iScience

Article

Computer-aided assessment for enlarged fetal heart with deep learning model

Highlights

- Enables early detection of fetal heart abnormalities during prenatal screenings
- Proposes a deep learning model for automating fetal heart enlargement assessment
- Introduces a custom segmentation model that integrates YOLO with a CBAM

Authors

Siti Nurmaini, Ade Iriani Sapitri, Muhammad Taufik Roseno, ..., Anggun Islami, Akhiar Wista Arum, Rio Bastian

Correspondence

sitinurmaini@gmail.com

In brief

Machine learning, artificial intelligence, and Cardiovascular medicine

Nurmaini et al., 2025, iScience *28*, 112288
May 16, 2025
https://doi.org/10.1016/j.isci.2025.112288

CellPress

iScience

CellPress
OPEN ACCESS

Article

Computer-aided assessment for enlarged fetal heart with deep learning model

Siti Nurmaini,[1,5,*] Ade Iriani Sapitri,[1] Muhammad Taufik Roseno,[2] Muhammad Naufal Rachmatullah,[1] Putri Mirani,[3] Nuswil Bernolian,[4] Annisa Darmawahyuni,[1] Bambang Tutuko,[1] Firdaus Firdaus,[1] Anggun Islami,[1] Akhiar Wista Arum,[1] and Rio Bastian[1]

[1]Intelligent System Research Group, Universitas Sriwijaya, Palembang, Indonesia
[2]Computer Science Department, Universitas Sumatera Selatan, Palembang, Indonesia
[3]Department of Obstetrics and Gynecology, Fetomaternal Division, Bunda Hospital, Palembang, Indonesia
[4]Department of Obstetrics and Gynecology, Fetomaternal Division, Dr. Mohammad Hoesin General Hospital, Palembang, Indonesia
[5]Lead contact
*Correspondence: sitinurmaini@gmail.com
https://doi.org/10.1016/j.isci.2025.112288

SUMMARY

Enlarged fetal heart conditions may indicate congenital heart diseases or other complications, making early detection through prenatal ultrasound essential. However, manual assessments by sonographers are often subjective, time-consuming, and inconsistent. This paper proposes a deep learning approach using the You Only Look Once (YOLO) architecture to automate fetal heart enlargement assessment. Using a set of ultrasound videos, YOLOv8 with a CBAM module demonstrated superior performance compared to YOLOv11 with self-attention. Incorporating the ResNeXtBlock—a residual network with cardinality—additionally enhanced accuracy and prediction consistency. The model exhibits strong capability in detecting fetal heart enlargement, offering a reliable computer-aided tool for sonographers during prenatal screenings. Further validation is required to confirm its clinical applicability. By improving early and accurate detection, this approach has the potential to enhance prenatal care, facilitate timely interventions, and contribute to better neonatal health outcomes.

INTRODUCTION

According to the data from the Global Burden of Disease Study 1990–2017, the prevalence of congenital heart disease (CHD) in Southeast Asia reached 1.6%–1.7%.[1] A study conducted at Bhumibol Adulyadej Hospital in Thailand found that among 145 diabetic pregnant women, the prevalence of fetal myocardial hypertrophy was 79.3%, highlighting a significant association between maternal diabetes and fetal heart enlargement.[1] Malaysia reported a birth prevalence of 6.7 per 1,000 live births, with a significant proportion of severe cases, indicating a rising trend in detecting heart anomalies.[2] Indonesia has a population growth rate of 1.39%, with 4.2–4.8 million babies born each year. Based on this, it is estimated that 71,400 to 81,600 babies are born with CHD in Indonesia annually.[3,4] Despite this, the cases of enlarged fetal heart remain limited.

Fetal heart enlargement, also known as fetal cardiomegaly, refers to an abnormal increase in the size of the fetal heart. This condition can arise from various causes, including congenital cardiac anomalies, twin-pregnancy-related complications, fetal dilated cardiomyopathies, abnormal shunting, fetal arteriovenous malformations, and placental chorioangioma.[2,4] The exact prevalence of fetal cardiomegaly is not well documented, as it is a relatively rare condition. It is often identified during routine prenatal ultrasounds, and its prevalence may vary depending on the population and the availability of prenatal care.[1] The enlargement typically involves the right-sided chambers, with the right atrium being involved in approximately 90% of cases.[3] Fetal heart enlargement can be diagnosed through prenatal ultrasound and echocardiography, which help identify the underlying causes and assess the severity of the conditions.[5–7] Early detection and intervention can significantly improve the prognosis for the fetus and potentially prevent complications during birth and beyond.[4]

Diagnosing an enlarged fetal heart typically involves measuring the size of the heart relative to the size of the chest on an ultrasound. Two ways of measuring the fetal heart condition are cardiothoracic ratio (CTR) and "cardiothoracic (C/T)" circumference ratio.[8] The CTR and the C/T circumference ratio both assess the relationship between the fetal heart and the chest size, but they measure different parameters.[7] The fetal CTR measures the ratio of the transverse diameter of the heart to the transverse diameter of the chest, whereas the fetal C/T circumference ratio measures the ratio of the circumference of the heart to the circumference of the chest.[7,9,10] However, the ratio of the cardiac circumference to the thoracic circumference is more easily measured on fetal ultrasound/echocardiography.[10] Classically, the C/T circumference ratio should be always <0.5 throughout gestation.[11] It slowly increases throughout gestation, such as ~11 weeks: 0.38 and ~17–20 weeks: 0.45.[11]

Measuring the fetal heart manually using the C/T circumference ratio is a particularly challenging task due to several factors[12,13]: (1) the fetal heart has intricate anatomy with small and closely packed structures, making it difficult to identify and trace the boundaries accurately; (2) shadows cast by the fetal sternum can obscure parts of the heart and chest, leading to incomplete or inaccurate measurements of the circumferences; (3) ultrasound images often contain speckle noise and artifacts, which can distort the appearance of anatomical structures and make it challenging to delineate the heart and chest accurately; (4) the fetus can move during the ultrasound examination, causing blurring and making it difficult to capture a clear and static image for measurement; and (5) the quality of ultrasound images can vary depending on the equipment used, the skill of the sonographer, and the specific conditions of the scan, affecting the accuracy of manual measurements. Hence, developing an automated system to accurately measure the fetal heart size is highly desirable.

Deep learning (DL) has been applied in the field of fetal echocardiography for various tasks, including standard plane identification from a sequence of fetal heart images,[14,15] detection of abnormal structures,[16,17] segmentation of cardiac structures,[18–21] and detection of prenatal CHD.[22] In particular, on heart enlargement conditions, the previous DL studies have proposed to support the clinical decision using ultrasound medical images.[23–25] However, all previously developed DL-based methods have been specifically designed for detecting abnormal conditions in adult cardiac structures. A DL model for assessing fetal heart enlargement has yet to be explored.

Recent research demonstrates that image processing analysis with attention module mechanisms can significantly enhance the performance of DL-based segmentation and detection models by focusing on relevant patterns within objects.[26–28] This is achieved by improving the feature extraction capabilities of convolutional networks, particularly on important features of small objects within large datasets.[28,29] One such attention module is the Convolution Block Attention Module (CBAM),[30,31] which indicates good performance to enhance important features in the complex background of medical imaging and is suitable for fetal heart ultrasound.[32,33] However, attention mechanisms like CBAM generally perform better with large datasets. Medical imaging often faces data limitations due to privacy constraints and labeling costs, which might restrict CBAM's potential if the dataset lacks diversity. To improve feature learning by facilitating gradient flow and enhancing feature extraction, the residual blocks specifically improve CBAM in some applications. Residual connections allow the network to learn residual mappings rather than directly trying to map inputs to outputs, making it easier to capture relevant features. CBAM, when combined with feature learning, can focus more on essential features and ignore irrelevant ones.[34–37] This study develops a custom architecture that combines the object-detection-based segmentation model and CBAM with residual block to improve model performance. By inserting CBAM with residual block into strategic locations within the YOLO network, we can enhance the ability to focus on relevant features, potentially improving ultrasound image detection performance. The contribution of our study is as follows:

(1) Developing a DL-based segmentation model for automatic identification of enlarged fetal heart.
(2) Proposing a custom architecture for fetal heart segmentation with an attention module with residual block to improve feature extraction.
(3) Evaluating the proposed model with unseen ultrasound images.

Related study

Prenatal ultrasound diagnosis of the fetal heart can assist in making clinical decisions and improve neonatal outcomes.[23] However, when identifying complex abnormal fetal heart anatomy,[6] detecting and localizing lesions precisely is difficult and time-consuming due to the activity of the fetus, the faster heart beating, the smaller heart size than adults, and the high requirement for expertise.[12,19,20,22] Moreover, in countries or regions lacking well-established healthcare systems like Indonesia, advanced echocardiographic equipment, and experienced technicians or specialists, prenatal CHD has a high rate of missed diagnosis, which can lead to delayed treatment and a poorer prognosis. The combination of artificial intelligence (AI) and traditional ultrasound is expected to alleviate the abovementioned problems.[38–41] AI techniques have made significant progress in assessing cardiac structure and function. DL, one of the AI approaches, has made remarkable progress in the field of medical image analysis because of its powerful ability to autonomously learn image features.[27,40,41] Such a method has been applied in the field of fetal echocardiography for various tasks.[21,23,24,29]

In recent years, DL techniques have made substantial progress in the assessment of fetal heart structure and function. To address the low prenatal detection rate of CHDs, a multi-scale gated axial-transformer network has been proposed to extract robust and discriminative representations from ultrasound images.[38] Comprehensive experimental results demonstrate that the proposed model effectively recognizes fetal CHD, achieving a precision of 95.92%, a recall of 94%, an accuracy of 95%, and an F1 score of 94.95% on the test set.[38] An intelligent feature-learning detection system with a multistage residual hybrid attention module has been developed to detect the four chambers of the heart. This system achieves a precision of 0.919, a recall of 0.971, an F1 score of 0.944, a mean average precision (mAP) of 0.953, and operates at 43 frames per second.[39] A stacked residual dense network model has been presented to generate an automatic echocardiographic interpretation by using segmentation of the entire region of the cardiac and classifying the defect positions. The classification of defect positions using unseen data had approximately 92.27% accuracy, 94.33% specificity, and 92.05% sensitivity.[40] An ensemble of neural networks based on supervised learning has been proposed to explore and identify specific cardiac views. Such a model used a separate model to differentiate between structurally normal hearts and complex CHDs. They were achieved with

Table 1. The performance results of the HC and CC instance segmentation for fetal heart

			Performance results (%)				
Set	Metrics	Class	YOLOv8	YOLOv8.2	YOLOv8 and CBAM	YOLOv9	YOLOv11
Validation	mAP	All	95.50	98.45	99.50	98.01	99.50
		HC	92.20	97.40	99.50	98.50	99.50
		CC	93.50	99.50	99.50	97.50	99.50
	IoU	All	67.53	73.38	80.52	68.41	72.22
		HC	66.03	74.24	78.69	66.31	66.69
		CC	69.04	72.53	82.36	70.51	77.76
Testing	mAP	All	67.70	70.40	84.10	68.41	88.10
		HC	81.70	98.88	89.50	82.80	76.80
		CC	53.70	41.90	78.70	74.20	99.50
	IoU	All	49.31	58.64	64.14	25.93	52.66
		HC	56.58	64.08	68.98	24.20	56.00
		CC	42.04	53.20	53.30	27.66	47.88

an area under the curve of 0.99, sensitivity of 95%, and specificity of 96%, which is comparable to expert human clinician performance.[41] As an alternative approach, Sulas et al. have investigated the use of neural networks to automatically interpret pulsed-wave Doppler traces on the left ventricular inflow and outflow. They postulate that this might be an alternative method of quantifying cardiac time intervals, given the difficulties in obtaining fetal electrocardiographic data, although the utility of this technique in identifying fetal disease has not been tested.[42]

Detecting CHDs using DL has shown significant progress, driven by the relatively higher prevalence of these conditions and the availability of extensive datasets for model training. In contrast, identifying fetal cardiomegaly poses distinct challenges due to its rarity and the variability in its presentation.[43] The limited number of cases means that DL models designed specifically for detecting cardiomegaly often lack sufficient data to achieve high performance and generalizability. Most studies have concentrated on heart segmentation in adult cases of cardiomegaly, but few have focused on cardiomegaly in fetuses. Fetal cardiomegaly is an exceptionally rare condition with a relatively low prevalence, as it is generally associated with specific underlying health issues. Although comprehensive population data on fetal cardiomegaly is scarce, estimates indicate that it affects approximately 0.05%–0.1% of pregnancies, varying by geographic location and diagnostic criteria. Undetected fetal cardiomegaly can lead to significant health complications for the child post-birth, highlighting the need for a thorough investigation. Therefore, an in-depth study of fetal heart enlargement is essential to enhance detection and improve outcomes.

RESULTS

YOLO with CBAM

This study has experimented with the YOLOv8, YOLOv8.2, YOLOv9, and YOLOv11 to make the result analysis comprehensive.[44] The YOLO algorithm excels in instance segmentation, which entails identifying and outlining individual objects within an image. There are 356, 48, and 24 fetal heart ultrasound images for training, validation, and testing (unseen) sets. The results, including the predictions generated by each model, their corresponding ground truths, and prediction labels, are displayed in Table 1. In the validation set for all classes, YOLOv8 with CBAM performed better than the existing models in terms of mAP—with scores of 95.50% (YOLOv8), 98.45% (YOLOv8.2), and 98.01% (YOLOv9). Also, YOLOv8 with CBAM performed better than the existing models in terms of IoU—with scores of 67.53%

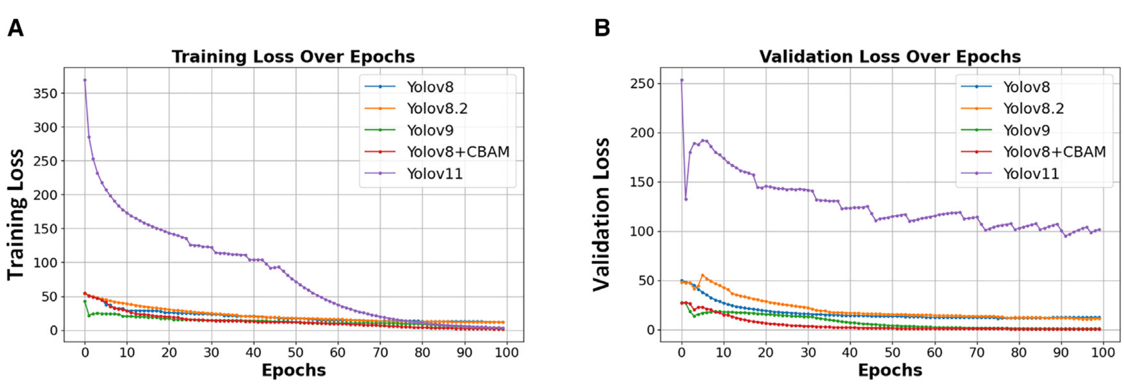

Figure 1. The training and validation loss of YOLOv8, YOLOv8.2, YOLOv8 with CBAM, YOLOv9, and YOLOv11

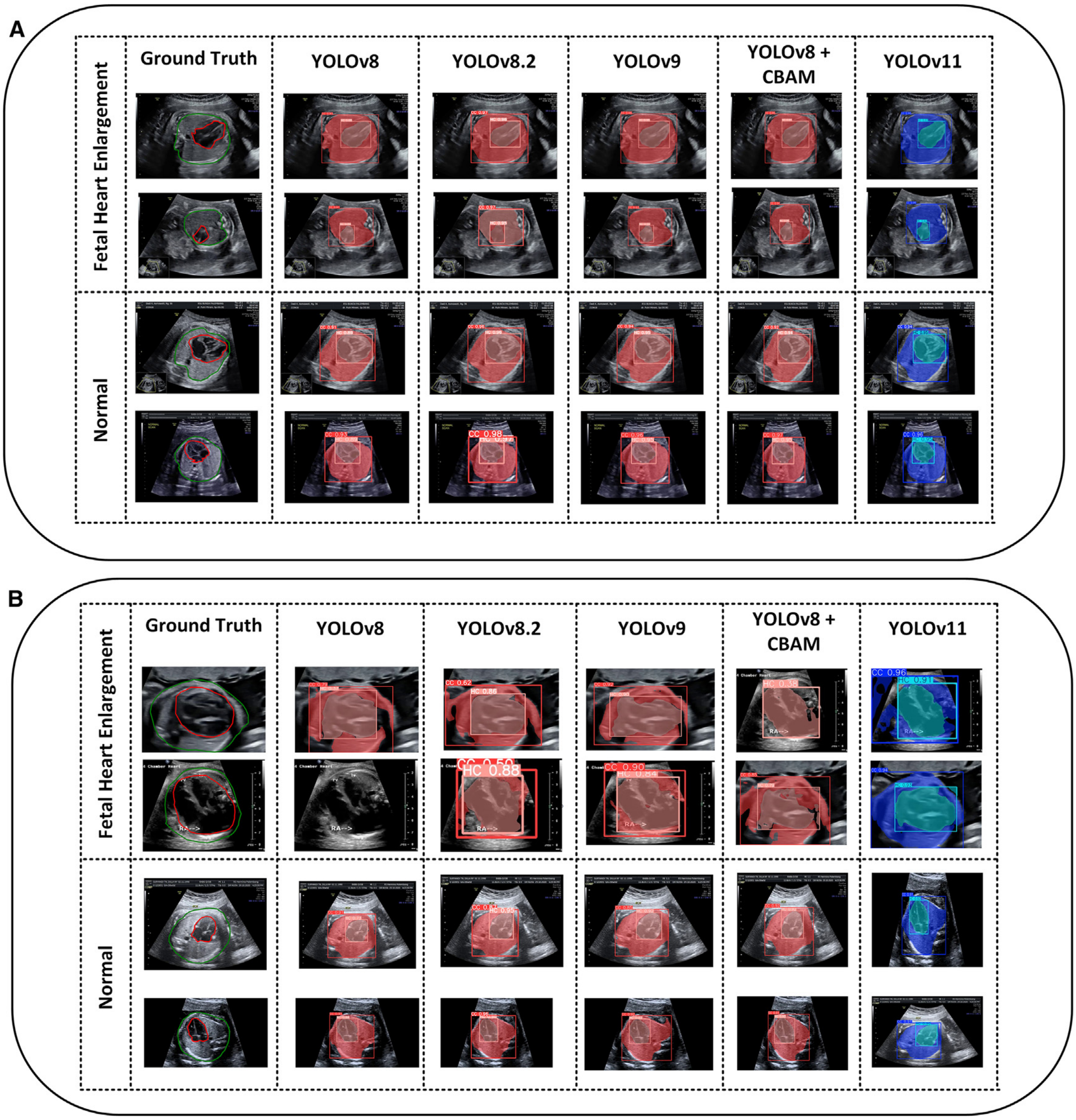

Figure 2. Representative examples of the HC and CC segmentation by the existing models
(A and B) validation set and (C and D) testing set.

(YOLOv8), 73.38% (YOLOv8.2), and 68.41% (YOLOv9). In addition, for the testing set, YOLOv8 with CBAM has also outperformed in terms of mAP and IoU. YOLOv8 with CBAM integrates channel and spatial attention mechanisms to efficiently highlight important features in fetal heart ultrasound images while minimizing unnecessary noise. This combined attention approach significantly improves detection accuracy and efficiency, particularly in intricate and changing ultrasonography, making CBAM a crucial component in advanced image-based fetal heart enlargement detection.

To evaluate its performance, the proposed model was also tested against YOLOv11. YOLOv11 is the latest series of real-time object detectors that enhances feature extraction capabilities for more precise object detection and complex task performance. Such model incorporates a self-attention mechanism known as the Cross-Stage Partial with Spatial Attention (C2PSA). This mechanism is designed to enhance feature representation by selectively focusing on spatially significant regions within the input data, improving the model's ability to detect and localize objects with higher precision. As shown in Table 1, although both models achieved an equivalent mAP of 99.50% during training, the proposed model's mAP was 4% lower during testing. However, YOLOv8 with CBAM outperformed YOLOv11 in IoU during both validation and testing, achieving 8% and 12% higher scores, respectively. These results indicate that the proposed model, with its higher IoU, produces segmentation results with greater accuracy in

Algorithm 1. Medical_Rules_Fetal_Heart_Enlargement_Condition

```
Input: chest_cavity_segment, heart_circle_segment
Output: "Fetal Heart Enlargement" or "Normal"
BEGIN
  # Function to calculate IoU between two bounding boxes function
calculate_IoU(box1, box2):
  x1_min, y1_min, x1_max, y1_max ← box1
  x2_min, y2_min, x2_max, y2_max ← box2
  x_inter_min ← max(x1_min, x2_min)
  y_inter_min ← max(y1_min, y2_min)
  x_inter_max ← min(x1_max, x2_max)
  y_inter_max ← min(y1_max, y2_max)
  inter_width ← max(0, x_inter_max - x_inter_min)
  inter_height ← max(0, y_inter_max - y_inter_min)
  inter_area ← inter_width * inter_height
  box1_area ← (x1_max - x1_min) * (y1_max - y1_min)
  box2_area ← (x2_max - x2_min) * (y2_max - y2_min)
  union_area ← box1_area + box2_area - inter_area
  IoU ← inter_area/union_area
  return IoU
  # Calculate IoU between CC and HC
  IOU ← calculate_IoU(chest_cavity_box, heart_circle_box)
  # Check if the IoU is greater than 0.3
  if IOU > 0.3 then
  return "Fetal Heart Enlargement"
  else
  return "Normal"
END
```

delineating object boundaries, leading to more precise and reliable predictions.

Figure 1 presented the learning curves for the training and validation loss in YOLOv8, YOLOv8.2, and YOLOv8 with CBAM, YOLOv9, and YOLOv11 models. A learning curve was derived from the training dataset, providing insight into the model's learning progress, and from a hold-out validation dataset, offering insight into the model's generalization capability. As seen in Figure 1, a training and validation loss for YOLOv8 with CBAM model decreases to a point of stability, with a minimal gap between the two final loss values. The performance of the proposed model changes gradually and consistently as it goes through training. Such model on validation loss curve shows a smooth learning curve that suggests the model is learning in a stable and consistent manner. The results of YOLOv8, YOLOv8.2, and YOLOv8 with CBAM, YOLOv9, and YOLOv11, along with the corresponding ground truths and prediction labels, are shown in Figure 2, for validation and testing set, respectively. The red circle of ground truth represents the head circumference (HC). YOLOv8 with CBAM mostly successfully detect the fetal heart enlargement, where IoU value between chest circumference (CC) and HC is more than 0.3. Analysis of instance segmentation results on fetal heart ultrasound images is effective for YOLOv8 with CBAM in detecting fetal heart enlargement.

To measure fetal heart enlargement, an IoU value can be used. By using Algorithm 1, the normal and abnormal condition is decided. The normal condition is achieved, if the centroids of the CC and HC are far apart; it generally indicates that the heart is appropriately sized and positioned within the chest cavity. Although the heart enlargement condition is decided, if the centroids of the CC and HC are close together, it suggests that the heart is enlarged, potentially indicating swelling or other abnormalities. This method provides a quantitative way to assess the relative size and positioning of the fetal heart within the chest cavity, helping to detect signs of fetal heart enlargement. It can be observed in Figure 3 that IoU of YOLOv8 with CBAM (Algorithm 2) reaches 0.33 and the distance between CC and HC centroid is close to around 13.81, and the other IoU is about 0.38 and the distance between CC and HC centroid is close to around 19.91. It means the fetal heart is enlarged, because the HC is close to the CC. Similarly, YOLOv8 with CBAM, which has an IoU value of 0.14 and a distance of approximately 60.92 between the CC and HC centroids, indicates that the fetal heart is normal. This is because the ratio involving HC is small, leading to a greater distance between the two centroids.

Some examples of frames that failed detection can be seen in Figure 4, clearly indicating that the decisions made by the proposed model were incorrect. Ultrasound, the primary imaging modality for fetal assessment, has limitations in resolution and image quality.[15,17] Factors like the position of the fetus, maternal body habitus, and the quality of the ultrasound machine can affect the clarity of the images.[6,45] Additionally, measuring the exact size of the heart on ultrasound can be challenging due to the small movement of the fetus.[46] Fetal heart enlargement can be a nonspecific finding and may be associated with a wide range of underlying conditions, such as fetal anemia, heart defects, or infections.[47,48] This makes it harder to detect and interpret as a singular finding, as it requires correlation with other clinical and ultrasound findings. The segmentation process on unseen data evaluates the proposed model's prediction accuracy for each frame, ensuring its ability to generalize beyond the training dataset. By incorporating the CBAM, the YOLOv8 model demonstrates enhanced performance on unseen inputs. Without CBAM, YOLOv8 correctly predicts 12 frames and incorrectly predicts 12, achieving an accuracy of 53.84%. With CBAM integration, the model correctly predicts 16 frames, increasing accuracy to 66.67% (Table 2).

To validate our approach, we also tested the latest YOLOv9 and YOLOv11 models. As shown in Table 2, YOLOv8 with CBAM outperformed these models in prediction accuracy and reliability. Specifically, YOLOv8 with CBAM achieved superior IoU, resulting in only eight false negatives (FNs) compared to 12 errors for both YOLOv9 and YOLOv11. The higher IoU enabled the model to more accurately delineate object boundaries, providing precise heart size measurements. Minimizing the FN rate is critical in detecting fetal heart enlargement, as missed diagnoses pose significant risks. The reduced FN rate of our proposed model reflects its robust capability to reliably identify cases of heart enlargement, thus improving diagnostic confidence. The results demonstrate that our model effectively generalizes to unseen data, learning new patterns in fetal heart images and offering reliable performance for clinical applications.

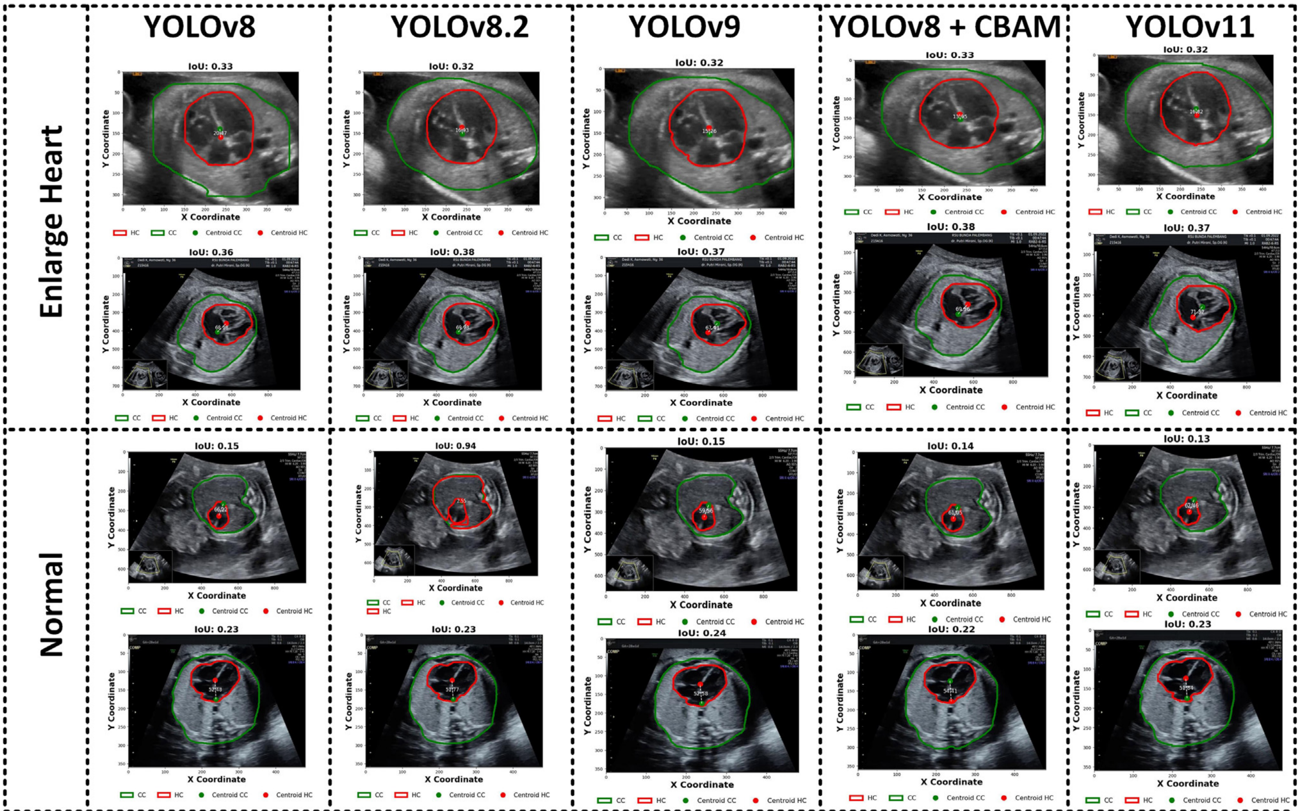

Figure 3. The visualization used to assist in making decisions about an enlarged fetal heart

YOLOv8 with CBAM demonstrates an improved ability to predict the fetal heart's enlargement, even in cases where YOLOv8, YOLOv8.2, YOLOv9, and YOLOv11 fail (Figure 5). However, under normal conditions, all models are successful in detection. CBAM enhances YOLOv8 by focusing on crucial channels and spatial regions, leading to improved detection accuracy. By concentrating on essential spatial features, CBAM also aids the model in better localizing objects, particularly useful for detecting small or overlapping objects. The difficulty in recognizing the fetal heart may be due to its small size and the presence of a complex, dark background.[15,17]

Improved CBAM with residual network block

In this study, we use the attention mechanism because it can be easily incorporated into YOLO architectures by wrapping convolutional blocks, allowing the model to learn which parts of the input are more relevant to the given task. However, this model produces errors in some images, which can reduce the performance of the proposed model (Table 2). To enhance the performance of the attention mechanism, a residual block (ResBlock) architecture is integrated with CBAM.[44,49,50] Now, the attention mechanism has a ResBlock. This helps prevent gradient vanishing and improves information flow in deeper networks.[44] This hybrid block leverages the residual structure while incorporating attention maps to refine feature representations, allowing the model to selectively enhance relevant parts of the input and suppress irrelevant details.[49] The custom architecture of our proposed model is named ResNeXt. The CBAM block structure in Figure 6 is replaced with ResNeXt block with the structure as seen in Figure 7 in YOLO backbone.

Algorithm 2. Convolutional Block Attention Module (CBAM)

Input:
- x: feature map (tensor) extracted from the input image
- channels: number of channels in the feature map x
- r: reduction ratio for the channel attention module

Output: "**Feature map with applied channel-wise and spatial attention**"

BEGIN

```
    function CBAM(x, channels, r)
  # Initialize modules
    cam = ChannelAttention(channels, r)
    sam = SpatialAttention()
  # Channel attention
    channel_attention_map = cam(x)
    x = x * channel_attention_map
  # Spatial attention
    spatial_attention_map = sam(x)
    x = x * spatial_attention_map
  return x
  end function
```

END

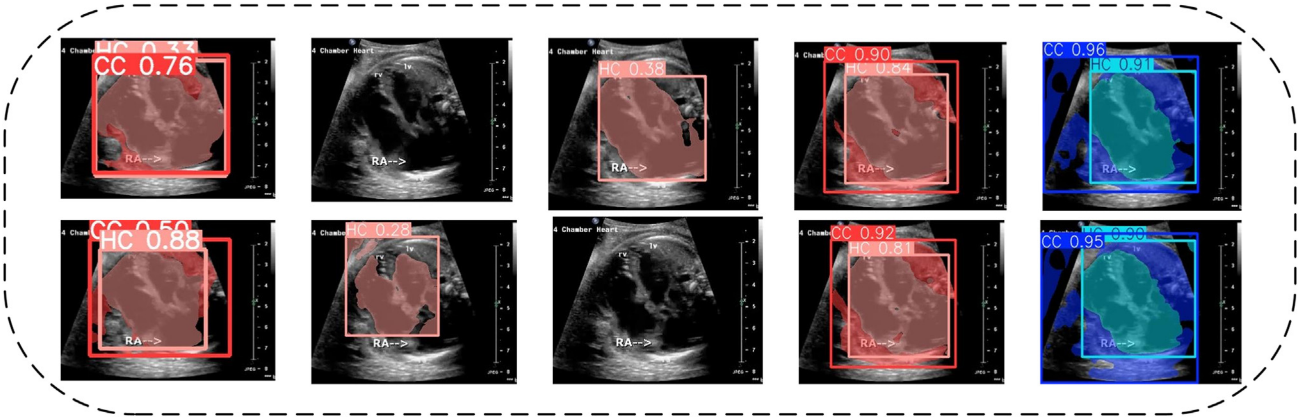

Figure 4. The sample frames with wrong decision

ResNeXt is an extension of the Residual Network architecture that introduces the concept of cardinality. In ResNeXt, features are processed by dividing them into multiple parallel network paths, each undergoing a transformation process through identical layer configurations but with distinct parameters. These paths are then recombined through summation.[31] This cardinal mechanism enhances learning capacity without substantially increasing computational complexity.[32] Each ResNeXt block incorporates three convolutional processes with different kernel sizes, and the outputs from these paths are aggregated and then added back to the input feature via a shortcut connection, similar to the residual mechanism in ResNet. This approach allows efficient data processing, leading to improved generalization capabilities.[32] By using ResNeXtBlock and CBAM, enhancing each block's capacity for selective feature attention, and improving the overall model performance. The results indicate significant improvements with the use of unseen data. The overall mAP performance using ResNeXtBlock and CBAM increased from 84.10% to 88.90%, whereas the IoU value improved from 64.14% to 77.40% (Table 3). Additionally, the performance of the Yolov8 with ResNeXtBlockCBAM model saw a reduction in prediction errors, decreasing from 50% to 25% (Table 4). The predictions result using the proposed model indicate that ResNeXt block contributes robust feature extraction, whereas CBAM helps filter out irrelevant details, both of which significantly enhance the accuracy in YOLO object detection tasks (Figure 8). By using such combination, this fine-grained attention helps improve accuracy by allowing the model to emphasize key object regions, especially in cluttered or high-density images like fetal heart object. It can be observed that five models failed to detect the presence of cardiomegaly condition, whereas YOLOv8 with ResNeXt-Block and CBAM successfully identified the condition with satisfactory performance, achieving confidence values above 90% (Figure 9).

DISCUSSION

Generalization ability

To ensure that the proposed "YOLOv8 and CBAM" model achieves good generalization ability, we evaluated it using K-fold validation, where the dataset was split into K subsets (folds) with K=10. The most appropriate comparison during testing was with YOLOv11, as it represents the latest YOLO architecture incorporating self-attention mechanisms. The model was trained multiple times, with the validation set rotating across different folds in each iteration. This approach helps in reducing bias and improving model reliability. The results indicate the YOLOv8 with CBAM model achieved an average AP of 68.12% for CC and 97.4% for HC, whereas the YOLOv11 model achieved a higher average AP of 72.99% for CC and 98.76% for HC (refer to Table 5). Despite the lower mAP, YOLOv8 with CBAM might still offer advantages in terms of IoU. The model attained IoU scores of 73.58% for HC and 47.42% for CC, PA values of 81.39% for HC and 65.40% for CC, and confidence scores of 89.95% for CC and 83.78% for HC. In contrast, YOLOv11 achieved slightly lower IoU scores of 71.09% for HC and 47.46% for CC, PA values of 79.25% for HC and 64.88% for CC, and confidence scores of 86.69% for CC and 77.24% for HC, despite achieving higher AP values (refer to Table 6). In fetal heart organ segmentation for cardiomegaly assessment, a

Table 2. Prediction results with unseen data

Notation	YOLOv8.0		YOLOv8.2		YOLOv8.0 and CBAM		YOLOv9		YOLOv11.0	
Frames	True	False	True	False	True	False	True	False	True	False
Validation	50	18	51	17	52	16	51	17	52	16
Testing	12	12	13	11	16	8	12	12	12	12

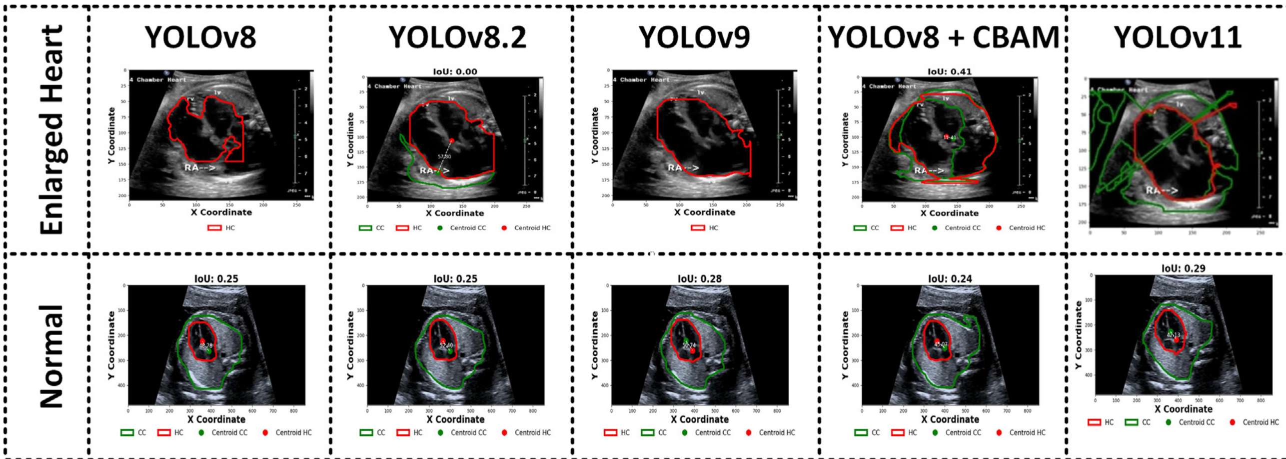

Figure 5. The decision with unseen data

high IoU is crucial due to the heart's small size and dynamic shape. The model must accurately detect the boundaries of the CC and HC areas with high precision. A low IoU may cause misalignment in the bounding box or mask, potentially leading to incorrect clinical decisions.

Benchmarking against other datasets

To evaluate the proposed model's performance across diverse environments, we tested it using an independent dataset comprising fetal ultrasound images of various organs. The data include 12,400 two-dimensional images from 896 pregnant women, which was divided into five objects, i.e., thorax, femur, abdomen, brain, and head. In our experiment, we compared several YOLO models, i.e., YOLOv7, YOLOv8, YOLOv8 with CBAM, YOLOv8 with ResBlockCBAM, YOLOv8 with ResNeXtBlockCBAM, YOLOv9, YOLOv9 with CBAM, and YOLOv11, including cross-stage partial with spatial attention module. It can be seen by using the proposed YOLOv8 and CBAM, it still produces good performance. The resulting mAP is 99.50% for fetal heart enlargement and 95.10% for fetal organ,

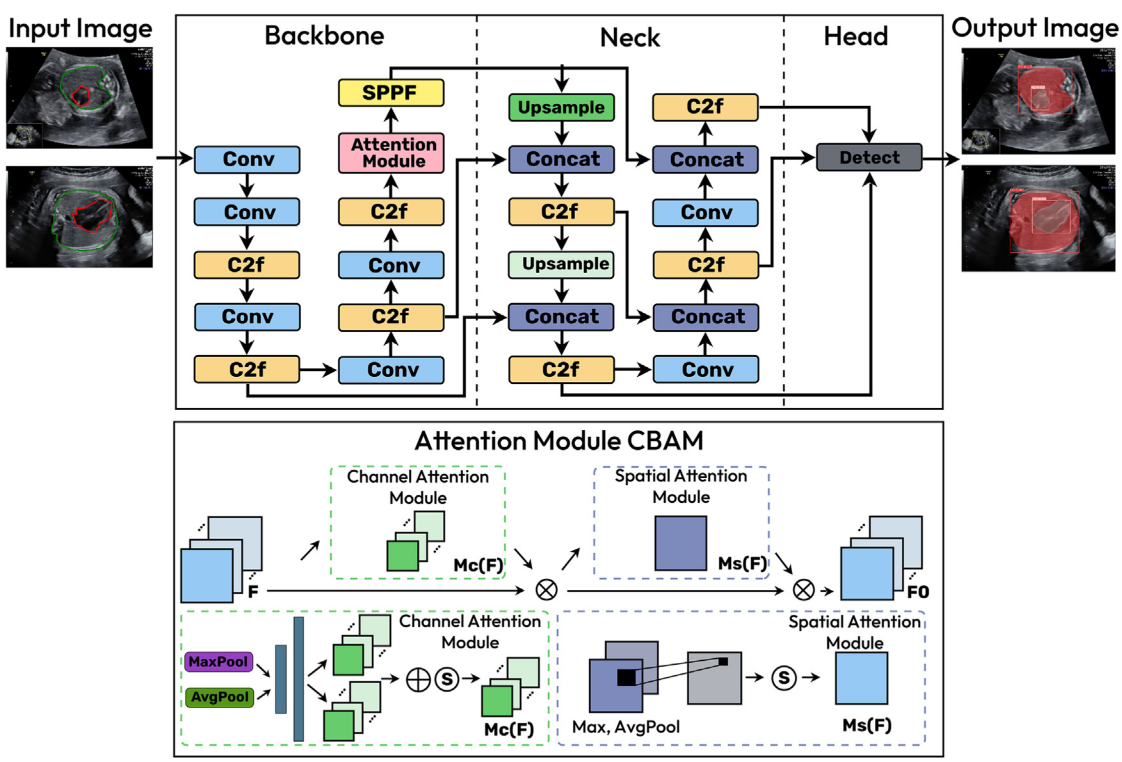

Figure 6. The proposed architecture

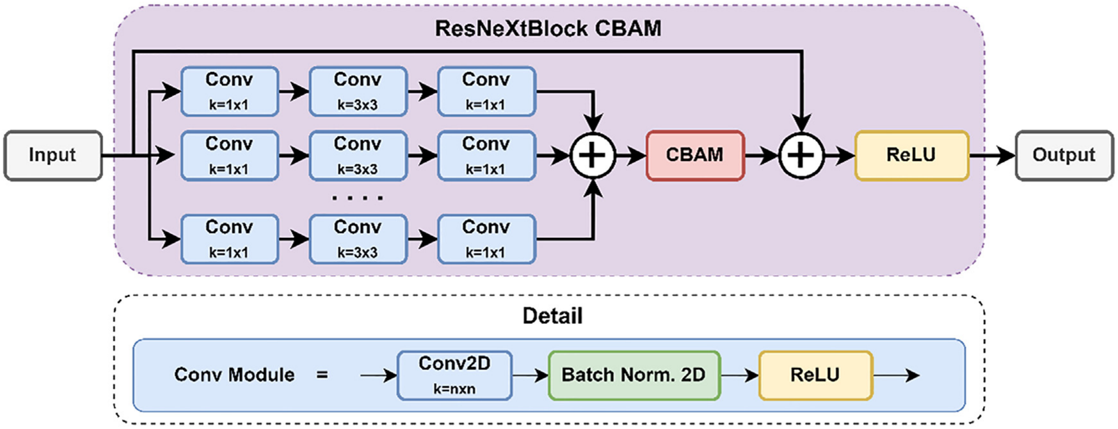

Figure 7. ResNeXtBlock and CBAM structure

meaning that the proposed model can identify or classify objects in the dataset very well (Table 7). The proposed YOLOv8 with CBAM model, compared to YOLOv11 with a self-attention module, achieved a nearly equivalent mAP value but produced an IoU score that was 2% higher. It means the model has more accurately localized the object within the image. These results suggest that YOLOv8 with CBAM provides a robust foundation for integrating computer vision into healthcare applications. By offering reliable predictions, this approach can support healthcare providers in making timely, informed decisions and improving patient outcomes.

Benchmarking with existing studies

Recent studies have focused on the segmentation of the fetal heart from ultrasound videos[49,51] in Table 8. Yan et al.[49] examined the segmentation of critical fetal heart structures, including the pulmonary artery, aorta, and superior vena cava, using the ultrasound three vessel view (3VV). They employed YOLOv5 and Deeplabv3 in their approach. The proposed framework demonstrated strong performance, achieving an average IoU value of 74.51% across all three vessels (3V). Both studies highlight the advancements in DL models for fetal heart segmentation, underscoring their potential to enhance prenatal care by providing more accurate and reliable assessments during ultrasound screenings. Balasubramani et al.[51] introduced the YOLOv8n segmentation model specifically for the left ventricle (LV) in echocardiogram images. This model achieved a mAP of 98.31% at a 50% IoU threshold (mAP50) and a mAP of 75.27% across IoU thresholds ranging from 50% to 95% (mAP50:95). The architecture of YOLOv8n-seg for LV segmentation was significantly improved through their work. In another study, however, these studies are somewhat limited in scope, focusing only on specific structures like the LV or the pulmonary artery, aorta, and superior vena cava. In this study, we propose an approach combining YOLOv8 with ResNeXtBlock and CBAM for segmenting the fetal heart from ultrasound videos and to assess enlargement. Our model achieved an mAP of 75.27% at a 50% IoU threshold (mAP50) and an IoU of 64.14%. These results indicate a robust performance in detecting fetal heart enlargement, offering a reliable tool to assist sonographers with initial computer-aided assessments during prenatal screenings. This proposed model not only builds upon the existing methodologies but also expands the scope of segmentation to include fetal heart enlargement and other fetal organ objects, thereby addressing the limitations of previous studies. By incorporating advanced techniques, YOLOv8 with ResNeXtBlock and CBAM, our model

Table 3. Performance comparison of YOLOv8 and CBAM with residual network

Metrics	Class	YOLOv8 with CBAM	YOLOv8 with ResBlockCBAM	YOLOv8 with ResNeXt BlockCBAM
Validation				
mAP	All	99.50	95.50	95.59
	HC	99.50	99.50	99.50
	CC	99.50	91.50	92.40
IoU	All	80.52	81.50	80.80
	HC	78.69	71.50	69.20
	CC	82.36	91.50	92.40
Testing				
mAP	All	84.10	86.70	88.90
	HC	89.50	89.10	91.30
	CC	78.70	84.40	86.60
IoU	ALL	64.14	76.10	77.40
	HC	68.98	71.80	72.60
	CC	53.30	80.40	82.20

Table 4. The prediction result

Prediction Result	YOLOv8 with CBAM		YOLOv8 with ResBlockCBAM		YOLOv8 with ResNeXt BlockCBAM	
Condition	T	F	T	F	T	F
Validation	42	6	43	5	46	2
Unseen	16	8	17	7	20	4

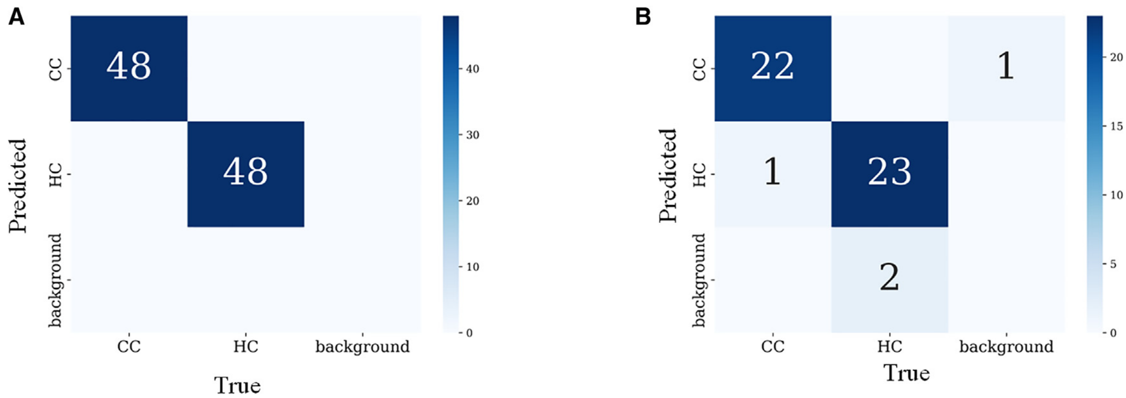

Figure 8. Confusion matrix of YOLOv8 with ResNeXtBlock and CBAM on validation and unseen dataset

demonstrates improved accuracy and reliability, showcasing significant potential for clinical applications in prenatal care.

Limitations of the study

In summary, recognizing structural abnormalities in fetal heart enlargement is particularly challenging.[6,8,15,47] The model's ability to predict these abnormalities is limited, so it must adhere to the guidelines established by obstetrics and gynecology consultants.[7,11,45] The identification of an enlarged fetal heart by these experts is crucial,[29,30] and we incorporated this rule when developing the YOLOv8 model with ResNeXtBlock and CBAM.[44,49] The DL algorithm demonstrated strong performance on community-acquired images from a low-risk population, including lesions it had not previously encountered. Additionally, when both the model and blinded human experts were limited to reviewing stored images—without the comprehensive set of images available to a clinician during a live scan—the model outperformed the human experts.[19] Although the results obtained were satisfactory, this study encountered several limitations such as the following:

(1) The medical datasets, especially those involving rare conditions like fetal heart enlargement, are often limited in size, leading to an imbalance in the dataset. YOLO might struggle to detect cardiomegaly accurately if the model is biased toward more common classes (e.g., normal heart) due to this imbalance. In this study, we attempted to balance normal data with abnormal conditions; however, adding more abnormal data would provide a better testing environment.
(2) The added parameters using CBAM with the residual network can increase the risk of overfitting, especially on small datasets. This happens as the model may become more specific to training features rather than generalizing to new data. CBAM may amplify certain noise patterns if they are mistakenly identified as important features. This could reduce YOLOv8's detection accuracy in noisy or low-quality images.
(3) DL models are often considered "black boxes" because their decision-making processes are not easily

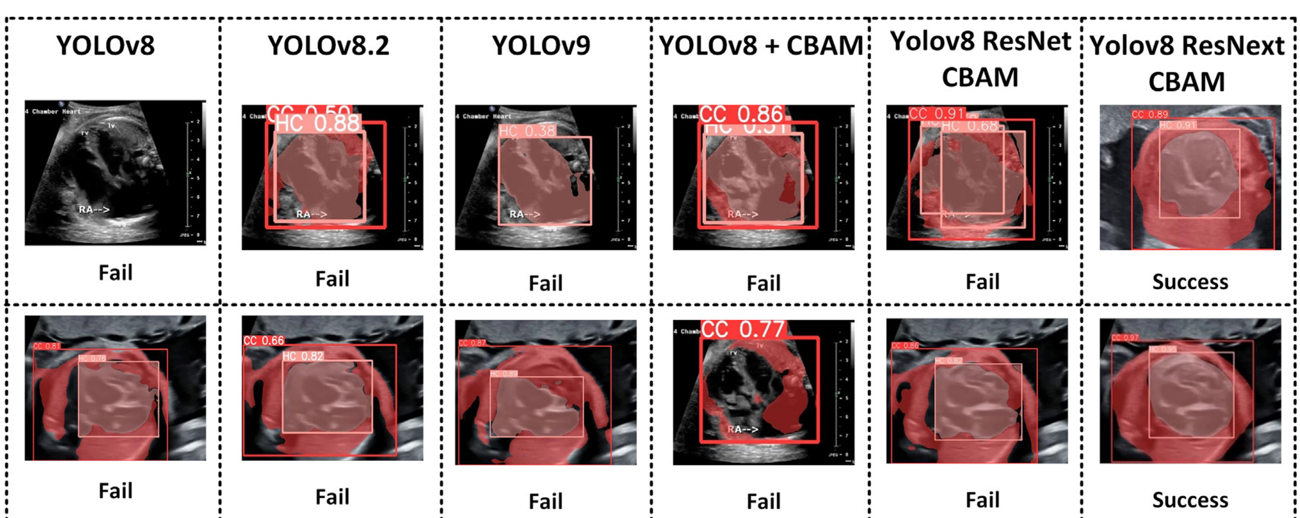

Figure 9. The sample of YOLOv8 with ResNeXtBlock and CBAM performance result

Table 5. mAP value for k = 10

	mAP (%)									
	YoloV8		YoloV8.2		YOLOv8-CBAM		YOLOv9		YOLOv11	
Fold	CC	HC	CC	HC	CC	HC	CC	HC	CC	HC
1	57.3	86.2	46	97.7	99.5	99.5	49.4	95.5	75.3	93.5
2	59.4	98.6	50	87.5	99.5	99.5	62.5	92.6	83.1	99.1
3	54.9	92.2	54.6	66.66	99.5	99.5	77.3	92	81.6	99.3
4	57.6	81.8	39.9	95.4	90.2	92.7	82.6	99.3	59.1	99.5
5	54.4	84.8	59.3	67	27.8	99.5	59.1	99.5	68.5	99.5
6	55.4	88.8	70.1	94.8	72	93.6	63.6	88.2	83.1	99.1
7	56.5	94.8	63	82.8	54.4	95.5	44.3	89.1	59.1	99.5
8	54.4	84.8	50	86.5	38.3	97.5	50	79.89	68.5	99.5
9	55.4	88.8	63.18	63.8	56.8	97.2	49.8	88.33	83.1	99.1
10	56.5	94.8	58.6	63.6	43.2	99.5	55.23	80.01	68.5	99.5
Average	56.18	89.56	55.468	80.576	68.12	97.4	59.383	90.443	72.99	98.76

interpretable. In medical applications, where understanding why a model made a particular decision is crucial; this lack of transparency can be a significant limitation for solving in the future research.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Siti Nurmaini (sitinurmaini@gmail.com).

Materials availability

This study did not generate new unique materials. The software and model are available upon request to Siti Nurmaini (sitinurmaini@gmail.com) or at https://github.com/ISySRGg/Fetal_Enlarge.

Data and code availability

- The dataset used in this study is available at https://github.com/ISySRGg/Fetal_Enlarge/tree/main/dataset.
- The deep learning model and code for Computer-aided Assessment for Enlarged Fetal Heart are available at https://github.com/ISySRGg/Fetal_Enlarge/tree/main/dataset or upon request.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported the Intelligent System Research Group (ISysRG) Universitas Sriwijaya. We gratefully acknowledge the financial support from Universitas Sriwijaya under Professional Research No. 0016/UN9/SK.LP2M.PT/2024, which made this research possible. Their generous funding provided the necessary resources and tools to develop and validate our DL model for the assessment of fetal heart enlargement. We also extend our gratitude to Dr. Mohammad Hoesin Central General Hospital and Bunda Hospital Palembang for their invaluable contributions and support throughout this project.

AUTHOR CONTRIBUTIONS

S.N., wrote the manuscript, funding acquisition, conceptualization, and formal analysis. A.I.S., M.T.R., and M.N.R., data curation, software

Table 6. Performance value for k = 10 Yolov8 and CBAM vs. Yolov11

YoloV8 + CBAM						Yolov11					
IoU (%)		PA (%)		Confidence (%)		IoU (%)		PA (%)		Confidence (%)	
CC	HC	CC	HC	CC	HC	CC	HC	CC	HC	CC	HC
74.56	51.81	84.68	65.42	74.56	51.81	72.75	48.31	81.61	64.74	78.96	85.84
78.38	48.19	84.84	65.16	78.38	48.19	72.75	48.31	81.61	64.74	79.96	85.84
77.14	47.07	84.54	65.44	77.14	47.07	72.51	47.80	81.58	64.51	79.16	86.03
75.76	47.71	84.48	65.49	75.76	47.71	71.04	46.56	78.69	64.91	77.38	86.14
70.32	44.70	75.48	65.77	70.32	44.70	71.04	46.56	78.69	64.91	77.38	86.14
66.70	50.36	76.07	65.35	66.70	50.36	71.04	46.56	78.69	64.91	77.38	86.14
66.72	46.73	75.61	65.22	66.72	46.73	68.83	48.68	78.16	65.13	73.74	89.25
69.98	48.22	78.81	65.08	69.98	48.22	71.04	46.56	78.69	64.91	77.38	86.14
77.46	44.63	84.69	65.61	77.46	44.63	68.83	48.68	76.16	65.13	73.74	89.25
78.80	42.80	84.66	65.50	78.80	42.80	71.04	46.56	78.69	64.91	77.38	86.14
Average											
73.58	47.42	81.39	65.40	83.78	89.95	71.09	47.46	79.25	64.88	77.24	86.69

Table 7. The benchmarking studies

Model	mAP (%) Fetal heart enlargement (two-classes)	Fetal Organ (five-classes)
YOLOv7	87.60	80.51
YOLOv8	95.50	94.40
YOLOv8 with CBAM	99.50	95.10
YOLOv8 with ResBlockCBAM	95.50	93.70
YOLOv8 with ResNeXtBlockCBAM	95.59	94.20
YOLOv9	98.01	91.90
YOLOv9 with CBAM	99.00	94.00
YOLOv11 with cross-stage partial with spatial attention module	99.50	95.40

analysis, resources, visualization preparation. P.M. and N.B., medical verification. A.D., data validation and editing the manuscript. B.T. and F.F., data curation. A.I., A.W.A., and R.B., software analysis and visualization preparation.

DECLARATION OF INTERESTS

The authors declare that they have no conflict of interest.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Fetal enlargement dataset
  - Data preparation
  - Data annotation
- METHOD DETAILS
  - The proposed architecture
  - Fetal heart decision
  - Platform
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

Received: August 29, 2024
Revised: November 20, 2024
Accepted: March 21, 2025
Published: March 27, 2025

Table 8. The benchmarking studies with the recent studies

Author	Model	Performance (%) mAP50	IoU
Yan et al.[49]	Yolov5 with Deeplabv3	–	74.51
Balasubramani et al.[51]	YOLOv8n-seg	98.31	–
This study	YOLOv8 with ResNeXtBlock and CBAM YOLOv8 and CBAM	95.59 99.50	80.80 80.52

REFERENCES

1. Sapanont, K., Luangdansakul, W., Pleankong, M., Smanchat, B., Bhamarapravatana, K., and Suwannarurk, K. (2023). Incidence of Fetal Myocardial Hypertrophy in Mother with Diabetes Mellitus by Using Cardio-Spatiotemporal Image Correlation (STIC) M-Mode. Int. J. Womens Health *15*, 703–710.
2. Mat Bah, M.N., Sapian, M.H., Jamil, M.T., Abdullah, N., Alias, E.Y., and Zahari, N. (2018). The birth prevalence, severity, and temporal trends of congenital heart disease in the middle-income country: a population-based study. Congenit. Heart Dis. *13*, 1012–1027.
3. Rahajoe, A.U. (2007). Management of patients with congenitally malformed hearts in Indonesia. Cardiol. Young *17*, 584–588.
4. Moerdijk, A.S., Claessens, N.H., van Ooijen, I.M., van Ooij, P., Alderliesten, T., and Grotenhuis, H.B.; FUTURE 2.0 consortium (2023). Fetal MRI of the heart and brain in congenital heart disease. Lancet. Child Adolesc. Health *7*, 59–68.
5. Tollenaar, L.S.A., Lopriore, E., Middeldorp, J.M., Klumper, F.J.C.M., Haak, M.C., Oepkes, D., and Slaghekke, F. (2020). Prevalence of placental dichotomy, fetal cardiomegaly and starry-sky liver in twin anemia–polycythemia sequence. Ultrasound Obstet. Gynecol. *56*, 395–399.
6. Salvi, S., D'Emidio, L., Roughton, M., De Carolis, S., Lanzone, A., Bhide, A., and Carvalho, J.S. (2021). Is Cardiomegaly an Indication of "Heart-Sparing Effect" in Small Fetuses? Fetal Diagn. Ther. *48*, 701–707.
7. Truszkiewicz, K., Poręba, R., and Gać, P. (2021). Radiological cardiothoracic ratio in evidence-based medicine. J. Clin. Med. *10*, 2016.
8. Sridevi, S., Murugesan, S., Sakthivel, V., and Kavitha, D. (2021). Computer-aided decision support system for symmetry-based prenatal congenital heart defects. In Advanced Machine Vision Paradigms for Medical Image Analysis (Elsevier), pp. 11–53.
9. Hassan, H.G.E.M.A., Habeeb, N.M.M., Ibrahim, A.S., Megahed, S.R., and Yassin, A. (2022). Feasibility study in assessment of congenital cardiovascular malformation by recent technique of fetal cardiac MR imaging. Egypt. J. Radiol. Nucl. Med. *53*, 98.
10. Pietrolucci, M., Maqina, P., Greco, S., Berti, G., Viviano, V., and Rizzo, G. (2023). Evaluation of an artificial intelligence algorithm (HeartAssist^{TM}) in the assessment of fetal cardiothoracic ratio: a prospective study. J Perinat Med *51*, 920–924.
11. Radswiki, T., and Campos, A. (2024). Fetal cardiothoracic circumference ratio [Internet]. https://radiopaedia.org/articles/15006.
12. Anuwutnavin, S., Rangseechamrat, P., Sompagdee, N., Ruangvutilert, P., and Viboonchard, S. (2024). Comparing three cardiothoracic ratio measurement techniques and creating multivariable scoring system to predict Bart's hydrops fetalis at 17–22 weeks' gestation. Sci. Rep. *14*, 8894.
13. Picazo-Angelin, B., Zabala-Argüelles, J.I., Anderson, R.H., and Sánchez-Quintana, D. (2018). Anatomy of the normal fetal heart: The basis for understanding fetal echocardiography. Ann. Pediatr. Cardiol. *11*, 164–173.
14. Rachmatullah, M.N., Nurmaini, S., Sapitri, A.I., Darmawahyuni, A., Tutuko, B., and Firdaus, F. (2021). Convolutional neural network for semantic segmentation of fetal echocardiography based on four-chamber view. Bulletin EEI. *10*, 1987–1996.
15. Nurmaini, S., Rachmatullah, M.N., Sapitri, A.I., Darmawahyuni, A., Tutuko, B., Firdaus, F., Partan, R.U., and Bernolian, N. (2021). Deep learning-based computer-aided fetal echocardiography: Application to heart standard view segmentation for congenital heart defects detection. Sensors *21*, 8007.
16. Torrents-Barrena, J., Piella, G., Masoller, N., Gratacós, E., Eixarch, E., Ceresa, M., and Ballester, M.Á.G. (2019). Segmentation and classification in MRI and US fetal imaging: recent trends and future prospects. Med.

Image Anal. *51*, 61–88. Available from: https://www.sciencedirect.com/science/article/abs/pii/S1361841518308478

17. Nurmaini, S., Partan, R.U., Bernolian, N., Sapitri, A.I., Tutuko, B., Rachmatullah, M.N., Darmawahyuni, A., Firdaus, F., and Mose, J.C. (2022). Deep Learning for Improving the Effectiveness of Routine Prenatal Screening for Major Congenital Heart Diseases. J. Clin. Med. *11*, 6454.
18. An, S., Zhu, H., Wang, Y., Zhou, F., Zhou, X., Yang, X., Zhang, Y., Liu, X., Jiao, Z., and He, Y. (2021). A category attention instance segmentation network for four cardiac chambers segmentation in fetal echocardiography. Comput. Med. Imaging Graph. *93*, 101983.
19. Dong, J., Liu, S., and Wang, T. (2019). ARVBNet: Real-Time Detection of Anatomical Structures in Fetal Ultrasound Cardiac Four-Chamber Planes. In Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting (Springer), pp. 130–137.
20. Cai, Q., Chen, R., Li, L., Huang, C., Pang, H., Tian, Y., Di, M., Zhang, M., Ma, M., Kong, D., and Zhao, B. (2022). The Application of Knowledge Distillation toward Fine-Grained Segmentation for Three-Vessel View of Fetal Heart Ultrasound Images. Comput. Intell. Neurosci. *2022*, 1765550.
21. Xu, L., Liu, M., Shen, Z., Wang, H., Liu, X., Wang, X., Wang, S., Li, T., Yu, S., Hou, M., et al. (2020). DW-Net: A cascaded convolutional neural network for apical four-chamber view segmentation in fetal echocardiography. Comput. Med. Imaging Graph. *80*, 101690.
22. Athalye, C., van Nisselrooij, A., Rizvi, S., Haak, M.C., Moon-Grady, A.J., and Arnaout, R. (2024). Deep-learning model for prenatal congenital heart disease screening generalizes to community setting and outperforms clinical detection. Ultrasound Obstet. Gynecol. *63*, 44–52.
23. Gupta, M., Singh, A., and Kumar, Y. (2024). Deep learning for prediction of cardiomegaly using chest X-rays. Neural Comput. Appl. *36*, 19383–19391.
24. Lee, M.S., Kim, Y.S., Kim, M., Usman, M., Byon, S.S., Kim, S.H., Lee, B.I., and Lee, B.D. (2021). Evaluation of the feasibility of explainable computer-aided detection of cardiomegaly on chest radiographs using deep learning. Sci. Rep. *11*, 16885.
25. Candemir, S., Rajaraman, S., Thoma, G., and Antani, S. (2018). Deep learning for grading cardiomegaly severity in chest x-rays: an investigation. In 2018 IEEE Life Sciences Conference (LSC) (IEEE), pp. 109–113.
26. Qing, B., Zhanquan, S., Kang, W., Chaoli, W., Shuqun, C., and Jiawei, Z. (2024). MPSA: Multi-Position Supervised Soft Attention-based convolutional neural network for histopathological image classification. Expert Syst. Appl. *253*, 124336.
27. Li, Y., Zhang, M., Zhang, C., Liang, H., Li, P., and Zhang, W. (2024). YOLO-CCS: Vehicle detection algorithm based on coordinate attention mechanism. Digit. Signal Process. *153*, 104632.
28. Ma, X., Li, Y., Yang, Z., Li, S., and Li, Y. (2024). Lightweight network for millimeter-level concrete crack detection with dense feature connection and dual attention. J. Build. Eng. *94*, 109821.
29. Huang, Z., Li, J., Mao, N., Yuan, G., and Li, J. (2024). DBEF-Net: Diffusion-Based Boundary-Enhanced Fusion Network for medical image segmentation. Expert Syst. Appl. *255*, 124467.
30. Woo, S., Park, J., Lee, J.Y., and Kweon, I.S. (2018). Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV) (Springer), pp. 3–19.
31. Luo, Y., and Wang, Z. (2021). An improved resnet algorithm based on cbam. In 2021 International Conference on Computer Network, Electronic and Automation (ICCNEA) (IEEE), pp. 121–125.
32. Yan, J., Zeng, Y., Lin, J., Pei, Z., Fan, J., Fang, C., and Cai, Y. (2024). Enhanced object detection in pediatric bronchoscopy images using YOLO-based algorithms with CBAM attention mechanism. Heliyon *10*, e32678.
33. Ji, Z., Wu, Y., Zeng, X., An, Y., Zhao, L., Wang, Z., and Ganchev, I. (2023). Lung Nodule Detection in Medical Images Based on Improved YOLOv5s (IEEE Access).
34. Zamri, F.N.M., Gunawan, T.S., Yusoff, S.H., Alzahrani, A.A., Bramantoro, A., and Kartiwi, M. (2024). Enhanced Small Drone Detection Using Optimized YOLOv8 with Attention Mechanisms (IEEE Access).
35. Yang, Z., Zhang, Q., Chen, W., and Chen, C. (2021). PolSAR Image Classification Based on Resblock Combined with Attention Model. In 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP) (IEEE), pp. 340–344.
36. Zheng, W., Ren, J., Ke, L.T., and Shen, Q.H. (2022). A New Visualization-guided Attention-ResNeXt Network for Trash classification. In 2022 41st Chinese Control Conference (CCC) (IEEE), pp. 6970–6977.
37. Huang, G., Gong, Y., Xu, Q., Wattanachote, K., Zeng, K., and Luo, X. (2020). A convolutional attention residual network for stereo matching. IEEE Access *8*, 50828–50842.
38. Qiao, S., Pang, S., Sun, Y., Luo, G., Yin, W., Zhao, Y., Pan, S., and Lv, Z. (2022). SPReCHD: Four-Chamber Semantic Parsing Network for Recognizing Fetal Congenital Heart Disease in Medical Metaverse (IEEE J Biomed Heal informatics).
39. Qiao, S., Pang, S., Luo, G., Pan, S., Chen, T., and Lv, Z. (2022). Flds: An intelligent feature learning detection system for visualizing medical images supporting fetal four-chamber views. IEEE J. Biomed. Health Inform. *26*, 4814–4825.
40. Nurmaini, S., Sapitri, A.I., Tutuko, B., Rachmatullah, M.N., Rini, D.P., Darmawahyuni, A., Firdaus, F., Mandala, S., Nova, R., and Bernolian, N. (2023). Automatic echocardiographic anomalies interpretation using a stacked residual-dense network model. BMC Bioinf. *24*, 365.
41. Arnaout, R., Curran, L., Zhao, Y., Levine, J.C., Chinn, E., and Moon-Grady, A.J. (2021). An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. Nat. Med. *27*, 882–891.
42. Sulas, E., Urru, M., Tumbarello, R., Raffo, L., and Pani, D. (2020). Automatic detection of complete and measurable cardiac cycles in antenatal pulsed-wave Doppler signals. Comput. Methods Programs Biomed. *190*, 105336.
43. Creasy, R.K. (2014). Creasy and Resnik's maternal-fetal medicine: principles and practice, 7th ed (Saunders/Elsevier). https://cir.nii.ac.jp/crid/1130000797420345088.
44. Yan, L., Ling, S., Mao, R., Xi, H., and Wang, F. (2024). A deep learning framework for identifying and segmenting three vessels in fetal heart ultrasound images. Biomed. Eng. Online *23*, 39.
45. van Beers, F., Lindström, A., Okafor, E., and Wiering, M. (2019). Deep neural networks with intersection over union loss for binary image segmentation. In Proceedings of the 8th international conference on pattern recognition applications and methods, pp. 438–445.
46. Balasubramani, M., Sung, C.W., Hsieh, M.Y., Huang, E.P.C., Shieh, J.S., and Abbod, M.F. (2024). Automated Left Ventricle Segmentation in Echocardiography Using YOLO: A Deep Learning Approach for Enhanced Cardiac Function Assessment. Electronics *13*, 2587.
47. (2023). YOLO by Ultralytics [Internet]. https://github.com/ultralytics/ultralytics.
48. Zion, D., Gillam, L., and Loff, B. (2000). The Declaration of Helsinki, CIOMS and the ethics of research on vulnerable populations. Nat. Med. *6*, 615–617.
49. Radswiki, T., and Elfeky, M. (2024). Fetal cardiomegaly [Internet]. https://radiopaedia.org/articles/15004.
50. Mallmann, M.R., Herberg, U., Gottschalk, I., Strizek, B., Hellmund, A., Geipel, A., Breuer, J., Gembruch, U., and Berg, C. (2020). Fetal cardiac intervention in critical aortic stenosis with severe mitral regurgitation, severe left atrial enlargement, and restrictive foramen ovale. Fetal Diagn. Ther. *47*, 440–447.
51. Padilla, R., Netto, S.L., and Da Silva, E.A.B. (2020). A survey on performance metrics for object-detection algorithms. In 2020 international conference on systems, signals and image processing (IWSSIP) (IEEE), pp. 237–242.

52. Melamed, N., Baschat, A., Yinon, Y., Athanasiadis, A., Mecacci, F., Figueras, F., Berghella, V., Nazareth, A., Tahlak, M., McIntyre, H.D., et al. (2021). FIGO (international Federation of Gynecology and obstetrics) initiative on fetal growth: best practice advice for screening, diagnosis, and management of fetal growth restriction. Int. J. Gynaecol. Obstet. *152*, 3–57.

53. Singh, Y., and McGeoch, L. (2016). Fetal Anomaly Screening for Detection of Congenital Heart Defects. J. Neonatal Biol. *05*, 1–5.

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Computer-Aided Assessment for enlarged fetal heart (Deep Learning Model)	This Paper	https://github.com/ISySRGg/Fetal_Enlarge
Python	Python software foundation	3.11
Tensor Flow	Google, Google Brain	2.14
PyTorch	Meta AI	2.5
GPU	NVDIA Corporation	GForce 2080 Ti 12 GB
CPU	Intel Corporation	i9-14900K
Memory	G.Skil Corporation	G.SKIL Trident Z GDDR 5 5600 Mhz
Operating System	Mircrosoft Corporation	Windows 11
IDE	Mircrosoft Corporation	Visual Studio Code
Numpy	Open Source Code (Developed by: Travis Oliphant and Matti Picus)	2.2
Pandas	Open Source Code (Developed by: Community)	2.2.3
ScikitLearn	Open Source Code (Developed by: David Cournapeau)	1.6
Matplotlib	Open Source Code (Developed by: Michael Droettboom)	3.10.1
Seaborn	Open Source Code (Developed by: Community)	0.13.2
Robo Flow	Roboflow	https://app.roboflow.com/
Ultrasound Fetal Heart Dataset	Mohammad Hoesin General Hospital, South Sumatera, Indonesia	https://github.com/ISySRGg/Fetal_Enlarge/tree/main/dataset
DICOM (Digital Imaging and Communications in Medicine)	NEMA (National Electrical Manufacturers Association)	Standard for medical imaging

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Fetal enlargement dataset

The sample size of the study was calculated using the Equation 1:[47]

$$m_k = 4 \times \frac{(1-k)}{W_k^2}\left((1-k)(1-2k) + \frac{k(2-k)}{2\pi_{Dis}(1-\pi_{Dis})}\right) z^2_{1-\frac{\alpha}{2^*}} \quad \text{(Equation 1)}$$

where, W_k = IK range, k = precision (10%), and z = degree of confidence. This study uses a significance level of 0.05 and a confidence interval of 95% ($Z\alpha$ = 1.96). Thus, a minimum of 60 samples is required for the study. At Dr. Mohammad Hoesin Hospital and Bunda Palembang, Indonesia, there were 1,537 deliveries in 2021, 1,628 in 2022, and 1,385 in 2023. With an estimated prevalence of congenital heart defects at 9.3 per 1,000 live births, approximately 16 cases of congenital heart defects are expected annually. This study will use consecutive sampling, inviting every potential respondent who meets the inclusion criteria to participate. Based on available data, we identified only 10 pregnant women with an enlarged fetal heart. To balance the dataset, we included 14 pregnant women with normal fetal hearts. The general characteristics of the research subjects in each group are presented in Table 9. However, due to ultrasound quality issues, we were only able to use 22 ultrasound videos, indicating that the number of pregnant women with an enlarged fetal heart condition is limited. All participants were pregnant and self-identifying as female. Participants were also not selected based on race, ethnicity, or ancestry. In addition, both male and female fetuses were included due to there is no inclusion or exclusion criterion for fetal gender. The dataset used in this study is available at https://github.com/ISySRGg/Fetal_Enlarge/tree/main/dataset.

Table 9. The general characteristic of the subject

Characteristics	Frequency (n)	Percentage (%)
Age (year)		
20–35	19	79.17
> 35	5	20.83
IMT		
Normal weight	5	20.83
Abnormal weight	19	79.17
Trimester		
Second	9	37.5
Third	15	62.5
Gestation		
1–4	23	95.83
>4	1	4.17
Parity		
0	9	37.5
1–4	15	62.5
> 4	0	0
Abortion history		
No	19	79.17
1	4	16.67
2	1	4.17
3	0	0
4	0	0

Data preparation

Ultrasound videos capturing fetal heart enlargement are scarce due to the difficulty in detecting such abnormalities in fetuses. Each ultrasound video in this study corresponds to a unique case. To maintain the integrity of the evaluation, all images extracted from a single video were exclusively assigned to one of three datasets: training, validation, or testing. A total of 22 ultrasound videos were used: 9 videos of normal cases and 9 videos of cardiomegaly cases for training and validation, while 2 videos of normal cases and 2 of cardiomegaly were reserved for testing. The full set of ultrasound videos was converted into 450 images, capturing both fetal heart enlargement and normal conditions. The dataset was then divided into 356 images for training, 48 images for validation, and the remaining samples for testing. This approach prevents potential data leakage and enables the model to be evaluated on completely unseen data from different cases, preserving the validity of the results. These ultrasound videos were collected from Dr. Mohammad Hoesin General Hospital and Bunda Hospital in Palembang, Indonesia. The ultrasound video recordings were saved in Digital Imaging and Communications in Medicine (DICOM) format, with durations ranging from 15 to 40 seconds and file sizes between two and 50 kilobytes (KB). All examinations were conducted using the GE Voluson V8 device, measuring cardiac size by dividing the heart area (HA) by the chest area (CA). A sample of the ultrasound fetal heart videos for both conditions is visualized in Figure 10.

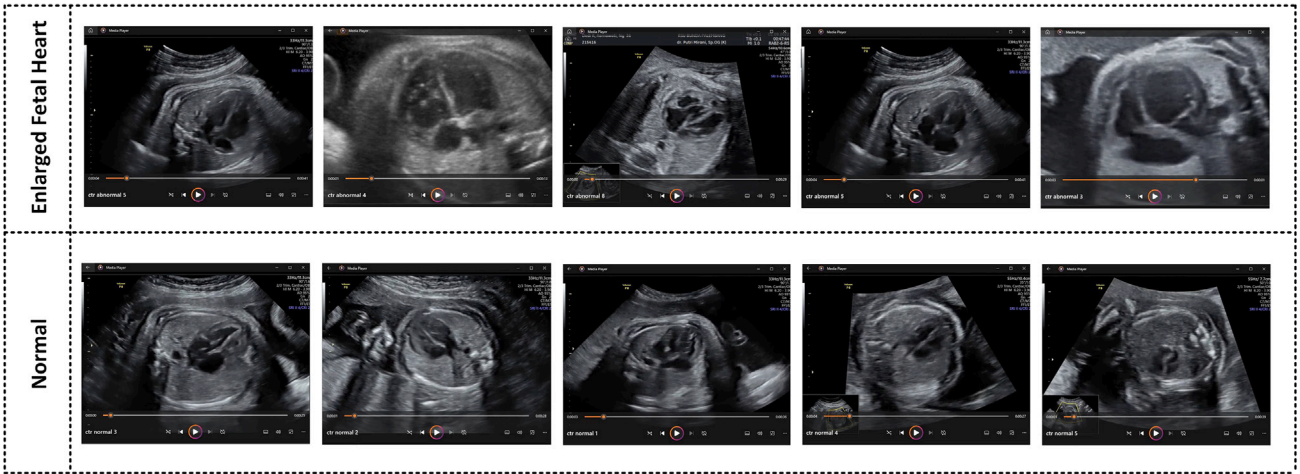

Figure 10. The sample of fetal cardiac ultrasound images
(A) fetal heart enlargement and (B) normal.

Data annotation

The ultrasound video is converted frame by frame using Python software to produce two-dimensional images. The labels consisted of heart circumference (HC) in red color and chest cavity (CC) in green color annotations (Figure 11). The process of generating annotation labels produces ground truth data, which involves using the Python LabelMe library to transform the images into semantic segmentation. The resulting labels are saved through image thresholding and then converted into JavaScript Object Notation (.json). All ground truth of ultrasound images is available in Joint Photographic Experts Group (.jpeg/jpg) format. The annotation process was supervised and validated by a fetomaternal specialist from Bunda Hospital, Indonesia, with over 15 years of experience. The fetal cardiothoracic (C/T) circumference ratio is used to assess fetal heart size by fetomaternal specialists. In this study, we estimate the size of the fetal heart concerning the CC. If the fetal heart experiences fetal heart enlargement, the heart size will increase near the CC.

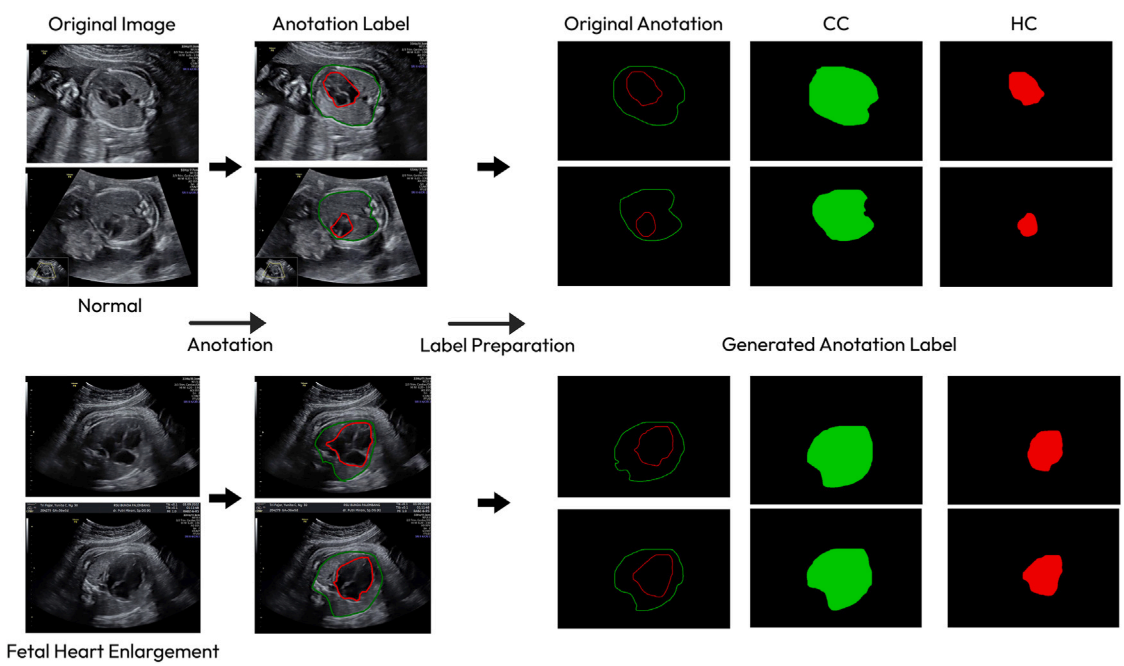

Figure 11. Fetal heart annotation
The green color is the CC, and the red color is the HC

METHOD DETAILS

The research methodology of this study describes the techniques and procedures used to identify and analyze fetal heart conditions based on the DL model. The whole explanation of the methodology is described as follows.

The proposed architecture

Regarding fetal heart ultrasound, the CC is crucial for evaluating the chest region, as it helps examiners identify the orientation and size of structures within the chest, which are essential for fetal prognosis. The CC and HC localization with instance segmentation is necessary, which it is not only provides the classes of the CC and HC objects but also provides the location of the CC and HC objects that have been classified. In this study, we experimented with and proposed YOLOv8 with Convolutional Block Attention Module (CBAM) to enhance the instance segmentation accuracy of the CC in fetal heart ultrasound images. The hyperparameter of the YOLOv8.0 model includes 200 epochs, a batch size of 4, and an SGD optimizer with a learning rate of 0.01 (Figure 6).

By delivering precise bounding boxes and accurate masks, YOLOv8 is an excellent choice for tasks demanding pixel-level analysis.[49] Detecting the CC and HC localization in the fetal heart presents a notable challenge for detection algorithms. To tackle this issue, the current study suggests incorporating the CBAM attention mechanism during the feature extraction phase of YOLOv8. CBAM integrates channel and spatial attention mechanisms, enabling effective identification of crucial image features while reducing irrelevant noise (refer to Algorithm 2). This dual attention mechanism significantly improves detection accuracy and efficiency, particularly for intricate and small fetal heart structures. Consequently, CBAM emerges as a crucial component in advanced fetal heart ultrasound object detection systems.

Fetal heart decision

The fetal heart enlargement and normal condition decision-making based on IoU value are developed based on medical knowledge consisting of (refer to Algorithm 1):

(1) Enlarged fetal heart: if the IoU value between CC and HC is more than 0.3, and
(2) If the knowledge-based decision is not fulfilled, then the class-predicted is a normal heart.

Platform

The experiment works on Intel(R) Core(TM) i9-14900K(32 CPUs), -3.2GHz RAM, and using NVIDIA GeForce RTX 2080 Ti. All experiments were run on Windows 11 Pro 64 Bit, Python code using VS Code, Pytorch, Numpy, Pandas, ScikitLearn, Matplotlib, Seaborn, and Roboflow were used (all details have mentioned in key resources table).

QUANTIFICATION AND STATISTICAL ANALYSIS

The evaluation metrics for an object detection model measure its proficiency in correctly identifying and pinpointing objects within an image. We have used intersection over union (IoU) and mean average precision (mAP) for the performance evaluation of CC and HC detection. The accuracy metric IoU (known as the Jaccard Index) evaluates the overlap between two bounding boxes—the predicted one and the ground truth box. IoU for comparing the similarity between two arbitrary shapes $A, B \subseteq \mathbb{S} \in \mathbb{R}^n$ is attained by Equation 2:[20,50]

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad \text{(Equation 2)}$$

where A is the ground truth, *and* B is the prediction of the output image. By comparing the IoU with a specified threshold t, We can determine whether a detection is correct or incorrect. If IoU $\geq t$, The detection is considered correct; if IoU $< t$, The detection is considered incorrect. The choice of IoU as the primary metric for segmentation in this study is based on its direct interpretation and effectiveness in evaluating the overlap of CC and HC in decision-making process based on medical rules. IoU measures the ratio of the intersection of the predicted and true segmentation areas to their union, making it a suitable metric for assessing how well the model captures the overall shape and area of the target region. The mAP is a metric used to measure the accuracy of object detectors across all classes. The mAP takes into account precision and recall for different IoU thresholds and object classes, with a higher mAP indicating better overall model performance, with the formula as follows[52]:

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \quad \text{(Equation 3)}$$

with AP_i being the average precision in the i-th class and N Is the total number of classes being evaluated.

ADDITIONAL RESOURCES

This study received approval from the Health Research Ethics Committee of Central General Hospital Dr. Mohammad Hoesin Palembang, Indonesia, under ethical certificate No. 18/keprsmh/2022. The procedures adhered to the principles of the Declaration of Helsinki and International Ethical Guidelines for Biomedical Research Involving Human Subjects.[53] Written informed consent to participate in the study was obtained from the parents, legal guardian or next of kin of the participants (men and women). Detailed information about the examination procedures and objectives was provided to all research subjects. Subsequently, their participation in the research was sought through the signing of a consent form, indicating their agreement to take part. It is important to note that research subjects participated voluntarily and retained the right to withdraw from the study at any point.