Deep Learning based

by Bambang Tutuko

Submission date: 02-Jan-2023 10:03AM (UTC+0700)
Submission ID: 1987849346
File name: Deep learning-based real time detection.pdf (7.94M)
Word count: 4928
Character count: 26178

Informatics in Medicine Unlocked 36 (2023) 101150

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Deep learning-based real time detection for cardiac objects with fetal ultrasound video

Ade Iriani Sapitri [a,b], Siti Nurmaini [a,*], Muhammad Naufal Rachmatullah [a], Bambang Tutuko [a], Annisa Darmawahyuni [a], Firdaus Firdaus [a], Dian Palupi Rini [c], Anggun Islami [a]

[a] Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya, Palembang, 30139, Indonesia
[b] Doctoral Program, Faculty of Engineering, Universitas Sriwijaya, Palembang, 30139, Indonesia
[c] Department of Informatics Engineering, Faculty of Computer Science, Universitas Sriwijaya, Palembang, 30139, Indonesia

ARTICLE INFO

Keywords:
Computer vision
YOLOv7
Object detection
Fetal echocardiography

ABSTRACT

Fetal cardiac anatomical structure interpretation by ultrasound (US) is a key part of prenatal assessment. Unfortunately, the numerous speckles in US video, the small size of fetal cardiac structures, and unfixed fetal positions make manual detection processes difficult. To alleviate such problems, a deep learning model was developed to fully automate the processing of fetal echocardiographs with 2-dimensional cross-sectional images. The weakness of such an approach is that a physician immediately interprets US video with 3-dimensional objects to diagnose patients in clinical practice. To improve fetal cardiac anatomical structure interpretation via US for accurate and real time diagnosis, this paper proposes a real-time fetal cardiac substructure detection using US video with the You Only Look Once (YOLO) framework. In YOLO, an end-to-end neural network makes predictions for cardiac substructure objects, boxes, and class probabilities simultaneously. To achieve reliable performance, 40 fetal echocardiography videos were trained with the new YOLOv7 architecture and fine-tuned to work optimally and run efficiently. We conducted with nine fetal cardiac substructure objects such as the left atrium, right atrium, left ventricle, right ventricle, tricuspid valve, pulmonary valve, mitral valve, aortic valve, and aorta. The results yielded the highest mean average precision of 82.10%, reaching 17 frames per second (FPS) for nine cardiac substructure objects in 0.3 ms. The main finding of our study is that with even a limited number of US videos, YOLOv7 can detect fetal cardiac substructure objects in real-time. Such a network can work efficiently to detect small fetal cardiac objects automatically in a rapid phase with the help of proper fine-tuning. This work mainly assists medical experts in the fetal cardiac anatomy diagnostic process.

1. Introduction

Automated cardiac video ultrasound (US) interpretation has the potential to transform clinical practice in multiple ways, including enabling serial assessment of cardiac function by nonexperts in primary care and rural settings [1]. With such an approach, for example, US interpretation in utero can allow early prenatal detection of most congenital heart diseases [2]. US interpretation varies at different stages of pregnancy and may evolve in utero with advancing gestational age [3]. Fetal US is a key aspect of prenatal health assessment across pregnancies worldwide due to its noninvasive nature, relatively low cost, and portability. As such, different aspects of the physiological assessment of a developing fetus rely upon an understanding of fetal US videos [4]. For fetal cardiac malformation to be detected accurately via US, the whole cardiac substructure should be recognized in normal anatomy. The normal cardiac substructure anatomy is divided into (i) the four main chambers, the left atrium (LA), right atrium (RA), left ventricle (LV), right ventricle (RV); (ii) the four valves, the tricuspid (TV), pulmonary (PV), mitral (MV), and aortic (AV); and (iii) one aorta (Ao) that distributes blood rich in oxygen throughout the body [1,2,5].

The process requires specific training based on theory and practical tests to manually produce accurate and precise cardiac substructure interpretation through US video. One of the greatest limits related to the US involves interpersonal variability, meaning that it depends on the skills of the examining physician and the patient's condition [6]. Moreover, analyzing these nine fetal cardiac substructures is extremely challenging due to several key factors, such as numerous speckles in US, the small size of the fetal heart, unfixed fetal positions, and category

* Corresponding author.
E-mail address: siti_nurmaini@unsri.ac.id (S. Nurmaini).

https://doi.org/10.1016/j.imu.2022.101150
Received 25 November 2022; Received in revised form 5 December 2022; Accepted 10 December 2022
Available online 12 December 2022
2352-9148/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

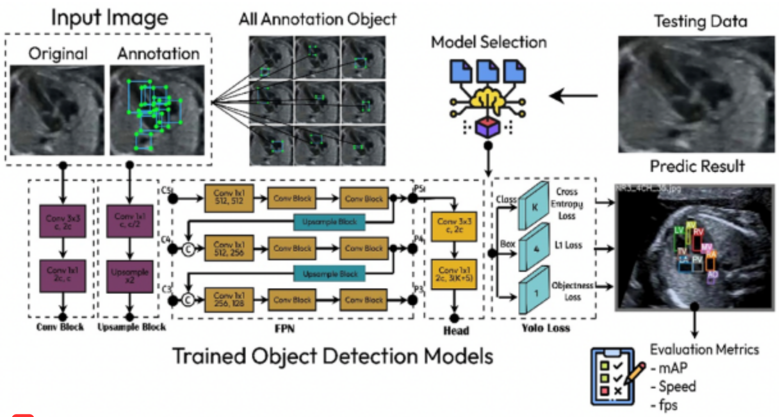


Fig. 1. The general methodology for real time detection of fetal cardiac substructure using US video.

indistinction caused by the similarity of cardiac chambers [7,8]. To relieve such problems, a computer-aided method that helps physicians automatically locate fetal cardiac objects has attracted much attention in recent years [8,9]. Such procedures can assist physicians in automatically diagnosing fetal cardiac structures and can contribute significantly to the early diagnosis of congenital heart diseases. In addition, a strategy with automated fetal US interpretation could involve handheld devices used by nonexperts at point-of-care locations (eg, primary care clinics).

Computer-aided deep learning (DL) method has recently been widely used in US imaging for fetal echocardiography analysis due to its strong ability to learn invariant features [10–17]. DL can employ devices to imitate the human cognitive process, such as learning, applying, and solving complex problems [4]. Convolutional neural networks (CNNs), as representative architectures of a DL model, have strong feature power and learn robust and discriminative features from medical images. In a recent work by Zhang et al. [1], researchers employed a DL model to fully automate the processing of echocardiographs, including disease detection, image segmentation, and structure and function quantification. In another work on fetal cardiac images by Arnaout et al. [18], they used an ensemble of neural networks to classify complex congenital heart disease with satisfactory results. However, the experiment was conducted on 2-dimensional cross-sectional cardiac US images, whereas in clinical practice, physicians immediately use US video with 3-dimensional objects to diagnose patients.

Feature information in 2D US image sequences can be weak and obscured due to speckling, acoustic shadowing and other artefacts [13], thereby reducing the accuracy of the detection process. To improve the detection rate, multiobject detection in US videos requires superior methods to integrate spatial and temporal linkages of salient features, which is the nature of simultaneously localizing anatomical structures [13]. Fetal cardiac object detection based on US video in real time would greatly facilitate analysis, leading to proper diagnoses [19–21]. Real-time detection and localization in freehand US was proposed with a DL model, but 2D fetal standard scan plane images were used [22]. One of the real-time object detection frameworks proposed based on You Only Look Once (YOLO) is a state-of-the-art framework [21]. A YOLO network directly predicts the location and class probabilities of objects, completely achieving end-to-end optimization of detection performance. However, there are three main drawbacks of such a YOLO network that make it highly challenging in terms of increasing accuracy and precision: (i) the accuracy of YOLO is not as high as that of Fast R–CNN [16] and that of Faster R–CNN [17]; (ii) the networks have difficulty detecting and localizing small objects in images that appear in groups [21,23]; and (iii) fetal US video with lower contrast and brightness and with speckle noise make YOLO computationally [23]. This can lead to overfitting, difficulty accurately localizing the fetal cardiac structure, and poor performance in the detection task due to the size of the fetal cardiac in utero is approximately 4–8 cm (the size will be smaller for cardiac substructure) with low quality video.

Hence, improvements must be made to the YOLO network model according to the characteristics of the fetal US videos dataset. The developed model is expected to achieve satisfactory performance. Motivated by the limitations of the state of the art in this area, the contributions of our work are as follows.

- A model for real-time US video detection of nine fetal cardiac substructure objects in normal anatomy is designed.
- The object detection performance of fetal echocardiography is improved with the YOLOv7 architecture
- The proposed model is validated and evaluated with six YOLO frameworks.

2. Material and method

The general methodology of our experiment is shown in Fig. 1. There are three main steps: image acquisition and labelling, object detection using the proposed algorithm with YOLOv7, and result validation using four metrics, mean average precision, frames per second, confusion matrix and a precision–recall curve.

2.1. Image acquisition and fetal cardiac annotation

Collecting a well-defined dataset is key to research on real-time fetal cardiac detection using US video. In our study, 60 US videos of normal structures were obtained from patients undergoing routine pregnancy tests in their second trimester at the obstetrics and genecology department of Mohammad Hoesin General Hospital, Indonesia. All US videos were identified with two experienced experts with waived consent for compliance. The inclusion criterion was fetuses of 18–24-weeks gestational age [15]. Specifically, the patient scans were recorded using a GE Voluson E6 instrument with a loop length of 2–20 s, and a file size of approximately 890 kB to 36.9 MB was used to generate the US images. Detection of fetal cardiac substructure objects is challenging for the model, possibly because of the smaller size of the fetal heart or the decrease in characteristic features [8–10]. The fetal cardiac normal anatomy was determined by reviewing the clinical report for each US video, and the experts annotated the selected fetal US for learning purposes. All US images were stored in the original Digital Imaging and Communication in Medicine (DICOM) format. Since US images do not

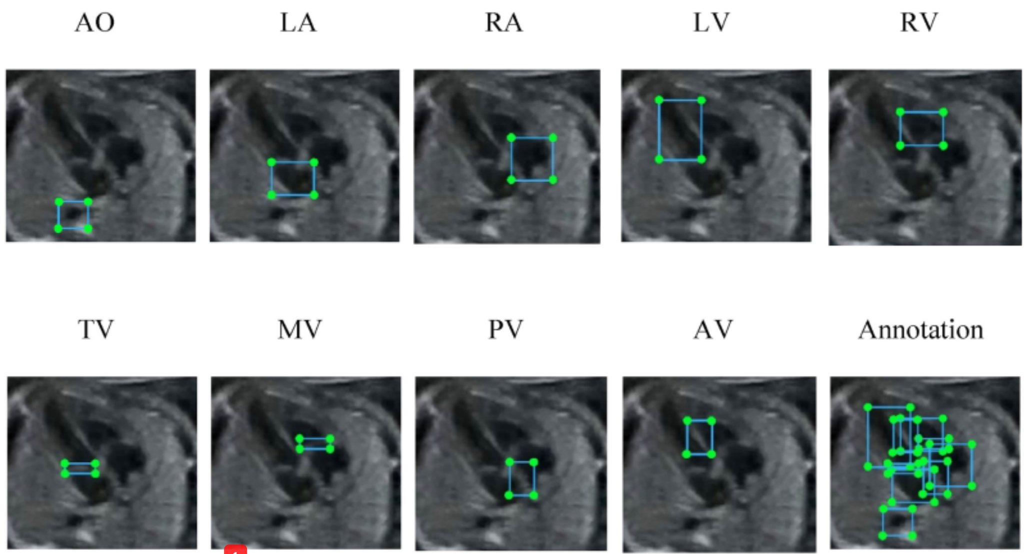


Fig. 2. Ground truth box of nine fetal cardiac substructure objects by expert.

Table 1
The YOLOv7 network structure.

	Type	Filter	Size	Output
Input	Layers	128	52 × 52	128 × 128
Backbone	Conv2D 32 × 3 x 3	32	416 × 416	128 × 128
	Residual Block 1 × 64	64	208 × 208	
	Residual Block 2 × 128	128	104 × 104	64 × 64
	Residual Block 8 × 256	256	52 × 52	
	Residual Block 8 × 512	512	26 × 26	64 × 64
	Residual Block 4 × 1024	1024	13 × 13	32 × 32
FPN	Conv2D 5 L	384	52 × 52	16 × 16
	Concat	384	52 × 52	
	Conv2D + Up2D	128	52 × 52	8 × 8
	Conv2D 5 L	1024	13 × 13	
Loss Function	Conv2D 3 × 3 + Con2D 1 × 1	255	52 × 52	8 × 8
	Conv2D 3 × 3 + Con2D 1 × 1	255	26 × 26	
	Conv2D 3 × 3 + Con2D 1 × 1	255	13 × 13	
	Avg Pool		Global	–
	Connected		1000	–
	Softmax		–	–

convey color information, the images were stored as grayscale bitmaps. Sample fetal cardiac annotations are depicted in Fig. 2. The anotated process for nine fetal cardiac substrutures is conducted object by object utilizing LabelMe software. All the data after the annotation process were named the ground truth box.

2.2. Proposed YOLOv7 for fetal cardiac substructure detection

In this work, we mainly explore and evaluate different YOLO architectures with different model training parameter values in the detection of nine substrutures in fetal cardiac anatomy. These YOLO architectures learn from the annotated set, which has major advantages over more traditional approaches that use hand-crafted features [21]. This model takes 2D images as input, making its adaptation to 3D medical images the first challenge. Every US scan slice was used as a single image for inference and training. Once the bounding boxes were found on every 2D image, 3D generalization of the non-maximum suppression algorithm was performed. This post-processing step groups the 2D boxes with threshold criteria corresponding to their intersection over union (IoU) to generate 3D bounding boxes. This study focused on the YOLOv7 framework, a new architecture for small objects with added instance segmentation to the YOLO arch (Fig. 1.) [24]. Many transformer backbones and arches were included,, as well as multitask training. As a simple and standard training framework for any detection and instance segmentation task, the main difference between YOLOv7 compares other architectures is the backbone of YOLOv7 is based on DarkNet53 with extended efficient layer aggregation network (E-ELAN) (Table 1). The E-ELAN structure use group convolution to increase the cardinality of the added features, and combine the features of different groups in a shuffle and merge cardinality manner. Such operation can enhance the features learned by different feature maps and improve the use of parameters and calculations [24].

Our previous works on fetal echocardiography analysis were based on instance segmentation; however, only 2D echocardiograms were used [8,16]. We wanted to upscale our model from frame (2D) to video (3D) to expand the real-time fetal echocardiography detection system. This paper is the first to study real-time fetal cardiac substructure detection using the YOLOv7 framework. All the networks were implemented and evaluated by the Pytorch 1.7.1 library and trained using a computer with the following specifications: an Intel® Core™ i9-9920X CPU @ 3.50 GHz with 490191 MB of RAM, a GeForce 2080 RTX Ti graphics card by NVIDIA Corporation GV102 (rev a1) and the Ubuntu 18.04.5 LTS operating system.

2.3. Model evaluation

Our model was validated based on the mean average precision (mAP) value. The AP is calculated by a formula based on the following submetrics: a confusion matrix (CM), the IoU, the recall (R), and the precision (P). The average precision, taken over all classes, is called the mean average precision (mAP) [8,21]. Another metric is the frames per second (FPS) value, which measures how fast our object detection model processes US video and generates the desired output [21]. The performances of YOLOv5n, YOLOv5s, YOLOv6n, YOLOv6s, YOLOv7, and YOLOv7-tiny were measured based on the mAP value. Detections were considered true positives or false positives based on the area of overlap with ground-truth bounding boxes. The mAP is a popular metric used to measure localization accuracy and calculate localization errors in object detection models. The mAP value should exceed 50% for measuring decision performance, where mAP is the overlap area between the

Table 2
Performance of the six YOLO models.

Object	mAP Performance (%)					
	v5n	v5s	v6s	v6n	v7	v7tiny
LA	88.50	89.50	56.78	67.56	90.40	85.70
RA	85.90	86.20	66.72	52.23	92.20	86.20
LV	91.90	88.30	73.56	43.21	94.80	81.20
RV	84.40	75.50	89.90	77.81	88.30	78.00
AO	89.80	88.60	90.08	83.51	87.40	80.30
TV	66.50	72.00	89.23	71.12	66.80	64.20
MV	69.70	66.70	78.98	66.45	69.20	51.11
PV	59.40	63.70	56.23	80.01	71.20	62.00
AV	64.60	76.00	66.54	86.32	78.40	63.90
Avrg. mAP	77.70	78.50	77.82	67.95	82.10	72.50

Table 3
Comparison of model evaluation.

Model	Size	FPS	mAP^{VAL} (%)	$Speed^{v100}$Fp 16 b32 (ms)	$Speed^{v100}$Fp32 b32 (ms)
YOLOv5n	640	15	77.70	1.40	1.30
YOLOv5s	640	25	78.50	1.50	1.10
YOLOv6s	640	17	77.82	0.38	1.10
YOLOv6n	640	18	67.95	0.33	1.07
YOLOv7	640	17	82.10	0.30	0.80
YOLOv7-tiny	640	17	72.50	0.30	0.70

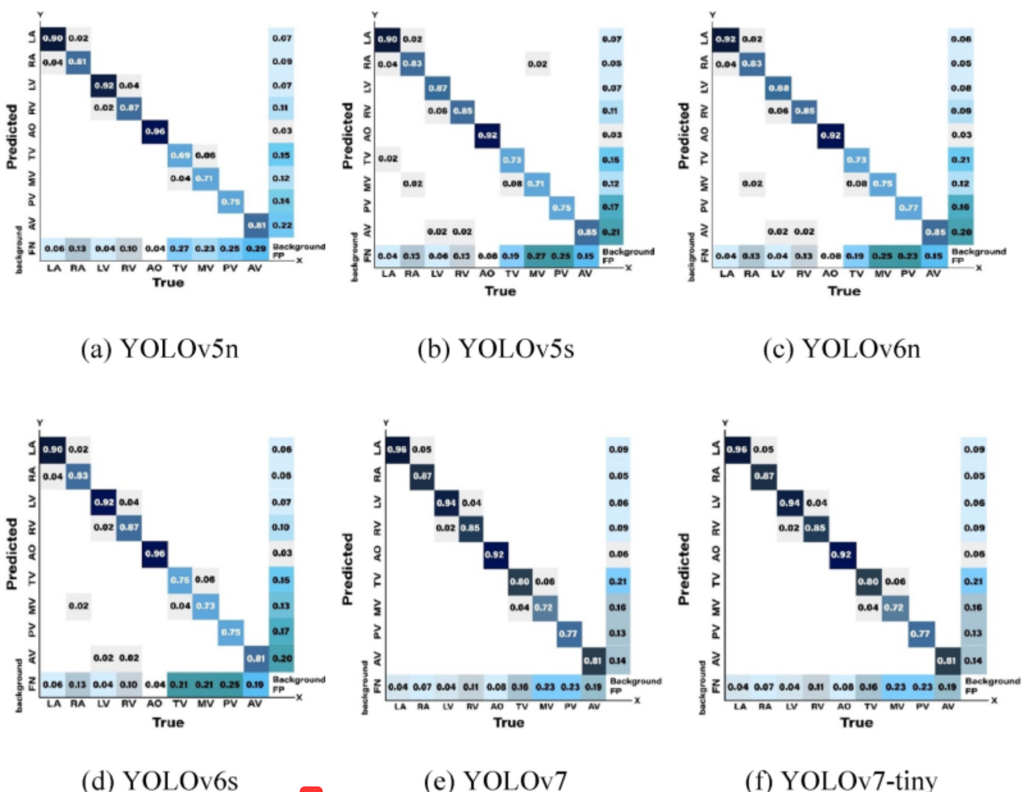
predicted bounding box and the ground-truth bounding box. An mAP value $>$ 0.5 was considered a positive prediction, while an mAP value $<$ 0.5 was considered a negative prediction.

3. Results and discussion

We developed a DL method to detect nine fetal cardiac substructure objects in normal US anatomy for accurate localization. Our approach is the first to focus on a YOLOv7 framework, which has not been used for fetal cardiac substructure detection. In a previous study, one possible reason for the delayed use of YOLO is its low accuracy in detecting small objects [21,23]. However, our model succeeded in detecting nine fetal cardiac objects, even though the fetal cardiac size is only approximately 4–8 cm in the 21st to 28th week of gestation. Table 2 shows the object detection performance achieved with YOLOv5n, YOLOv5s, YOLOv6n, YOLOv6s, YOLOv7, and YOLOv7-tiny architectures. YOLOv7 produced satisfactory performance, with an mAP of 82.10% and 17 FPS for 0.3 ms. Our propose real-time object detection model achieved fewer trainable parameters and fast processing and was accurate (Table 3).

In this work, the mAP threshold was set at 50%, meaning that the overlap between the ground-truth bounding box and the detected box returns a 0.5 score. The higher the score is, the more accurate the model detection. As the result shows, the proposed object detection model produced LA, RA, LV, RV, AO, TV, MV, PV, and AV performances with an mAP exceeding 50%. From the six YOLO models, YOLOv7 reached the best value for nine objects, with an mAP of 82.10%. In addition, the YOLOv7 evaluation model produced high FPS, mAPVAL, Speedv100 Fp 16 b32, and Speedv1 00 Fp 32 b32 scores (Table 3). This metric indicates the performance of YOLOv7 in terms of the object detection processing time for all frames in the validation model.

The CM for all YOLO networks is shown in Fig. 3. The IoU values reflects the resulting matches between the ground truth and detected box. For each ground truth box, the YOLO algorithm generates an IoU with every detected box. A match is found if both boxes have an IoU



(a) YOLOv5n (b) YOLOv5s (c) YOLOv6n (d) YOLOv6s (e) YOLOv7 (f) YOLOv7-tiny

Fig. 3. The confusion matrix for six YOLO architectures.

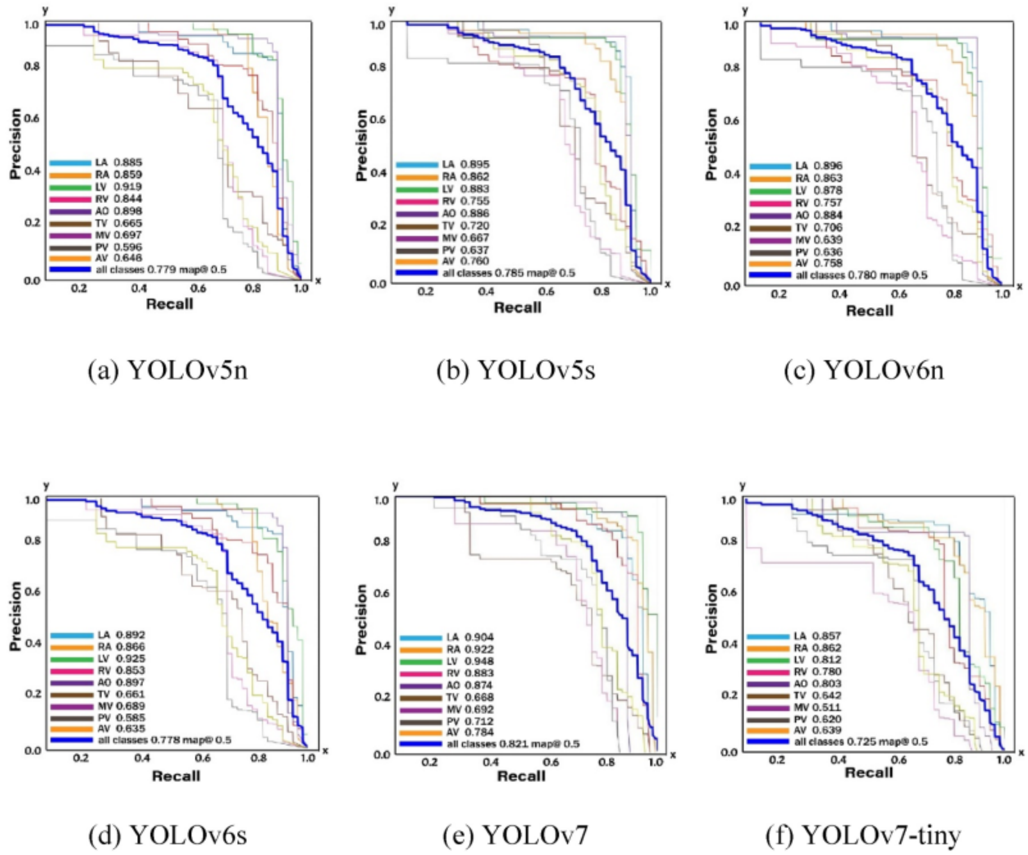


Fig. 4. PR curve for six YOLO architectures.

greater than or equal to 0.5. Only detections with a IoU score greater than or equal to 0.5 are considered, and anything under this value is discarded. YOLOv7 achieved a higher IoU value for each class, indicating that it outperformed other YOLO networks (Fig. 3). This means that our proposed model can detect nine classes of fetal cardiac object with an IoU over 0.5 with satisfactory results: the IoU values are 0.96, 0.87, 0.94, 0.85, 0.92, 0.80, 0.72, 0.77, and 0.81 for the LA, RA, LV, RV, Ao, TV, MV, PV and AV, respectively. Objects that are part of the ground truth but were not detected are counted in the last column of the matrix (in the row corresponding to the ground-truth class) as false positives (FPs). Objects that are detected but are not part of the confusion matrix are counted in the last row of the matrix (in the column corresponding to the detected class) as false negatives (FNs).

The YOLO algorithm should have both high precision (P), and high recall (R). These two metrics are shown in the PR curve. A PR curve plots the precision against recall for different confidence threshold values, but most algorithms often involve a trade-off between the two. A good precisison–recall (PR) curve has a higher area under curve (AUC). Based on the experiment, YOLOv7 achieves achieve high precision and recall, an average AUC of 0.821 with mAP of 0.5, and the AUC of each class has a higher value outperforms the other YOLO models (Fig. 4).

The sample of detection results for the nine fetal cardiac substructures based on the ground truth and detected bounding boxes as depicted in Fig. 5. The mAP for the four cardiac chamber objects (LA, RA, RV, and LV) reached over 88%; however, the mAP for the four valve objects (TV, MV, PV, and AV) only reached over 66%. This occurs because the cardiac valve can open and close, changing the shape of the heart. It is important to adjust the size of the region of interest (RoI) to ensure that the whole valve can be detected. However, the overall performance of YOLOv7 reached mAP values above 50% for all detected objects, even though all the objects were extremely small. This indicates that the proposed YOLO model is not problematic for small object sizes, which was a weakness of the YOLO model in previous studies [21,23].

A sample of the validation results based on the proposed YOLOv7 model obtained using a dataset of nine fetal cardiac substructure objects from patients with normal cardiac anatomy as depicted in Fig. 6. The bounding boxes detected with YOLOv5 to YOLOv7-tiny have mAP values. The fetal cardiac score threshold was set to 50% (0.5) for the mAP. Nine fetal cardiac substructure (LA, RA, RV, LV, TV, MV, PV, AV, and AO) were detected by the proposed model using such a threshold. Although the ground truth had a speculated perimeter, the proposed YOLOv7 model accurately defined the boundary of the nine fetal cardiac objects. The proposed architecture successfully detected each object with a high object score of approximately 91%; in contrast, false-positive detection was associated with a much lower object score of approximately 30%.

The benchmark performance of our proposed YOLOv7 model compared to the current state-of-the-art method for detecting fetal cardiac substructures in US videos is depicted in Table 4. Unfortunately, the number of previous studies are limited, especially for real-time detection with YOLO in fetal echocardiography; thus, comparisons were only made with two experiments by Qiao et al. [25] and Komatsu et al. [26]. As a result, the proposed YOLOv7 achieved the best performance among all the models proposed by Refs. [25,26]. This YOLOv7 network

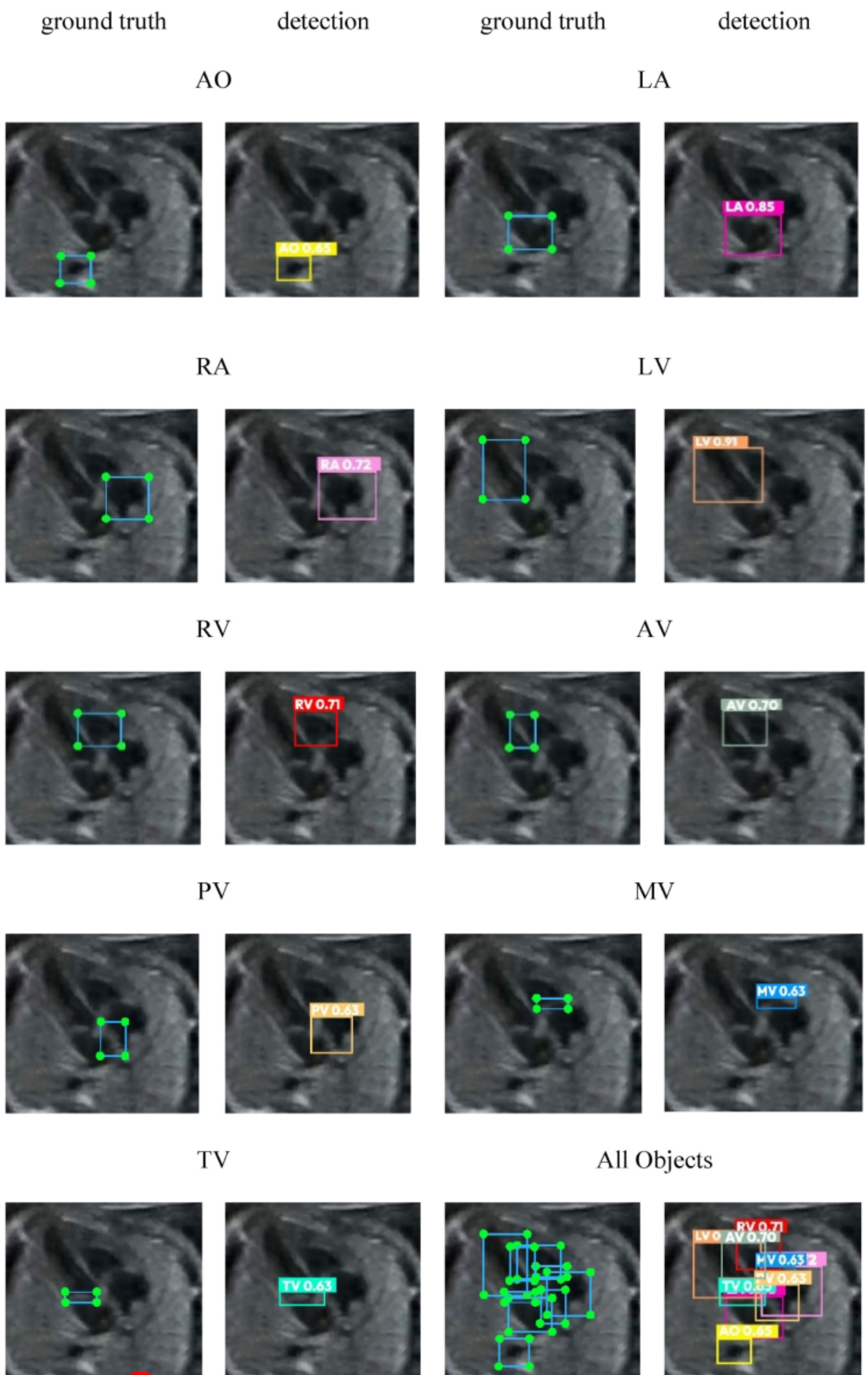


Fig. 5. The localization results. Ground truth and detection box of nine different anatomical cardiac substructure.

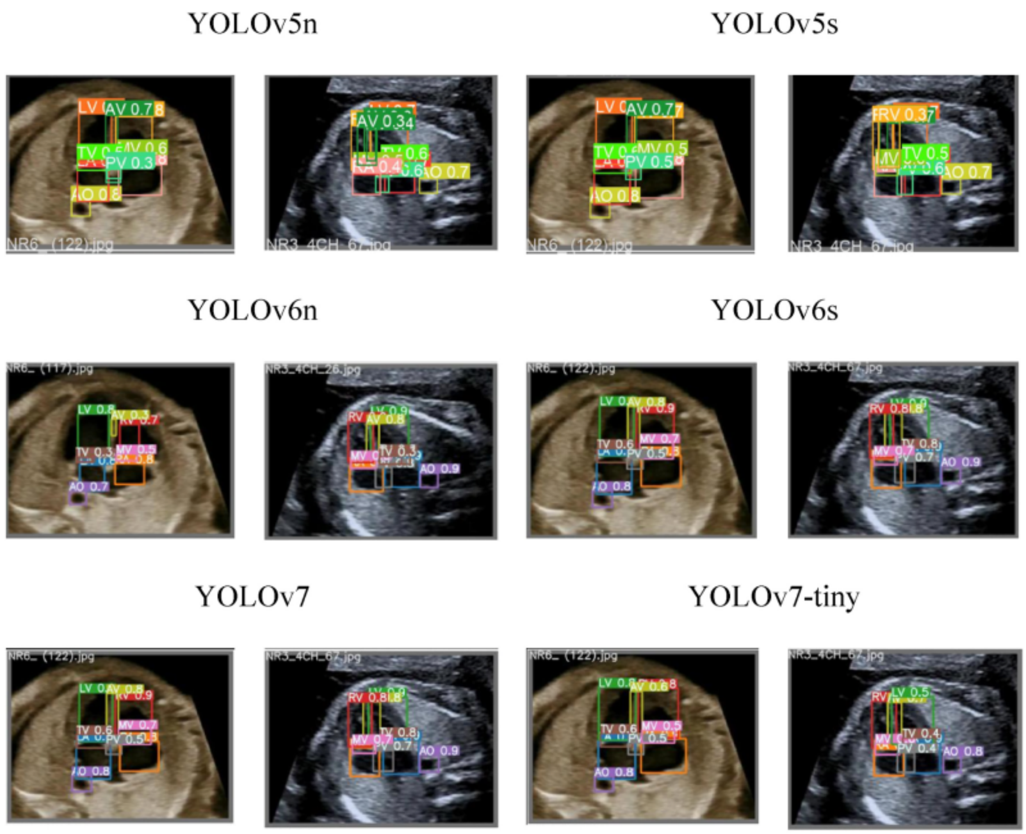


Fig. 6. The sample of real time detection for nine fetal cardiac substructure objects using US video with YOLOv5n, YOLOv5s, YOLOv6n, YOLOv6s, YOLOv7, and YOLOv7-tiny.

Table 4
Comparison with the state-of-the-art methods in fetal cardiac substructure detection.

Model	mAP	FPS	Number of objects
YOLOv4Slim [25]	0.75	80	4 cardiac substructure
YOLOv4 [25]	0.82	74	
CBAM-YOLOv4Slim [25]	0.90	30	
MRHAM-YOLOv4Slim [25]	0.91	43	
SONO-YOLOv2 [26]	0.685	–	18 cardiac substructure
Proposes YOLOv7	0.92	30	4 cardiac substructure
	0.82	17	9 cardiac substructure

demonstrates the effectiveness of our proposed model in detecting nine cardiac substructures. As illustrated in Table 4, the proposed model achieves a state-of-the-art mAP of 92% for four cardiac substructures and 82% for nine cardiac substructures while maintaining a comparable testing speed. In detecting the fetal cardiac substructures, the challenge is recognizing the valve due to the fetal cardiac chambers being small and suffering from unfixed positions and category indistinction caused by the similarity of cardiac chambers. However, the proposed model was able to localize each object of the fetal cardiac anatomy. The DL-based YOLOv7 architecture has proven to be a state-of-the-art for real time object detection system that supports this process. Such a method can assess the quality of fetal cardiac substructure detection based on the relative size of the region of interest and the key anatomical structures within it.

Despite the promise of these study, the limitations of our approach must be noted. This approach is limited as it was not tested on US videos of fetal cardiac anatomy with abnormalities. A validation model based on more realistic fetal cardiac US videos (i.e., videos with extra noise, morphological variations, and normal and abnormal structures) will be essential to fully understand the rationale and robustness of the proposed model. In addition, YOLOv7 added errors to fetal cardiac object localization, and heavily pathological fetal cardiac slices were not classified properly; therefore, the bounding boxes generated were not perfectly aligned with the objects of interest.

4. Conclusion

An automated echocardiographic interpretation approach can potentially enable the democratization of health care. Studies taking place earlier in a disease course and in geographic areas with limited specialized expertise should be facilitated. In this paper, we propose a YOLOv7 model to accurately locate nine important fetal cardiac substructures, the RA, LA, RV, LV, TV, PM, MV, AV and Ao in a real-time. The YOLOv7 model showed promising results for fetal cardiac detection, suggesting that its promise in fetal cardiac organ recognition, even with few US scans as a training set. The validation results demonstrate the ability of the model to generalize to a broad range of fetal cardiac anatomies. The detection of nine important fetal cardiac substructures is the first step in congenital heart disease research. With promising fetal cardiac detection precision, the proposed model has excellent potential for clinical application. In addition to clinical use, such a method could also facilitate research and discovery by standardizing and accelerating analysis of the millions of echocardiograms archived within medical systems.

Funding

This work was funded by "Hibah Doktor" the Ministry of education, culture, research and technology Indonesia, No. 0145.004/UN9.3.1/PL/2022 and under Professional Grants by Universitas Sriwijaya.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the Intelligent System Research Group (ISysRG) Universitas Sriwijaya, Indonesia, for their full support of the research infrastructure.

References

[1] Zhang J, et al. Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy. Circulation 2018;138(16):1623–35.
[2] Van Nisselrooij AEL, et al. Why are congenital heart defects being missed? Ultrasound Obstet \& Gynecol 2020;55(6):747–57.
[3] Yagel S, et al. Congenital heart defects: natural course and in utero development. Circulation 1997;96(2):550–5.
[4] Shen Y-T, Chen L, Yue W-W, Xu H-X. Artificial intelligence in ultrasound. Eur J Radiol 2021;139:109717.
[5] Picazo-Angelin B, Zabala-Argüelles JI, Anderson RH, Sánchez-Quintana D. Anatomy of the normal fetal heart: the basis for understanding fetal echocardiography. Ann Pediatr Cardiol 2018;11(2):164.
[6] Boccatonda A. Emergency ultrasound: is it time for artificial intelligence? Journal of Clinical Medicine 2022;11(13):3823. MDPI.
[7] Pinheiro DO, et al. Accuracy of prenatal diagnosis of congenital cardiac malformations. Rev Bras Ginecol e Obs 2019;41:11–6.
[8] Nurmaini S, et al. Deep learning-based computer-aided fetal echocardiography: application to heart standard view segmentation for congenital heart defects detection. Sensors 2021;21(23):8007.
[9] Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view classification of echocardiograms using deep learning. npj Digit Med 2018;1(1):6. https://doi.org/10.1038/s41746-017-0013-1.
[10] Sundaresan V, Bridge CP, Ioannou C, Noble JA. Automated characterization of the fetal heart in ultrasound images using fully convolutional neural networks. In: Biomedical imaging (ISBI 2017), 2017 IEEE 14th international symposium on; 2017. p. 671–4. https://doi.org/10.1109/ISBI.2017.7950609.
[11] Hsu W-Y. Automatic left ventricle recognition, segmentation and tracking in cardiac ultrasound image sequences. IEEE Access 2019;7:140524–33.
[12] Cameli M, Pastore MC, Henein MY, Mondillo S. The left atrium and the right ventricle: two supporting chambers to the failing left ventricle. Heart Fail Rev 2019;24(5):661–9.
[13] Patra A, Noble JA. Multi-anatomy localization in fetal echocardiography videos. In: 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019); 2019. p. 1761–4.
[14] Xu L, et al. DW-Net: a cascaded convolutional neural network for apical four-chamber view segmentation in fetal echocardiography. Comput Med Imaging Graph 2020;80:101690.
[15] Leclerc S, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. IEEE Trans Med Imaging 2019;38(9):2198–210.
[16] Nurmaini S, et al. Accurate detection of septal defects with fetal ultrasonography images using deep learning-based multiclass instance segmentation. IEEE Access 2020;8:196160–74. https://doi.org/10.1109/ACCESS.2020.3034367.
[17] Nurmaini S, et al. An improved semantic segmentation with region proposal network for cardiac defect interpretation. Neural Comput Appl 2022:1–14.
[18] Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view classification of echocardiograms using deep learning. npj Digit Med 2018;1(1):1–8. https://doi.org/10.1038/s41746-017-0013-1.
[19] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 779–88.
[20] Yang Y, et al. Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China. medRxiv 2020. https://doi.org/10.1101/2020.02.10.20021675.
[21] Du Z, Yin J, Yang J. Expanding receptive field yolo for small object detection. Journal of Physics: Conference Series 2019;1314(1):12202.
[22] Baumgartner CF, et al. SonoNet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. IEEE Trans Med Imaging 2017;36(11):2204–15. https://doi.org/10.1109/TMI.2017.2712367.
[23] Liu C, Hu S-C, Wang C, Lafata K, Yin F-F. Automatic detection of pulmonary nodules on CT images with YOLOv3: development and evaluation using simulated and patient data. Quant Imaging Med Surg 2020;10(10):1917.
[24] Wang C-Y, Bochkovskiy A, Liao H-YM. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv Prepr arXiv220702696 2022. https://doi.org/10.48550/arXiv.2207.02696.
[25] Qiao S, et al. Automatic detection of cardiac chambers using an attention-based YOLOv4 framework from four-chamber view of fetal echocardiography. arXiv Prepr arXiv201113096 2020. https://doi.org/10.48550/arXiv.2011.13096.
[26] Komatsu M, et al. Detection of cardiac structural abnormalities in fetal ultrasound videos using deep learning. Appl Sci 2021;11(1):371.

Deep Learning based

ORIGINALITY REPORT

89%	31%	88%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Ade Iriani Sapitri, Siti Nurmaini, Muhammad Naufal Rachmatullah, Bambang Tutuko et al. "Deep learning-based real time detection for cardiac objects with fetal ultrasound video", Informatics in Medicine Unlocked, 2023 Publication	87%
2	Submitted to Sriwijaya University Student Paper	1%
3	eprints.ncl.ac.uk Internet Source	1%
4	Yi-Qi Huang, Jia-Chun Zheng, Shi-Dan Sun, Cheng-Fu Yang, Jing Liu. "Optimized YOLOv3 Algorithm and Its Application in Traffic Flow Detections", Applied Sciences, 2020 Publication	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		