fetal echocardiography. The study is original.

The title is fully representative of the content of the article.

The abstract is concise and summarizes the essential information of the paper.

Reply: Thank you.

Reviewer B

I have the following comments:

1. Lines 67 to 68: "But due to busy work for fetal heart detection and time-consuming measurement, few examiners would like to remember one more value. "This is not an accurate statement because it takes less than 30 seconds to make the measurements of the atrial and ventricular length and obtain the ratio.

Reply 1: Thank you for the recommendation. We have modified our text as "However, due to busy work for fetal heart detection, few examiners would like to remember one more value." (See Page 4, line 78-80)

2. Line 107: Awkward sentence. "All images including the videos were stored in the database finally."

Reply 2: Thank you for the recommendation. We have modified our text as "All images were stored in the database finally." (See Page 6, line 129)

3. Figure 3. Which of the images belong to the Link-net, MA-net and U-net networks?

Reply 3: We have added the names of three models and related diseases to the picture.

(See Page 29, Figure 3)

4. Lines 138 to 148. Data Processing. The authors should describe in detail as a supplement how they accomplished the steps described in this section. This is so others can duplicate what the authors did. Were the steps to accomplish what they describe using another program such as PHOTOSHOP in which the pixels can be adjusted or were they using a program specific to each of the 3 networks.

Reply 4: I am sorry that this part of what I wrote is not clear enough. This process was realized by using Python. Firstly, all images in different pixels were standardized into 384×512 pixels. And then the patients' privacy area was removed and the final pixels for each image were 340×450. Thirdly, to unify the grayscale of all images to the same interval, the images were normalized according to the formula: $x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$, where max is the maximum value of the sample data, and min is the minimum value of the sample data.
See the changes in line 170-180.

5. Lines 150 to 158. Model Building. The authors should provide a video as a supplement of these steps for those readers who are unfamiliar with the mode building.

Reply: Thank you for the question. However, recording a video only shows the running process, which is not very helpful for understanding the details of work operation. It is better to refer to the links involved in our usage process. Our models were implemented based on the Pytorch programming language and the relevant links of the model building for three models were as follow: https://pypi.org/project/segmentation-models-pytorch/.

6. Lines 160 to 170 Training Procedure. Again, a video describing these steps would be useful for the reader who is unfamiliar with this process.

Reply: Thank you for the question. Similar to the previous question, it is better to refer to the links involved in our usage process. Regarding the RMSprop optimizer, you can refer to https://pytorch.org/docs/stable/generated/torch.optim.RMSprop.html . Regarding MSE as a loss function, you can refer to https://pytorch.org/docs/stable/generated/torch.nn.MSELoss.html.

7. The authors illustrate the usefulness of the three different types of software to identify the landmarks of the atrial and ventricular lengths to diagnose AVSD. However, these

measurements can be made at the time of the examination of the fetal heart quite easily. While the study is of interest to illustrate that the software used is interesting, it application as a resource to be implemented on ultrasound machines would be limited unless that authors could find other “measurement” applications that could be used to measure the end-diastolic images of the 4-chamber view that have been proposed as screening tools for detecting the risk for other malformations such as tetralogy-of Fallot and D-transposition of the great arteries. See references below.

1. Abnormalities of the Width of the Four-Chamber View and the Area, Length, and Width of the Ventricles to Identify Fetuses at High-Risk for D-Transposition of the Great Arteries and Tetralogy of Fallot
J Ultrasound Med. 2023 Feb;42(3):637-646. doi: 10.1002/jum.16060.Epub 2022 Jul 13.
Authors
Greggory R DeVore 1 2 3 , Bettina Cuneo 4 , Mark Sklansky 5 , Gary Satou 5
Affiliations
1 Fetal Diagnostic Centers of Pasadena, Tarzana, and Lancaster, Los Angeles, California, USA.
2 Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, David Geffen School of Medicine, UCLA, Los Angeles, California, USA.
3 Department of Obstetrics and Gynecology, Wayne State University, Detroit, Michigan, USA.
4 The Heart Institute and the Colorado Fetal Care Center, Departments of Pediatrics and Obstetrics, Children's Hospital Colorado, University of Colorado School of Medicine, Aurora, Colorado, USA.
5 Division of Pediatric Cardiology, Department of Pediatrics, UCLA Mattel Children's Hospital, David Geffen School of Medicine at UCLA, Los Angeles, California, USA.
PMID: 35822424
DOI: 10.1002/jum.16060
Abstract
Objectives: The prenatal detection of D-Transposition of the great arteries (D-TGA) and tetralogy of Fallot (TOF) has been reported to be less than 50% to as high as 77% when adding the outflow tracts to the four-chamber screening protocol. Because many examiners still struggle with the outflow tract examination, this study evaluated whether changes in the size and shape of the heart in the 4CV as well as the ventricles occurred in fetuses with D-TGA and TOF could be used to screen for these malformations.
Methods: Forty-four fetuses with the pre-and post-natal diagnosis of D-TGA and 44 with TOF were evaluated between 19 and 36 weeks of gestation in which the 4CV was imaged. Measurements of the end-diastolic width, length, area, and global sphericity

index were measured for the four-chamber view and the right and left ventricles. Using z-score computed values, logistic regression was performed between the 88 study and 200 control fetuses using the hierarchical forward selection protocol.
Results: Logistic regression identified 10 variables that correctly classified 83/88 of fetuses with TOF and TGA, for a sensitivity of 94%. Six of 200 normal controls were incorrectly classified for a false-positive rate of 3%. The area under the receiver operator classification curve was 98.1%. The true positive rate for D-TGA was 93.2%, with a false-negative rate to 6.8%. The true positive rate for TOF was 95.5%, with a false negative rate of 4.5%.
Conclusions: Measurements of the 4CV and of the RV and LV may help identify fetuses at risk for D-TGA or TOF.
Keywords: D-Transposition; congenital heart defect; fetal echocardiography; four-chamber screening; logistic regression; speckle tracking; tetralogy of Fallot.

2. Evaluation of Fetal Cardiac Size and Shape: A New Screening Tool to Identify Fetuses at Risk for Tetralogy of Fallot
J Ultrasound Med. 2021 Dec;40(12):2537-2548. doi: 10.1002/jum.15639.Epub 2021 Jan 27.
Authors
Greggory R DeVore 1 , Gary M Satou 2 , Yalda Afshar 1 , Danielle Harake 2 , Mark Sklansky 2
Affiliations
1 Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, David Geffen School of Medicine, UCLA, California, Los Angeles, USA.
2 Division of Pediatric Cardiology, Department of Pediatrics, UCLA Mattel Children's Hospital, David Geffen School of Medicine at UCLA, California, Los Angeles, USA.
PMID: 33502041
DOI: 10.1002/jum.15639
Abstract
Objective: Prenatal detection rates for tetralogy of Fallot (TOF) vary between 23 and 85.7%, in part because of the absence of significant structural abnormalities of the 4-chamber view (4CV), as well as the relative difficulty in detection of abnormalities during the screening examination of the outflow tracts. The purpose of this study was to evaluate whether the 4CV and ventricles in fetuses with TOF may be characterized by abnormalities of size and shape of these structures.

Methods: This study retrospectively evaluated 44 fetuses with the postnatal diagnosis of TOF. Measurements were made from the 4CV (end-diastolic length, width, area, global sphericity index, and cardiac axis) and the right (RV) and left (LV) ventricles (area, length, 24-segment transverse widths, sphericity index, and RV/LV ratios). Logistic regression analysis was performed to identify variables that might separate fetuses with TOF from normal controls.
Results: The mean gestational age at the time of the last examination prior to delivery was 28 weeks 5 days (SD 4 weeks, 4 days). The mean z-scores were significantly lower in fetuses with TOF for the 4CV and RV and LV measurements of size and shape. Logistic regression analysis identified simple linear measurements of the 4CV, RV, and LV that had a sensitivity of 90.9 and specificity of 98.5% that outperformed the 4CV cardiac axis (sensitivity of 22.7%) as a screening tool for TOF.
Conclusions: Measurements of the 4CV, RV, and LV can be used as an adjunct to the outflow tract screening examination to identify fetuses with TOF.
Keywords: Tetralogy of Fallot; fetal echocardiography; fetal heart; prenatal diagnosis; speckle tracking.

Reply: Thank you for your question. 1. This study was a continuation of our team's previous research. In previous study, although speckle tracking technology was not utilized, we analyzed 113 quantitative data of fetal echocardiography and finally selected AVLR in end diastolic period as the critical parameter for prenatal diagnosis of AVSD. Relevant literature could be found in: https://doi.org/10.32604/chd.2023.029060.

2. The neural network models used in this study could not be implemented on ultrasound machines so far, and the application was limited indeed. However, On the one hand, the main contribution of our study was the discovery that applying landmark detection technology to measure cardiovascular structures could achieve precise measurements. Compared to segmentation technology, landmark detection technology applied landmarks labeling which was not only shorten the doctors' time in annotation process, but also measures cardiac structures accurately on fetal echocardiography. On the other hand, we will put multi-center images into the model and use the AVLR cutoff value to classify the images to screen out AVSD cases in the next study.

Reviewer C

The article is of interest and well written in English

My comments are:

Line 80: Please underline the importance of prenatal detection of congenital heart disease and its consequences on Central Nervous system DOI: 10.3390/jcm11226740

Reply: Thank you for the recommendations. Congenital cardiovascular defects affect the outcomes of newborns or infants and the quality of life of individuals and families. In addition, Vena and the colleagues show that obstructive lesions of the left-side and right-side cardiac system are associated with fetal neurodevelopmental abnormalities. Therefore, it is important to make prenatal detection of congenital heart disease, which could stratify the risk of the disease, provide a basis for management and decision-making prenatally as well as reasonable treatment and intervention postnatally, improve the treatment effects, reduce the mortality, and improve the quality of life of children.

Add the details in line 92-100.

Line 223: Please cite how also other articles measured e atrial length (AL) to ventricular length (VL) ratio for detection of other congenital heart defects. It is possible to use artificial intelligence also in these cases? DOI: 10.1002/uog.22008

Reply: 1. Thank you for the recommendations. But I am sorry, we don't know where to cite the literature in this sentence.

2. Yes, the neural networks could be used for prenatal screening of the normal fetuses, fetuses with suspected coarctation of the aortic arch but normal after birth, and fetuses diagnosed with aortic coarctation finally mentioned in the article. However, this is relevant with our next research, so we regret that we could not go into detail here.

Reviewer D

Congratulations on your article. Although data is still limited in the scientific literature for the application of Neural Networks, works like yours will allow you to add more and more data on new approaches to prenatal diagnosis and especially to fetal echocardiography.