

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—Measuring fetal head circumference (HC) accurately from ultrasound images is important for tracking fetal growth and identifying possible health problems during pregnancy. However, manual measurement takes a lot of time and may differ between doctors. This study explores an automatic method to estimate fetal head circumference using deep learning. A convolutional neural network (CNN) is trained on the public HC18 dataset to predict head circumference values directly from ultrasound images. The performance of the model is evaluated using Mean Absolute Error (MAE). The results show that this end-to-end regression approach is possible, but they also reveal challenges such as model underfitting and differences in the dataset.

I. INTRODUCTION

Fetal biometry plays an important role in prenatal care, and head circumference is one of the main indicators used to assess fetal growth and gestational age. Ultrasound imaging is commonly used because it is safe, non-invasive, and provides real-time results. However, measuring head circumference manually depends heavily on the experience of the operator and can lead to inconsistent results.

With recent advances in machine learning, deep learning models, especially convolutional neural networks, have shown strong results in medical image analysis. This study investigates a deep learning regression approach that directly predicts fetal head circumference from ultrasound images without the need for explicit image segmentation.

II. DATASET

This study uses the HC18 dataset, which is a publicly available benchmark for fetal head circumference measurement. The dataset contains two-dimensional ultrasound images with ground truth head circumference values measured 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 the performance of the model, the labeled data is divided 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 the regression task. The model includes three convolutional layers with ReLU activation and max pooling, followed by an adaptive average pooling layer and a fully connected regression layer. The network produces a single continuous value that represents 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 chosen as the loss function because it is easy to interpret in a clinical context. The training process runs 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.

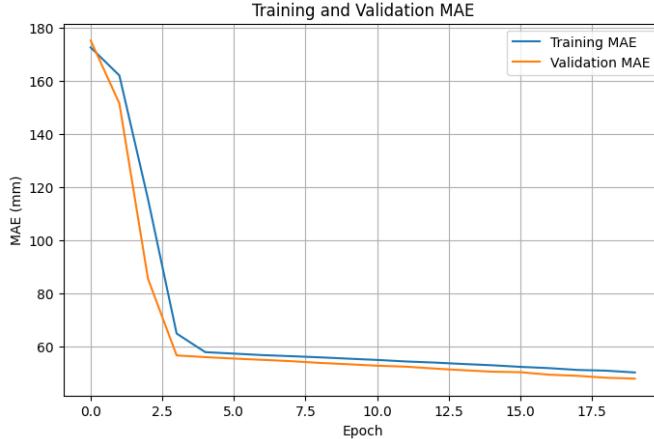


Fig. 1. Training and validation MAE over epochs.

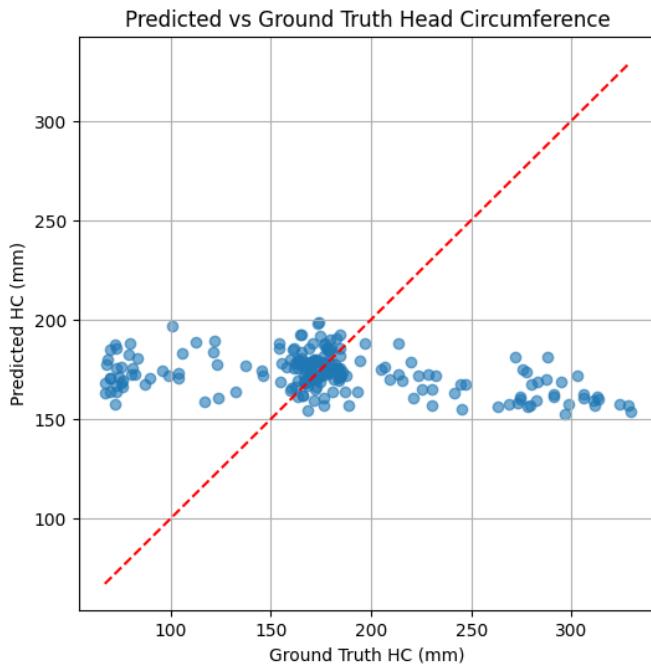


Fig. 2. Predicted versus ground truth head circumference values.

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.

V. DISCUSSION

The experimental results show that the proposed CNN can learn basic relationships between the appearance of ultrasound images and head circumference. However, the consistently high MAE and the clustering of predictions suggest that the model is underfitting.

Possible reasons for this include limited model capacity, large differences in ultrasound image quality, and the lack of

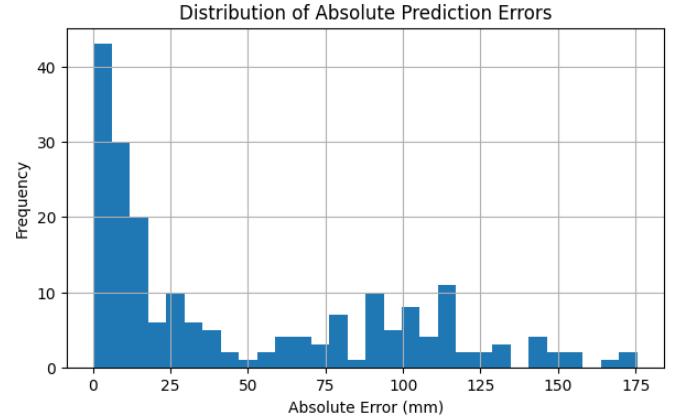


Fig. 3. Distribution of absolute prediction errors.

explicit anatomical constraints. Performance may be improved by using data augmentation, deeper network architectures, or combining segmentation with regression.

VI. CONCLUSION

This study presents a deep learning regression approach for automatically estimating fetal head circumference from ultrasound images. Although the current model shows limited accuracy, it provides a baseline for future research. Future work will focus on improving the model architecture and preprocessing methods to reach clinically acceptable performance.

VII. REFERENCES

REFERENCES

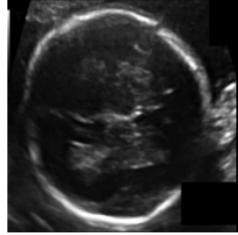
- [1] T. L. A. van den Heuvel, D. de Bruijn, C. L. de Korte, and B. van Ginneken, “Automated measurement of fetal head circumference using 2D ultrasound images,” HC18 dataset, Zenodo, DOI: 10.5281/zenodo.1327317, 2018. [Online]. Available: <https://zenodo.org/records/1327317>. :contentReference[oaicite:2]index=2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

Example Fetal Head Circumference Predictions

GT: 166.2 mm
Pred: 171.6 mm



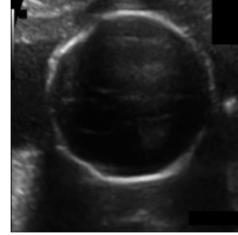
GT: 213.6 mm
Pred: 172.7 mm



GT: 116.8 mm
Pred: 158.8 mm



GT: 159.5 mm
Pred: 164.3 mm



GT: 186.9 mm
Pred: 163.5 mm



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