Machine learning in Medical

Ultrasound: Measurement of Fetal head circumference using Regression method

Tran Trung Kien - 22BI13227

Abstract—The fetus head circumference (HC) is a key biometric to monitor fetus growth during pregnancy, which is estimated from ultrasound (US) images. The standard approach to automatically measure the HC is to use a segmentation network to segment the skull, and then estimate the head contour length from the segmentation map via ellipse fitting, usually after post-processing. In this lab work, another method: estimate directly the HC with a regression network is carried out and evaluated.

1. Introduction

Fetal head *tau class* circumference (HC) measurement is an important biometric parameter used in prenatal assessments. Traditionally, this measurement is performed manually by healthcare professionals on ultrasound images, which is both time-consuming and dependent on operator expertise.

Recent advances in deep learning have enabled automated estimation of fetal HC from ultrasound images, either through segmentation-based or regression-based approaches. In this labwork, we focus on implementing a regression-based method to estimate the head circumference directly from images, without requiring manual annotations or segmentation masks.

The objective of this lab is to reproduce and evaluate a regression model based on EfficientNetB2, as proposed in the study "Segmentation-Based vs. Regression-Based Biomarker Estimation: A Case Study of Fetus Head Circumference Assessment from Ultrasound Images." Our implementation includes dataset preparation, data augmentation, model training, and performance evaluation on the HC18 challenge dataset.

2. Dataset: HC18

2.1. Data description

The dataset used in this lab-work is the HC18 Challenge dataset, which consists of fetal ultrasound images aimed at evaluating automatic fetal head circumference (HC) measurement techniques. The dataset includes a total of 999 training images and 335 test images.

Each training sample is provided with its corresponding head circumference ground truth, expressed in millimeters, and pixel spacing information. In addition, binary segmentation masks that outline the fetal head are available for each training image. However, in this lab-work, the main focus is the regression task which uses the head circumference values as labels for model training.

The dataset is structured into a training folder, containing both the ultrasound images and their annotations, and a testing folder, which contains ultrasound images without annotations. Two CSV files, train.csv and test.csv, provide the corresponding metadata, including the filename, pixel size, and head circumference (only for training data).

2.2. Data pre-processing

The ultrasound pictures were resized from their original resolution (800×540 pixels) to a fixed size of 224×224 pixels. This allowed the images to match the expected input shape of the EfficientNetB2. Pixel intensity values were normalized to the [0, 1]. To improve the generalization ability of the model, data augmentation techniques were applied to the training set: each image was augmented through random rotations within a $\pm$10-degree range and horizontal flipping.

3. Model

3.1. Model initialization

The regression model implemented in this lab-work is based on the EfficientNetB2 architecture, which serves as the feature extraction backbone. e model was initialized with weights pre-trained on the ImageNet dataset.

On top of the EfficientNetB2 backbone, a regression head was added to predict the fetal head circumference (HC) value directly from the input ultrasound images. The regression head consists of a Global Average Pooling (GAP) layer, followed by a Dropout layer with a dropout rate of 0.7. The final layer is a Dense layer with a single output node and a linear activation function to the predict the head circumference of the infants.

All layers of the EfficientNetB2 backbone were set as trainable to allow fine-tuning on the HC18 dataset.

3.2. Model training

As suggested in the paper, the regression model was trained using the Mean Absolute Error (MAE) loss function, which directly optimizes the difference between the predicted and actual values of head circumference. The Adam optimizer was used with a learning rate of 1e-4 to update the model weights.

The training process was conducted over the maximum of 100 epochs. EarlyStopping callback was employed, monitoring the validation loss (MAE) with a patience of 10 epochs. The best model was achieved after 31 epochs.

4. Result

4.1. Training result

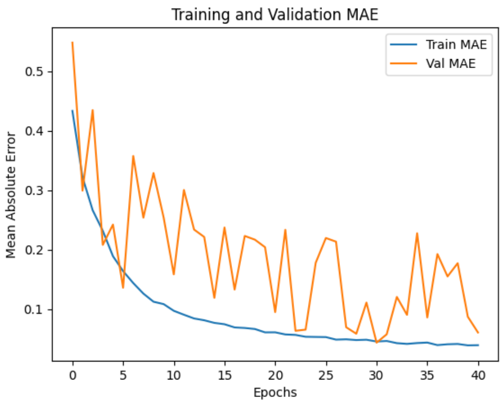

Figure 1. Training and Validation MAE Curve

The training MAE shows a consistent downward trend. However, the validation MAE exhibits significant fluctuations throughout the training process, without a clear and stable convergence.

4.2. Testing result

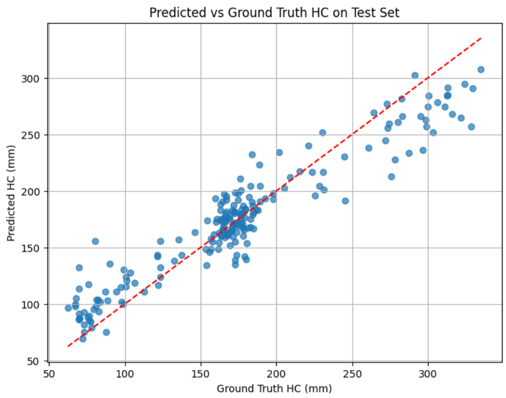

Figure 2. Predicted vs. Ground Truth HC Plot

Figure 2 presents a scatter plot of the predicted head circumference (HC) values versus the ground-truth HC values on the test set. The red dashed line represents the ideal scenario where the predicted values perfectly match the ground truth.

Most predictions align with the general trend of ground truth. However, the deviations from the idea line at the 2 end of the line is significant, suggesting that the model is struggling with extreme values.

The model achieved a MAE of **17.50 mm** and a PMAE of **11.81** percent on the test set. Although the model was able to learn from the training data and predict head circumference values, its accuracy was lower than expected, especially compared to the results reported in the referenced paper (MAE = 1.83 mm).

