CS631T Computer Vision Project

Segmentation of the Tibia Bone from MR Images using Deep Convolutional Neural Networks U-net

Author: Junbin Sun, Lixin Chen
Seidenberg School of CSIS, Pace University, Pleasantville, NewYork 10570
Email: {js09442p, lc34664n} @pace.edu

Abstract

- MRI is gradually replacing CT for bone and joint examination. Many diagnostic applications critically depend on the successful localization of bone structure. U-net is a semantic segmentation network based on FCN, which is suitable for segmentation of medical images.Many split networks are based on FCNs for improvements, including Unet.

- In this project, we have trained a machine learning model (a convolutional neural network called U-Net) to perform bone segmentation on 3D knee magnetic resonance imaging (MRI). The goal is to develop a method that can automatically detect the bone area from a knee MRI image.

Introduction

U-net is a semantic segmentation network based on FCN, which is suitable for segmentation of medical images.
The U-net architecture is as shown below: It consists of contraction path and expansion path.

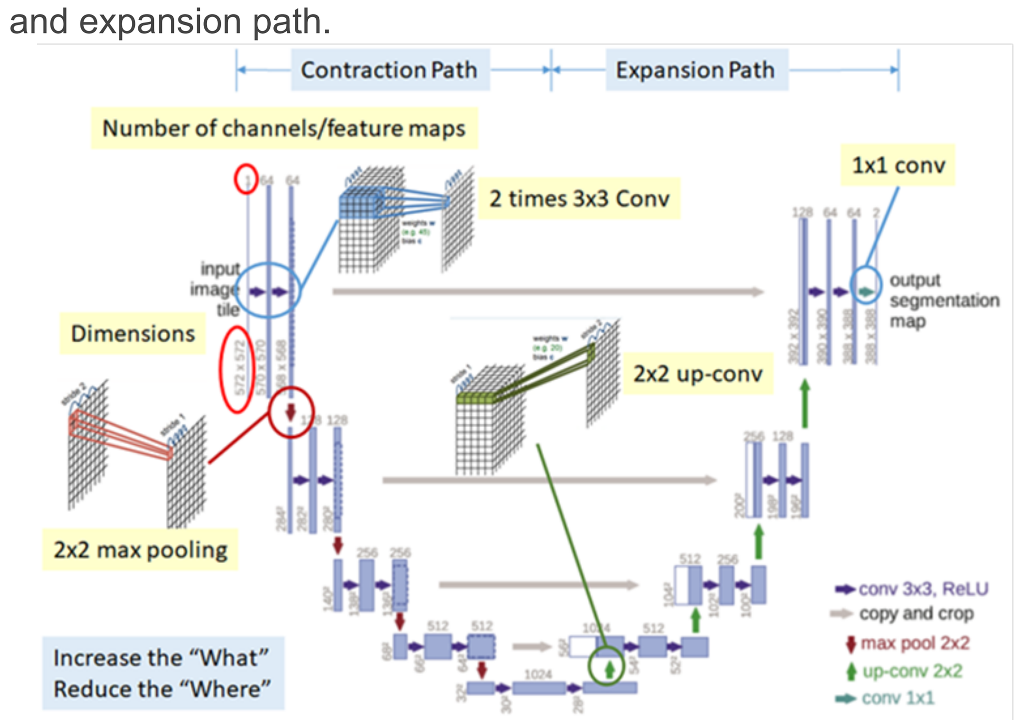

Figure 1. The U-net architecture

Contraction path :
Consecutive of two times of 3*3 Conv and 2*2 max pooling is done.
Expansion path :
Consecutive of 2*2 Up-conv and two times of 3*3 Conv is done to recover the size of segmentation map. However, the above process reduces the "where" though it increases the "what".
That means, we can get advanced features, but we also loss the localization information.
Thus, after each up-conv, we also have concatenation of feature maps(gray arrows) that are with the same level.
This helps to give the localization information from contraction path to expansion path.

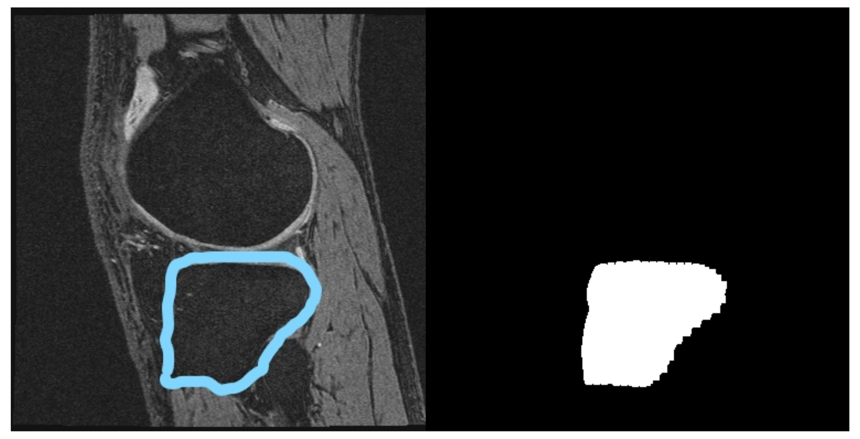

Figure 2. If an image is given. The detection is to frame the object out, and the image segmentation separates the exact contour of an object. Also consider this, give an image I, the problem is to find a function, mapping from I to Mask. There are many ways to find this function. We can see this picture, the image is given on the left, the tibia bone can be seen, and the result is segmented on the right.

Method

Data Labeling
- Using BML-Baseline Software to Label the Tibia Bone.Generate txt files.

Data preparation
- Generate binary masks from txt files , Modify and run Bone_Onefolder_edit2.m
- Convert all DICOM images into PNG format
- Change name of images for data.py and Divide images into three sets

Model Training
- Using Google Colab to run code.

Similarity Coefficient
- Given two sets A, the B jaccard coefficient is defined as the ratio of the size of the intersection of A and B to the size of the union. The larger the jaccard value, the higher the similarity, $J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$

Model Evaluation
The evaluation for segmentation accuracy commonly hires DICE coefficient and similarity.
DICE_COEF = (2 * sum(g_truth .* predict) + 0.00001) / [(sum(g_truth) + sum(predict) + 0.00001)]
SIM = sum(abs(g_truth .* predict)) / (sum(abs(g_truth) + abs(predict)) - sum(abs(g_truth .* predict)))
(g_truth : ground truth matrix   predict : detection matrix from detector)

Experiment Result

- Experiment Result (Accuracy Evaluation on Test, Validate, Train Samples)

	10 cases data-set 61-70(6 2 2)	Enlarge data-set(40 cases)
Sample count (train / valid / test)	659 / 119 / 224	2155 / 1130 /1090
Epochs until completed	44 / 100	30 / 100
Training time	57 min	2 hours
Accuracy for Test samples	loss: -0.8974272337343011 dice_coef: 0.8974272337343011 similarity: 0.8210569231637886	loss: -0.8916647493292432 dice_coef: 0.8916647493292432 similarity: 0.8148831796208653
Accuracy for Validate samples	val_loss: -0.8677 val_dice_coef: 0.8677 val_similarity: 0.7847	val_loss: -0.8680 val_dice_coef: 0.8680 val_similarity: 0.7782
Accuracy for Train samples	loss: -0.9415 dice_coef: 0.9415 similarity: 0.8906	loss: -0.9439 dice_coef: 0.9439 similarity: 0.8971

Table 1. Experiment Result

- First Data Model: 10 cases data-set 61-70(6 cases Train, 2 cases Validation, 2 cases Test)

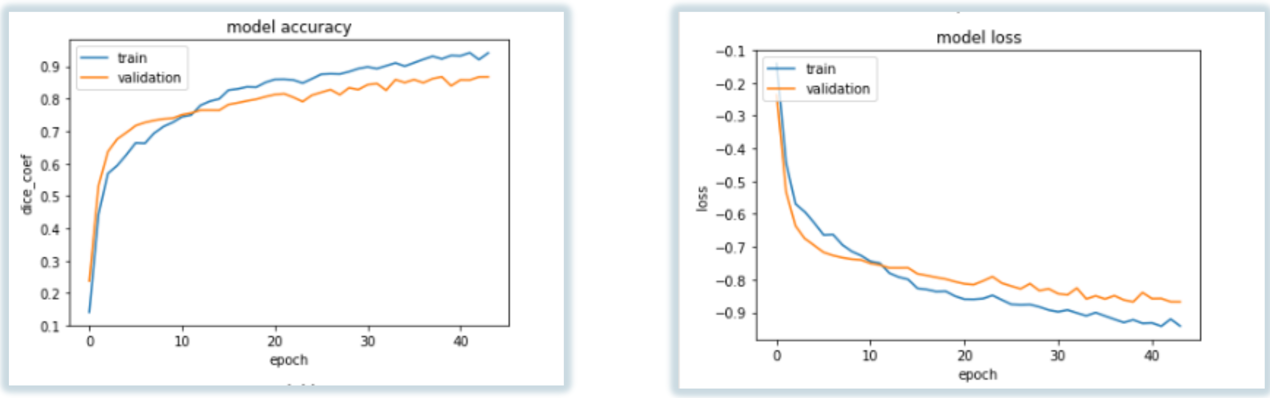

Figure 3. First data model dataset training model and validation model accuracy plot and loss function plot.

- Second Enlarged Data Model: 40 cases data-set 51-90(20 cases Train, 10 cases Validation, 10 cases Test)

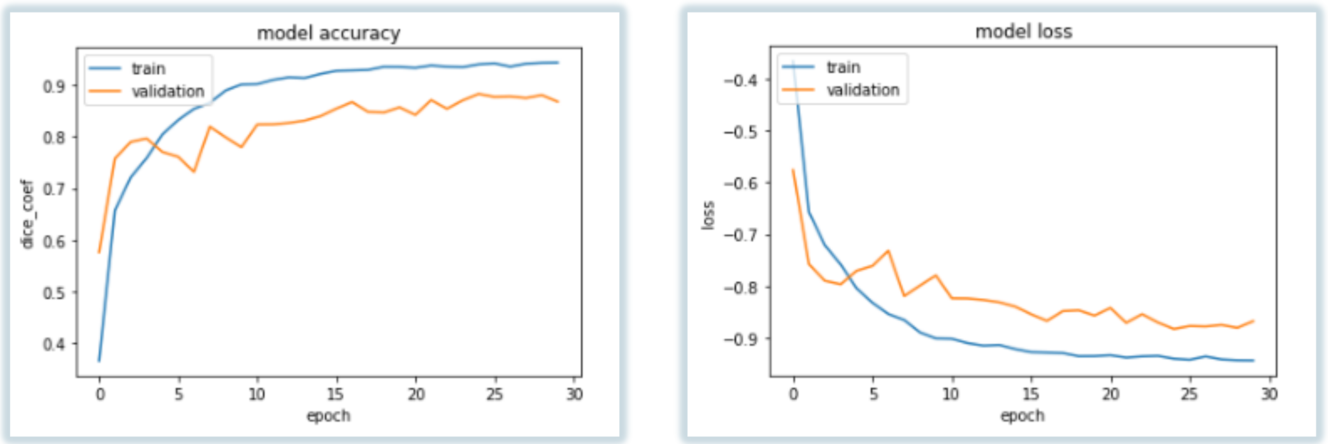

Figure 4. Enlarged dataset training model and validation model accuracy plot and loss function plot.

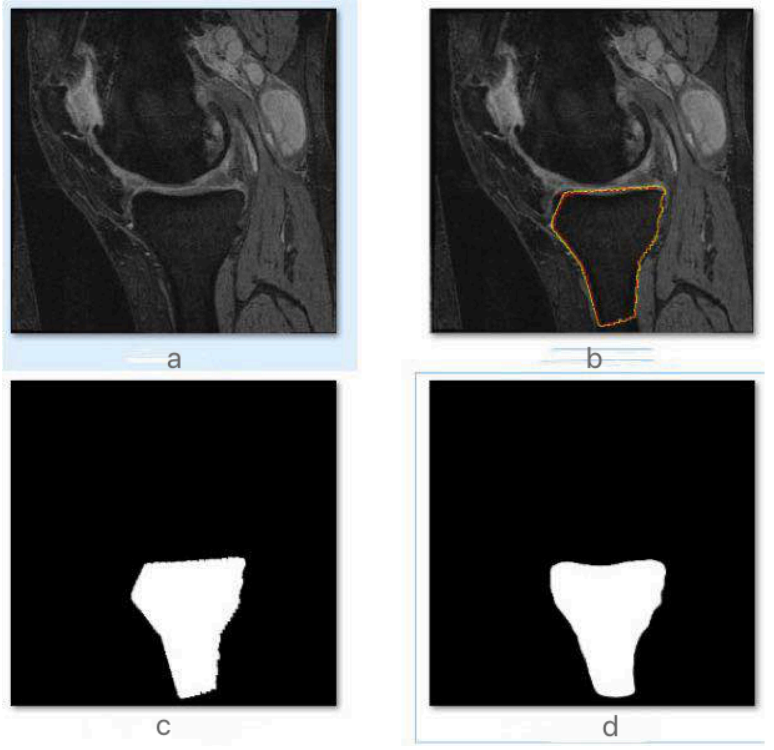

Figure 5. (a) Raw Image. (b) Manual Segmentation. (c) Mask Image Generated from Manual Segmentations. (d) Segmentation Result from U-net.

Conclusion

- The model from large dataset gives nearly the same segmentation result than one form small dataset.
- U-Net convolution network can give good segmentation result by training with even small amount of samples.
- The final training accuracy for the 10 cases dataset training sample is 94%, the validation set is 87%, the test set is 90%.
- The final training accuracy for the Enlarge 40 cases dataset training sample is 94%, the validation set is 87%, the test set is 89%.
- According to the plot, we can clearly notice the training dataset curve is more smooth when we increase the training sample sizes.

Reference

[1] https://en.wikipedia.org/wiki/U-Net
[2] https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/
[3] https://towardsdatascience.com/review-u-net-biomedical-image-segmentation-d02bf06ca760
[4] Cai, Jinzheng, et al. "Pancreas segmentation in MRI using graph-based decision fusion on convolutional neural networks." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2017.
[5] Wang, Chengjia, et al. "A two-stage 3D Unet framework for multi-class segmentation on full resolution image." arXiv preprint arXiv:1804.04341 (2018).[6]
[6] Seoud, Lama, et al. "Training a CNN to robustly segment the human body parts in range image sequences." Optical Data Science II. Vol. 10937. International Society for Optics and Photonics, 2019.