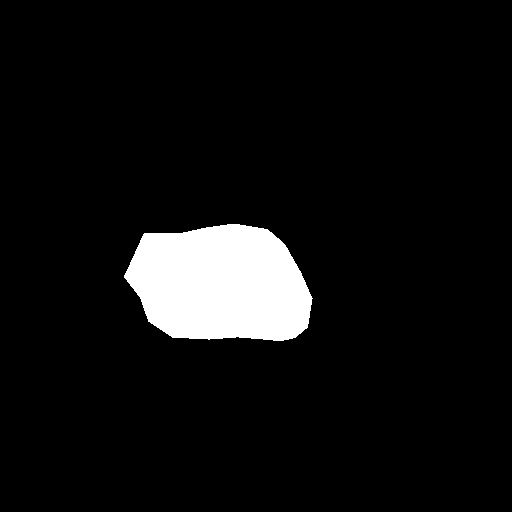
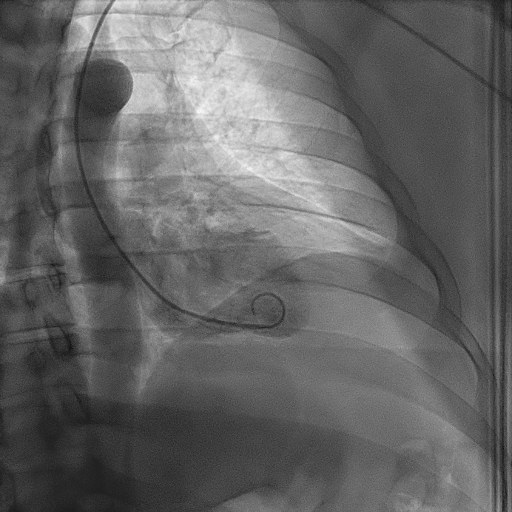
Method

Data collection

In this study, we gathered 42 patients of cardiac whole volume liver x-ray imaging data from Chi Mei Medical Center. The total collected data included 1717 slices of cardiac images with both jpg format x-ray images and mask images. Mask images were drawn by doctors (by a group of annotators under the doctor’s supervision), which later became binary masks for each image(shown below). We randomly divide data into training(95%) and validation(5%) in the training section.



Data arguementation and preprocess

Arguementation is used to mimic the condition of a real medical image influenced by patient position or individual difference, while preprocessing is used to normalize and enhance the x-ray image contrast.

Arguementation follows the 7 steps, each of them has a probability of 0.3, and the final probability of arguementation is 0.8:

1. GridDistortion
2. ElasticTransform
3. Affine (rotate, rescale, shift)
4. GaussNoise
5. Blur
6. Downscale (reduce the image quality)
7. RandomBrightnessContrast

Preprocessing: CLAHE (enhance the contrast of image)

The result of training data (argementation+preprocessing):

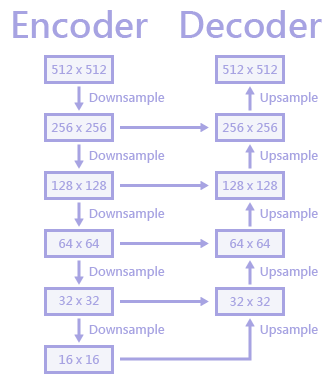
The result of validation data (argementation+preprocessing):

The result of validation data without argementation (preprocessing only):

(I’ll paste the result later, please remind me if I forget)

Models

We choose Unet as the main architecture for the segmentation task because it’s the most commonly used segmentation model in medical images. Unet is composed of two major parts: encoder and decoder.



The encoder part involves a series of downsample steps, and it’s the conventional architecture for most CNN models. So we conducted a series of experiments to find out which CNN model is the most suitable model for the encoder of Unet in our task. We have three candidate models: ResNet, DenseNet, EfficientNet. For fair comparison, we choose the subtypes of these models with similar numbers of parameters.

ResNet-50 : 23 M # parameters

DenseNet-201: 18 M # parameters

EfficientNet-B4 : 17M # parameters

Besides model structure itself, we also try to answer the question: Can we benefit from pretrained weights learned from Imagenet dataset even in medical image tasks? So, we further separate each model into two groups: pretrained or non-pretrained for the discussion.

ResNet-50

DenseNet-201

EfficientNet-B4

We are using ResNet-50, EfficientNet-B4, and DenseNet-201 to compare the results in order to get the best segmentation mask for cardiac image. The three models have a similar number of parameters.

Performance evaluate

To evaluate the model’s performance, we use DiceBCELoss as a validation index. Dice loss could measure the similarity of the two contours, and calculate the overlap region between prediction and mask to calculate the loss function. By combining DICE and BCE loss functions(DiceBCELoss), the model training could be improved. And each model runs for ten times to check the stability of the accuracy.

Results

We test three models, ResNet-50, EfficientNet-64, and DenseNet-20, with two conditions, pretrained/non-pretrained. In Fig.1, all the results show mean DiceBCELoss under 0.055. The lowest means of DiceBCELoss is EfficientNet-64 with a pretrained model. It shows mean DiceBCELoss = 0.0389 in train condition, 0.0391 in validation condition. Fig.2 shows trends in 30 epochs of each model; the pre-trained model converges fine in all six conditions. Fig.3 shows the concentration of 10 runs in each model in validation with augmentation, EfficientNet with a pre-trained model shows lowest DiceBCELoss and lowest DiceBCELoss standard deviation in 10 runs. Pre-trained models do not always have the lowest DiceBCELoss in our results.

