Report on HC18 dataset

Tran The Trung

March 19, 2024

1 Introduction

During pregnancy of every women, people have tried to make the most of ultrasound image. Among the measurements made by this tool, fetal head circumference or HC is utilized to make prediction about gestational age and monitor growth of the fetus. Simply speaking, HC refers to the measurements of a person's head and is usually calculated in the widest point above the ears and eyebrows using a flexible measuring tape.

The dataset comes from a challenge which serve as the benchmark for comparing differnt algirthms for automated measurement of fetal head circumference. It contains of a total 1334 two-dimensional ultrasound images of the standard plane. For more details, the data has already been divided into two subsets: 999 images for training while a third of them for testing. All the images are grayscale and have the same resolutions which is 800 by 540 pixels with a variation of pixel size ranging from 0.052 to 0.326 mm. They also include a manual annotation for each image. The annotation is a .png file that draw the boundary of the fetal head and it has already been smoothen make it resemble the shape of an eclipse. Additionally, some ultrasound images were made during the same echoscopic examination and have therefore a very similar appearance.

2 Method

2.1 Architecture

For the implementation, we use a traditional model called U-Net that's primarily used for segmentation problem. The model is a simple encode-decode Convolutional Neural Network (CNN) model. It was first designed for biomedical image segmentation tasks. The reason it was named U-Net is because of its architecture making of two symmetric branches which resemble the U-shape.

Looking at the figure, UNet architecture is basically composed of two sections, contraction on the left and expansion on the right. Each will be in charge of

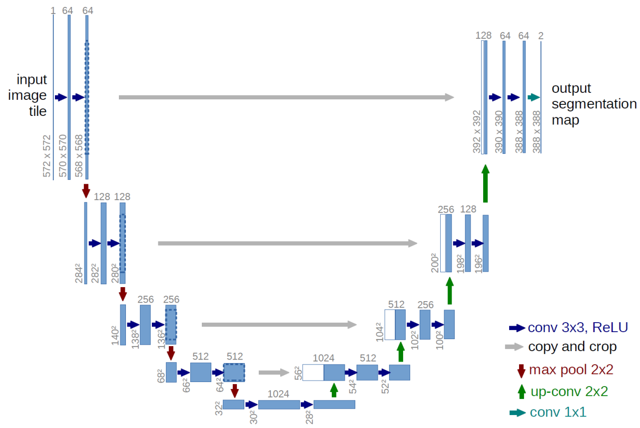

Figure 1: U-Net architecture

different tasks:

Contraction Extract features or patterns to get more information related to the image. It plays the same role to that of an Encoder which is a Deep CNN network that will act as a feature extractor. The reason for this branch to be named as contraction is because the length and width dimensions of the layers gradually increase from an input size of 572x572 to 32x32. Meanwhile, the inverse trend can be captured regarding the number of channels.

Expansion Composed of symmetric layers with what we have in the contracting path. The upsampling process being applied help increase the size of the layers gradually and in the end, we receive a mask that encode the label for each object at a pixel level. Another noticable feature is that UNet will not contain any fully-connected layer, a common layer appear in many end-to-end Deep Learning network, this allow the model to receive input at flexible resolutions.

Also, the model integrates skip-connection from downsampling path to the upsampling one.

2.2 Loss function

For this problem, we will implement different loss function: Mean Square Error loss (MSE) and Binary Cross Entropy loss (BCE)

- MSE loss is traditionally used in regression tasks while in segmentation, it is used when the task involves predicting pixel-wise intensities or values by measuring the average squares difference between the predicted and the ground-truth values

- BCE loss seems to be more suitable for this task for its purpose of classifying pixel into two classes, whether it is fetal head or not. It measures the dissimilarity between the predicted distribution and the ground truth one.

2.3 Data preprocessing

Since the annotation for the fetal head only comes as the boundary. This might not be sufficient enough for the UNet model to work well and we might need to fill out the inner part of the annotation. So first, we will find the contour which is relatively easy since the annotation has already been nicely drawn. And then we simply fill out the rest which comes in ellipse shape. The result can be shown below

Figure 2: Preprocessing

We will evaluate the performance on both of the data and see the differnce. Also, for better performance, we will conduct resize the image to 256x256 resolution and then normalize the image to 0-1 range.

2.4 Data postprocessing

Then after the model have outputed the segmentation of the whole fetal head, we will now extract the boundary of the result for evaluation process simply by reverse the process. The result is as follow:

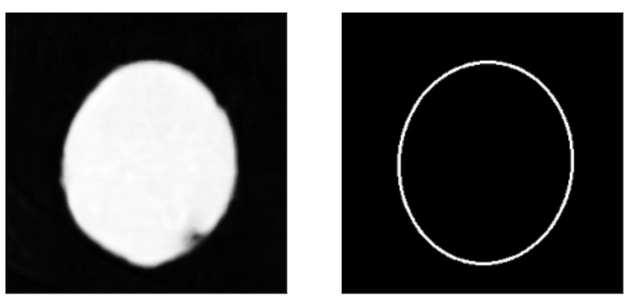

Figure 3: Postprocessing

	Without masking		With masking	
	MSE Loss	BCE Loss	MSE Loss	BCE Loss
RMSE	0.09	0.16	0.11	0.2

3 Evaluation

3.1 Metrics

For the evaluation of the problem, we will use Root Mean Square Error (RMSE) to compute the difference between the prediction and the result of the model

3.2 Result

We can see that the MSE loss have a better RMSE score than BCE. This can be explained simply due to the fact that the loss is align more with the evalution compared to its counterpart. Another point to mention is that the masking make the score becomes worse. The error may lie in the postprocess step where the generated boundary is not close enough to the ground truth image.

4 Discussion

Overall, we can see that the UNet performs pretty well on the dataset and gives out an acceptable result visually. Still, the output is not that good and in the future we can make use of some other advanced model like YOLO for segmentation.