Automatic Measurement of Fetal Head Circumference from Ultrasound Images Using Deep Learning

Pham Minh Duc (22BA13083
Machine Learning in Medicine 2026
University of Science and Technology of Hanoi (USTH)
Email: ducpm.22ba13083@usth.edu.vn

***Abstract*—Accurate measurement of fetal head circumference (HC) from ultrasound images is essential for monitoring fetal growth and detecting potential abnormalities during pregnancy. Manual measurement is time-consuming and subject to inter-observer variability. In this study, we investigate an automated approach for estimating fetal head circumference using deep learning regression. A convolutional neural network (CNN) is trained on the HC18 public dataset to directly regress head circumference values from ultrasound images. Model performance is evaluated using Mean Absolute Error (MAE). Experimental results demonstrate the feasibility of end-to-end regression from ultrasound images while highlighting challenges related to model underfitting and dataset variability.**

I. Introduction

Fetal biometry is a fundamental component of prenatal care, with head circumference serving as a key indicator of fetal development and gestational age. Ultrasound imaging is widely adopted due to its non-invasive nature and real-time capabilities. However, manual head circumference measurement is dependent on operator expertise and is subject to variability.

With the advancement of machine learning, deep learning models—particularly convolutional neural networks—have demonstrated strong performance in medical image analysis. This work explores a deep learning-based regression approach that directly estimates fetal head circumference from ultrasound images without requiring explicit segmentation.

II. Dataset

This study uses the HC18 dataset, a publicly available benchmark for fetal head circumference measurement. The dataset consists of two-dimensional ultrasound images with corresponding ground truth head circumference values provided in millimeters.

The dataset includes:

- A labeled training set with ultrasound images and head circumference annotations.
- A separate unlabeled test set used for inference.

To evaluate model performance, the labeled data is split into training (80%) and validation (20%) subsets.

III. Methods

A. Preprocessing

All ultrasound images are converted to grayscale, normalized to the range [0,1], and resized to 256×256 pixels. Each image is represented as a single-channel tensor.

B. Model Architecture

A convolutional neural network is designed for regression. The architecture consists of three convolutional layers with ReLU activation and max pooling, followed by an adaptive average pooling layer and a fully connected regression head. The network outputs a single continuous value representing the fetal head circumference.

C. Training Procedure

The model is trained using the Adam optimizer with a learning rate of 10^{-4}. Mean Absolute Error (MAE) is used as the loss function due to its clinical interpretability. Training is conducted for 20 epochs with a batch size of 8.

IV. Results

A. Training and Validation Performance

Figure 1 shows the training and validation MAE across epochs. The relatively flat curves indicate limited learning progress, suggesting underfitting.

B. Prediction Accuracy

Figure 2 presents a scatter plot of predicted versus ground truth head circumference values on the validation set. The red dashed line represents the ideal prediction ($y = x$). Predictions cluster around a narrow range, indicating regression toward the mean.

C. Error Distribution

The distribution of absolute prediction errors is shown in Figure 3. While some predictions exhibit low error, a long tail of large errors is observed.

D. Qualitative Results

Figure 4 displays example ultrasound images from the validation set along with their ground truth and predicted head circumference values. The model performs better on mid-range values but struggles with extreme cases.

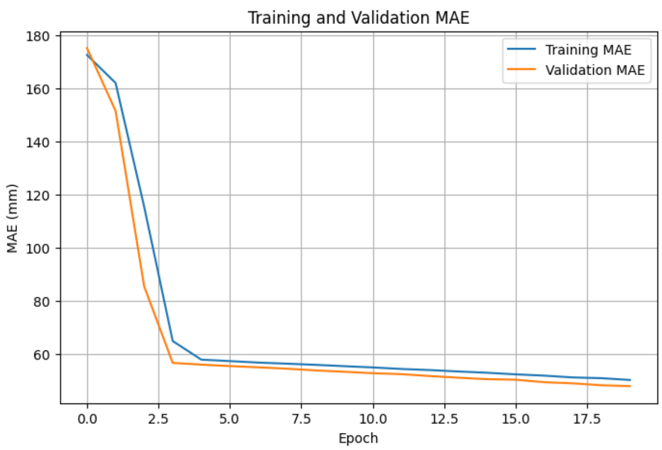

Fig. 1. Training and validation MAE over epochs.

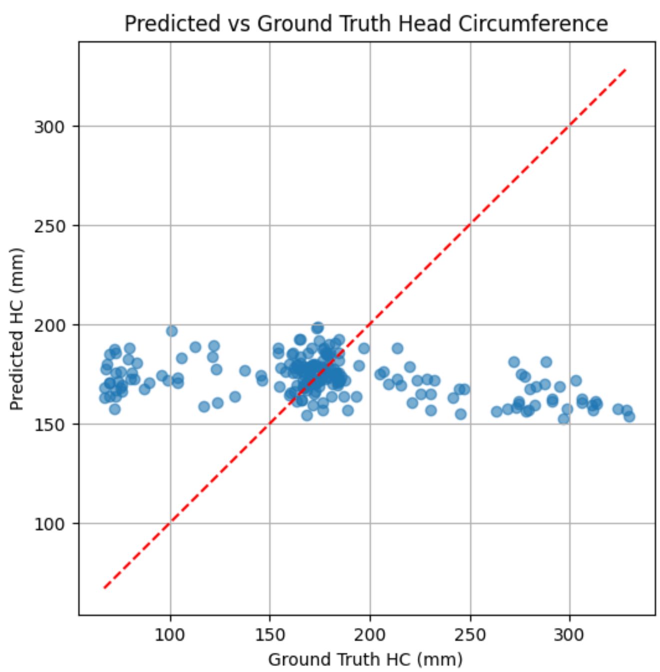

Fig. 2. Predicted versus ground truth head circumference values.

V. Discussion

The experimental results reveal that the proposed CNN is able to learn coarse relationships between ultrasound image appearance and head circumference. However, the consistently high MAE and clustered predictions indicate underfitting.

Potential reasons include limited model capacity, high variability in ultrasound image quality, and the absence of explicit anatomical constraints. Incorporating data augmentation, deeper architectures, or combining segmentation with regression may improve performance.

VI. Conclusion

This study presents a deep learning-based regression approach for automatic fetal head circumference estimation from

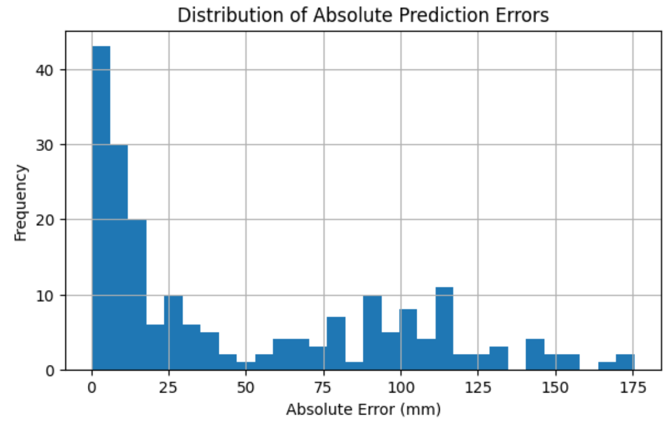

Fig. 3. Distribution of absolute prediction errors.

ultrasound images. Although the current model demonstrates limited accuracy, it establishes a baseline for further research. Future work will focus on architectural improvements and enhanced preprocessing to achieve clinically acceptable performance.

VII. References

References

[1] T. L. A. van den Heuvel, D. de Bruijn, C. L. de Korte, and B. van Ginneken, “Automated measurement of fetal head circumference using 2D ultrasound images,” HC18 dataset, Zenodo, DOI: 10.5281/zenodo.1327317, 2018. [Online]. Available: https://zenodo.org/records/1327317. :contentReference[oaicite:2]index=2

[2] HC18 Grand Challenge, “Automated measurement of fetal head circumference,” Grand-Challenge.org. [Online]. Available: https://hc18.grand-challenge.org/. Accessed: Oct. 25, 2025. :contentReference[oaicite:3]index=3

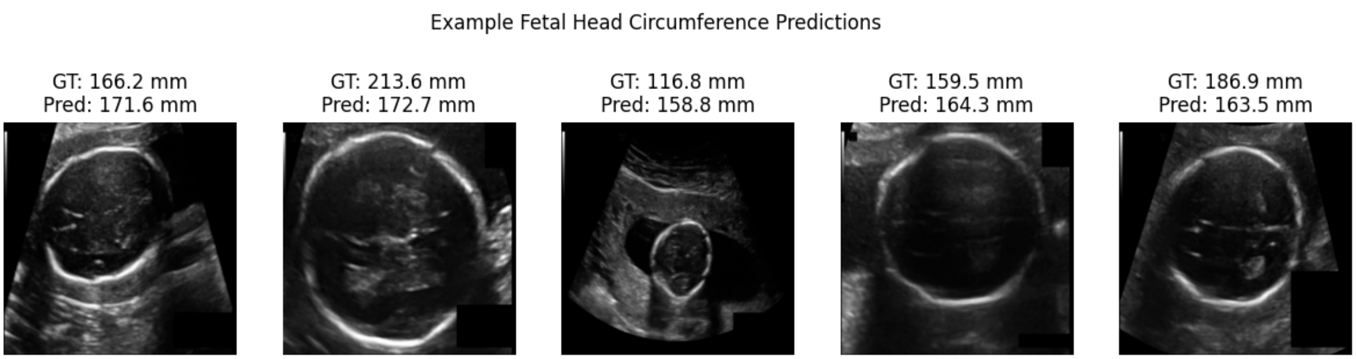

Fig. 4. Example fetal head circumference predictions on ultrasound images.