Analyzing the Effect of Attention Gates in U-Net for Knee Recess Distention Segmentation

Jingyu (Nana) Ye , Supervisor: Professor Pascal Tyrrell

Department of Statistical Science, University of Toronto

Abstract

In the medical domain, the **convolutional neural network (CNN)** is the dominant choice for **image segmentation** due to its excellent representational power, fast inference, and filter-sharing properties. U-Net is a commonly used architecture because of its accurate performance and efficient GPU memory usage. However, recent studies have proposed that encoding attention gates to standard U-Net can improve its accuracy in a large variety of datasets without a significant cost of computational resources. Attention U-Net is proposed by Oktay and colleagues in 2018 which incorporates additive soft attention in standard U-Net. This paper explores the effect of **attention gates** in U-Net for knee recess distention segmentation with the aim to improve segmentation accuracy of ultrasound images by highlighting the salient areas while trimming unnecessary information. 3750 ultrasound images of the knee joint for patients with recess distention were used for training, validating, and testing both models. Although it is hoped in our hypothesis that the attention gates would bring more robust improvement, we only see a 1.3105% improvement on dice coefficient and 1.7431% on IOU for attention U-Net. Moreover, attention U-Net showed smoother and more complete segmentation which may be explained by the benefit of attention gates in obtaining global information features of recess distention and suppressing irrelevant background noise.

Introduction

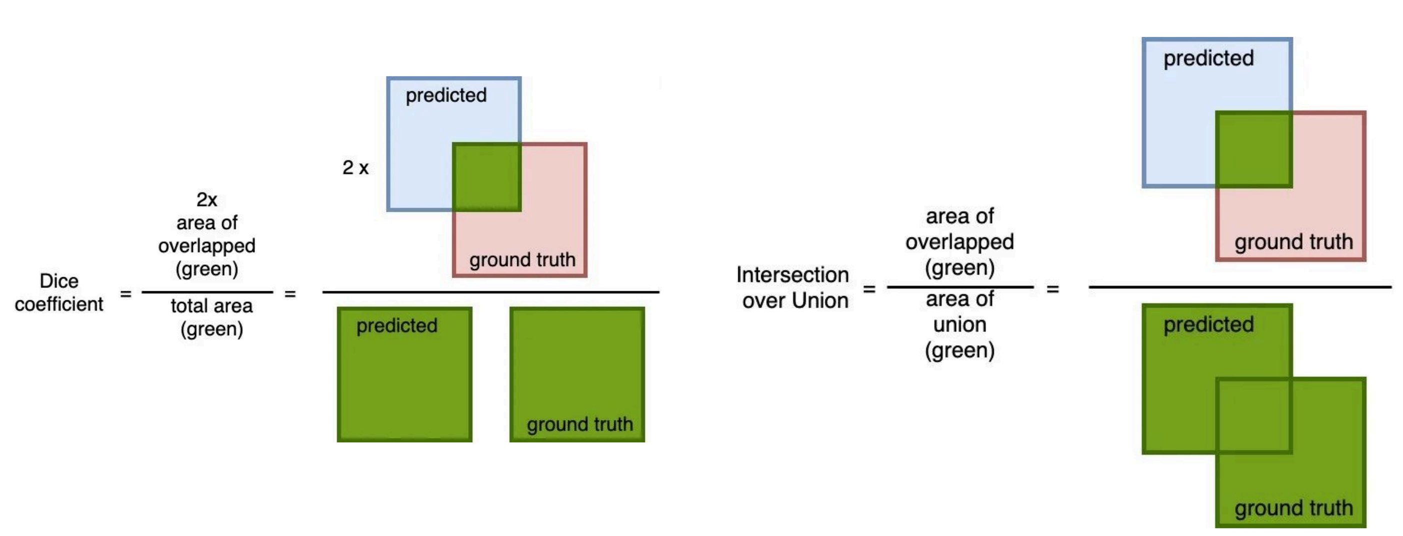
Purpose: To compare the performance of Attention U-Net and Traditional U-Net on knee recess distention segmentation.

Hypothesis: The additive attention gate in standard U-Net architecture would improve segmentation performance on knee recess distention ultrasound dataset. Attention U-Net would outperform Traditional U-Net on knee recess distention segmentation.

Objectives:

1. To implement and describe Attention U-Net for knee recess distention ultrasound dataset.
2. To compare the performance of Attention U-Net and Classical U-Net on knee recess distention ultrasound dataset.
3. To propose and recommend when the addition of attentional layers would be beneficial.

Method



- Dataset comprises 3750 ultrasound images of the knee joint that shows recess distention was used.
- **Attention U-Net** and **U-Net** trained for 200 epochs with a batch size of 30 on the twined NVIDA GeForce GTX 1080Ti GPU.
- 556 images and masks were held out for testing.
- Training dataset split into 70:30 for training and validation.

Results

Figures 1 and 2 show the learning curve of dice coefficients for both models. Both models have a low training and validation dice coefficient at the beginning which gradually increases upon adding more epochs and flattens gradually indicating the addition of more epochs won't improve the models' performance on unseen data.

The dice coefficient of both models will almost always be lower for validation dataset than for the training dataset. This gap between the models' performance in training and validation dataset is referred to as the generalization gap.

From the behaviour, it can be concluded that both Attention U-Net and standard U-Net trained for 200 epochs were neither undercutting nor overfitting and the models are both training well and generalizing well.

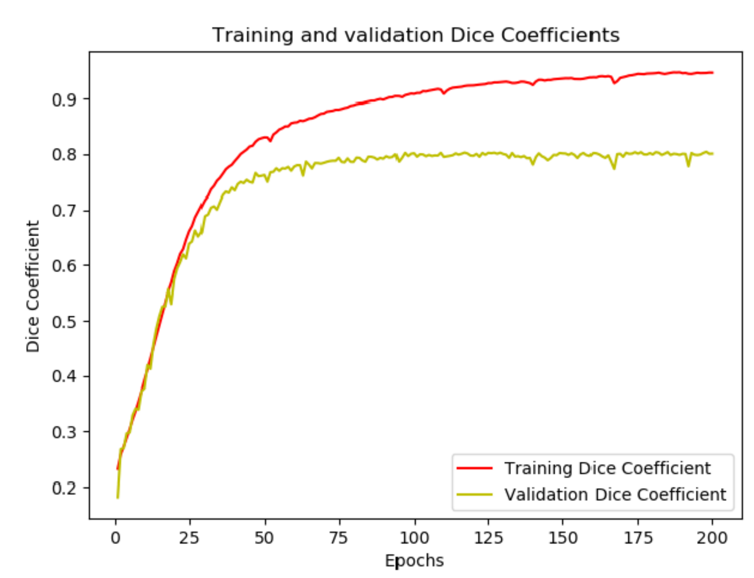


Figure 1. Attention U-Net Training and Validation Dice coefficient

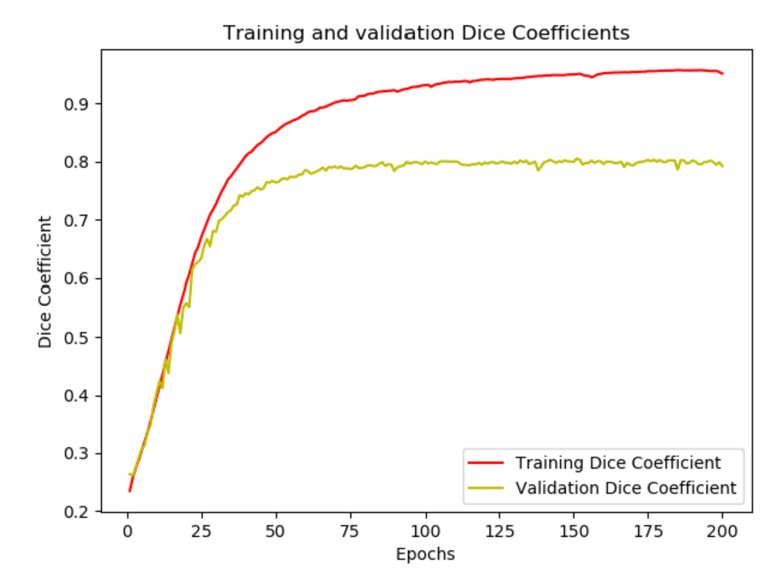


Figure 2. U-Net Training and Validation Dice coefficient

From the overlaid images of predicted segmentation for Attention U-Net, standard U-Net, and ground truth, we can visualize the difference in predicted segmentation for Attention U-Net, U-Net, and the ground truth.

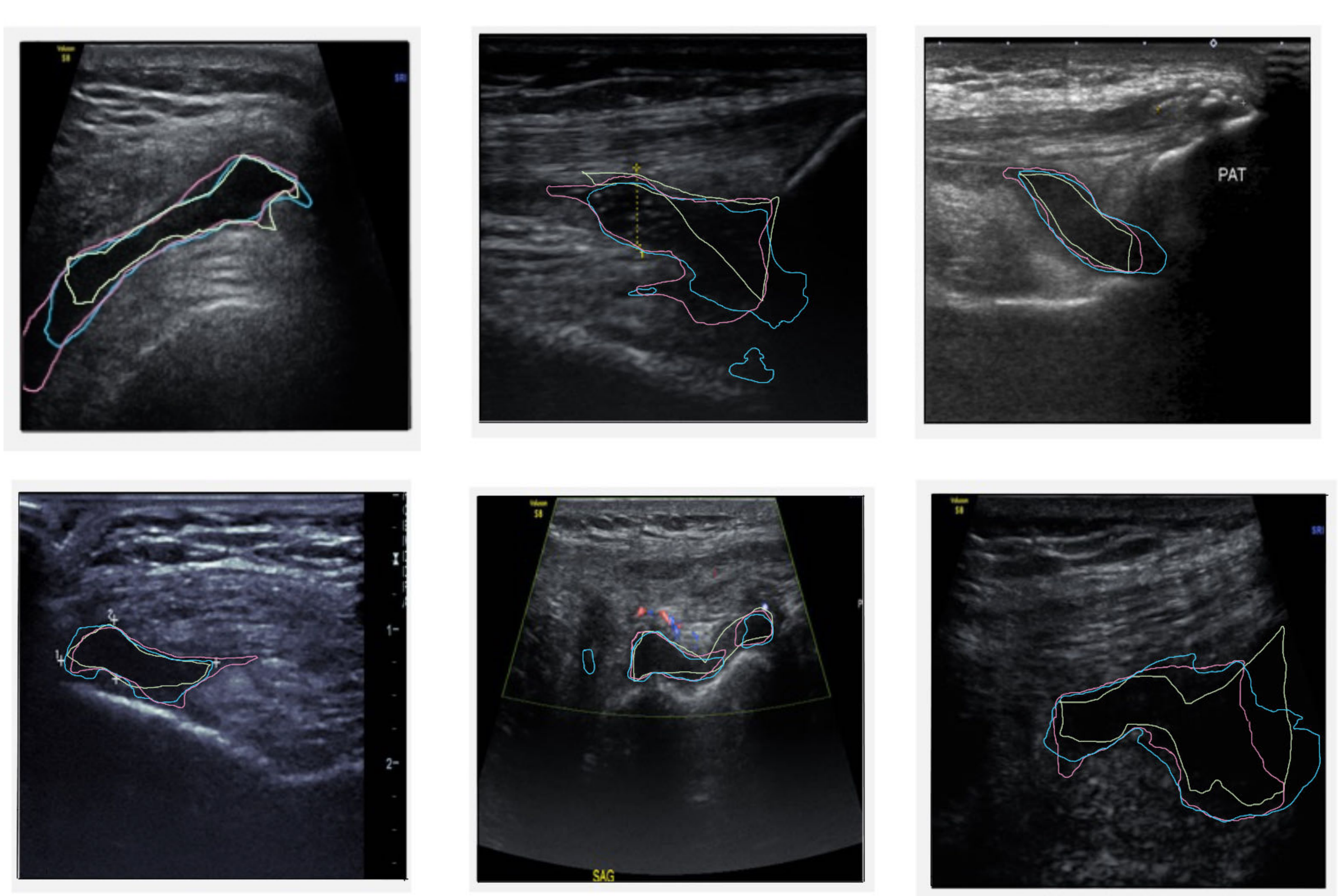


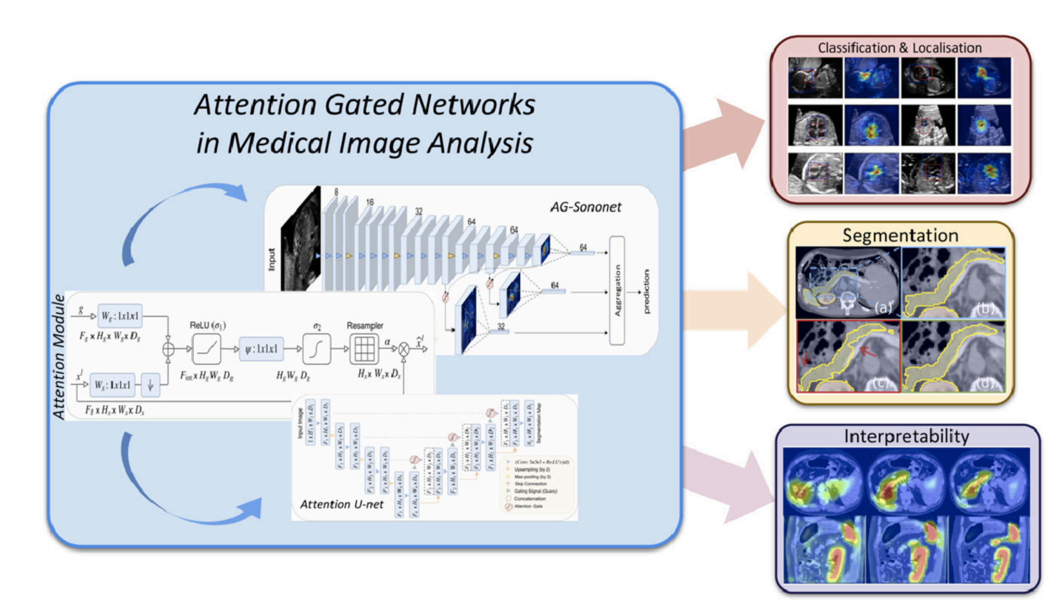
Figure 3. Knee recess distention segmentation results
yellow tracing: Ground Truth, Pink tracing: Attention U-Net, Blue tracing: U-Net

Results

	Attention U-Net	U-Net
mean Dice coefficient	0.7817	0.7686
mean IOU	0.6430	0.6256
mean % coverage difference from ground truth	1.3258%	1.0863%
mean number of segmentation pieces	1.1696	1.5088

Conclusion

- Compared to standard U-Net, Attention U-Net achieved a **1.3105%** improvement on **Dice coefficient** and **1.7431%** on **IOU**. Although hoping for more robust improvement, we only see very little and could not conclude whether it is statistically significant.
- Both standard U-Net and Attention U-Net tend to output more segmentation area in their prediction compared to the ground truth which may be due to the higher uncertainty of annotations around the boundary of recess distention.
- Attention U-Net showed smoother and more complete segmentation than standard U-Net. Standard U-Net showed less continuity in its segmentation (often broken down into multiple pieces) as the ground truth is all in one piece.
- Without requiring a large number of model parameters, Attention U-Net showed a slight improvement in knee recess distention segmentation than standard U-Net. Therefore, these results do suggest that future research in the Attention Gates and medical imaging is worthwhile.



Acknowledgement

I would like to thank Professor Tyrrell, Atsuhiro Hibi, Mauro Mendez Mora for their guidance, encouragement, and suggestions during my research work.