

Fetal head segmentation

Nguyen Xuan Minh Vu

I. INTRODUCTION

Nowadays, before giving birth to a baby, the mother must do many prenatal check-ups for health monitoring. This ensures that the baby will be born healthy and detect any problems soon. One of the most common procedures for prenatal check-ups is ultrasound. The ultrasound works by emitting very high-frequency signals (or sounds) to build and reconstruct the image of structure inside the body.

Ultrasound is used to investigate many body parts of the fetus, especially the fetal head. The fetal head is egg-shaped, being broader posteriorly and symmetric without irregularity of contour. Ossification of the skull vault is complete by 12 weeks gestation with the sutures remaining visible throughout pregnancy. By doing the ultrasound, the doctor can recognize any abnormality in the fetus' head.

The segmentation task is an important procedure in ultrasound as it can accurately capture the fetal head position inside the mother's body using a 2D ultrasound image. Usually, it is done manually by the doctor; however, this may cost time when there are too many images. It is also complicated to utilize modern machines as the data must be uploaded handily instead of segmenting automatically. Therefore, during the scope of this report, I introduce a segmentation model using convolutional neural networks (CNN) to detect the shape and position of the fetus's head, which serves the works of analyzing and investigating abnormalities later.

This project will use the Unet model and the HC18 dataset.

II. BACKGROUND

A. Head circumference (HC)

During pregnancy, ultrasound imaging measures fetal biometrics. One of these measurements is the fetal head circumference (HC). The HC can be used to estimate the gestational age and monitor the growth of the fetus. The HC is measured in a specific cross-section of the fetal head, which is called the standard plane.

B. Convolutional Neural Network (CNN)

CNN excels at processing data like images, speech, and audio signals. This is due to their unique architecture composed of three main layers: The convolutional layer is to extract information. The pooling layer reduces the dimensionality of the data and captures spatial variance. Lastly, the fully connected layer combines the extracted features from earlier layers and makes the final classification or prediction.

III. DATASET

In this project, I will use the HC18 dataset, provided by Grand Challenge. This dataset includes 999 images for the training set and 335 images for the test set. Those images are captured by ultrasound, and present the real fetus's head inside the mother's body. With each image in the training dataset, there is an annotation dataset, which bounds the shape of the fetus's head. This is used as the label of the corresponding ultrasound image. On the other hand, the test set is independently used to evaluate the model so there is no annotation.

IV. METHOD

A. Data preprocessing

There are 999 data samples in the training set and most of them are in the shape of 540x800. However, some images have different shapes so I decided to cut them off. Therefore, the number of data for training is 975. They are resized to (572x572) for the ultrasound images and to (388x388) for the corresponding annotation images.

B. CNN architecture

The Unet model is in the "U" shape. The left side will undertake the task of extracting features while the right side will combine them to produce the result of a segmented image. The special of this architecture is that there is no fully connected layer, the right side of the Unet model will undertake it. Therefore, the input size can be varied. The output size is in a particular shape (388x388) to duplicate the annotation image size.

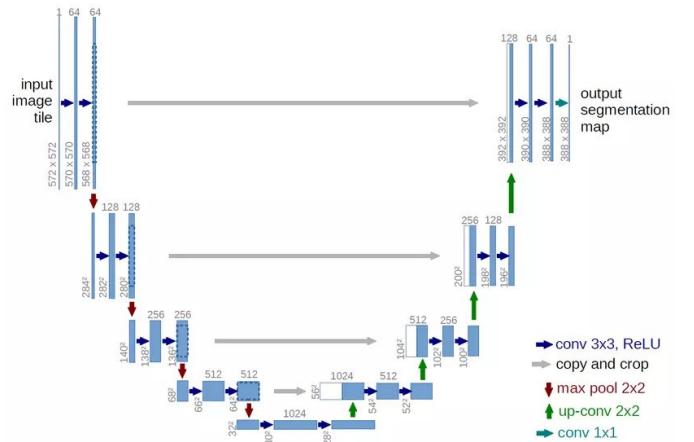


Fig. 1. Overall Unet model

C. Training

For training, I use the RMSProp optimizer algorithm with a learning rate = 0.0001, momentum = 0.999, and weight decay = 0.0000001. For the loss function, I use the Cross Entropy Loss. The batch size is set as 16 in this work. The model runs through 10 epochs.

D. Data splitting

The training dataset is divided into a smaller training set and a validation set. From 975 images, I use 750 samples for the training set and 225 samples for the validation set.

V. EVALUATION

Dataset
Metric
Result
Discussion

VI. CONCLUSION