Self-Supervised Pretraining with Synthetic Data For Improving Ultrasound Medical Image Segmentation

Yuzhi Tang, Supervisor: Professor Pascal Tyrrell

University of Toronto Research Opportunity Program

Background

- A significant limitation in training **machine learning** models in the field of medical imaging is the lack of annotated data due to the high cost of annotation (Ker et al., 2018).
- **Self-supervised learning** methods such as DenseCL (Wang et al., 2020) mitigate this issue by using a pretext task to first pretrain a network to learn useful image representations from unannotated data. The learned representations can then be transferred to a downstream model to be fine-tuned using annotated data.

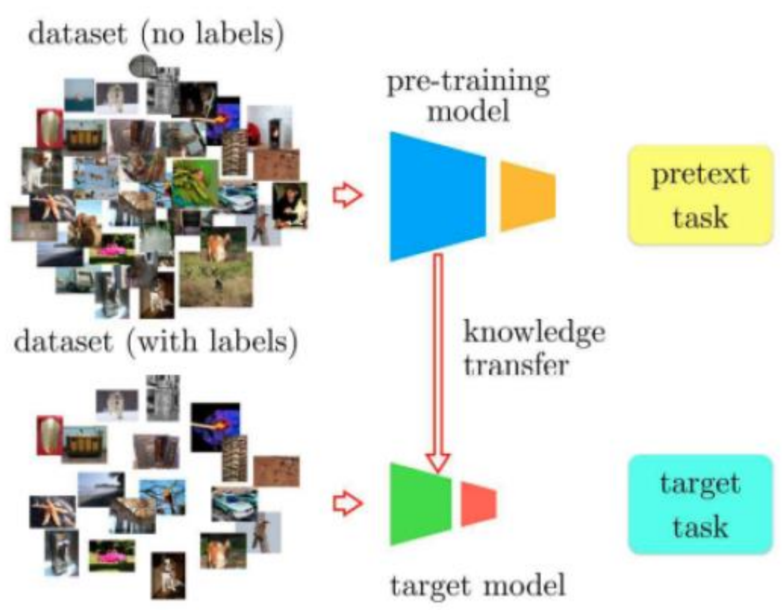

Fig 1. Illustration of a typical self-supervised learning framework

- **DenseCL** pretrains a network by using the dense correspondence pretext task. DenseCL pretraining demonstrates superior downstream performance on natural image segmentation while remaining comparatively computationally efficient (Wang et al., 2020).

Fig 2. In a dense correspondence task, the network tries to map corresponding pixels between images of different views

Purpose and Objectives

Purpose:
This project aims to provide and test a method for improving the segmentation performance of ultrasound medical images when the amount of annotated data is scarce.

Objectives:

1. Adapt the DenseCL pretraining method for segmentation of ultrasound medical images.
2. Compare the segmentation performance of a model with DenseCL pretraining and baseline (Kaiming) weight initialization.
3. Visualize the image representations of a model with DenseCL pretraining and baseline weight initialization.

Methodologies

- 3750 annotated **ultrasound knee recess distension** images at the MiData lab and 20k unannotated knee ultrasound images obtained by a synthetic simulation approach (Mauro et al., 2023) were used for DenseCL pretraining. The 3750 annotated data was then used to fine-tuning the segmentation model.

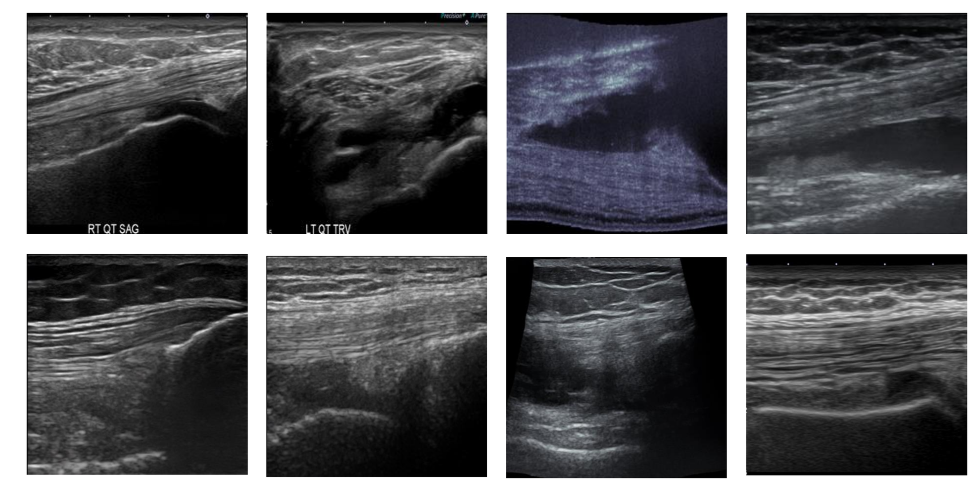

Fig 3. (Top row) Example images from the MiData lab, (bottom row) generated synthetic image

- A **ResNet-50** was pretrained using DenseCL for 500 epochs with a batch size of 32. The pretrained weights were transferred to the downstream model, which was then fine-tuned for 40k iterations with a batch size of 16 for 3 independent trials.

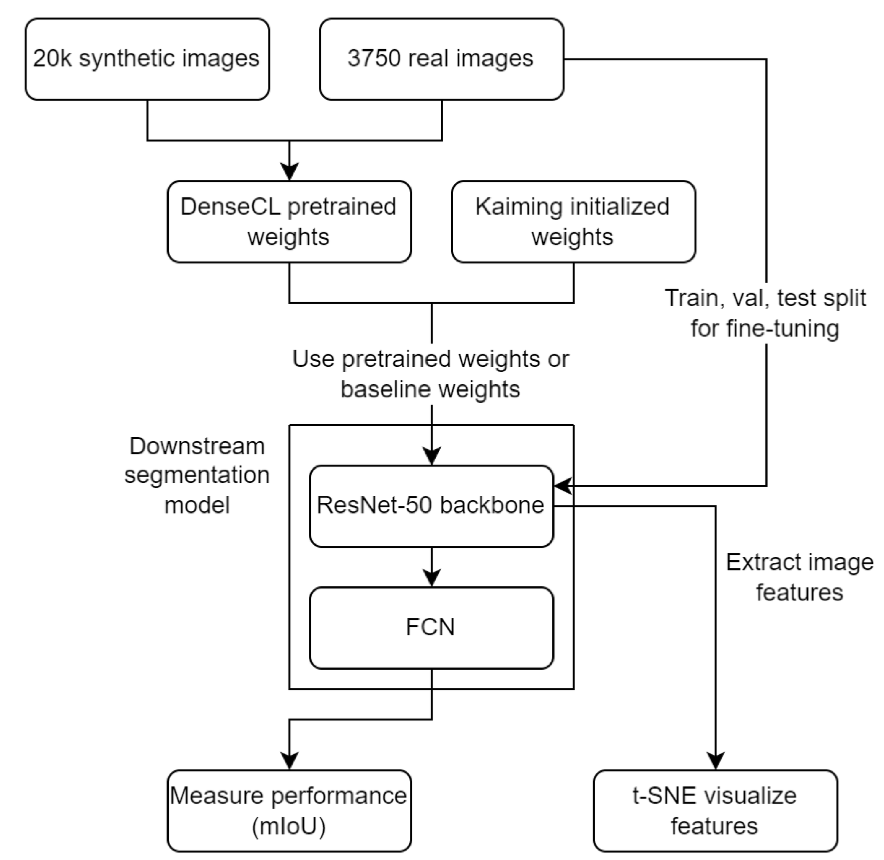

Fig 4. Project flowchart

Results

Experiment 1: Comparison of segmentation performance of downstream model with and without DenseCL pretraining

Trial	DenseCL Pretrained	Kaiming initialized
1	78.97	75.77
2	79.02	75.76
3	79.35	74.21
Average	79.11	75.25

Table 1: Performance comparison (in mIoU) between DenseCL pretrained and Kaiming initialized models across different trials

Results

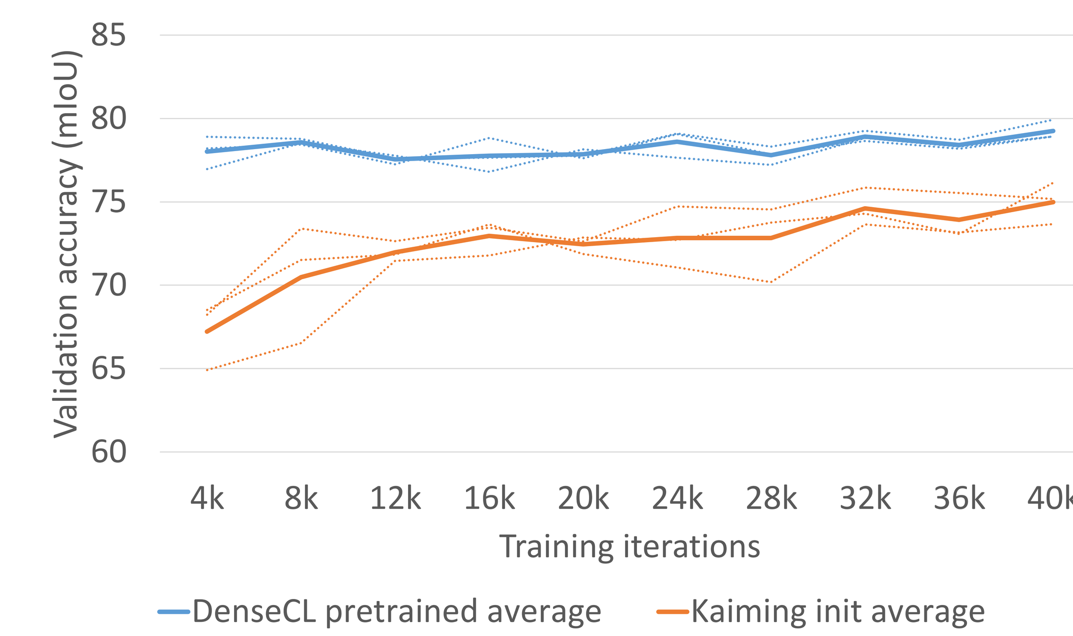

Figure 5: Performance comparison between DenseCL pretrained and Kaiming initialized models on different training iterations

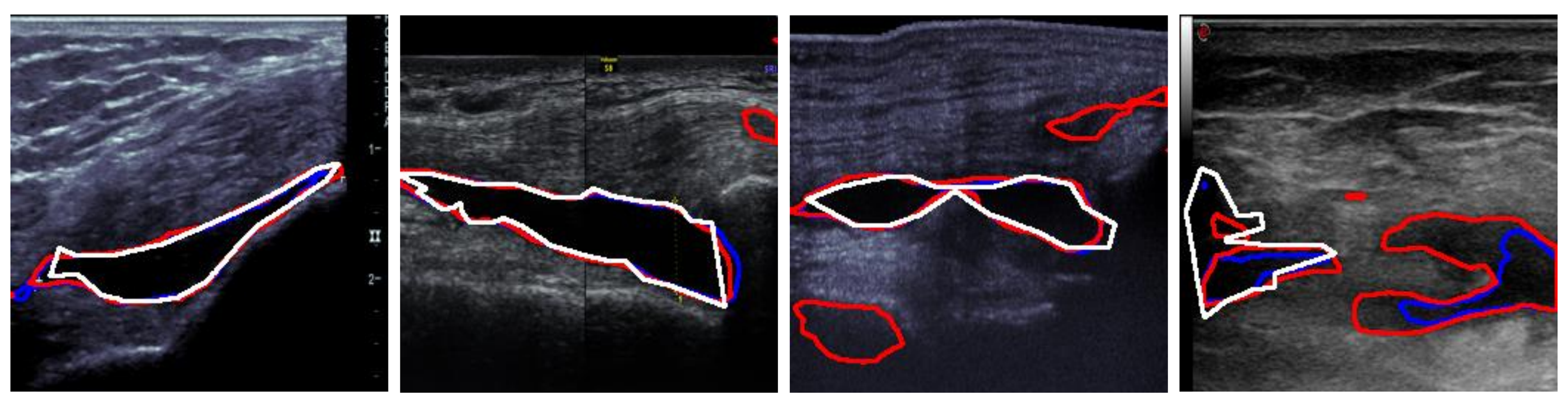

Figure 6: Example segmentation results comparison between DenseCL pretrained and Kaiming initialized models (white: groundtruth, blue: DenseCL pretrained, red: Kaiming initialized)

Experiment 2: Visualization of image representations of downstream model with and without DenseCL pretraining

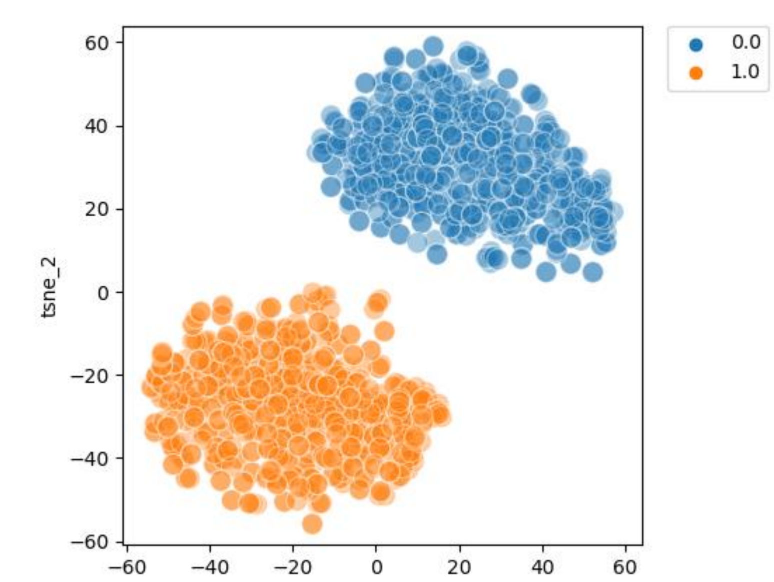

Figure 7: t-SNE visualization of extracted image representations between DenseCL pretrained and Kaiming initialized models (blue: pretrained, orange: Kaiming initialized)

Averaging the 3 independent training trials, there was a **3.86 mIoU improvement** for the model pretrained using DenseCL compared to a Kaiming initialized model. The DenseCL pretrained model was able to obtain good performance **much faster** (achieving around 78 mIoU on the validation set in 4k training iterations) compared to the baseline. A DenseCL pretrained model had **vastly different** image representations compared to a Kaiming initialized model.

References

- Ker, J., Wang, L., Rao, J., & Lim, T. (2018). Deep Learning Applications in Medical Image Analysis. *IEEE Access*, 6, 9375–9389. https://doi.org/10.1109/ACCESS.2017.2788044
- Wang, X., Zhang, R., Shen, C., Kong, T., & Li, L. (2020). *Dense Contrastive Learning for Self-Supervised Visual Pre-Training*. https://doi.org/10.48550/ARXIV.2011.09157
- Mauro, et al., 2023 is currently under review