Fetal Head Circumference Prediction Using Machine Learning

Nguyen Minh Dat
Machine Learning in Medicine
USTH University
Email: datnm.23bi14089@usth.edu.vn

Abstract—**Fetal head circumference (HC) is an important indicator of fetal development in prenatal ultrasound. This study proposes a U-Net–based segmentation approach to estimate HC from 2D ultrasound images automatically. Two annotation representations, contour-based annotations and filled binary masks, are evaluated. The results show that contour-based annotations produce broken masks and inaccurate HC estimation, while filled binary masks enable more complete segmentation and more reliable HC measurements. These findings highlight the importance of annotation representation and demonstrate the potential of deep learning for automated fetal head circumference estimation.**

I. INTRODUCTION

During pregnancy, monitoring fetal development plays a crucial role in ensuring fetal health. Head Circumference (HC) is a key biometric measurement used to assess whether the fetal head size is within the normal range for a given gestational age, and it is considered an essential parameter in prenatal ultrasound examinations.

With the rapid development of modern computational techniques, advanced scientific models have been increasingly adopted to assist clinicians in achieving faster and more accurate diagnoses. In this context, this study focuses on applying machine learning models to predict fetal head circumference, aiming to support medical decision-making and improve diagnostic efficiency.

II. DATA DESCRIPTION

The dataset used in this study is entitled *Automated Measurement of Fetal Head Circumference Using 2D Ultrasound Images*. It consists of two subsets: a training set and a testing set. The training set contains 999 fetal ultrasound images along with 999 corresponding annotation images, while the testing set includes 355 ultrasound images. All ultrasound images are provided in grayscale format with a fixed spatial resolution. The annotation images provide ground-truth information required for supervised learning.

In addition to the image data, the dataset provides several associated features that are used for head circumference (HC) measurement and evaluation. These features are summarized in Table I.

TABLE I
DESCRIPTION OF DATASET FEATURES

Feature	Description
Filename	Name of the ultrasound image file corresponding to each sample.
Pixel size (mm)	Physical size represented by each pixel in millimeters, used to convert pixel-based measurements to real-world units.
Head Circumference (mm)	Ground-truth fetal head circumference measured in millimeters, provided by expert annotation.

III. METHODOLOGY

A. Data Preparation

Since the image directory contains both ultrasound images and corresponding annotation images, they were first separated into two individual folders before the training process.

Image Preprocessing: All ultrasound images and annotation masks were resized to a fixed resolution of 256×256 pixels to ensure consistent input dimensions for the model. The ultrasound images were normalized to the range $[0, 1]$ to stabilize the training process.

Mask Preprocessing: The annotation images were converted into binary masks, where pixel values of 0 represent the background, and 1 represent the fetal head region.

B. Annotation Representation

In this study, two types of annotation representations were used. The first type is contour-based annotations, which provide only the boundary of the fetal head. The second type is a filled binary mask, where pixels inside the head contour are labeled as foreground, and pixels outside are labeled as background.
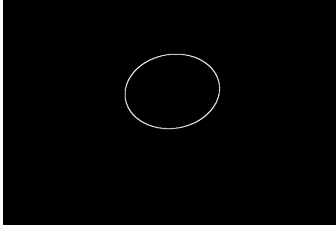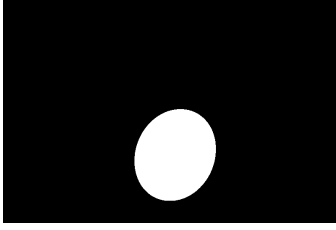
Fig. 1. Contour-based annotations

Fig. 2. Filled binary mask

C. Model Selection

The U-Net architecture was selected due to its strong performance in medical image segmentation tasks, particularly in scenarios that require accurate boundary and shape detection. In the context of fetal head circumference measurement, precise extraction of the head region from ultrasound images is a critical step.

U-Net effectively captures global contextual information while preserving fine-grained spatial details through skip connections between the encoder and decoder, making it well-suited for segmenting the fetal head region at the pixel level.

IV. RESULTS

A. Traning 1: contour-based annotations

For the first training experiment, the original annotations provided by the dataset were used. After 50 training epochs, the model loss gradually decreased from 1.29 at epoch 1 to 0.26 at epoch 50.

The trained model was then used to predict the head circumference (HC), which was calculated using the following formula:

$$HC_{\text{mm}} = \text{Perimeter}_{\text{pixel}} \times \text{PixelSize}_{\text{mm/pixel}}$$

TABLE II
PREDICTED HEAD CIRCUMFERENCE ON THE TEST SET

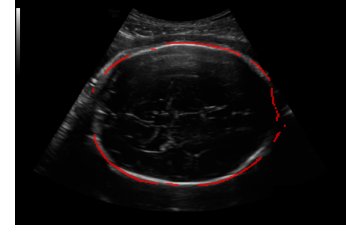| Filename	Pixel size (mm/pixel)	HC_{pred} (mm)
000_HC.png	0.2359	265.60
001_HC.png	0.0680	8.74
002_HC.png	0.1652	82.31
003_HC.png	0.0967	134.44
004_HC.png	0.2022	292.58
005_HC.png	0.1167	155.66

Fig. 3. prediction of training 1

Based on the visual results and quantitative evaluation, the predicted masks are not fully closed, which leads to incomplete extraction of the head region. Because the mask boundaries are not continuous, the head circumference is computed as the sum of the marked pixels, which introduces errors. As a result, the estimated head circumference values are not yet optimal and cannot be considered fully accurate.

B. Training 2: Filled Binary Mask

In this training experiment, to provide a more comprehensive evaluation of the training process, the dataset was split into 80% for training and 20% for validation.

TABLE III
TRAINING AND VALIDATION PERFORMANCE OVER EPOCHS FOR THE
FILLED BINARY MASK SETTING

| Epoch	Train Loss	Train Dice	Train IoU	Val Loss	Val Dice	Val IoU
1	1.101	0.405	0.280	0.979	0.582	0.414
2	0.834	0.673	0.511	0.839	0.686	0.528
3	0.765	0.701	0.544	1.175	0.418	0.266
4	0.730	0.719	0.566	0.965	0.521	0.355
5	0.682	0.740	0.591	0.761	0.687	0.528
6	0.651	0.752	0.607	0.668	0.744	0.599
7	0.615	0.768	0.628	0.771	0.685	0.526
8	0.599	0.774	0.636	1.262	0.369	0.229
9	0.589	0.779	0.642	0.600	0.776	0.641
10	0.561	0.789	0.657	0.730	0.681	0.521
11	0.544	0.795	0.664	0.624	0.774	0.638
12	0.518	0.806	0.679	0.914	0.632	0.466
13	0.509	0.811	0.685	0.633	0.777	0.642
14	0.496	0.815	0.693	0.573	0.792	0.663
15	0.479	0.822	0.701	0.551	0.780	0.643
16	0.476	0.823	0.702	0.510	0.808	0.684
17	0.459	0.829	0.713	0.476	0.820	0.700
18	0.456	0.830	0.713	0.497	0.812	0.688
19	0.431	0.841	0.729	0.570	0.798	0.671
20	0.441	0.836	0.723	0.460	0.825	0.707
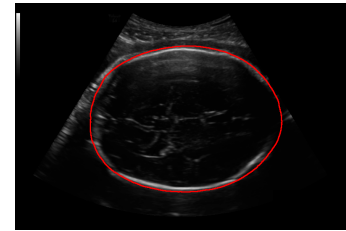
Fig. 4. prediction of training 2

The model shows a stable learning trend, as the training loss steadily decreases from 1.101 to 0.441. At the same time, the Train Dice and Train IoU continuously increase, reaching 0.836 and 0.723 at epoch 20, respectively. This indicates that

the model is able to effectively learn segmentation features from the training data.

For the validation set, although there are strong fluctuations at some epochs (especially at epochs 3 and 8 where the validation loss increases), the model still achieves good performance in the later stage. At epoch 20, the model obtains the best results with a Val Dice of 0.825 and a Val IoU of 0.707, suggesting that the generalization ability has significantly improved.

TABLE IV
PREDICTED HEAD CIRCUMFERENCE (HC) MEASUREMENTS

Filename	Pixel size (mm)	HC_{pred} (mm)
000_HC.png	0.2359	300.40
001_HC.png	0.0680	71.77
002_HC.png	0.1652	212.62
003_HC.png	0.0967	168.51
004_HC.png	0.2022	320.22
005_HC.png	0.1167	169.47

In this experiment, visual inspection shows that the predicted masks are drawn more accurately, with closed, continuous boundaries rather than broken lines. Therefore, the head circumference (HC) estimation in this training run is more accurate compared to the first training experiment.

V. CONCLUSION

This study applies a U-Net–based segmentation approach to estimate fetal head circumference from 2D ultrasound images. The results show that annotation representation plays an important role in segmentation quality and HC accuracy. Contour-based annotations tend to produce broken and discontinuous boundaries, whereas filled binary masks allow the model to generate more complete and continuous annotations. As a result, the second training experiment demonstrates stable learning, improved generalization, and more reliable HC measurements.