I. Abstract

Accurate measurement of fetal head circumference (HC) from ultrasound images is essential for monitoring fetal growth and detecting developmental abnormalities. Manual measurement is time-consuming and operator-dependent, motivating the development of automated approaches. This study presents a simple end-to-end convolutional neural network (CNN) for direct regression of fetal head circumference from 2D ultrasound images. Unlike segmentation-based methods, the proposed approach predicts HC values directly from image appearance while incorporating pixel size information to preserve physical scale. Experiments conducted on the HC18 dataset demonstrate that the proposed model achieves stable convergence and reasonable prediction accuracy, establishing a strong baseline for automated fetal biometry.

II. Introduction

Fetal head circumference is a key biometric indicator used in prenatal care to assess fetal growth and identify potential abnormalities. In clinical practice, HC is typically measured manually by sonographers using ellipse fitting on ultrasound images. This process is subject to inter- and intra-observer variability and requires considerable expertise.

Recent advances in deep learning have enabled automated analysis of medical images, including fetal ultrasound. Many existing approaches rely on segmentation-based pipelines, where the fetal head boundary is first extracted and geometric measurements are subsequently computed. While effective, such methods often require complex architectures, such as U-Net, and precise annotations. In this work, we explore a simpler alternative: an end-to-end CNN that directly regresses head circumference from ultrasound images. By incorporating pixel size as an auxiliary numerical input, the model maintains awareness of physical scale without explicitly performing segmentation. The objective of this study is to evaluate whether a lightweight CNN can serve as a reliable baseline for automated HC estimation.

III. Dataset

A. HC18 Dataset

Experiments were conducted using the HC18 dataset, which contains 2D fetal ultrasound images annotated for head circumference measurement. Each sample consists of: A grayscale ultrasound image and a corresponding annotation mask indicating the head contour

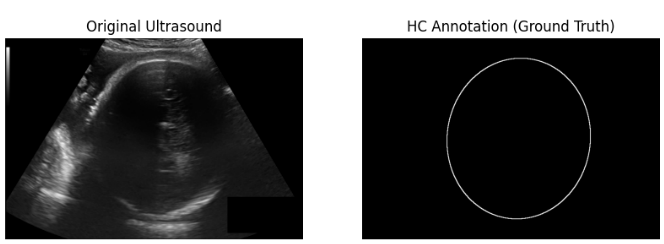

Fig. 1. Data sample

	filename	pixel size(mm)	head circumference (mm)
0	000_HC.png	0.069136	44.30
1	001_HC.png	0.089659	56.81
2	002_HC.png	0.062033	68.75
3	003_HC.png	0.091291	69.00
4	004_HC.png	0.061240	59.81

Total rows: 999

A CSV file providing pixel size (mm per pixel) and ground truth head circumference (mm)

In this study, only the ultrasound images and numerical metadata were used for model training, while annotation masks were reserved for visualization and future segmentation-based extensions.

B. Data Split

The dataset was divided into three disjoint subsets:

Training set: 70 %

Validation set: 15 %

Test set: 15 %

The split was performed at the sample level to ensure no overlap between sets.

IV. Data Preprocessing

Each input sample was preprocessed using a consistent pipeline: Grayscale loading: Ultrasound images were loaded in grayscale format.

Resizing: Images were resized to 256×256 \$$times256256pixels$.

Normalization: Image intensities were normalized to the range [0,1][0, 1][0,1].

Channel expansion: A single channel dimension was added to match CNN input requirements.

Auxiliary input extraction: Pixel size (mm) and ground truth HC (mm) were read directly from the CSV file.

Annotation masks were binarized and resized for inspection purposes but were not used as model inputs.

V. Methodology

Model Architecture The proposed model is a dual-input CNN consisting of two branches: Image branch: Extracts visual features from ultrasound images using convolutional and max-pooling layers.

Pixel size branch: Processes pixel size information through fully connected layers.

The extracted features from both branches are concatenated and passed through additional dense layers to predict a single continuous output representing head circumference in millimeters. This design allows the model to learn image-based representations while explicitly accounting for physical scaling.

A. Training Configuration

The model was trained using the following settings: Optimizer: Adam

Learning rate: 1×1041 $\times 10^{-4}1104$

Loss function: Mean Squared Error (MSE)

Evaluation metric: Mean Absolute Error (MAE)

Batch size: 8

Number of epochs: 30

Training and validation were performed using data generators to efficiently load and preprocess samples on-the-fly.

VI. Experimental Results

A. Learning Curves

Training and validation loss curves demonstrate stable convergence over epochs, with no significant divergence between training and validation performance. This indicates that the model does not suffer from severe overfitting despite the limited dataset size.

Similarly, the MAE curves show a consistent decrease during training, suggesting effective learning of the regression task.

B. Predicted vs Ground Truth Head Circumference

A scatter plot of predicted versus ground truth HC values on the test set shows that predictions generally align with the identity line. Most samples exhibit small deviations, indicating that the model captures the relationship between ultrasound appearance, pixel size, and head circumference.

C. Quantitative Evaluation

Performance was evaluated using Mean Absolute Error on the test set. The achieved MAE, measured in millimeters, demonstrates that the proposed simple CNN provides reasonable accuracy for automated HC estimation and serves as a strong baseline for further improvements.

VII. Discussion

The results indicate that direct regression of head circumference from ultrasound images is feasible using a relatively simple CNN architecture. Incorporating pixel size as an auxiliary input is crucial, as it provides the model with physical scale information that cannot be inferred reliably from image appearance alone. Compared to segmentation-based approaches, the proposed method offers several advantages: Reduced architectural complexity

Faster training and inference

No reliance on precise contour extraction

However, the lack of explicit geometric constraints may limit accuracy in challenging cases, such as low image quality or partial head visibility. Future work could combine this regression approach with segmentation-based methods to improve robustness and interpretability.

VIII. Conclusion

This study presented a simple end-to-end CNN for automated estimation of fetal head circumference from 2D ultrasound images. By directly regressing HC values and incorporating pixel size information, the proposed method avoids complex segmentation pipelines while achieving stable and accurate predictions. The model provides a solid baseline for automated fetal biometry and can be extended with more advanced architectures or multi-task learning strategies in future research.