Knee MRI Bone Segmentation Using Machine Learning Model

By Reem AlRowaili, Mingming Qiu

CS 631T Computer Vision Final Project, Seidenberg School of CSIS, Pace University

Abstract

Image segmentation has an important role in different medical imaging applications. It is used to automate or assist in finding the anatomical structure and regions of interest.

The need of reliable medical image segmentation is desired to support physicians and radiologists in providing faster service for their patients.

Many diagnostic applications critically depend on the successful localization of bone structure. In this project we show the use of machine learning model to do segmentation on the tibia bone in a 3D knee MRI images. Background and introduction to the segmentation process are first presented. The methodology we used is then explained in detail. Results and experiments are then showed with line charts of the experimentation that has been done. Finally, we conclude with our findings and suggestions to help improve the prediction of our model.

Introduction

- Automation of image-processing and analysis techniques is compulsory to assist physicians in treatment planning and clinical diagnosis. [3]
- Consistent algorithms are mandatory for the delineation of regions of interest and the anatomical structures. [3]
- MRI is one of the popular imagery in medical imaging the domain. See figure 1.
- The principal goal of the segmentation process is to partition an image into regions that are homogeneous with respect to one or more characteristics or features. [2]
- Segmentation is important for feature extraction, image measurements, and image display. In some applications it may be useful to classify image pixels into anatomical regions, such as bones, muscles, and blood vessels, while others into pathological regions, such as cancer, tissue deformities, and multiple sclerosis lesions. [2]
- In some studies, the goal is to divide the entire image into subregions such as the white matter, and cerebrospinal fluid spaces of the brain, whereas in others one specific structures has to be extracted, for example, breast tumors from MRI.[2]
- In this project we have 160 knee MRIs in dicom format in 50 folders. Each folder represents one patient. We are going to train a model on these images to automate the segmentation of the tibia bone.

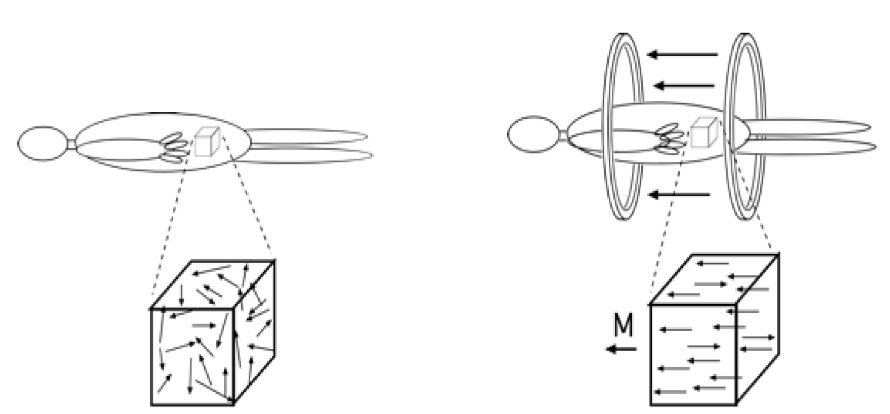

Figure 1: Atomic nuclei and hydrogen nuclei, 1H, in particular, have a magnetic moment. Moments tend to become aligned to applied field. Creates magnetization, m(x,y,z) (a tissue property). MRI makes images of m(x,y,z)

Methodology

- Before starting the training process on the images, we need to manually label the tibia bone in the images, and then convert the labels to an image format readable by the python script.
- The image format is tiff. We also need to convert the original MRIs to tiff image format.
- The original images and their labels need to be distributed into train, validate, and test folders for the training process.
- The model used is a U-net model. The structure of the U-net model can be seen in the following image:

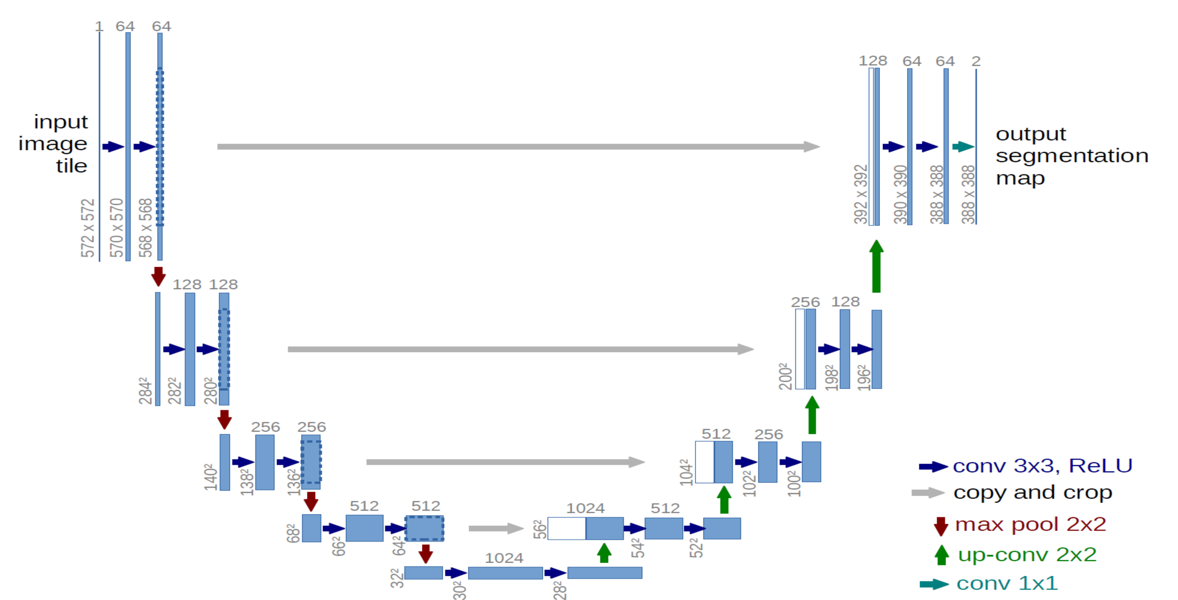

Figure 2: U-net architecture

- The evaluation metric used is Dice similarity coefficient. It is used to compute the spatial overlap accuracy of the segmented MRIs. The higher the value the closer it is in finding the region of interests. As shown in figure 3

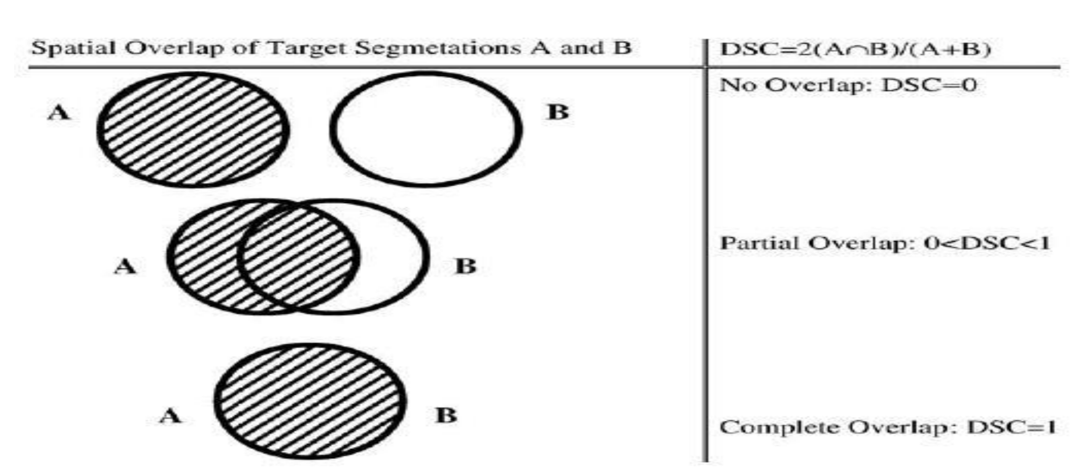

Figure 3: Dice similarity coefficient

Experiments & Results

We did two experiments 10 cases and 40 cases:

- First experiment trained the model on 10 cases (51-60) and it took approximately an hour. The results were 95% training accuracy and 78% for validation accuracy:

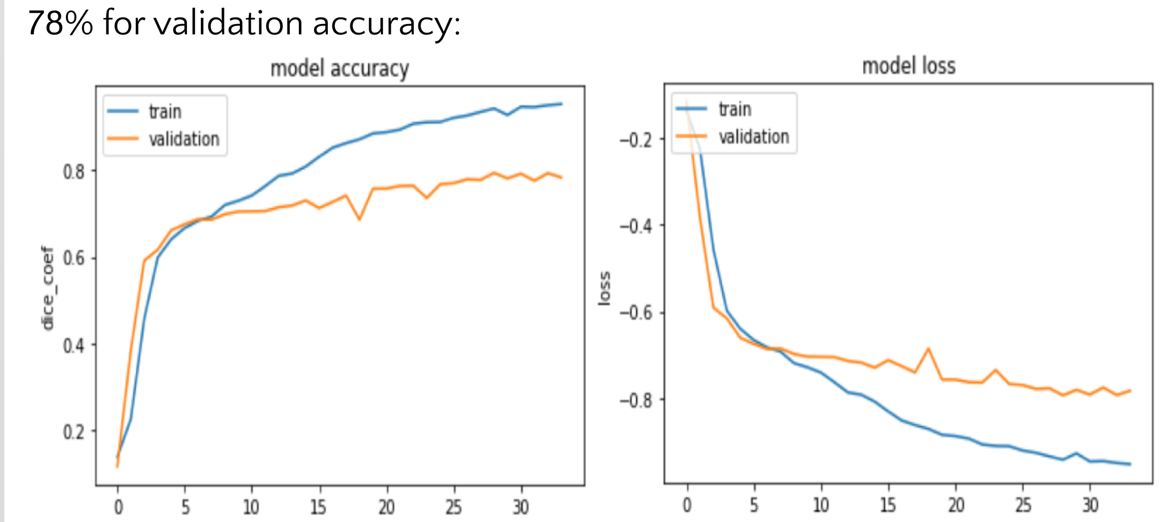

Figure 4: Accuracy and loss for training and validation on 9 cases

Experiments & Results

```
Evaluating the model on test data...
--------------------------------
103/103 [==============================] - 4s 42ms/step

loss:  -0.8884580951292538
dice_coef:  0.8884580951292538
similarity:  0.8056045116730106
```

Figure 5: Evaluation on testing dataset – – first experiment

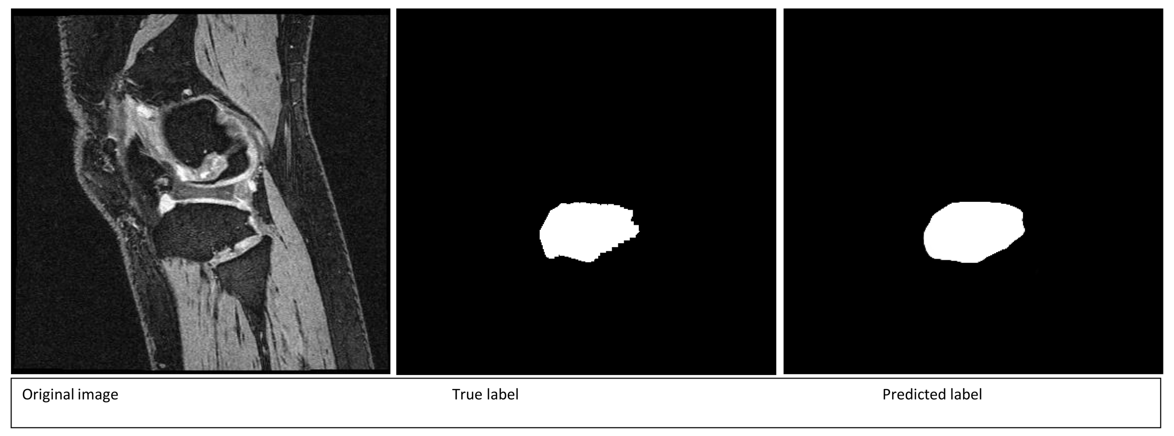

Figure 6: Sample of prediction – first experiment

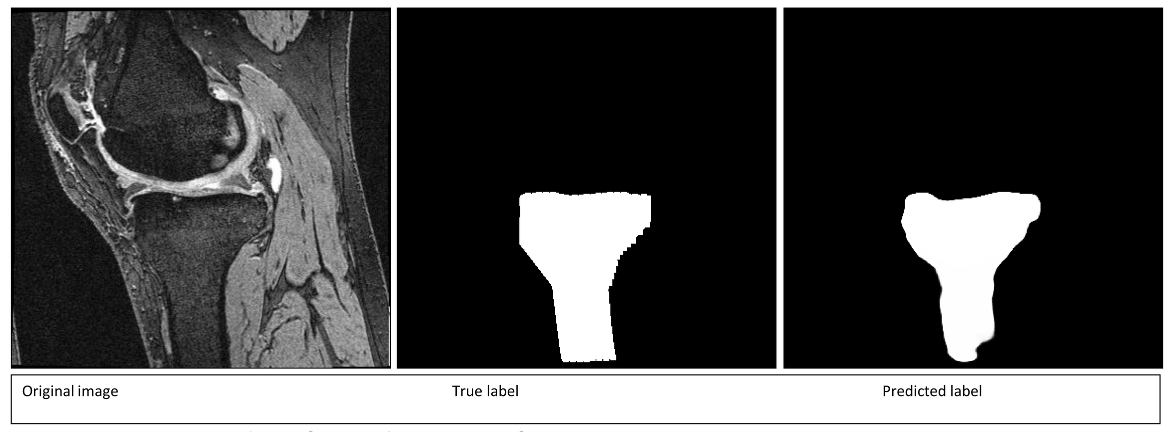

Figure 7: Sample of prediction – first experiment

- Second experiment trained the model on 40 cases, it took approximately 4 hours. The results were 94% for training accuracy and 89% for validation accuracy:

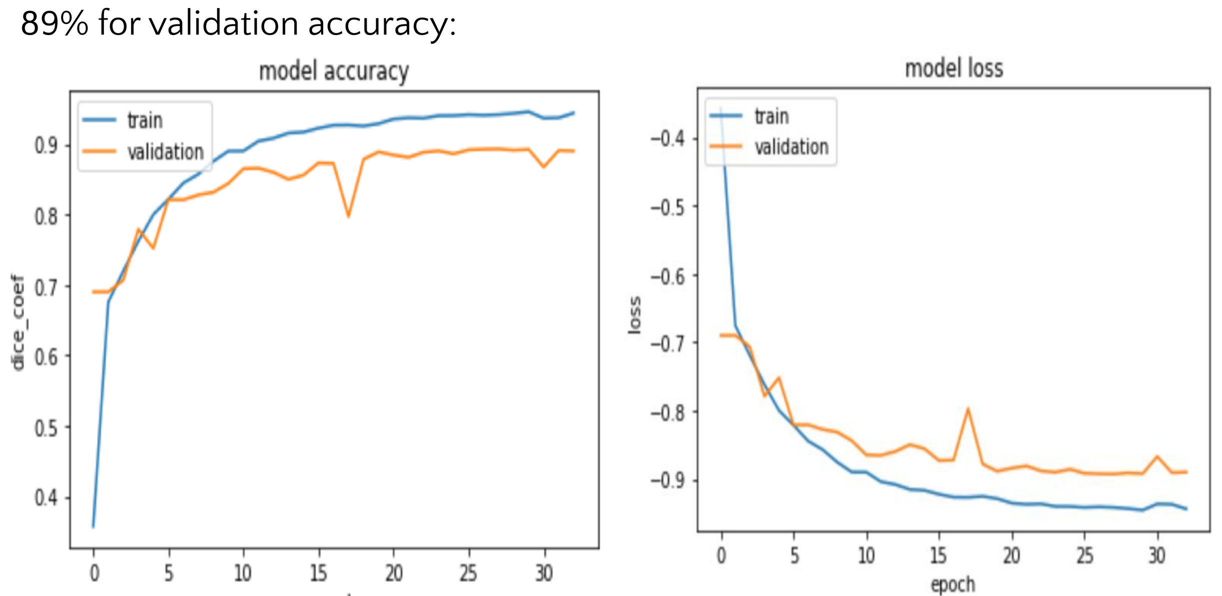

Figure 8: Accuracy and loss for training and validation on 38 cases

```
Evaluating the model on test data...
--------------------------------
219/219 [==============================] - 34s 156ms/step

loss:  -0.9198015258736807
dice_coef:  0.9198015332221985
similarity:  0.8532705747917907
```

Figure 9: Evaluation on testing dataset – second experiment

Figure 10: Sample of prediction – second experiment

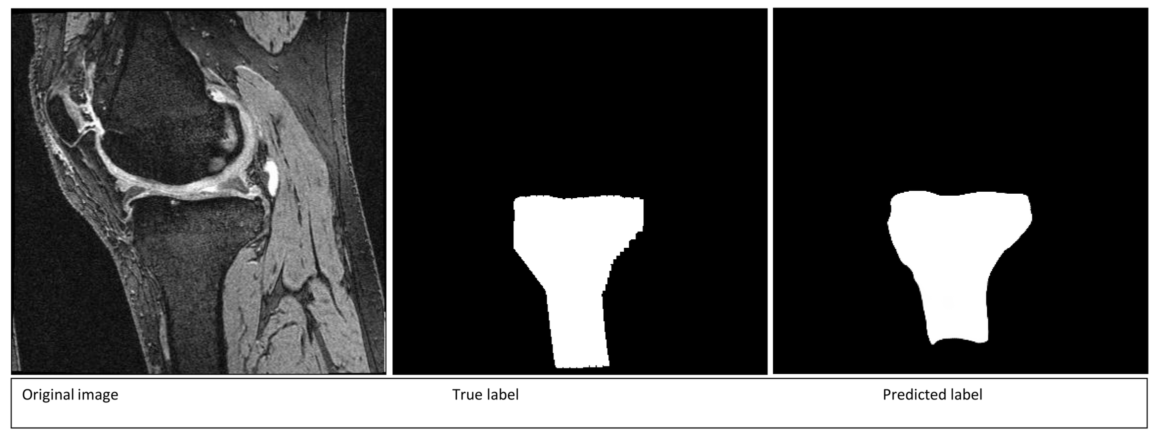

Figure 11: Sample of prediction – second experiment

Conclusion

- As seen in the results, both experiments are trained for approximately 30 epochs.
- Having more cases to experiment on improves the accuracy as seen in the validation part in the charts.
- The convergence is faster in the 40 cases more than 10 cases.
- The test examples in the 10 cases experiment shows that the segmentation is happening but not very well but having the model trained on 30 cases the segmentation improves without any cutoffs.
- The more data we have to train on the better accuracy the model will be.
- The manual labeling part is important because if we have accurate labeling we will get better training results.
- Introducing other metrics (cosine coefficient) for evaluating the segmentation by the model, will be a good way in assessing the model from another perspective.
- We could use image enhancement technics to enhance the contrast of the original images, to make it easier for the model to detect the pixel features for segmentation.

References

1. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
2. Bankman, I. (Ed.). (2008). Handbook of medical image processing and analysis. Elsevier.
3. Dey, N., Ashour, A. S., Shi, F., & Balas, V. E. (2018). Soft Computing Based Medical Image Analysis. Academic Press.
4. Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M., Kaus, M. R., Haker, S. J., ... & Kikinis, R. (2004). Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. Academic radiology, 11(2), 178-189.