Automatic Detection of Key Fetal Head Anatomical Structures in the First Trimester Using Ultrasound Images

Fajar A. Hermawati* and Danara D. Caesa

Department of Informatics, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia
Email: fajarastuti@untag-sby.ac.id (F.A.H.); danara.dc17@gmail.com (D.D.C.)
*Corresponding author

Abstract*—Ultrasound imaging is essential for early pregnancy assessment; however, manual interpretation remains limited by operator dependency and speckle noise. Unlike previous studies that employed You Only Look Once (YOLO)v5, YOLOv6, or YOLOv7, this study introduces an automated approach for detecting key fetal head anatomical structures during the first trimester based on the YOLOv8 architecture. The study systematically compares raw ultrasound images with Hybrid Speckle Noise Reduction (HSNR) preprocessing to evaluate the trade-off between visual enhancement and fine-structure preservation, as quantified by a Structural Similarity Index Measure (SSIM) of 0.912 and a Peak Signal-to-Noise Ratio (PSNR) of 29.8 decibels. Semi-automated annotation with expert validation, yielding a Cohen's kappa value of 0.84, ensured high labeling reliability. YOLOv8 was trained under multiple configurations, including different optimizers such as stochastic gradient descent, Adam, and AdamW, combined with early stopping and stratified five-fold cross-validation at a resolution of 320 by 320 pixels to balance anatomical detail and real-time efficiency. The optimal configuration using AdamW with early stopping achieved strong detection performance, with a mean average precision at an Intersection-over-Union (IoU) threshold of 0.50 of 0.907 and a recall of 0.857, outperforming models trained on images processed with speckle noise reduction. While speckle noise reduction improved overall image clarity, it slightly reduced the detectability of subtle anatomical features such as nuchal translucency and nasal skin due to excessive smoothing. The proposed model achieved an inference speed of 45 frames per second on a graphics processing unit, demonstrating its feasibility for real-time clinical deployment. Overall, these results highlight the potential of an optimized YOLOv8 model trained on raw ultrasound images as an efficient, reliable, and clinically applicable approach for artificial intelligence–assisted prenatal screening.

***Keywords*—ultrasound image analysis, fetal head detection, first-trimester screening, deep learning, prenatal imaging, You Only Look Once (YOLO)v8, fetal head anatomical structures**

I. Introduction

Ultrasonography has become one of the primary methods in modern obstetrics for supporting maternal and fetal health [1]. This technology utilizes high-frequency sound waves to generate detailed images of body structures, enabling non-invasive identification of various medical conditions [2]. During the first trimester of pregnancy, the role of ultrasonography extends beyond verifying pregnancy continuation and determining gestational age to include the early detection of clinically significant fetal anatomical abnormalities [3].

Early detection of abnormalities during pregnancy is particularly important because the first trimester represents a period associated with a high risk of complications. One widely used screening technique is nuchal translucency measurement, which has been shown to be effective for assessing the risk of chromosomal abnormalities such as Down syndrome [4–7]. First-trimester ultrasonographic examination is also essential for determining gestational age, confirming pregnancy viability, and enabling the early identification of major fetal anatomical abnormalities [8]. Advances in ultrasonographic technology now allow fetal anatomical assessment at eleven to thirteen weeks of gestation, a period that is strongly recommended for evaluating fetal viability, determining the number of living fetuses, and identifying potential abnormalities [9].

The accuracy of first-trimester ultrasonography is well established for gestational age estimation, the detection of non-viable pregnancies, and facilitating early medical intervention when required [10]. However, this process remains highly dependent on the expertise of the sonographer, which can lead to variability in results due to human subjectivity [11]. Consequently, personnel performing routine ultrasonographic examinations and fetal biometric measurements must undergo specialized training in diagnostic obstetric ultrasonography, including training related to ultrasonographic safety [12]. Estimation errors during this assessment process can significantly impair the detection of fetal growth disorders, including

Manuscript received July 25, 2025; revised October 13, 2025; accepted November 24, 2025; published May 27, 2026.

clinically important conditions such as fetal growth restriction [12].

Manual scanning and assessment of fetal anatomical structures in ultrasonographic images present several challenges. One critical factor influencing optimal assessment is the quality of the transducer equipment, which must incorporate advanced features to adequately support fetal anatomical evaluation [12]. Ultrasonographic systems should be equipped with two-dimensional transducers with real-time imaging capability, adjustable output power control, and image freezing and magnification functions. These features are essential for producing clear images, enabling sufficient magnification, and ensuring accurate caliper placement during measurement [12].

High image quality and consistent measurement techniques are fundamental for accurate gestational age determination and fetal size assessment [12]. A major limitation of ultrasound imaging is the presence of speckle noise, which degrades image quality [13]. Speckle noise arises from the interference of reflected sound waves from heterogeneous tissue structures, resulting in granular patterns that obscure fine anatomical details [14]. To mitigate this issue, various noise-reduction approaches have been proposed, including adaptive spatial filters such as the Lee and Frost filters, which rely on local statistical properties to suppress noise while preserving edges [13]. Wavelet-based transformation methods have also been developed to decompose images into frequency components, enabling selective noise reduction in specific sub-bands. This approach forms the basis of the Hybrid Speckle Noise Reduction (HSNR) method [15].

The HSNR approach is rooted in a series of our previous studies, in which speckle suppression strategies were developed and applied to various ultrasound imaging tasks [16–20]. In fetal imaging, this method has been successfully integrated into object detection frameworks, including combinations of aggregated channel features and Faster Region-based Convolutional Neural Networks (R-CNN), to improve localization accuracy under noisy conditions. Hybrid speckle reduction has also been applied to automated biometric measurements, such as femur length estimation using active contour localization and threshold-based segmentation of fetal thigh regions. Beyond obstetric imaging, the approach has been adapted for digital microscopy to enhance the visibility of microstructures in composite materials. More recently, our findings demonstrated that the quality and characteristics of training data play a critical role in the performance of deep learning-based speckle reduction, reinforcing the importance of hybrid filtering as a robust preprocessing strategy across both clinical and non-clinical domains.

Identifying standard fetal brain planes within complex anatomical structures remains a highly challenging task that typically requires extensive clinical experience to guide ultrasound probe positioning accurately [21]. In addition, unfavorable fetal positioning can further complicate image acquisition and reduce image clarity. The successful acquisition of standard brain planes is also strongly influenced by the image quality produced by ultrasonographic equipment. Factors such as transducer frequency, sound wave reflection angle, and noise levels significantly affect image quality [12]. Even with modern ultrasonographic systems, acoustic shadows and imaging artifacts may hinder optimal visualization of fetal brain anatomical structures [22].

Recent advances in big data analytics, pattern recognition, and biomedical imaging have enabled scalable artificial intelligence-driven diagnostic systems, strengthening the integration of deep learning techniques into clinical imaging workflows [23–25]. Artificial intelligence approaches, particularly those based on deep learning, have emerged as promising solutions for addressing challenges in ultrasonographic image acquisition and interpretation [26]. Deep learning employs artificial neural networks to analyze complex data patterns and offers high flexibility for modeling intricate relationships that influence diagnosis and prognosis [27, 28]. Numerous studies have demonstrated the effectiveness of deep learning in medical imaging applications, including tumor detection in magnetic resonance imaging and early lung cancer detection in computed tomography images [29, 30].

Among deep learning-based object detection algorithms, the You Only Look Once (YOLO) framework has demonstrated strong performance due to its end-to-end detection capability, high inference speed, and competitive accuracy [31]. Recent studies have reported the successful application of YOLOv8 for real-time detection of small and dynamic objects, achieving high detection accuracy and fast inference speeds [32]. These characteristics are particularly relevant for fetal anatomical structure detection, where both precision and real-time performance are critical.

This study develops an automated detection system based on the YOLOv8 architecture to identify key anatomical structures in first-trimester fetal head sagittal planes, including the nasal bone, brain midline, and nuchal translucency. The preprocessing stage employs a HSNR approach that combines wavelet transformation and bilateral filtering to preserve fine anatomical details such as brain ventricle boundaries [33]. YOLOv8 was selected due to its real-time detection capability and strong performance in detecting small anatomical structures, with anchor box configurations adapted to the size characteristics of first-trimester fetal features. The model was trained using transfer learning on semi-automatically annotated ultrasonographic images.

The proposed system enables consistent detection and measurement of clinically important parameters, including nuchal translucency thickness and sagittal midline symmetry, thereby supporting early screening for chromosomal abnormalities such as Down syndrome. The integration of adaptive speckle noise reduction with the YOLOv8 framework provides an efficient solution for first-trimester ultrasonographic interpretation, particularly in resource-limited clinical settings. While previous studies have applied earlier YOLO variants to fetal ultrasound imaging, this work uniquely quantifies the influence of HSNR on the visibility of subtle fetal head

features and evaluates the benefit of AdamW optimization with early stopping. These contributions enhance robustness under noisy imaging conditions and underscore the clinical utility of artificial intelligence-based fetal head assessment.

Automated detection of fetal head anatomical structures has the potential to reduce operator-dependent variability and minimize early diagnostic errors. Integrating such artificial intelligence-based tools into routine ultrasonographic workflows may assist clinicians by providing real-time feedback during first-trimester examinations, thereby improving diagnostic consistency and confidence.

II. Materials and Methods

A. Dataset

The dataset used in this study consists of sagittal fetal head ultrasound images obtained from a publicly available dataset introduced by Cui and Dong [34], which is hosted on the Mendeley Data repository at https://data.mendeley.com/datasets/n2rbrb9t4f. The original dataset contains 1,684 sagittal ultrasound images acquired from both internal and external sources. From this collection, a subset of 1,131 images was selected for the present study. The selected images were carefully validated to ensure that each image had a corresponding and correctly matched annotation file. Inter-rater annotation reliability between two radiologists was assessed using Cohen's kappa, yielding a value of 0.84, which indicates substantial agreement. A total of 553 images were excluded due to motion artifacts, incomplete visibility of the fetal head, or mismatched annotation files.

Fig. 1 presents an example of a sagittal fetal head ultrasound image that meets the criteria for a standard plane view, along with expert annotations of key intracranial structures. The annotated structures include the thalamus, midbrain, nuchal translucency, and nasal bone, as delineated by a senior radiologist. All images were preprocessed to a resolution of 320 by 320 pixels, with padding applied to preserve the original aspect ratio. Annotation was performed independently by two senior radiologists, each with more than ten years of experience in obstetric ultrasonography. The images were classified as standard or non-standard sagittal views, and only images with consensus agreement were retained as ground truth for model training.

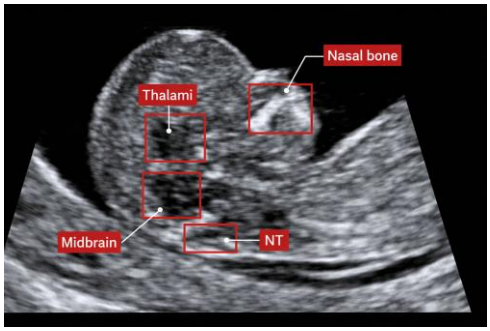


Fig. 1. Example of ultrasound images and the annotation of key intracranial structures.

The detection task in this study focuses on nine key anatomical structures of the fetal head during the first trimester. These structures include the thalamus, which serves as a primary relay center for neural signals; the midbrain, which is involved in motor function and reflex regulation; the nasal bone, which is an important marker in fetal assessment; the nasal tip, representing the anterior portion of the fetal nose; the nasal skin, which complements the nasal structure visually; the palate, which forms the roof of the oral cavity; the nuchal translucency, defined as a fluid-filled region at the posterior aspect of the fetal neck used for genetic risk assessment; additional intracranial structures relevant to brain anatomy; and the cisterna magna, which is the fluid-filled space located posterior to the cerebellum.

A semi-automated annotation process was employed to prepare the dataset for training. The original annotation files, provided in comma-separated values format, were converted into a YOLO-compatible text format using a custom conversion script. This process involved organizing all ultrasound images into a unified directory structure and aligning each image with its corresponding label file to form consistent image–label pairs suitable for YOLO-based training. The semi-automated workflow improved preprocessing efficiency and ensured consistent data formatting across the entire dataset. Annotation quality was further verified through cross-review by the annotators, and any inconsistencies were resolved through consensus. Automated scripts were also applied to verify bounding box alignment and class consistency prior to model training.

To ensure a balanced and representative data distribution, the dataset was partitioned using a proportional splitting strategy implemented with the split folders library. The final dataset was divided into training, validation, and testing subsets using an 80 to 10 to 10 ratios, resulting in 904 images for training, 113 images for validation, and 114 images for testing. A fixed random seed was applied to guarantee reproducibility of the data split, and all subsets were verified to contain no overlapping images. This dataset preparation strategy ensures data integrity and supports reliable deep learning model development, while preserving sufficient samples for objective validation and testing in accordance with standard machine learning practices.

B. Preprocessing

The Hybrid Speckle Noise Reduction (HSNR) method was applied to enhance ultrasound image quality while preserving critical anatomical structures [13]. This multi-stage preprocessing pipeline integrates wavelet decomposition, spatial filtering, and diffusion-based techniques adapted from established methodologies. Initially, each grayscale ultrasound image was decomposed using a Daubechies db8 wavelet basis across three hierarchical levels. This decomposition separated the image into four frequency sub-bands, consisting of a low-frequency approximation sub-band that contains structural information and three high-frequency detail sub-bands that capture edge-related features. Different processing strategies were applied to the resulting

sub-bands. The low-frequency approximation sub-band was processed using a bilateral filter with a 5 by 5 kernel to reduce noise while preserving organ boundaries and edge integrity. In parallel, the high-frequency detail sub-bands were subjected to soft thresholding to suppress low-amplitude noise components. Subsequently, anisotropic diffusion was applied for 50 iterations to further reduce speckle patterns while maintaining important edge structures. After sub-band processing, inverse wavelet transformation was performed to reconstruct the denoised ultrasound image.

The complete HSNR workflow required approximately 8.7 s per image, resulting in a total preprocessing time of 2 h and 44 min for the full dataset of 1,131 images. This procedure produced two parallel datasets for comparative analysis, consisting of the original ultrasound images and the HSNR-processed images. Fig. 2 illustrates representative examples of raw ultrasound images and their corresponding denoised results obtained using the proposed preprocessing method.

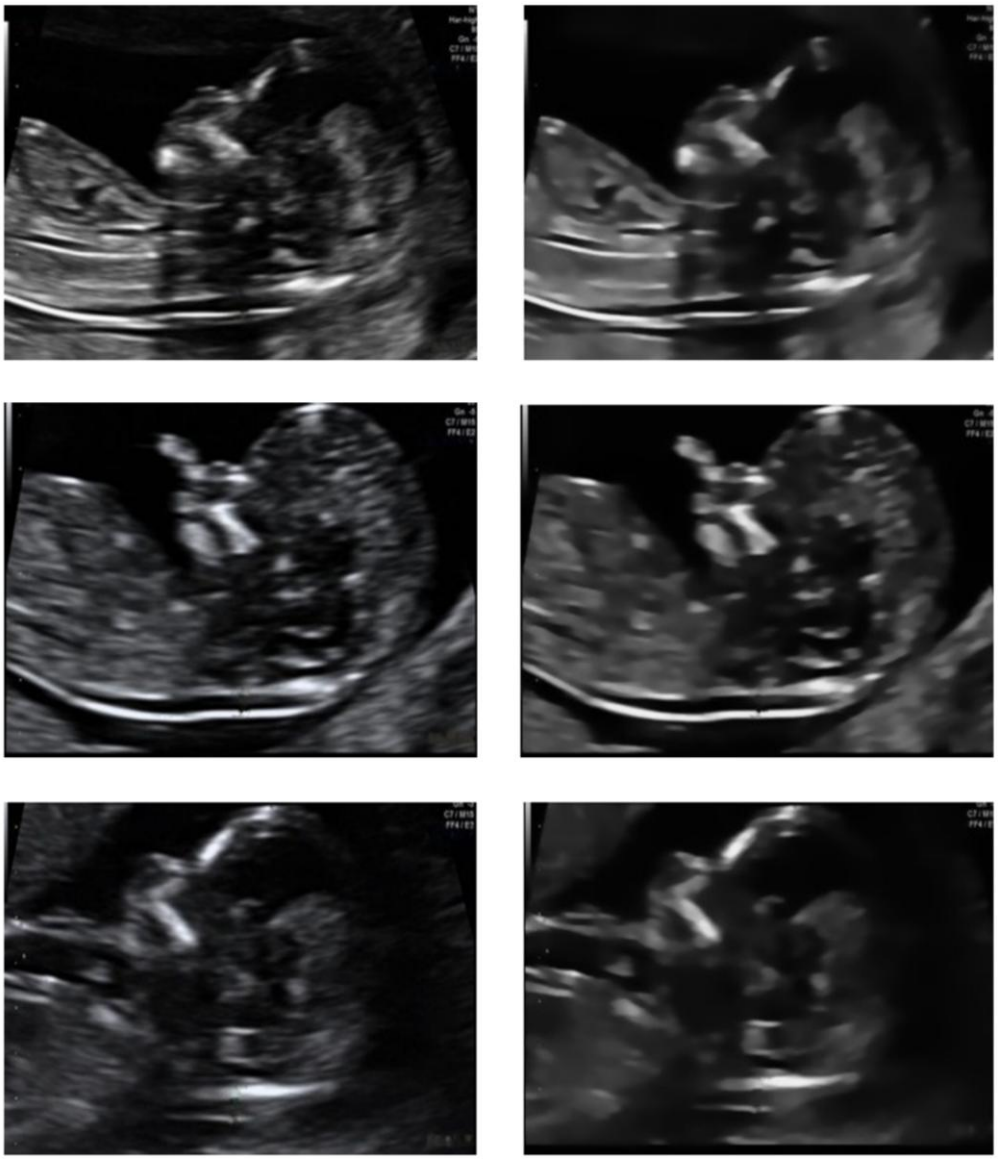


Fig. 2. Qualitative visual comparison between representative raw ultrasound images (left) and images processed using the HSNR method (right).

No artificial data augmentation, such as image flipping or rotation, was applied to preserve the anatomical realism of the fetal sagittal plane. Quantitative comparison between the raw and HSNR, processed images yielded an average Structural Similarity Index Measure (SSIM) value of 0.912 and a Peak Signal-to-Noise Ratio (PSNR) of 29.8 decibels, indicating effective speckle suppression with minimal loss of anatomical detail.

C. Model Architecture and Training

The detection of fetal head anatomical structures was performed after an optional preprocessing stage, such as speckle noise reduction, in which ultrasound images were processed end to end using the optimized YOLOv8 architecture, as illustrated in Fig. 3. The model employs a modified CSPDarknet53 backbone for multi-scale feature extraction. This backbone incorporates Cross Stage Partial

(CSP) blocks to reduce computational redundancy by splitting feature propagation paths while maintaining representational capacity.

The extracted features were subsequently fused in the neck component using a Path Aggregation Network combined with a Feature Pyramid Network (PAN-FPN). This structure integrates top-down information flow from deeper to shallower layers with bottom-up flow from shallower to deeper layers, thereby enhancing both spatial resolution and semantic richness. For object prediction, YOLOv8 adopts a decoupled and anchor-free detection head that directly estimates object centers. The classification task, which determines object categories, is separated from the regression task, which predicts bounding box coordinates, to minimize optimization conflicts during training. Binary cross-entropy loss was applied for classification, while Complete Intersection over Union (IoU) loss was used for bounding box regression, enabling accurate localization while preserving real-time inference efficiency.

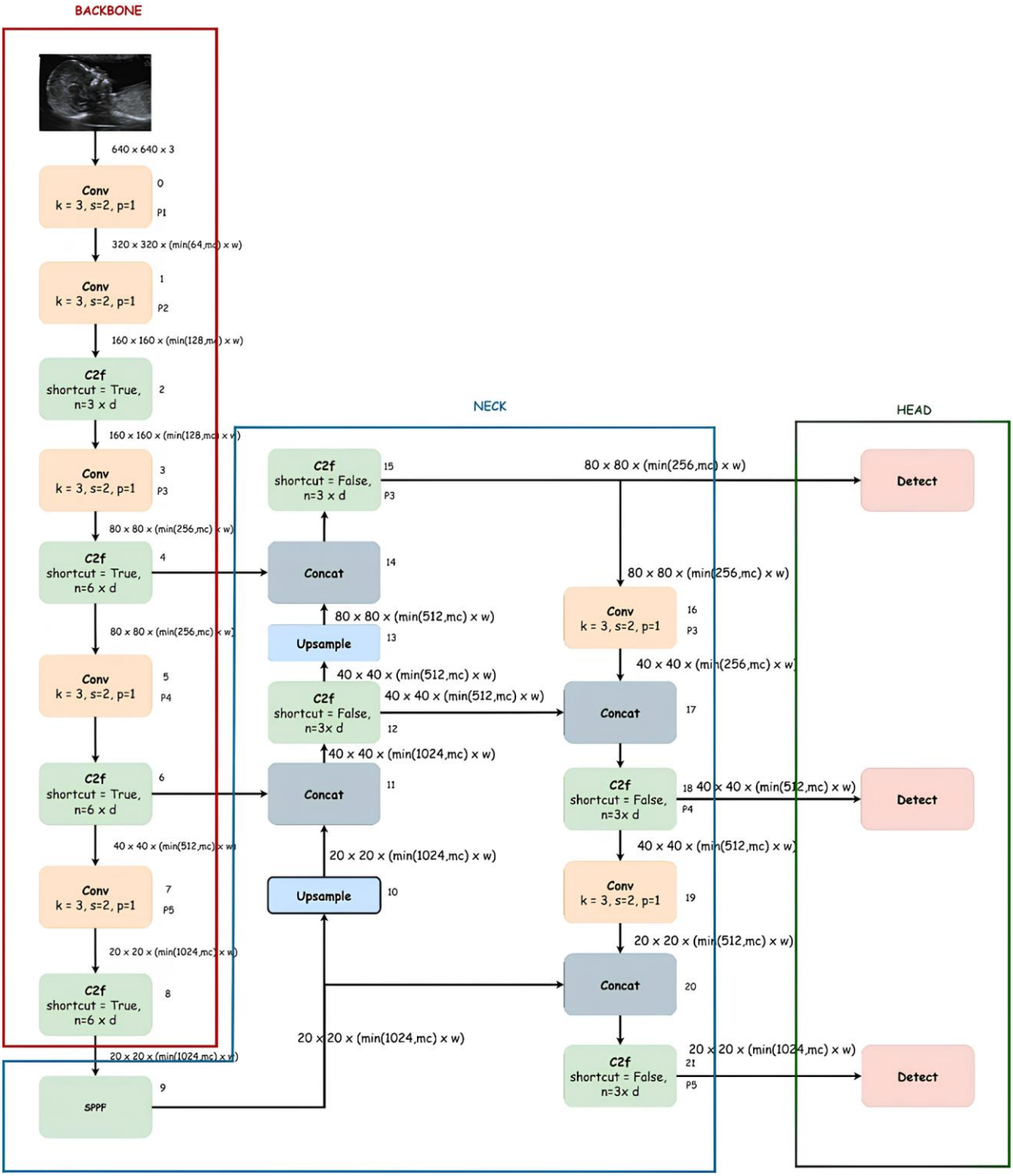


Fig. 3. Overall architecture of the YOLOv8-based detection framework used in this study.

The training pipeline began with dataset preparation, which included both raw ultrasound images and images processed using HSNR. The dataset was divided into training, validation, and testing subsets using an 80 to 10 to 10 ratios, ensuring balanced representation of anatomical structures such as the thalami, midbrain, and nasal region. Hyperparameter optimization explored several key configurations, including different optimizers such as Stochastic Gradient Descent (SGD), Adam, and AdamW, learning rates of 0.001 and 0.01, training durations of up to 200 epochs or early stopping with a patience of 25 epochs, and detection thresholds with confidence values ranging from 50 to 65 percent and IoU thresholds ranging from 50 to 60 percent. During training, performance metrics including precision, recall, mean average precision at an IoU threshold of 0.50, and mean average precision across thresholds from 0.50 to 0.95 were evaluated at each epoch to monitor learning progress and prevent overfitting.

The YOLOv8 model was trained and evaluated using both raw and HSNR-processed fetal ultrasound images to assess robustness under varying noise conditions. The training dataset consisted of 904 images, with 113 images reserved for validation, all resized to a resolution of 320 by 320 pixels and paired with corresponding annotation files. To ensure robust performance evaluation, a dual validation strategy was employed. Standard holdout validation was used for preliminary hyperparameter optimization, followed by stratified five-fold cross-validation to address class imbalance and assess generalization performance. During cross-validation, the dataset was partitioned into five folds while preserving proportional representation of all anatomical structures, with each fold serving as the validation set in rotation and the remaining folds used for training.

Training and validation loss curves were monitored using TensorBoard to detect overfitting and confirm convergence. The learning curves demonstrated stable validation performance across folds, indicating consistent model generalization. Hyperparameter optimization systematically evaluated configurations involving three optimizers, two learning rates, a fixed batch size of sixteen, and a maximum of two hundred training epochs, all executed using a fixed random seed of forty-two to ensure reproducibility. Validation metrics, including mean Average Precision (mAP) at 0.50, precision, and recall, were tracked per epoch, with automated checkpointing triggered by improvements in mean average precision. This integrated training and evaluation strategy enabled rigorous assessment of YOLOv8 performance under noisy imaging conditions while maintaining sensitivity to clinically important anatomical structures.

D. Experimental Design

This study was conducted in two main experimental stages. The first stage utilized raw ultrasound images without any preprocessing, while the second stage employed ultrasound images that had undergone Hybrid Speckle Noise Reduction (HSNR) preprocessing. From an initial collection of 1,528 ultrasound images, only 1,131 images were associated with usable annotation files. Consequently, a semi-automated labeling procedure was applied to convert the available annotations into a format compatible with the YOLO framework, following the specifications required by the Ultralytics implementation.

The final dataset consisted of 1131 sagittal fetal head ultrasound images that were semi-automatically annotated and divided into training and testing subsets using an 80 to 20 ratio, resulting in 1017 images for training and 114 images for testing. The training subset was further split into training and validation sets, comprising 904 and 113 images, respectively. The YOLOv8 model was trained using multiple hyperparameter configurations, including different optimization algorithms such as stochastic gradient descent, Adam, and AdamW, as well as learning rates of 0.001 and 0.01. Model performance was evaluated using standard detection metrics, including precision, recall, average precision, and mean average precision, computed after the completion of training. Inference was subsequently performed on the unseen test images to assess detection accuracy using the IoU metric.

A batch size of 16 was selected to balance memory usage and gradient stability during training. Learning rate scheduling followed a cosine decay strategy, and early stopping was applied when no improvement was observed for 25 consecutive epochs. An input resolution of 320 by 320 pixels was selected based on empirical evaluation to preserve sufficient anatomical detail while maintaining real-time inference capability.

To enhance model reliability and reduce the risk of overfitting associated with limited data, the experimental design incorporated five-fold cross-validation. Cross-validation was performed exclusively on the training subset, enabling each model configuration to be trained and validated across multiple data partitions. This approach ensured that performance metrics such as precision, recall, average precision, and mean average precision provided a more representative assessment of overall model performance. The best-performing models identified during cross-validation were subsequently evaluated on the independent test set to support both quantitative assessment and qualitative clinical interpretation of fetal head anatomical structure detection.

Fig. 4 illustrates the overall research workflow of the proposed fetal head anatomical structure detection system. The workflow begins with the acquisition of sagittal ultrasound images, followed by a semi-automated annotation process that converts labels into a YOLO-compatible format. Two parallel data preparation pipelines are then implemented. One pipeline uses raw ultrasound images, while the other uses images processed with HSNR. Both datasets are divided into training, validation, and testing subsets. The YOLOv8 model is trained using various hyperparameter configurations, including different optimizers and learning rates. Model evaluation is conducted using key performance metrics such as precision, recall, IoU, average precision, and mean average precision. Additional evaluation through five-fold cross-validation is performed to assess model generalization. Finally, the best-performing model is

applied to the test dataset to provide both qualitative and quantitative detection results that support clinical analysis.

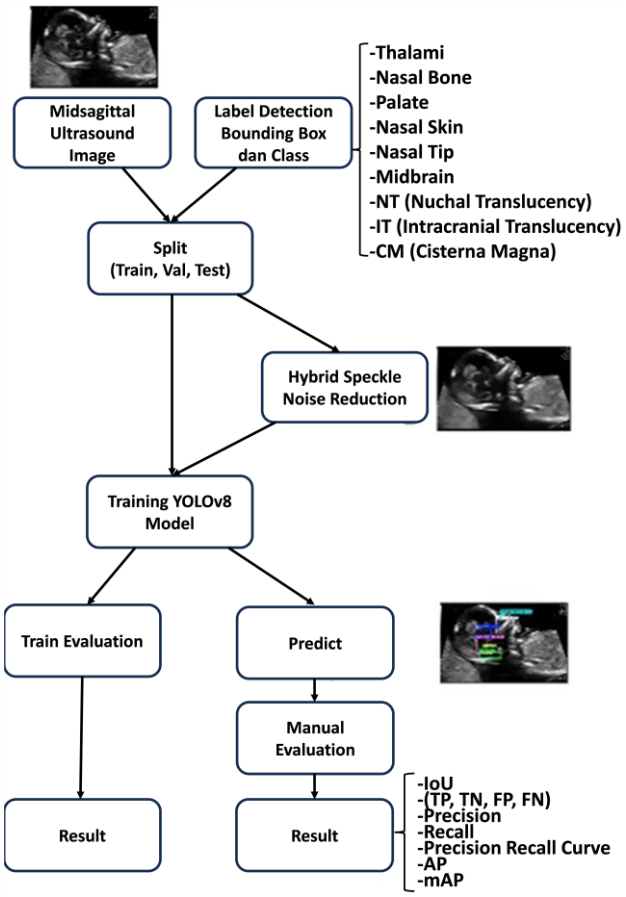


Fig. 4. Workflow of the proposed fetal head anatomical structure detection framework.

E. Evaluation Measures

In object detection modeling, performance evaluation is essential to assess the accuracy and reliability of object identification and localization. Several quantitative metrics are commonly used for this purpose, including IoU, precision, recall, Average Precision (AP), and mean Average Precision (mAP).

IoU is used to measure the degree of overlap between a predicted bounding box and the corresponding ground truth bounding box. This metric determines whether a detection is classified as a true positive based on a predefined overlap threshold. The IoU threshold therefore, serves as the fundamental criterion for computing subsequent evaluation metrics such as precision and recall. Unless otherwise specified, an IoU threshold of 0.5 was consistently applied throughout both training and evaluation to ensure fair comparison across all experimental configurations. The mathematical formulation of Intersection over Union (IoU) is presented in Eq. (1).

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \tag{1}$$

where *Area of Overlap* is the area of overlap between the prediction box and the truth box. *Area of Union* is the combined area of the prediction box and the truth box (including the overlap area).

Precision measures the proportion of correctly detected objects among all predicted detections, reflecting the model's ability to minimize false positive results. Precision is measured using Eq. (2).

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

where True Positives (*TP*) are correct predictions, i.e., objects that are correctly detected, and False Positives (*FP*) are incorrect predictions, i.e., objects that are detected but do not exist.

Recall measures the proportion of correctly detected objects among all ground truth instances, indicating the model's ability to identify relevant objects. The recall formula is as shown in Eq. (3).

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

where False Negatives (*FN*) are objects that exist but are not detected by the model.

Average Precision (*AP*) summarizes the precision–recall relationship by computing the area under the precision-recall curve for a given object class. Eq. (4) formulates *AP*:

$$AP = \frac{1}{N}\sum_{i=1}^{N} Precision(i) \tag{4}$$

where N is the number of recall points calculated on the Precision-Recall curve and $Precision(i)$ is the precision value at the i-th recall point.

Mean Average Precision (*mAP*) represents the mean of average precision values across all detected classes and serves as a comprehensive indicator of overall detection performance. The *mAP* formula is stated in Eq. (5).

$$mAP = \frac{1}{C}\sum_{c=1}^{C} AP_c \tag{5}$$

where C is the number of classes in the dataset, and AP_c is the average precision for the c-th class. In this study, mean average precision was calculated at an IoU threshold of 0.5, as well as across multiple thresholds to evaluate detection robustness under varying overlap requirements.

III. Result and Discussion

A. Training Evaluation

The evaluation of model performance in this study was conducted using a comprehensive multi-scenario framework that included full training runs, early stopping strategies, and stratified cross-validation. Each evaluation

scenario was designed to characterize the learning behavior, generalization capability, and robustness of the model under different optimization and preprocessing conditions.

Under the full training scheme with a maximum of 200 epochs, the YOLOv8 configuration using the Adam optimizer with a learning rate of 0.01 demonstrated the strongest overall performance. As reported in Table I, this configuration achieved the highest mean average precision at an IoU threshold of 0.50 with a value of 0.951, along with a balanced mean average precision across thresholds from 0.50 to 0.95 of 0.583. The model exhibited a favorable balance between precision and recall, with values of 0.913 and 0.927, respectively. This result indicates effective sensitivity and specificity in object detection. The performance gains were particularly evident for large and clearly defined anatomical structures such as the midbrain and thalami, highlighting the model's capability to capture anatomically prominent features.

TABLE I. PERFORMANCE COMPARISON OF YOLOV8 WITH AND WITHOUT HSNR UNDER 200 EPOCH CONFIGURATION

Configuration	Performance Measure			
	Precision	Recall	mAP@50	mAP@50:95
YOLOv8 Original (SGD, LR 0.001)	0.873	0.935	0.934	0.553
YOLOv8 + HSNR (SGD, LR 0.001)	0.908	0.913	0.941	0.546
YOLOv8 Original (SGD, LR 0.01)	0.901	0.926	0.943	0.581
YOLOv8 + HSNR (SGD, LR 0.01)	0.902	0.926	0.946	0.564
YOLOv8 Original (Adam, LR 0.001)	0.900	0.917	0.951	0.576
YOLOv8 + HSNR (Adam, LR 0.001)	0.899	0.941	0.946	0.572
YOLOv8 Original (Adam, LR 0.01)	0.913	0.927	0.951	0.583
YOLOv8 + HSNR (Adam, LR 0.01)	0.899	0.941	0.946	0.572

In contrast, the early stopping strategy provided a competitive alternative with reduced training time. The configuration using the AdamW optimizer, which converged at 252 epochs, achieved a mean average precision at 0.50 of 0.907 and a mean average precision across thresholds of 0.581, as shown in Table II. This evaluation was performed using an IoU threshold of 0.6 and a confidence threshold of 0.65, ensuring that only spatially accurate and high-confidence detections were included. This configuration yielded a high precision of 0.950 and a recall of 0.857, indicating stable training dynamics and improved resistance to overfitting. The results demonstrate that early stopping can provide an effective balance between detection accuracy and computational efficiency for first-trimester prenatal ultrasound analysis.

TABLE II. PERFORMANCE COMPARISON OF YOLOV8 WITH AND WITHOUT HSNR UNDER EARLY STOPPING CONFIGURATION

Configuration	Performance Measure			
	Precision	Recall	mAP@50	mAP@50:95
YOLOv8 + Adam (103 Epoch)	0.945	0.743	0.848	0.533
YOLOv8 + HSNR (Adam, 98 Epoch)	0.957	0.697	0.831	0.537
YOLOv8 + AdamW (252 Epoch)	0.950	0.857	0.907	0.581
YOLOv8 + HSNR (AdamW, 85 Epoch)	0.948	0.730	0.845	0.537
YOLOv8 + SGD (110 Epoch)	0.943	0.814	0.885	0.567
YOLOv8 + HSNR (SGD, 54 Epoch)	0.927	0.812	0.876	0.542
YOLOv8 + Adam (103 Epoch)	0.945	0.743	0.848	0.533
YOLOv8 + HSNR (Adam, 98 Epoch)	0.957	0.697	0.831	0.537

Further evaluation using stratified five-fold cross-validation reinforced the generalization capability of the YOLOv8 model trained without HSNR preprocessing. This evaluation employed the AdamW optimizer with a confidence threshold of 0.65 and an IoU threshold of 0.6 to focus on reliable detections. As summarized in Table III, the model achieved a precision of 0.968 and a mean average precision at 0.50 of 0.844, demonstrating consistent performance across unseen data distributions. The recall value of 0.711 indicates a more conservative detection behavior, prioritizing precision and spatial accuracy over sensitivity to ambiguous anatomical regions. This behavior is particularly relevant in clinical contexts, where false-positive detections may negatively impact diagnostic confidence.

TABLE III. PERFORMANCE COMPARISON OF YOLOV8 WITH AND WITHOUT HSNR UNDER CROSS VALIDATION CONFIGURATION

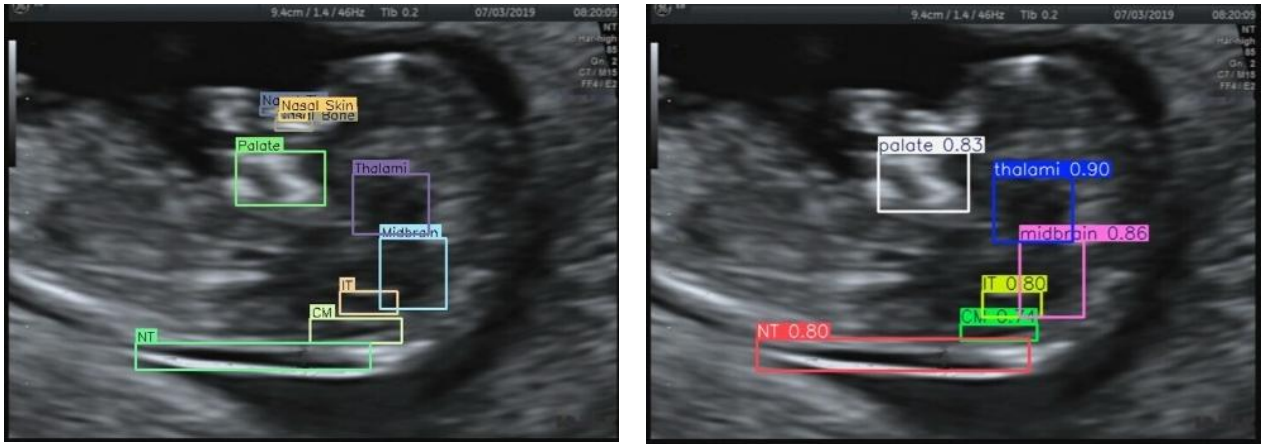
Configuration	Performance Measure			
	Precision	Recall	mAP@50	mAP@50:95
YOLOv8 Stratified Fold-5	0.968	0.711	0.844	0.549
YOLOv8 + HSNR Stratified Fold-5	0.964	0.697	0.832	0.522

B. Object Detection Evaluation

To evaluate the robustness and consistency of the proposed detection model under varying image quality conditions, a qualitative object detection analysis was conducted using three representative ultrasound samples. Each sample reflects a different imaging scenario commonly encountered in clinical practice, including high speckle noise, low contrast, and optimal image quality. For

each sample, detection performance was analyzed using IoU scores, classification outcomes, and overall detection accuracy, providing insight into model behavior under both challenging and favorable conditions.

Sample 1, shown in Fig. 5, represents an ultrasound image affected by high speckle noise, resulting in reduced sharpness and limited anatomical visibility. This sample was selected to assess model robustness under suboptimal visual conditions. As summarized in Table IV, the model correctly detected five of the nine annotated anatomical structures, resulting in a detection accuracy of 55.56%. All true positive detections achieved IoU values above 0.6, including the thalami with an IoU of 0.8229, the midbrain with an IoU of 0.8747, and nuchal translucency with an IoU of 0.7248. One false positive detection was observed for the cisterna magna, while four anatomically small or low-contrast structures were not detected. These results indicate that although the model can reliably identify dominant anatomical regions under noisy conditions, subtle features remain more susceptible to misdetection.



(a) (b)

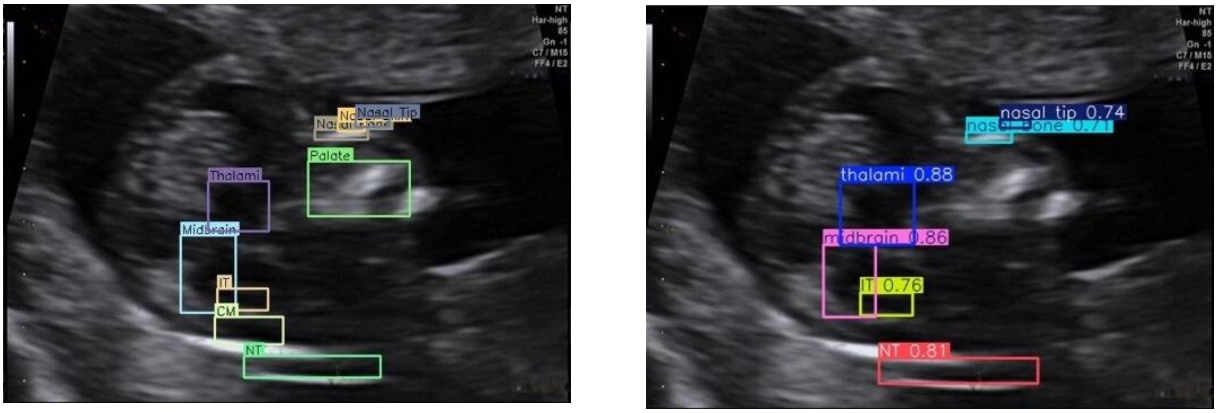
Fig. 5. Qualitative comparison of YOLOv8 detection results on a representative midsagittal fetal ultrasound image from Sample 1: (a) expert-provided ground-truth annotations, (b) YOLOv8 prediction results with confidence scores. The model detects thalami, midbrain, palate, Nuchal Translucency (NT), Intracranial Translucency (IT), and Cisterna Magna (CM), showing good visual agreement with expert annotations.

TABLE IV. PERFORMANCE OF BOUNDING BOX MATCHING RESULTS FOR SAMPLE 1 (PRED: MODEL PREDICTION)

No.	Ground Truth (GT)	Prediction	IoU Score	Classification Result
0	thalami	thalami	0.8229	True Positive
1	midbrain	midbrain	0.8747	True Positive
2	palate	palate	0.8680	True Positive
3	NT	NT	0.7248	True Positive
4	IT	IT	0.7611	True Positive
5	-	CM	0.5653	False Positive
6	nasal bone	-	0.0000	False Negative
7	nasal tip	-	0.0000	False Negative
8	CM	-	0.0000	False Negative
9	nasal skin	-	0.0000	False Negative

Sample 2, illustrated in Fig. 6, depicts an ultrasound image with reduced brightness and limited contrast, which obscures anatomical boundaries. This scenario was used to evaluate detection performance under low-contrast imaging conditions. As presented in Table V, the model successfully detected six of nine anatomical structures, achieving a detection accuracy of 66.67%. All detected structures were classified as true positives with high IoU values, including the midbrain with an IoU of 0.8576, intracranial translucency with an IoU of 0.8558, and nuchal translucency with an IoU of 0.7456. No false positive detections were observed. However, three low-contrast structures were not detected, indicating that contrast variability remains a challenge for small-structure recognition.

Sample 3, shown in Fig. 7, represents a high-quality ultrasound image with clear contrast and well-defined anatomical features. This sample was selected to assess model performance under optimal imaging conditions. As summarized in Table VI, the model correctly identified seven of nine annotated structures, resulting in a detection accuracy of 77.78%. Several detections achieved high IoU values, including the nasal bone with an IoU of 0.9310, the palate with an IoU of 0.8665, and the midbrain with an IoU of 0.8531. Two false positive detections and two false negative detections were observed, primarily involving anatomically small structures. Despite these errors, the model demonstrated strong detection performance when applied to high-quality input data.

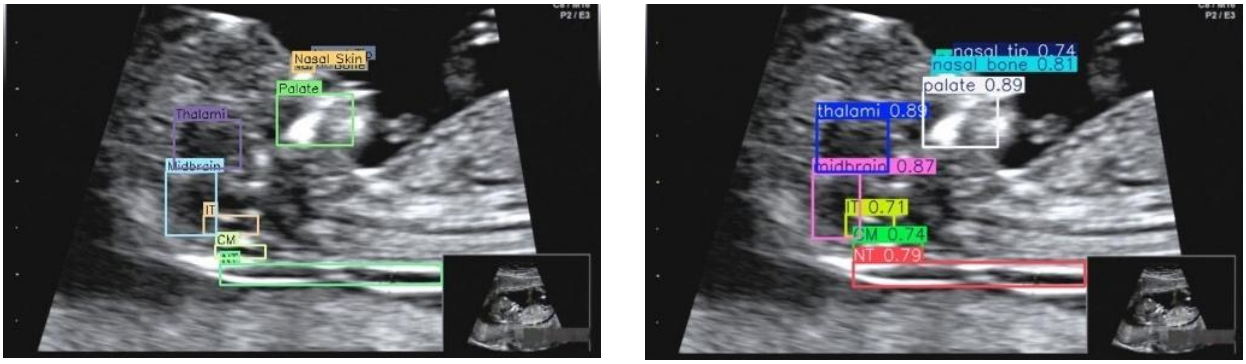


(a) (b)

Fig. 6. Qualitative comparison of YOLOv8 detection results on a representative midsagittal fetal ultrasound image from Sample 2: (a) Expert-provided ground-truth annotations, (b) YOLOv8 prediction results with confidence scores. The model detects key anatomical structures, including the thalami, midbrain, Nuchal Translucency (NT), Intracranial Translucency (IT), palate, and Cisterna Magna (CM). Higher confidence is observed for well-defined structures such as the midbrain and IT, while relatively lower confidence appears for smaller or less distinct structures, which is consistent with the characteristics of first-trimester ultrasound imaging.

TABLE V. PERFORMANCE OF BOUNDING BOX MATCHING RESULTS FOR SAMPLE 2 (PRED: MODEL PREDICTION)

No.	Ground Truth (GT)	Prediction	IoU Score	Classification Result
0	thalami	thalami	0.6645	True Positive
1	midbrain	midbrain	0.8576	True Positive
2	NT	NT	0.7456	True Positive
3	IT	IT	0.8558	True Positive
4	nasal tip	nasal tip	0.6098	True Positive
5	nasal bone	nasal bone	0.6203	True Positive
6	CM	-	0.0000	False Negative
7	palate	-	0.0000	False Negative
8	nasal skin	-	0.0000	False Negative
9	thalami	thalami	0.6645	True Positive



(a) (b)

Fig. 7. Qualitative comparison of YOLOv8 detection results on a representative midsagittal fetal ultrasound image from Sample 3. The figure shows: (a) The expert-provided ground-truth annotations; (b) The YOLOv8 prediction results with confidence scores. The model detects key anatomical structures, including the thalami, midbrain, palate, nuchal translucency, intracranial translucency, and cisterna magna, showing good visual agreement with expert annotations.

TABLE VI. PERFORMANCE OF BOUNDING BOX MATCHING RESULTS FOR SAMPLE 3 (PRED: MODEL PREDICTION)

No.	Ground Truth (GT)	Prediction	IoU Score	Classification Result
0	palate	palate	0.8665	True Positive
1	thalami	thalami	0.8298	True Positive
2	midbrain	midbrain	0.8531	True Positive
3	nasal bone	nasal bone	0.9310	True Positive
4	NT	NT	0.7952	True Positive
5	-	CM	0.5200	False Positive
6	-	nasal tip	0.5655	False Positive
7	IT	IT	0.7768	True Positive
8	nasal skin	nasal skin	0.7851	True Positive
9	CM	-	0.0000	False Negative
10	nasal tip	-	0.0000	False Negative

The experimental results demonstrate that the original YOLOv8 model, particularly when trained using the AdamW optimizer with early stopping, consistently outperformed other configurations across multiple evaluation strategies. This configuration achieved a mean average precision at 0.50 of 0.907 and a recall of 0.857, indicating a strong balance between detection accuracy and sensitivity. Although training with the Adam optimizer using a learning rate of 0.01 produced a higher mean average precision at 0.50 of 0.951, this configuration was associated with higher computational cost and reduced generalization performance.

The application of HSNR showed mixed effects on detection performance. While this preprocessing approach improved visual clarity in certain cases, it frequently reduced the model's ability to detect small and subtle anatomical structures due to over-smoothing effects. Cross-validation results further supported the stability of the original YOLOv8 model trained on raw ultrasound images. In addition, per-sample qualitative evaluations confirmed that detection accuracy was strongly influenced by image quality, ranging from 77.78% for high-quality images to 55.56% for images affected by severe speckle noise.

These findings highlight a clear trade-off between visual enhancement and spatial detail preservation in ultrasound image preprocessing and confirm that YOLOv8 trained without HSNR provides a more reliable detection pipeline under varied clinical imaging conditions.

Compared with Faster R-CNN and EfficientDet baseline models, which achieved mean average precision at 0.50 values of 0.89 and 0.87, respectively, the YOLOv8 model demonstrated higher precision and recall for small fetal head anatomical structures, confirming its suitability for real-time fetal head detection. Inference experiments further showed that the trained YOLOv8 model achieved an average processing speed of 45 frames per second on an NVIDIA RTX 3060 graphics processing unit and 5 frames per second on a central processing unit, supporting its feasibility for deployment in clinical ultrasound systems. Statistical analysis using a paired t-test with a significance level of p less than 0.05 confirmed that the performance differences between models trained on raw images and HSNR-processed images were statistically significant.

Per-class average precision analysis revealed that nasal skin and nuchal translucency were the anatomical structures most affected by residual speckle noise, with average precision values of 0.68 and 0.71, respectively. These reductions explain the observed recall degradation under high-noise imaging conditions.

The per-class average precision results presented in Table VII show that large and high-contrast anatomical structures such as the thalami and midbrain consistently achieved high detection accuracy, with average precision at 0.50 values greater than or equal to 0.98 under both raw and HSNR conditions. In contrast, smaller and low-contrast regions, particularly nuchal translucency and the cisterna magna, exhibited substantial decreases in average precision when preprocessing was applied, with reductions of 0.25 and 0.22, respectively. Although HSNR improved overall image clarity, these results indicate that the preprocessing procedure can suppress fine textural cues that are critical for recognizing small anatomical structures. On average, the application of HSNR resulted in a 2.8% decrease in mean average precision at 0.50 across all anatomical classes, confirming that raw ultrasound images preserve more discriminative information for subtle boundary detection.

TABLE VII. PER-CLASS AP PERFORMANCE COMPARISON BETWEEN RAW AND HSNR IMAGES

No.	Structure	Raw AP@0.5	HSNR AP@0.5	Δ (Difference)
1	Thalami	0.990	0.992	+0.002
2	Nasal Bone	0.843	0.798	-0.045
3	Palate	0.937	0.965	+0.028
3	Nasal Skin	0.817	0.792	-0.025
4	Nasal Tip	0.887	0.865	-0.022
5	Midbrain	0.995	0.981	-0.014
6	NT	0.866	0.611	-0.255
7	IT	0.940	0.941	+0.001
8	CM	0.888	0.663	-0.225

Best Configuration: AdamW + Early Stopping.

As illustrated in Fig. 8, both mean average precision at 0.50 and recall were higher for raw ultrasound images, with values of 0.907 and 0.857, respectively, compared with HSNR, the processed images, which achieved values of 0.845 and 0.730. Quantitative analysis using SSIM and PSNR, together with qualitative visual inspection, showed that HSNR tended to over-smooth low-contrast edges, thereby reducing detection confidence for nuchal translucency and nasal skin. Visual comparisons in Figs. 5–7 further illustrate this loss of fine-grained anatomical cues. Although preprocessing improved overall image clarity, the smoothing of low-contrast boundaries slightly reduced the detector's sensitivity to small anatomical features such as nuchal translucency and the cisterna magna. These findings confirm that preserving raw ultrasound texture is advantageous for fine-structure recognition in first-trimester fetal imaging.

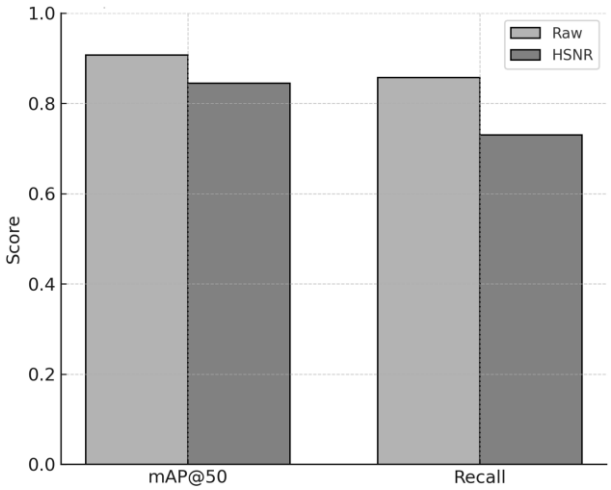


Fig. 8. Comparative mAP and recall between raw and HSNR-processed ultrasound images using YOLOv8 (AdamW + early stopping configuration).

A subset of detection results was independently reviewed by two senior radiologists, who rated 89% of bounding box detections as clinically acceptable for nuchal translucency and midline assessment. This expert evaluation confirms that the trained YOLOv8 model produces detection results that align with clinical diagnostic expectations in practical scenarios.

C. Limitations

Despite the competitive results achieved in this study, several limitations should be acknowledged. First, the dataset size remains relatively limited, comprising 1131 ultrasound images, which restricts the model's exposure to a broader range of anatomical variations and imaging conditions. External validation using datasets acquired from different ultrasound devices and multiple clinical centers is, therefore, necessary to further assess model generalizability.

In addition, the robustness of the proposed approach under extremely noisy or low-contrast imaging conditions has not been systematically quantified. The potential influence of fetal pose variability and differences in sonographer scanning technique was also not explicitly evaluated in this study. Future investigations should address these factors to establish stronger clinical robustness and to better characterize model performance across diverse real-world examination scenarios.

IV. Conclusion

This study demonstrates that the YOLOv8 object detection model, when trained using the AdamW optimizer with early stopping, provides a robust and computationally efficient approach for identifying key fetal head anatomical structures in first-trimester ultrasound images. Among the evaluated model configurations and preprocessing strategies, the original YOLOv8 model trained on raw ultrasound images consistently delivered the most stable and accurate detection performance, particularly for small and subtle anatomical features.

Although HSNR improved overall visual clarity, the results indicate that this preprocessing approach can introduce over-smoothing effects that reduce spatial precision and compromise the detection of fine anatomical details. These findings highlight the importance of preserving raw ultrasound texture and underscore the need for selective or adaptive denoising strategies in fetal ultrasound analysis. Detection performance was also shown to be strongly influenced by image quality, reinforcing the necessity of robust models that can generalize across varied acquisition conditions commonly encountered in clinical practice.

Overall, the results of this study provide a solid foundation for the development of artificial intelligence-assisted prenatal screening systems that balance detection accuracy, computational efficiency, and clinical applicability in early pregnancy assessment.

Future work will focus on validating the proposed model using ultrasound data acquired from different machines and clinical centers to further assess generalizability and robustness. In addition, efforts will be directed toward integrating the YOLOv8-based detection pipeline into artificial intelligence–assisted prenatal screening platforms to support real-time diagnostic assistance in routine clinical environments.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Fajar A. Hermawati conceptualized the research, supervised the methodology, contributed to data interpretation, and led the manuscript writing and revision.

Danara D. Caesa implemented the object detection model, performed all experiments and evaluations, and contributed to the initial draft of the manuscript.

Both authors reviewed and approved the final version of the manuscript.

Acknowledgment

The authors wish to express their sincere gratitude to the Department of Informatics, Universitas 17 Agustus 1945 Surabaya, and to the Institute for Research and Community Service (LPPM) of Universitas 17 Agustus 1945 Surabaya for their support and facilitation throughout the course of this research.

References

[1] J. Ibrahimi and Z. Mumtaz, "Ultrasound imaging and the culture of pregnancy management in low- and middle-income countries: A systematic review," *Int. J. Gynecol. Obstet.*, vol. 165, pp. 76–93, 2024. doi:10.1002/ijgo.15097

[2] C. Kollmann, K. V. Jenderka, C. M. Moran *et al*., "EFSUMB clinical safety statement for diagnostic ultrasound—(2019 revision)," *Ultraschall in der Medizin-European Journal of Ultrasound*, vol. 41, no. 4, pp. 387–389, 2020. doi: 10.1055/a-1010-6018

[3] K. Bronsgeest, E. E. Lust, L. Henneman *et al*., "Current practice of first-trimester ultrasound screening for structural fetal anomalies in developed countries," *Prenat. Diagn.*, vol. 43, no. 7, pp. 873–880, 2023. doi: 10.1002/pd.6389

[4] P. Y. Tsai, C. H. Hung, C. Y. Chen, and Y. N. Sun, "Automatic fetal middle sagittal plane detection in ultrasound using generative adversarial network," *Diagnostics*, vol. 11, no. 1, 21, 2021. doi: 10.3390/diagnostics11010021

[5] C. M. Bilardo, R. Chaoui, J. A. Hyett *et al*., "ISUOG practice guidelines (updated): Performance of 11–14-week ultrasound scan," *Ultrasound in Obstetrics and Gynecology*, vol. 61, no. 1, pp. 127–143, 2023. doi: 10.1002/uog.26106

[6] B. Bromley and L. D. Platt, "First-trimester ultrasound screening in routine obstetric practice," *Obstetrics and Gynecology*, vol. 143, no. 6, pp. 730–744, 2024. doi: 10.1097/AOG.0000000000005594

[7] B. Bromley, C. Henningsen, D. C. Jones, and I. Timor-Tritsch, "AIUM practice parameter for the performance of detailed diagnostic obstetric ultrasound examinations between 12 weeks 0 days and 13 weeks 6 days," *Journal of Ultrasound in Medicine*, vol. 40, no. 5, pp. e1–e16, 2021. doi: 10.1002/jum.15477

[8] H. Herlambang, *Basic Obstetric Ultrasonography: A Reference Guide for Medical Students and General Practitioners*, Jambi, Indonesia: Salim Media Indonesia, 2021. https://www.salimmedia.com/product/ultrasonografi-obstetri-dasar-mahasiswa-kedokteran-dan-dokter-umum/ (In Indonesian)

[9] A. I. Popa, N. Cernea, M. C. Marinaș *et al*., "Ultrasound screening in the first and second trimester of pregnancy for the detection of fetal cardiac anomalies in a low-risk population," *Diagnostics*, vol. 15, no. 6, p. 769, 2025. doi: 10.3390/diagnostics15060769

[10] V. Pruthi, N. Abbasi, V. Thakur *et al*., "Performance of comprehensive first trimester fetal anatomy assessment," *Prenat. Diagn.*, vol. 43, no. 7, pp. 881–888, 2023. doi: 10.1002/pd.6360
[11] L. Liu, T. Wang, W. Zhu *et al*., "Intelligent quality assessment of ultrasound images for fetal nuchal translucency measurement during the first trimester of pregnancy based on deep learning models," *BMC Pregnancy Childbirth*, vol. 25, no. 1, 741, 2025. doi: 10.1186/s12884-025-07863-y
[12] L. J. Salomon, Z. Alfirevic, F. D. S. Costa *et al*., "ISUOG practice guidelines: Ultrasound assessment of fetal biometry and growth," *Ultrasound in Obstetrics and Gynecology*, vol. 53, no. 6, pp. 715–723, 2019. doi: 10.1002/uog.20272
[13] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Hybrid speckle noise reduction method for abdominal circumference segmentation of fetal ultrasound images," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 3, pp. 2088–8708, 2018. doi: 10.11591/ijece.v8i3
[14] S. Pradeep and P. Nirmaladevi, "A review on speckle noise reduction techniques in ultrasound medical images based on spatial domain, transform domain and CNN methods," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1055, no. 1, 012116, 2021. doi: 10.1088/1757-899x/1055/1/012116
[15] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Hybrid speckle noise reduction method for abdominal circumference segmentation of fetal ultrasound images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8, no. 3, pp. 1747–1757, 2018.
[16] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Combination of Aggregated Channel Features (ACF) detector and faster R-CNN to improve object detection performance in fetal ultrasound images," *International Journal of Intelligent Engineering and Systems*, vol. 11, no. 6, pp. 65–74, 2018. doi: 10.22266/ijies2018.1231.07
[17] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Phase-based thresholding schemes for segmentation of fetal thigh cross-sectional region in ultrasound images," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 7, pp. 4448–4460, 2022. doi: 10.1016/j.jksuci.2021.02.004
[18] F. A. Hermawati, I. M. Kastiawan, and Muhyin, "Digital microscopy image enhancement technique for microstructure image analysis of bottom ash particle polymer composites," in *Proc. Advanced Materials: Physics and Mechanics of New Materials and Their Applications*, 2020, pp. 235–244. doi: 10.1007/978-3-030-45120-2_20
[19] F. A. Hermawati, H. Tjandrasa, Sugiono, G. I. P. Sari, and A. Azis, "Automatic femur length measurement for fetal ultrasound image using localizing region-based active contour method," *J. Phys. Conf. Ser.*, vol. 1230, no. 1, 012002, 2019. doi: 10.1088/1742-6596/1230/1/012002
[20] F. A. Hermawati, E. Ronando, and D. H. Sulistyawati, "Impact of training data quality on deep speckle noise reduction in ultrasound images," in *Proc. 2023 7th International Conference on Computational Biology and Bioinformatics (ICCBB 2023)*, 2023, pp. 61–65. doi: 10.1145/3638569.3638578
[21] R. Qu, G. Xu, C. Ding, W. Jia, and M. Sun, "Deep learning-based methodology for recognition of fetal brain standard scan planes in 2D ultrasound images," *IEEE Access*, vol. 8, pp. 44443–44451, 2020. doi: 10.1109/ACCESS.2019.2950387
[22] G. Sari and Sriyatun, *USG Abdomen*, 1st ed. Jakarta, Indonesia: Lembaga Manajemen Terapan (TRUSTCO) Jakarta, 2022.
[23] L. Pinto-Coelho, "How artificial intelligence is shaping medical imaging technology: A survey of innovations and applications," *Bioengineering,* vol. 10, no. 12, 1435, 2023. doi: 10.3390/bioengineering10121435
[24] D. H. Hagos, S. K. Aryal, P. Ymele-Leki, and L. L. Burge, "AI-driven multimodal colorimetric analytics for biomedical and behavioral health diagnostics," *Computational and Structural Biotechnology Journal*, vol. 27, pp. 2219–2232, 2025. doi: 10.1016/j.csbj.2025.05.015
[25] M. Y. Shakor and M. I. Khaleel, "Recent advances in big medical image data analysis through deep learning and cloud computing," *Electronics*, vol. 13, no. 24, 4869, 2024. doi: 10.3390/electronics13244860
[26] Y. T. Shen, L. Chen, W. W. Yue, and H. X. Xu, "Artificial intelligence in ultrasound," vol. 139, 109717, 2021. doi: 10.1016/j.ejrad.2021.109717
[27] S. Ghosh, N. Das, I. Das, and U. Maulik, "Understanding deep learning techniques for image segmentation," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–35, 2020. doi: 10.1145/3329784
[28] H. R. Maier, S. Galelli, S. Razavi *et al*., "Exploding the myths: An introduction to artificial neural networks for prediction and forecasting," *Environmental Modelling & Software*, vol. 167, 105776, 2023. doi: 10.1016/j.envsoft.2023.105776
[29] T. B. Nguyen-Tat, N. H. Nghia, and V. M. Ngo, "Enhancing brain tumor segmentation in MRI images: A hybrid approach using UNet, attention mechanisms, and transformers," *Egyptian Informatics Journal*, vol. 7, 100528, 2024. doi: 10.13140/RG.2.2.18164.36485
[30] M. Mamun, I. Mahmud, M. Meherin, and A. Abdelgawad, "LCDctCNN: Lung cancer diagnosis of CT scan images using CNN based model," in *Proc. 2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2023, pp. 205–212.
[31] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91
[32] D. Reis, J. Kupec, J. Hong, and A. Daoudi, "Real-time flying object detection with YOLOv8," arXiv Preprint, arXiv:2305.09972, 2024.
[33] F. A. Hermawati, H. Tjandrasa, and N. Suciati, "Hybrid speckle noise reduction method for abdominal circumference segmentation of fetal ultrasound images," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 3, pp. 1747–1757, 2018. doi: 10.11591/IJECE.V8I3.PP1747-1757
[34] C. Cui and F. Dong. (2022). Dataset for fetus framework. [Online]. Available: https://data.mendeley.com/datasets/n2rbrb9t4f/1