Biomedical Signal Processing and Control 82 (2023) 104535

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

# Ellipse guided multi-task network for fetal head circumference measurement

Jinting Wang, Zhiwen Fang *, Sheng Yao *, Feng Yang **

*School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China*
*Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou, 510515, China*
*Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Southern Medical University, Guangzhou, 510515, China*

ARTICLE INFO

*Keywords:*
Ultrasound imaging
Fetal head circumference measurement
Multi-task learning
Ellipse shape constrain

ABSTRACT

Precise fetal head circumference measurement by ultrasound imaging is of great significance for prenatal examination. However, missing or blurring boundaries caused by artifacts and noises challenge measurement accuracy. The inconsistency between the segmentation pseudo-label and the ellipse contours also generates measurement errors. To improve the measurement performances of fetal head circumference, in this study, we propose an ellipse-guided multi-task network that measures the fetal head circumference according to detected ellipse boundary pixels. In the proposed network, an region segmentation branch is designed to learn region features of the fetal head, and a feature fusion module is applied to combine region features with boundary features, which contribute to exploring more context information about the fetal head and locating boundary pixels in boundary missing or blurring regions. A loss function is also designed in the network to ensure the boundary estimation in an ellipse shape. Experiments are conducted on both the public fetal head circumference measurement dataset HC-18 and a self-built ultrasonic phantom dataset. The experimental results show that the proposed method achieves DSC 97.97%, HD 1.22 mm, and ADF 1.85 mm, which demonstrates that the proposed method achieves excellent performance to compete with other state-of-the-art methods in fetal head circumference measurement.

## 1. Introduction

Ultrasound imaging is widely used for monitoring fetal growth and development during prenatal examination due to the advantages of being non-invasive, non-radioactive, painless, and cost-effective [1]. In prenatal examinations, fetal head circumference (HC) is an important physiological indicator, which is required to be precisely measured for gestational age prediction [2]. Besides, Measurement of the HC is employed in many commonly used weight equations for the estimated fetal weight (EFW) [3]. Nevertheless, currently, clinical HC measurement based on ultrasound imaging requires well-trained and experienced sonographers since ultrasound images are not intuitive and HC measurement results are operator-dependent and machine-specific that leading to inter- and intra-observer variability [4,5]. Therefore, an automatic measurement system of fetal ultrasound images is needed, which not only can reduce the variability but also reduce the workload of clinicians.

One of the commonly used approaches in automatic HC measurement is ellipse estimation, as shown in Fig. 1(b). Random forest classifier and Hough transform are applied in [4] to measure the fetal HC. However, the traditional machine learning model is poor to against noise. Specifically, deep-learning-based measurement methods [6–10] have been presented to regress the ellipse parameters directly from the original ultrasound images. Zhang et al. [6] directly measure the head circumference by a CNN to localize and identify the head contour, without having to resort to handcrafted features or manually labeled segmented images. Li et al. [7] present a network to automatically measure the fetal head circumference (HC), biparietal diameter (BPD), and occipitofrontal diameter (OFD) length from 2D ultrasound images. Rong et al. [8] propose using GVF as a reference to train a CNN to derive an external force for the parametric active contour models for curve evaluation. A multi-task deep convolutional neural network is proposed for automatic segmentation and estimation of HC ellipse in [9]. A regression CNN is designed to accurately delineate the HC in [10]. However, these methods with a regression branch suffer from high measurement errors in the ellipse parameter estimation since ultrasound imaging contains various artifacts, such as motion blurring, missing boundaries, acoustic shadows, and speckle noises, which may cause boundary missing (the region in the red bounding box in Fig. 1(a)) or blurring (the region in the orange bounding box in Fig. 1(a)) of the fetal head. To solve this problem, HC measurement

* Corresponding authors.
** Corresponding author at: School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China.
*E-mail addresses:* fzw310@smu.edu.cn (Z. Fang), yaosheng@smu.edu.cn (S. Yao), yangf@smu.edu.cn (F. Yang).

https://doi.org/10.1016/j.bspc.2022.104535
Received 17 May 2022; Received in revised form 28 October 2022; Accepted 18 December 2022
Available online 2 January 2023

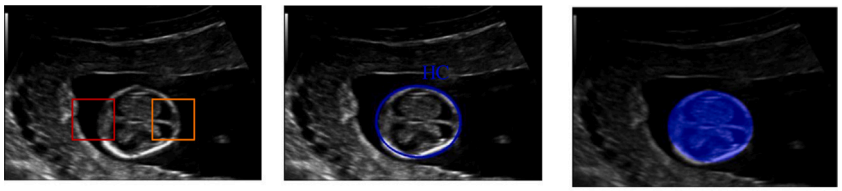


**Fig. 1.** (a) is an original ultrasound image and (b) is an ellipse annotation produced by the sonographer, (c) is a segmentation pseudo-label generated by filling the ellipse annotation showed in (b). The boundary missing region is illustrated in the red bounding box and the boundary in the orange bounding box is blurring.

approaches based on the segmentation contour of fetal head area are proposed [11–14]. A cascaded Fully Convolutional network is proposed to exploit feature extractions and distinguish the anatomy with a dense prediction map in [11]. A probabilistic deep learning approach for estimation of fetal HC is developed in [12]. Al-Bander et al. [13] propose to incorporate object localization mechanism in the segmentation framework for improving fetal head contour detection. Multi-scale inputs are employed to improve the localization of the fetal head in [14]. With the abundant context information of the segmentation results, these segmentation-based methods can improve the performance of evaluating the parameters of the fetal head region to a certain extent. Nevertheless, the segmentation results are unreliable, because the annotation provided by the clinician is fetal head elliptical contour (shown in Fig. 1(b)) rather than the ground truth of the segmentation of the real fetal head area. To train segmentation models [11–14], the whole ellipse is treated as the pseudo-label of the segmentation (displayed in Fig. 1(c)). As shown in Fig. 1, the shape of the fetal head region is not a standard ellipse, which will lead to the model errors between the fetal head region and the pseudo-label of the segmentation in the training phase. This problem will weaken the performance of segmentation-based measurement methods. Obviously, for the irregular fetal head shape in ultrasound images, it is much easier to detect the fetal head boundary pixels and fit an ellipse that is as consistent as possible with the boundary than to obtain an accurate ellipse segmentation. Since the elliptical contours are provided in this task, we argue that an elliptical boundary, which can mostly cover the fetal head region, would achieve a reliable measurement of fetal head circumference.

In this paper, we propose an ellipse-guided multi-task network to boost the performance of HC measurement in ultrasound images. Specifically, a boundary detection branch is developed to locate fetal head boundary pixels considering that HC measurement is based on the contours composed of these boundary pixels. Nevertheless, the existence of artifacts and maternal uterus, whose texture structure is similar to fetal head boundary, would lead to missing or blurring boundaries in ultrasound images, thus weakening the accuracy of boundary detection. Therefore, the region features of the head area are extracted to assist boundary detection, which can contribute to locating the boundary of the fetal head and effectively remove false boundary detection. Moreover, to improve the robustness of the boundary detection, a feature fusion module is designed to integrate the boundary features and region features. More useful context information can be extracted from the fused features, which improve the capability of boundary detection in low contrast ultrasound images and even eliminates boundary leakages in the boundary missing or blurring regions. As the shape prior to the fetal head is an ellipse, an ellipse-shaped loss function is proposed to constrain the results of boundary detection. Since the annotations are standard ellipses, the ellipse-shaped loss can effectively guide the detected boundary contours close to the elliptical annotations as much as possible. This is particularly important for the improvement of performance for HC measurement.

The proposed method is evaluated on the public dataset (HC-18 [4]) and a self-built simulated phantom ultrasound dataset. The experimental results show that the proposed method outperforms the state-of-the-art methods. The main contributions of this work are summarized as follows:

•An ellipse-guided multi-task network is proposed for end-to-end fetal head circumference measurement.

•The fetal head boundaries are obtained by the boundary detection branch without having to resort to segmentation, which reduces errors caused by the inconsistency between the segmentation pseudo-label and the fetal head area.

•The localization information of the fetal head boundaries in the region features from the region segmentation branch and the ellipse information provided by an ellipse-shape loss are applied to improve the measurement performance.

The remainder of this paper is organized as follows: In Section 2, we describe the published fetal head measurement methods based on a single model and multi-task learning. Then the details of our method are illustrated and introduced in Section 3. The experimental results and discussions on the HC-18 dataset are conducted in Section 4. The experimental results and discussions on our self-built dataset are conducted in Section 5. Finally, Section 6 and Section 7 discuss the results and concludes the whole paper.

## 2. Deep learning-based method for fetal head circumference measurement

Recently, deep learning methods have won great success in HC measurement. The work in HC [12] develops two probabilistic CNN methods: Monte Carlo Dropout during inference and probabilistic UNet. The generated set of segmentation masks is used to reject acquired images that produce sub-optimal HC measurements. In [11], a cascaded FCN is applied for fetal head segmentation. Approaches based on a single segmentation task such as [11,12] may ignore the potential information in related tasks that can improve measurement performance. As well-known that multi-task learning has been used successfully across all applications of machine learning (e.g., [15–18]). The multi-task learning strategy also has been used in HC measurement (e.g., [7,19], and [9]). A regression branch is added in an FCNN, and the localization information from the segmentation mask is applied to improve the HC measurement in [7]. The work in [19] combines the classification information of the fetal pregnancy cycle to improve the segmentation performance. For multi-task learning, the higher the correlation between tasks, the greater the promotion effect on the main task [20]. Based on the simplification of fetal head circumference into an ellipse circumference as mentioned in Section 1, ellipse information is much useful to improve the measurement performance. Regressing ellipse parameters from the segmentation features is employed as an auxiliary task in a multi-task neural network [9], which can be trained by minimizing a compound cost function composed of segmentation dice score and mean square error (MSE). The localization information of the fetal head is also considered for HC measurement performance improvement. A combined fetal head localization and fetal head segmentation method based on Mask R-CNN is proposed in [13].

Most of these above approaches take segmenting the whole fetal head as a preliminary step for HC measurement. However, some works [21,22] have been proposed that aim at directly extracting biomarkers from medical images. The goal is to avoid intermediate steps, such as segmentation, that may be computationally expensive

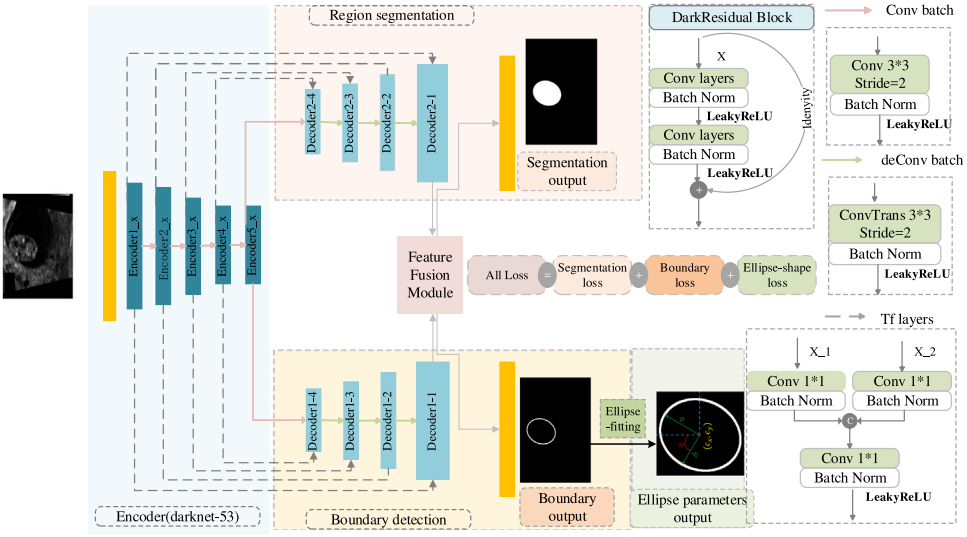


**Fig. 2.** The architecture of our network. The network consists of a Darknet-53 encoder path and two decoder paths. One decoder path is the head region segmentation branch, the other is the ellipse detection branch. The features from the segmentation branch and ellipse detection branch are fed into a feature fusion module, which can help the ellipse detection branch to accurately locate the head boundary. The segmentation loss is computed by the segmentation output with the region ground truth, the boundary loss is computed by boundary output and the contours labeled by experts, and the ellipse loss is computed by ellipse parameter output and the ellipse parameter label, which is designed to do shape constraints on boundary detection results.

and prone to errors. A similar approach is proposed for HC measurement in [6], where a regression CNN is applied to estimate the HC from fetal head ultrasound images without having to resort to segmentation. Considering that the localization of the fetal head is beneficial to improving HC measurement performance, a region proposal CNN for head localization and centering is used in [10]. The experimental results demonstrate that it can improve the HC delineation performance of the regression CNN.

According to these considerations, in this study, we design a direct approach to estimate the HC by fetal head boundary detection with a multi-task learning network. Different from the above methods, the fetal head ellipse parameters are fitted on the detected boundaries in our method, which does not rely on the edge of the head region segmentation. For improvement of detection performance, the fetal head region features from the region segmentation branch are used as localization information of the fetal head boundary. And the ellipse information is introduced into the network by an ellipse-shaped loss which guides the detected boundary close to the ellipse shape.

## 3. Improving ellipse boundary detection with region features fusion for fetal head circumference measurement

Although the fetal head boundaries are incomplete due to artifacts and noise interference in ultrasound images, a good fetal head circumference method should be able to provide an accurate ellipse. It means that the model should not only avoid the interference caused by artifacts and noise but also obtain the ellipse contours from these incomplete boundary pixels. The proposed method in this work extracts the region features of the fetal head area for localization and guides the output of the network close to the ellipse shape by an ellipse-shape loss. These make our method has superior performance under the interference of artifacts and noise.

Fig. 2 shows the architecture of our proposed network, the proposed multi-task ellipse-guided network includes a framework for region segmentation and an ellipse detection. Inspired by [23,24], and [25], which show that coarse segmentation results can infer the context information for performance improvement, We introduce segmentation features into the boundary detection branch to dynamically extract the context information contained in the segmentation, which helps the boundary detection branch to distinguish boundary pixels from background pixels. The region features from the segmentation branch and boundary features from the detection branch are integrated and further selected to explore more context information for boundary detection. The prior shape information is also incorporated into the ellipse detection branch, which can guide the feature extraction in the model. Fit the output of the boundary detection with an ellipse, and an ellipse-shaped loss is designed to supervise the ellipse parameters. Based on the output ellipse parameters, the HC length is computed. More details will be introduced in this section.

### 3.1. Ellipse detection and region segmentation framework

Generally, the context information is useful for improving the performance in many visual applications (e.g., [26–28]). As mentioned above, the information in region segmentation features is also useful for distinguishing boundary pixels and noise. For enhancing the ability of ellipse detection of the fetal head, the context information in fetal head region features is introduced into boundary detection for an accurate boundary. A feature fusion module is used to integrate information between the region segmentation features and the ellipse detection features, which contributes to accurate ellipse detection due to the provision of boundary and region information of the fetal head.

In this study, we modify the Darknet-53 [29] as an encoder path to extract detailed context features, the two decoder paths are the region segmentation branch and the ellipse detection branch, respectively. As shown in Fig. 2, each the $i_ih$ Encoder block contains $x_i$ DarknetResidual block (where $x_i \in \{1, 2, 8, 8, 4\}$, and i$\in \{1, 2, 3, 4, 5\}$). We apply Conv batch instead of Maxpooling layer to do downsampling since some significant spatial information is lost by Maxpooling operations [30], which is important for boundary detection. The Conv batch block contains a convolutional layer (kernel size = 3, stride = 2). The convolution is followed by batch normalization and a LeakyRelu with a constant multiplier $\alpha$, equal to 0.1 to control the slope of the activation function for negative values. Similarly, the Tf layers presented in Fig. 2 are also applied both on the two decoders to concatenate the correspondingly feature map from the encoder path. Each decoder block contains 1 DarknetResidual Block. A deConv batch block is used to do upsampling.

To alleviate the influence of boundary missing and blurring regions for the ellipse detection, the framework in our method simultaneously learns the region information in the segmentation features and boundary information in the detection features. And the region information

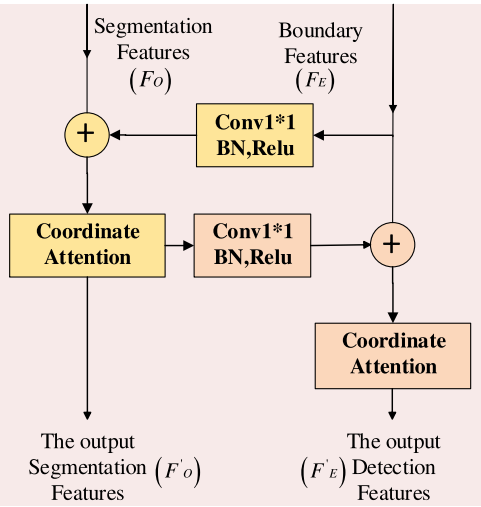


**Fig. 3.** Ellipse boundary and segmentation feature fusion module.

can be treated as boundary localization in boundary missing or blurring regions. For exploring more context information on the fetal head boundary, a feature fusion module is designed to achieve feature integration between region features and boundary features. The details of the feature fusion module will be presented in the next section.

### 3.2. Ellipse boundary and segmentation feature fusion module

For the information integration between the region and the boundary, a feature fusion module is designed to fuse context features in the two decoding paths. The fusion module can assist the ellipse detection branch in accurately locating boundaries according to the region information. The feature fusion module is shown in Fig. 3. The region segmentation features $F_O$ contain rich context information, while the boundary features $F_E$ contain rich shape information. The fusion block can be formulated as follows:

$$F'_E = \alpha(f(F_O) + F_E) + F_E \tag{1}$$

$$F'_O = \alpha(f(F'_E) + F_O) + F_O \tag{2}$$

where $F'_E$ denotes the output detection features, and $F'_O$ denotes the output segmentation features, $f$ means $1 \times 1$ convolution with BN and Relu. $\alpha$ means coordinate attention module [31]. Some visualization examples are given in Fig. 4. From the comparison, we can see that the processed features by feature fusion module (illustrated as Fig. 4(d) and Fig. 4(i)) contain more accurate boundary information in boundary missing or blurring regions than original features (presented as Fig. 4(b) and Fig. 4(g)), which improves the boundary detection performance, as shown in Fig. 4(c) and Fig. 4(e), Fig. 4(h) and Fig. 4(j). Obviously, the boundary detection capability of the ellipse detection framework is get improved so that weak boundaries in boundary missing or blurring regions can be correctly identified. It means that the feature fusion of the segmentation context information is necessary.

### 3.3. Ellipse-shaped loss

To keep the boundary close to the fetal head ellipse shape and achieve higher performance, ellipse fitting is performed on the output of the boundary detection branch, and an ellipse-shaped loss is designed (as shown in Fig. 2) considering the shape constraint [32,33], which can be defined as

$$L_S = MSE(\hat{y}(a, b, x, y, \theta), y(a, b, x, y, \theta)) \tag{3}$$

**Table 1**
The distribution of HC-18 dataset.

| Period | Train set | Validation set |
|---|---|---|
| First trimester | 165 | 55 |
| Second trimester | 693 | 233 |
| Third trimester | 141 | 47 |
| Total | 999 | 335 |

**Table 2**
The dataset distribution of ablation experiments.

| Period | Training set | Test set |
|---|---|---|
| First trimester | 150 | 15 |
| Second trimester | 623 | 70 |
| Third trimester | 125 | 16 |
| Total | 898 | 101 |

where MSE is the Mean Square Error function. $a, b, x, y, \theta$ are the ellipse parameters, $a$ and $b$ are major and minor semi-axes of ellipse, $x$ and $y$ are center coordinates of ellipse, $\theta$ is rotation angle of ellipse, as shown in Fig. 2. $\hat{y}(a, b, x, y, \theta)$ is the prediction of ellipse fitted by detected boundary pixels, $y(a, b, x, y, \theta)$ is the provided label. The least-square fitting is used to do ellipse fitting on the detected boundary pixels' coordinates. The ellipse-shaped loss $L_S$ can guide the training of our model in an end-to-end fashion.

With the obtained ellipse-shaped loss in Eq. (3), the final loss function of our multi-task network can be defined as:

$$L = L_{dice}(\hat{y_S}, y_S) + L_{dice}(\hat{y_E}, y_E) + L_{MSE}(\hat{y}(a, b, x, y, \theta), y(a, b, x, y, \theta)) \tag{4}$$

where the segmentation loss $L_{dice}(\hat{y_S}, y_S)$, the boundary detection loss $L_{dice}(\hat{y_E}, y_E)$ and the ellipse-shaped loss $L_{MSE}(\hat{y}(a, b, x, y, \theta), y(a, b, x, y, \theta))$. $\hat{y_S}$ is segmentation prediction, $y_S$ is the segmentation pseudo-label. And $\hat{y}_E$ is the boundary detection results, $y_E$ is the fetal head ellipse contour. With the final loss, the multi-task model can eventually learn the features of fetal head boundary and region under the constraint of ellipse shape. Here, dice loss is a general objective for region segmentation [34]. For boundary detection loss, we also employ dice loss as objective rather than cross-entropy, because the dice loss is insensitive to the number of foreground/background pixels, thereby alleviating the class-imbalance problem [35].

## 4. Experiments on HC-18 dataset

### 4.1. Experimental setup

In experiments, we evaluated the performance of our method on a publicly available dataset HC-18. Table 1 shows the distribution of HC-18 datasets from the first to the third trimester. Since the ground truth was not available in the original HC-18 test set, we divided 90% annotated images of the HC-18 training set into our train set and the rest images into validation set, as listed in Table 2. The ablation experiments were studied by using 10-fold cross-validation. Annotated images of HC-18 had the resolution of $800 \times 540$, with the pixel size from 0.0520 mm to 0.326 mm. In our experiments, all images were padded to $800 \times 800$ pixels with zero, then resized to $416 \times 416$ pixels. Online data augmentation such as rotation, gray transformation, and horizontal flip was randomly applied at each training iteration. The batch size was set as 8 and the epochs were 200. In the training step, Adam optimizer with an initial learning rate of 0.0001 was used to minimize the multi-task loss in Eq. (4). All experiments were implemented using the deep learning framework Pytorch on a computer with NVIDIA GTX 1080Ti GPUs.

To evaluate the HC measurement performance on HC-18, we chose three metrics, including Dice similarity coefficient (DSC) [%], Absolute

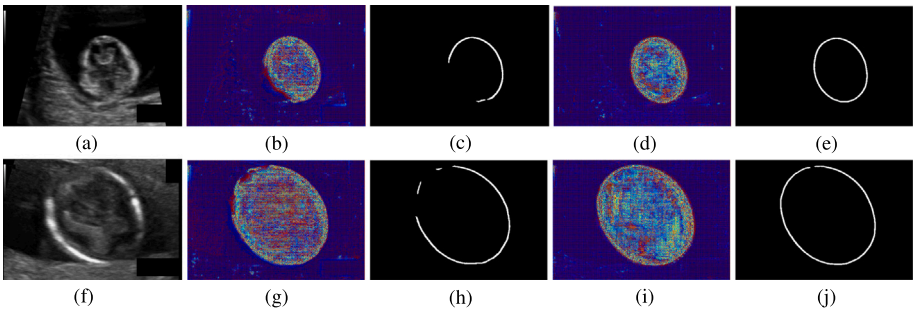


**Fig. 4.** Comparison between with and without feature fusion module. (a) is an ultrasound image with a blurring boundary. (b) is the output feature without the feature fusion module in the model. (c) is the detected head boundary without the feature fusion module. (d) is the output feature with the feature fusion module in the model. (e) is the detected head boundary with the feature fusion module. (f) is an ultrasound image with boundary missing. (g) is the output feature without the feature fusion module in the model. (h) is the detected head boundary without the feature fusion module. (i) is the output feature with the feature fusion module in the model. (j) is the detected head boundary with the feature fusion module. It can be seen that more details about head boundary can be found by the context information from the segmentation branch.

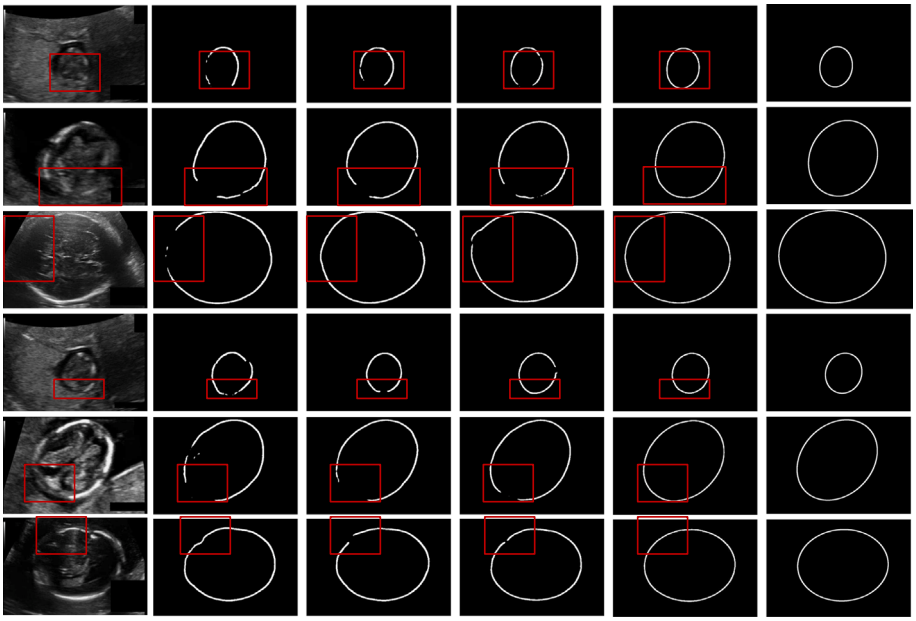

**Fig. 5.** Qualitative evaluation of boundary detection ablation studies on the HC-18 training set with 10-fold cross-validation. The first and the fourth row show examples with poor fetal head boundary contrast to the background. The images in the second and the fifth row are examples of fetal head boundary-blurring regions. The third and the sixth row presents examples with boundary missing regions. Each column shows the original ultrasound images (first column), the corresponding ellipse annotations (sixth column), the ellipse detection results obtained by Model 1 (second column), Model 2 (third column), Model 3 (fourth column), Model 4 (fifth column), respectively.

Difference (ADF) [mm], and Hausdorff Distance (HD) [mm]. They are defined as:

$$DSC = \frac{2(Y_p \cap Y_g)}{|Y_p| + |Y_g|} \tag{5}$$

$$ADF = |HC_p - HC_g| \tag{6}$$

$$\begin{aligned} HD &= max(h(Y_g, Y_p), h(Y_p, Y_g)) \\ h(Y_p, Y_g) &= max_{Y_g \in Y_p} min_{Y_p \in Y_g} \parallel Y_p - Y_g \parallel \\ h(Y_g, Y_p) &= max_{Y_p \in Y_g} min_{Y_g \in Y_p} \parallel Y_g - Y_p \parallel \end{aligned} \tag{7}$$

where $Y_g$ denotes the fetal head area delimited by the ellipse contours delineated by the clinician, $Y_p$ is the fetal head area delimited by the ellipse obtained with our proposed method. $HC_p$ represents the circumference measured from the ellipse detection branch, and $HC_g$ is the manual annotation.

### 4.2. Ablation experiment on the HC-18 training set

To show the effectiveness of different components in our model, we present an ablation experiment on the training set with 10-fold cross-validation. In summary, the following models are compared:

**Table 3**
Ablation results on the HC-18 training set with 10-fold cross-validation (mean ± standard deviation).

| Methods | DSC (%) | HD (mm) | ADF (mm) |
|---|---|---|---|
| Model 1 | 97.86 ± 0.85 | 1.20 ± 0.98 | 2.07 ± 1.54 |
| Model 2 | 97.90 ± 1.13 | 1.18 ± 1.26 | 1.94 ± 1.15 |
| Model 3 | 98.05 ± 0.77 | 1.14 ± 0.82 | 1.88 ± 1.34 |
| Model 4 (ours) | **98.08 ± 0.96** | **1.08 ± 1.02** | **1.82 ± 1.21** |

**Model 1**: consist of the Darknet-53 encoder and a decoder for boundary detection.

**Model 2**: combine a region segmentation branch with the **Model 1**.

**Model 3**: add feature fusion module based on **Model 2**.

**Model 4 (ours)**: add the ellipse-shape loss to introduce shape constraint to the **Model 3**.

The quantitative analyses are shown in Table 3 and the qualitative of boundary detection and region segmentation are presented in Figs. 5 and 6. Note that all evaluation criteria in Table 3 are obtained by averaging the 10-fold cross-validation, from which it can be seen that the lowest mean AD (2.07 mm) is obtained with the Model 1 (boundary detection only), which shows the challenge of boundary missing or

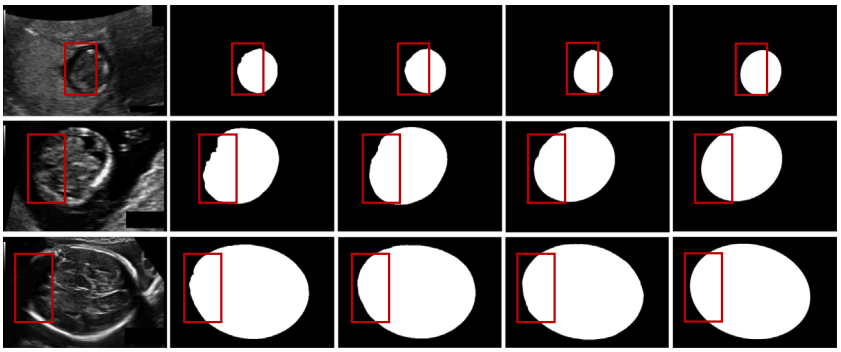

**Fig. 6.** Qualitative evaluation of region segmentation ablation studies on the HC-18 training set with 10-fold cross-validation. From top to bottom, there are examples in the first trimester, second trimester, and third trimester, respectively. From left to right, each column shows the original ultrasound images (first column), the corresponding ground truth (last column), and the segmentation results obtained by Model 2 (second column), Model 3 (third column), and Model 4 (fourth column), respectively. .

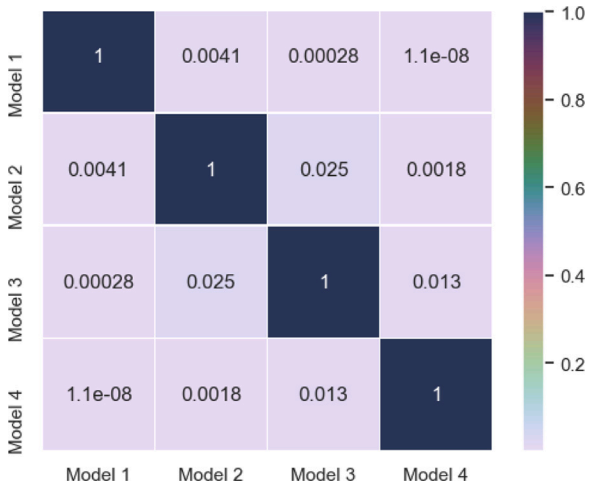


**Fig. 7.** The t-test results of statistical significance of the difference between ablation study on each component in our framework. The smaller the *p*-value, the more significant the difference.

blurring regions for boundary detection. However, it can be observed that Model 2 yields improvements in terms of ADF (0.13 mm). We can see from the second column and the third column in Fig. 5, that after adding the region segmentation branch, more accurate boundary pixels can be detected in boundary missing or blurring regions noted in the red bounding box. This achievement can be explained by learning region features which encourage the network to explore more information to distinguish the fetal head area from the background. Benefiting from the feature fusion module, the performance of DSC, HD, and ADF in Table 3 has been further improved, increasing by 0.15%, 0.04 mm, and 0.06 mm, respectively. As shown in the fourth column in the figure, the wrong boundary pixels are corrected, and the detected boundaries of the fetal head are more complete. This means that the designed feature fusion module can select better context information in the boundary features and region features, thereby facilitating the boundary detection branch to locate the head boundary. With the ellipse shape loss added in Model 4, the detected boundary is closer to the elliptical shape, and the detected boundary is more complete. And according to the Table 3, DSC value, HD value and ADF value increased by 0.03%, 0.06 mm, 0.06 mm respectively. The ellipse-shaped loss does well in shape constraint and guiding the detected boundary contours and boundary of region segmentation much closer to the ellipses covering the fetal head area, as shown in Figs. 5 and 6.

In addition, we also have reported the complexities of the proposed method in Table 4, which includes the parameter amounts for each component in our network and the computational complexity (GFLOPs). It can be seen from Table 4 that our feature fusion module costs only 0.0028 M parameters, which has significantly improved the measurement performance as shown in Table 3 and Fig. 5.

To further explore the robustness of each component in our framework, we perform t-tests on the results of the ablation study on each component, the p-values are presented in Fig. 7. It can be seen that there are significant differences in performance between the models with dual decoders and model 1, which is with one decoder for boundary detection. It is also evident that the boundary localization information contained in the region segmentation features contributes to improving boundary detection. Moreover, Model 4 with ellipse shape loss to do shape constraint also gains a performance gap in ADF scores compared with those without the ellipse shape loss. From Fig. 5, it can be observed that the detected boundary of the fetal head becomes more and more accurate, which reveals the powerful ability of each component in our method.

### *4.3. Generalizing analysis of cross-trimester validations on the HC-18 train set*

For generalizing analysis of our proposed network, the cross-trimester validations are performed on the HC-18 train set with 10-fold cross-validation. The results are listed in Table 5. For the former, the model is trained on the single-trimester training set and tested on the all-trimester testing set. It can be seen that the metrics decrease slightly when training with single-trimester data and testing on all trimesters, which demonstrates that the proposed network is with good generalizing ability. Additionally, when the model is trained on the third trimester data, the metrics decline minimally, indicating that accurate measurement in the third trimester has a large impact on the performance.

### *4.4. Ablation experiment on the HC-18 test set*

Combined with our best settings in the deep neural network, we experimented with our proposed HC measurement method on the HC-18 test set. The quantitative and qualitative evaluation results are illustrated in Table 6 and Fig. 8, respectively. It can be seen from that all the components of our model have improved the performance on the HC-18 test set. Compared with Model 1, Model 2 shows better performance with DSC and ADF increased by about 0.23% and 0.11 mm, respectively. The feature fusion module has further improved performance in Model 3, with DSC, HD, and ADF increased by 0.16%, 0.13 mm, and 0.09 mm, respectively. This reflects that the context information contained in the region features is beneficial to the boundary detection in boundary missing or blurring regions. Under the supervision of ellipse shape loss in Model 4, DSC, AD, and HD have been improved by 0.07%, 0.01 mm, and 0.06 mm, respectively. This achievement can be explained by limiting the detected boundary pixels to the contour of the ellipse as much as possible by the ellipse loss function.

**Table 4**
The complexity analysis of the proposed network.

| Parameters (M) | | | | | FLOPs (G) |
|---|---|---|---|---|---|
| Encoder | Decoder 1 | Decoder 2 | ESFF | Total | GFLOPs |
| 14.89 | 2.79 | 2.79 | 0.0028 | 20.47 | 45 |

ESFF denotes the ellipse boundary and segmentation feature fusion module.

**Table 5**
The generalizing performance results of cross-trimester validations (mean ±standard deviation).

| Train data | Test data | Metrics | | |
|---|---|---|---|---|
| | | DSC (%) | HD (mm) | ADF (mm) |
| First trimester | All trimester | 97.65 ± 2.06 | 1.25 ± 1.56 | 1.90 ± 1.64 |
| Second trimester | All trimester | 97.73 ± 1.12 | 1.18 ± 0.96 | 1.87 ± 1.35 |
| Third trimester | All trimester | 97.79 ± 1.03 | 1.15 ± 0.67 | 1.85 ± 1.46 |
| All trimester | All trimester | **98.08 ± 0.96** | **1.08 ± 1.02** | **1.82 ± 1.21** |

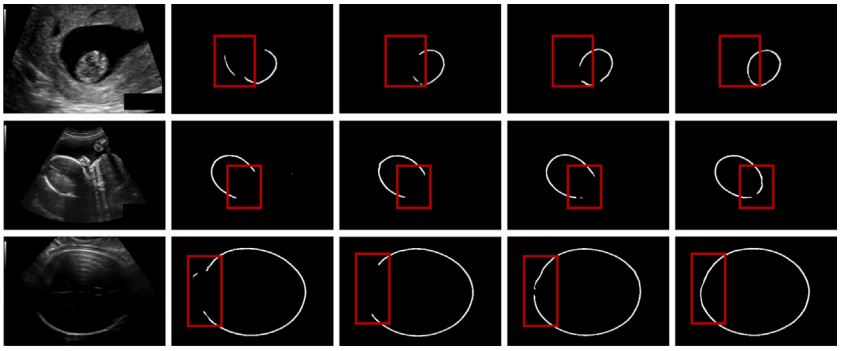

**Fig. 8.** Qualitative evaluation of ablation study on the HC-18 test set. From top to bottom, there are examples in the first trimester, second trimester, and third trimester, respectively. From left to right, each column shows the original ultrasound images (first column), and the boundary detection results obtained by Model 1 (second column), Model 2 (third column), Model 3 (fourth column), and Model 4 (fifth column), respectively. .

**Table 6**
Ablation studies on the HC-18 test dataset. The bold number indicates the best performance in each column (mean ± standard deviation).

| Methods | DSC (%) | HD (mm) | ADF (mm) |
|---|---|---|---|
| Model 1 | 97.97 ± 1.09 | 1.25 ± 0.71 | 2.11 ± 1.97 |
| Model 2 | 97.74 ± 1.24 | 1.36 ± 1.94 | 2.00 ± 1.94 |
| Model 3 | 97.90 ± 1.26 | 1.23 ± 0.69 | 1.91 ± 1.79 |
| Model 4 (ours) | **97.97 ± 1.15** | **1.22 ± 0.77** | **1.85 ± 1.96** |

### *4.5. Comparison with previous studies on the HC-18 test set*

Performance comparisons are conducted with different state-of-the-art methods, where the results are listed in Table 7. All these methods were developed and tested using the HC-18 [4] dataset. The relative performance metrics, exhibited in Table 7, are extracted from the corresponding published papers. The fetal head measurement method is based on a single model [8] which employs the active contour model to obtain the fetal head boundaries and achieve the worst performance. Although they have the same ideas as us that employing the direct least-squares fitting of ellipses method to fit the fetal head boundaries obtained models, and do not rely on the fetal head region segmentation, our method can obtain about 2.48%, 1.22 mm, and 0.6 mm improvements in terms of DSC, HD, and ADF, respectively. This can be explained by the reason that we apply a multi-task learning strategy in our method, which improves the boundary detection performance by introducing the head region information and ellipse shape prior information to the boundary detection branch. The work in [14] is also based on a single model, thus, they obtain improvements of 0.58 mm in ADF index than [8], by using multi-scale inputs in the segmentation network. The ability of a single-task supervised model to extract features is poor, and the head region segmentation method [14] or boundary detection method [8] based on a single model needs to be improved in the performance of fetal head circumference measurement. The work in [10,13] both employ object localization to improve the HC measurement accuracy. In [13], the Mask R-CNN is adopted to improve the fetal head segmentation accuracy by combining object localization and segmentation in a framework, which helps them obtain a 1.2 mm improvement in ADF metric than [33]. Differently, a region-proposal CNN is used for head localization and centering in [10], which improves the accuracy of a regression CNN for HC delineation. They get significant improvements of 0.55 mm and 0.32 mm in ADF score than [8,14], respectively. And the work in [9] creatively introduces estimation of HC ellipse in a framework developed for segmentation purposes to improve the segmentation accuracy. The results improved by 0.33 mm and 0.1 mm in terms of ADF. Compared with these methods, our method achieves the best performance in all three metrics, which benefit from the ellipse-shape constraint by ellipse-shape loss and the boundary localization information provided by fetal head region features. And we tackle the problem of inconsistency between segmentation pseudo-label and ellipse annotation mentioned in Fig. 1 by employing the ellipse fitting on the detected fetal head boundaries directly.

### *4.6. Loss curves*

In this section, the loss curves of training phase is provided in Fig. 9. It can seen that the training loss and validation loss reduce gradually. The convergence of the loss curves demonstrates the effectiveness of the proposed method.

## 5. Experiments on our self-built dataset

### *5.1. Experimental setup*

As mentioned, the main idea of our method is doing the measurement on the detected boundary pixels, which is helpful in reducing

**Table 7**
Comparison with other methods on HC-18 test dataset. The bold number indicates the best performance in each column (mean ± standard deviation).

| Methods | DSC (%) | HD (mm) | ADF (mm) | Params (M) |
|---|---|---|---|---|
| Heuvel et al. [4] | 97.00 ± 2.80 | 2.00 ± 1.60 | 2.80 ± 3.30 | – |
| Rong et al. [8] | 95.49 ± 4.11 | 2.44 ± 1.96 | 2.45 ± 2.55 | – |
| Al-Bander et al. [13] | 97.73 ± 1.32 | 1.39 ± 0.82 | 2.33 ± 2.21 | 345 |
| Sobhaninia et al. [14] | 92.46 ± N.A. | 3.40 ± N.A. | 2.22 ± N.A. | – |
| Sobhaninia et al. [9] | 96.84 ± 2.89 | 1.72 ± 1.39 | 2.12 ± 1.87 | 26.91 |
| Fiorentino et al. [10] | 97.76 ± 1.32 | 1.32 ± 0.73 | 1.90 ± 1.77 | 50.54 |
| **Ours** | **97.97 ± 1.15** | **1.22 ± 0.77** | **1.85 ± 1.96** | 20.47 |

The params are calculated based on the network structure described in the corresponding work.

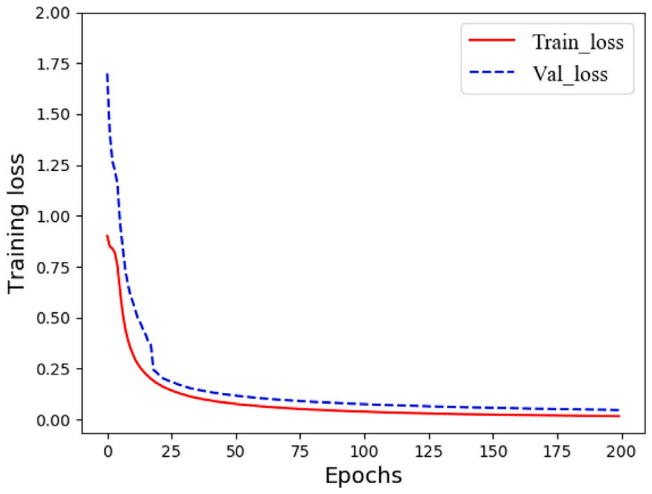


**Fig. 9.** Loss curves of training and validation set.

errors caused by the inconsistency between the segmentation pseudo-label and the ellipse annotations. And the boundary pixels in missing or blurring boundary regions can be more easily localized by the proposed method. To validate the effectiveness and robustness of the proposed method, experiments were conducted on a self-built dataset.

We used hand-held ultrasound equipment to collect phantom data, which is regarded as the simulation of fetal head ultrasound images. The boundary in the ultrasound phantom image was randomly occluded to simulate boundary-missing cases (For each image, we used cut image blocks to randomly occlude the boundary of $\frac{1}{10}$ circumference). And measuring the circumference of the circular area in the ultrasound image was used to simulate the HC measurement. Here we designed four groups of experiments on our self-built dataset to verify the effectiveness and robustness of our method for cases with boundary missing or blurring cases: ultrasound images with 5 MHz center frequency (data 1, high-resolution images with clear boundaries), ultrasound images with 3.5 MHz frequency center frequency (data 2, low-resolution images with blurring boundaries), ultrasound images with 5 MHz center frequency and boundary occlusion (data 3), ultrasound images with 3.5 MHz and boundary occlusion (data 4). Data 1 and data 2 were conducted as a comparative experiment to study the boundary detection ability of our method with different blurriness images. Data 3 and data 4 were used to study the measurement accuracy of our method with the boundary-missing images. We compared the experiment results by UNet++ [36], and Sobhaninia et al. [9]. One is a measurement method based on a single segmentation model, and the other is a measurement method based on multi-task learning. The experimental parameters and software and hardware environment were the same as those experiments in Section 4, except the epochs were set at 30 for our self-built dataset.

**Table 8**
Equipment parameters for ultrasound image acquisition.

| Parameters | Machine index | Gain | Dynamic range | Center frequency |
|---|---|---|---|---|
| data1 | 0.7 | 80 dB | 80 dB | 5 MHz |
| data2 | 0.6 | 88 dB | 90 dB | 3.5 MHz |

### 5.2. Ultrasonic data acquisition and processing

The handheld ultrasound equipment we used was Sonoster UProbe (a digital ultrasonic imaging diagnostic instrument with 128 arrays convex probe shown in Fig. 10(a)), and the ultrasound phantom is CIRS 040GSE (Fig. 10(b)). And ultrasonic images with different resolutions were obtained with the center frequencies of 5 MHz and 3.5 MHz, respectively (an example as Fig. 10(c), the equipment parameters of ultrasonic image acquisition are listed in Table 8.) Placed the probe directly on the surface of the phantom and changed the parameters of the probe to obtain ultrasound images of the phantom with different resolutions. Images were stored in MP4 format, and each MP4 file contained 100 images. After obtaining the phantom ultrasound image, the image was clipped according to the position of the circular area, and the clipped sizes were 395 × 278 pixels to 222 × 166 pixels. The pixel size of these images is 0.0036 mm. Then we labeled the clipped image with LabelMe (an online open annotation tool). The data was augmented by rotation (90°, 180°, 270°) and gamma transformation (gamma factor $\gamma \in \{0.6, 0.7, 0.8, 0.9, 1.1, 1.2, 1.3, 1.4\}$). The private simulated phantom ultrasound dataset included 700 images for training and 100 images for testing. The size of images in this dataset was from 395 × 278 pixels to 222 × 166 pixels, resized to 256 × 256 pixels with zero padding.

### 5.3. Results on our self-built dataset

We now compare our method with two published studies [9,36] on our self-built dataset. For the fairness of the results, the same experimental parameters and experimental data were used for these three methods. With results shown in Fig. 11, it can be observed that: our proposed method consistently outperforms these two methods on our self-built dataset in terms of DSC, HD, and ADF metrics for all four experiments. For ultrasound images with 5 MHz or 3.5 MHz center frequency, which means these images with different blurriness boundaries, the results by UNet++ [36] drop 0.0003 mm, the results by Sobhaninia et al. [9] drop 0.0007 mm, and the results of our method drop 0.0001 mm. Our method achieves the least accuracy reduction compared with UNet++ [36] and Sobhaninia et al. [9] in ADF values. These results further validate the effectiveness of our proposed method for boosting the measurement performance in boundary-blurring cases. Our proposed method and the work [9] both apply the multi-task learning strategy, but the results of our method are better. The intermediate steps, such as segmentation in HC measurement, are prone to errors.

For ultrasound images with boundary occlusion, which means these images with boundary missing regions. The results show that our proposed method achieves the best performance in terms of DSC, HD, and ADF metrics. The measurement method based on UNet++ [36] is most affected, and the accuracy of the three metrics decreased


**Fig. 10.** The equipment of our experiments. (a) is the handheld ultrasound equipment, (b) is the ultrasound phantom, (c) is an ultrasound image with 5 MHz center frequency of the probe.

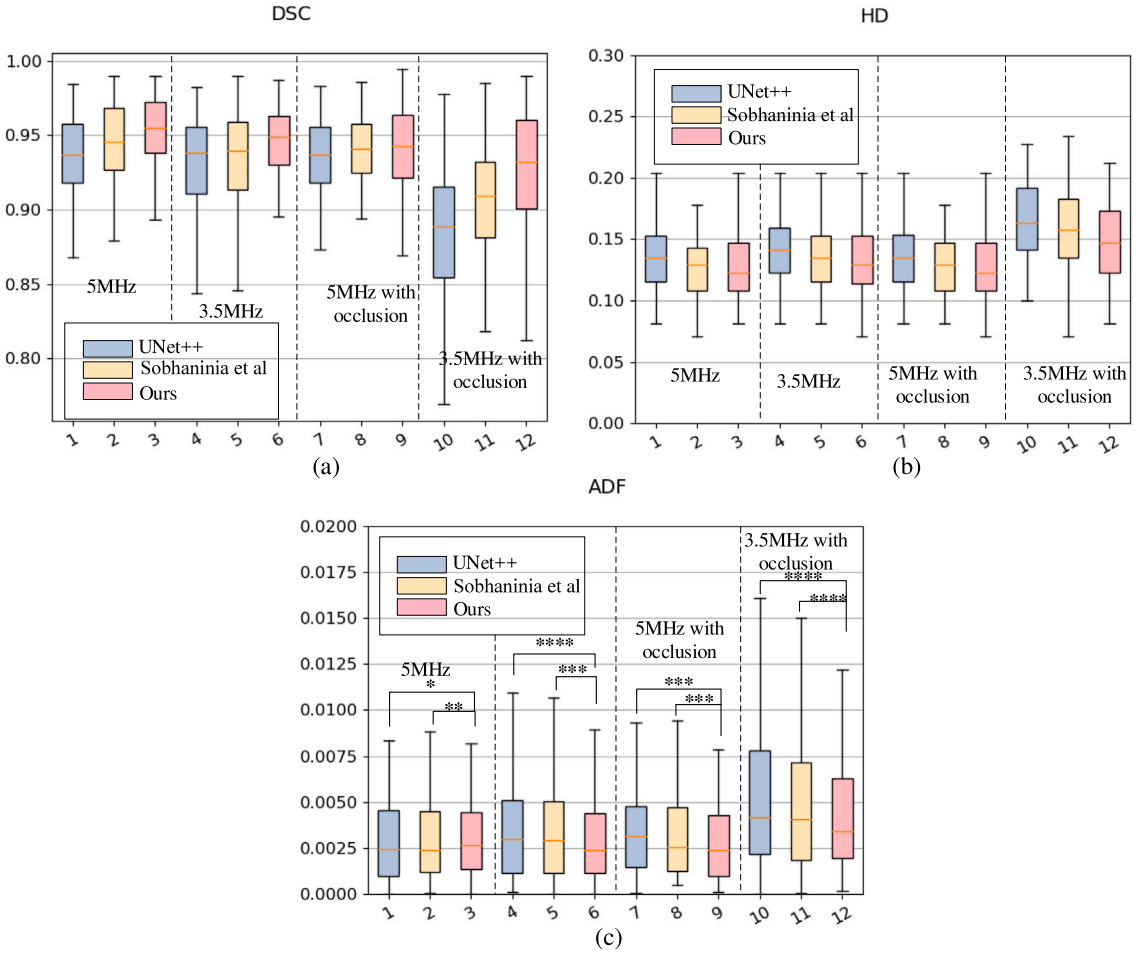


**Fig. 11.** Results of our self-built dataset. (a)–(c) are DSC (%), HD (mm), ADF (mm) respectively. We compared our method with the other two (UNet++ [36], Sobhaninia et al. [9]) on the four experiments (5 MHz denotes ultrasound images with 5 MHz center frequency (data 1), 3.5 MHz denotes ultrasound images with 3.5 MHz center frequency (data 2), 5 MHz with occlusion denotes ultrasound images with 5 MHz center frequency and boundary occlusion (data 3), 3.5 MHz with occlusion denotes ultrasound images with 3.5 MHz center frequency and boundary occlusion (data 4). The t-tests are conducted on the results in terms of ADF as shown in (c): $*$ denotes the $p$-value $< 0.05$; $**$ denotes the $p$-value $< 0.01$; $***$ denotes the $p$-value $< 0.001$; and $****$ denotes the $p$-value $< 0.0001$.

sharply. Sobhaninia et al. [9] apply a multi-task strategy to improve the accuracy of boundary-blurring cases, but their method does not work in boundary-missing cases, with a reduction by 0.6% (DSC), 0.0039 mm (HD), 0.0054 mm (ADF) on data 3, and reduction by 4.38% (DSC), 0.0276 mm (HD), 0.0035 mm (ADF) on data 4, respectively. They cannot effectively guide the segmentation in these boundary-missing pixels by regressing elliptic parameters on the segmentation features. Differently, we add ellipse-shaped loss on boundary detection results, which avoids the difficulty of network training caused by the fluctuation of shape constraints and can guide the prediction results closer to the ellipse more effectively. The results of the proposed method slightly drop 0.1% (DSC), 0.0006 mm (HD), 0.0054 mm (ADF) on data 3, and 0.19% (DSC), 0.0014 mm (HD), 0.0016 mm (ADF) on data 4, respectively.

Fig. 11(c) presents the T-tests are performed on the results of different methods in terms of ADF, it can be seen that there are significant differences in performance between the proposed method and the other two. Especially for the images with more noise and boundary occlusion, our proposed method gains a performance gap compared with U-Net++ [36] and Sobhaninia et al. [9] And the results presented above show that the proposed method achieves the most superior performance in terms of ADF and its average value and variance for the designed four experiments. That convincingly demonstrates the effectiveness and

robustness of our proposed method in handling cases with boundary missing or blurring regions for fetal head circumference measurement.

## 6. Discussion

While a lot of deep learning methods have been proposed so far, it is still a challenging task for us to find an effective and robust solution to the HC measurement task in ultrasound images. Compared with other medical imaging technologies, one of the main challenges for tasks in ultrasound images is the low contrast between the pixels of the target boundary and background due to extensive artifacts and noise during imaging. Some HC measurement methods based on a single model to segment the fetal head region or detect the fetal head boundary cannot achieve good performance in the first trimester (e.g., [4]) since the fetal skull is relatively soft in the first trimester, it does not always appear brighter than the inside of the fetal head. Therefore, it is sometimes very difficult to detect the boundary of the fetal head in the first trimester, especially when it lies close to the wall of the uterus. For this reason, some methods employ a multi-task strategy to introduce fetal head localization for the improvement of fetal head boundary detection. However, they obtain lower performance for images from the third trimester (e.g., [10]). The primary reason for these results from the fact that the higher pixel dimension in images from the third semester with respect to images from the other trimesters, thus the advanced reason is that error caused by inconsistency between the head contour labeled by experts and the actual fetal head shape. To handle the challenges mentioned above, a region segmentation branch is designed to learn region features of the fetal head, which contributes to exploring more context information and locating boundary pixels in boundary missing or blurring regions. And we do ellipse fitting on the detected fetal head boundary pixels rather than on the extracted boundary of segmentation results, which avoids errors caused by the inconsistency of fetal head shape and segmentation pseudo-labels. Besides, an ellipse-shape loss is employed in our method that supervises the detected boundaries close to the ellipse shape. It can be seen that the proposed framework provided the highest performance among the state-of-the-art methods tested on the same dataset in Table 6, which obtain encouraging results with a mean AD of 1.85 mm and low standard-deviation value ($\pm$ 1.96). The value obtained is two orders of magnitude lower with respect to the HC length (mean HC in the training set = 174.38 mm), thus showing great potential for clinical practice applications. The proposed pipeline achieves real-time inference using a powerful GPU, which costs about 0.02s. At the same time, when using a less powerful computational resource (Intel (R) Core (TM) i7-7700 CPU), we achieved an inference time lower than 0.5s. We believe this may be suitable for clinical applications, considering that the manual measurement of the US images by clinicians could take more time. Moreover, we also apply the proposed framework to the self-built dataset, which is collected by a digital ultrasonic imaging diagnostic instrument with 128 arrays convex probe (Sonoster UProbe). The results in Fig. 11 are a powerful validation of the generalization ability of the proposed work, we believe this may be compatible with existing sonography.

Although our proposed method has shown promising results, there are still limitations that existed. To analyze the limitations of the proposed method, we give two visualization samples of failure cases in Fig. 12. Due to the joint noise and artifacts covering most boundary pixels of the fetal head, the performance of our method will be weakened, and the fitted ellipses will be erroneous. In these failure cases, long-range dependence is needed for predicting the missing boundary pixels according to the few that existed. Future directions of this work include investigating more advanced modules such as Transformer (e.g., [37,38]) for improving the boundary detection performance. For the threats-to-validity of the experimental results, we think that the prior ellipse information may not be the best shape loss because the ellipse cannot fit the fetal head well. Considering that most existing methods [4,9], and [10] are based on Ground Truth with the ellipse shape, we also choose the prior ellipse information as the shape loss for a fair comparison. In the future, we will do further research to discover better shape loss for this task. Additionally, considering the measurement of HC is an important parameter for EFW, on the basis of this study, the fetal weight estimation will be further explored in the future.

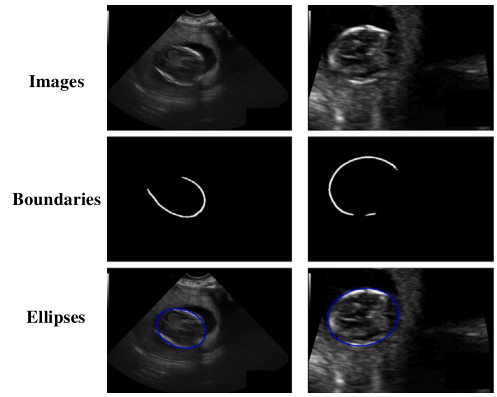


**Fig. 12.** Failure cases of proposed method on the HC-18 test set.

## 7. Conclusion

In this work, we propose an ellipse-guided network for fetal head circumference measurement in ultrasound imaging, which focuses on pixel prediction in boundary missing or blurring regions. The core idea is to introduce region features by the feature fusion module and detect the fetal head boundary under the ellipse shape supervision to improve the measurement performances. Both the public dataset HC-18 and a self-built dataset are used to validate the effectiveness of our proposed method. The experimental results show that the proposed method achieves DSC 97.97%, HD 1.22 mm, and ADF 1.85 mm, which are better than the performances of the other state-of-the-art approaches. It is also worth mentioning that this work is among the first attempts to directly detect the fetal head boundaries for fetal head circumference measurement without resorting to segmentation edges, which provides great support to sonographers in the clinical practice of prenatal examination.

## CRediT authorship contribution statement

**Jinting Wang:** Writing – original draft, Methodology, Software. **Zhiwen Fang:** Conceptualization, Supervision. **Sheng Yao:** Writing – review & editing. **Feng Yang:** Funding acquisition, Supervision.

## Declaration of competing interest

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

## Data availability

No data was used for the research described in the article.

## Acknowledgments

This work was supported in part by the National Nature Science Foundation of China (NSFC) (61771233).

## References


[1] J. Jang, Y. Park, B. Kim, S.M. Lee, J.-Y. Kwon, J.K. Seo, Automatic estimation of fetal abdominal circumference from ultrasound images, IEEE J. Biomed. Health Inf. 22 (5) (2017) 1512–1520.

[2] U. Schmidt, D. Temerinac, K. Bildstein, B. Tuschy, J. Mayer, M. Sütterlin, J. Siemer, S. Kehl, Finding the most accurate method to measure head circumference for fetal weight estimation, Eur. J. Obstet. Gynaecol. Reprod. Biol. 178 (2014) 153–156.

[3] U. Schmidt, D. Temerinac, K. Bildstein, B. Tuschy, J. Mayer, M. Sütterlin, J. Siemer, S. Kehl, Finding the most accurate method to measure head circumference for fetal weight estimation, Eur. J. Obstet. Gynaecol. Reprod. Biol. 178 (2014) 153–156.

[4] T.L. van den Heuvel, D. de Bruijn, C.L. de Korte, B.v. Ginneken, Automated measurement of fetal head circumference using 2D ultrasound images, PLoS One 13 (8) (2018) e0200412.

[5] I. Sarris, C. Ioannou, P. Chamberlain, E. Ohuma, F. Roseman, L. Hoch, D. Altman, A. Papageorghiou, International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st), Intra-and interobserver variability in fetal ultrasound measurements, Ultrasound Obstet. Gynecol. 39 (3) (2012) 266–273.

[6] J. Zhang, C. Petitjean, P. Lopez, S. Ainouz, Direct estimation of fetal head circumference from ultrasound images based on regression CNN, in: Medical Imaging with Deep Learning, PMLR, 2020, pp. 914–922.

[7] P. Li, H. Zhao, P. Liu, F. Cao, Automated measurement network for accurate segmentation and parameter modification in fetal head ultrasound images, Med. Biol. Eng. Comput. 58 (11) (2020) 2879–2892.

[8] Y. Rong, D. Xiang, W. Zhu, F. Shi, E. Gao, Z. Fan, X. Chen, Deriving external forces via convolutional neural networks for biomedical image segmentation, Biomed. Opt. Express 10 (8) (2019) 3800–3814.

[9] Z. Sobhaninia, S. Rafiei, A. Emami, N. Karimi, K. Najarian, S. Samavi, S.R. Soroushmehr, Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning, in: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, IEEE, 2019, pp. 6545–6548.

[10] M.C. Fiorentino, S. Moccia, M. Capparuccini, S. Giamberini, E. Frontoni, A regression framework to head-circumference delineation from US fetal images, Comput. Methods Programs Biomed. 198 (2021) 105771.

[11] L. Wu, Y. Xin, S. Li, T. Wang, P.-A. Heng, D. Ni, Cascaded fully convolutional networks for automatic prenatal ultrasound image segmentation, in: 2017 IEEE 14th International Symposium on Biomedical Imaging, ISBI 2017, IEEE, 2017, pp. 663–666.

[12] S. Budd, M. Sinclair, B. Khanal, J. Matthew, D. Lloyd, A. Gomez, N. Toussaint, E.C. Robinson, B. Kainz, Confident head circumference measurement from ultrasound with real-time feedback for sonographers, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2019, pp. 683–691.

[13] B. Al-Bander, T. Alzahrani, S. Alzahrani, B.M. Williams, Y. Zheng, Improving fetal head contour detection by object localisation with deep learning, in: Annual Conference on Medical Image Understanding and Analysis, Springer, 2019, pp. 142–150.

[14] Z. Sobhaninia, A. Emami, N. Karimi, S. Samavi, Localization of fetal head in ultrasound images by multiscale view and deep neural networks, in: 2020 25th International Computer Conference, Computer Society of Iran, CSICC, IEEE, 2020, pp. 1–5.

[15] X. Li, Y. Guo, F. Jiang, L. Xu, F. Shen, Z. Jin, Y. Wang, Multi-task refined boundary-supervision U-Net (MRBSU-net) for gastrointestinal stromal tumor segmentation in endoscopic ultrasound (EUS) images, IEEE Access 8 (2020) 5805–5816.

[16] Y. Ruan, D. Li, H. Marshall, T. Miao, T. Cossetto, I. Chan, O. Daher, F. Accorsi, A. Goela, S. Li, MB-FSGAN: Joint segmentation and quantification of kidney tumor on CT by the multi-branch feature sharing generative adversarial network, Med. Image Anal. 64 (2020) 101721.

[17] E.Z. Chen, X. Dong, X. Li, H. Jiang, R. Rong, J. Wu, Lesion attributes segmentation for melanoma detection with multi-task u-net, in: 2019 IEEE 16th International Symposium on Biomedical Imaging, ISBI 2019, IEEE, 2019, pp. 485–488.

[18] X. Liu, H. Wang, AdvNet: Multi-task fusion of object detection and semantic segmentation, in: 2019 Chinese Automation Congress, CAC, IEEE, 2019, pp. 3359–3362.

[19] P. Liu, H. Zhao, P. Li, F. Cao, Automated classification and measurement of fetal ultrasound images with attention feature pyramid network, in: Second Target Recognition and Artificial Intelligence Summit Forum, vol. 11427, International Society for Optics and Photonics, 2020, p. 114272R.

[20] S. Ruder, An overview of multi-task learning in deep neural networks, 2017, arXiv preprint arXiv:1706.05098.

[21] X. Zhen, S. Li, Towards direct medical image analysis without segmentation, 2015, arXiv preprint arXiv:1510.06375.

[22] X. Zhen, Z. Wang, A. Islam, M. Bhaduri, I. Chan, S. Li, Direct volume estimation without segmentation, in: Medical Imaging 2015: Image Processing, vol. 9413, International Society for Optics and Photonics, 2015, p. 94132G.

[23] L. Yu, H. Chen, Q. Dou, J. Qin, P.-A. Heng, Automated melanoma recognition in dermoscopy images via very deep residual networks, IEEE Trans. Med. Imaging 36 (4) (2016) 994–1004.

[24] I.G. Díaz, Incorporating the knowledge of dermatologists to convolutional neural networks for the diagnosis of skin lesions, 2017, arXiv preprint arXiv:1703.01976.

[25] Y. Xie, J. Zhang, Y. Xia, C. Shen, A mutual bootstrapping model for automated skin lesion segmentation and classification, IEEE Trans. Med. Imaging 39 (7) (2020) 2482–2493.

[26] G. Wang, W. Li, S. Ourselin, T. Vercauteren, Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks, in: International MICCAI Brainlesion Workshop, Springer, 2017, pp. 178–190.

[27] P. Mlynarski, H. Delingette, A. Criminisi, N. Ayache, 3D convolutional neural networks for tumor segmentation using long-range 2D context, Comput. Med. Imaging Graph. 73 (2019) 60–72.

[28] D. Guo, L. Wang, T. Song, G. Wang, Cascaded global context convolutional neural network for brain tumor segmentation, in: International MICCAI Brainlesion Workshop, Springer, 2019, pp. 315–326.

[29] J. Redmon, A. Farhadi, Yolov3: An incremental improvement, 2018, arXiv preprint arXiv:1804.02767.

[30] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[31] Q. Hou, D. Zhou, J. Feng, Coordinate attention for efficient mobile network design, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 13713–13722.

[32] A. Biswas, P. Bhattacharya, S. Maity, 3D segmentation of liver and its lesions using optimized geometric contours, Procedia Comput. Sci. 133 (2018) 240–247.

[33] A. Saito, S. Nawano, A. Shimizu, Joint optimization of segmentation and shape prior from level-set-based statistical shape model, and its application to the automated segmentation of abdominal organs, Med. Image Anal. 28 (2016) 46–65.

[34] S. Krishnamoorthy, Y. Zhang, S. Kadry, W. Yu, Framework to segment and evaluate multiple sclerosis lesion in MRI slices using VGG-unet, Comput. Intell. Neurosci. 2022 (2022).

[35] R. Deng, C. Shen, S. Liu, H. Wang, X. Liu, Learning to predict crisp boundaries, in: Proceedings of the European Conference on Computer Vision, ECCV, 2018, pp. 562–578.

[36] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2018, pp. 3–11.

[37] Y. Zhang, H. Liu, Q. Hu, Transfuse: Fusing transformers and cnns for medical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2021, pp. 14–24.

[38] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A.L. Yuille, Y. Zhou, Transunet: Transformers make strong encoders for medical image segmentation, 2021, arXiv preprint arXiv:2102.04306.