PELTARION

KTH VETENSKAP OCH KONST

/16 November, 2021

Resource-efficient image segmentation using self-supervision and active learning

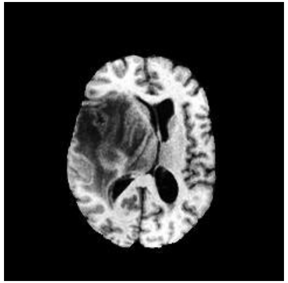
Muriel Max

Peltarion
Holländargatan 17
SE-111 60 Stockholm
Sweden

Introduction

01

Task



02

Large Datasets required for Deep Learning

03

Labelling - Expert Consensus

Introduction

Alleviate data problems by learning from *unlabeled* samples and only label the most informative samples:

SELF-SUPERVISED LEARNING AND ACTIVE LEARNING

TO THE RESCUE

Background

1

Image Segmentation

2

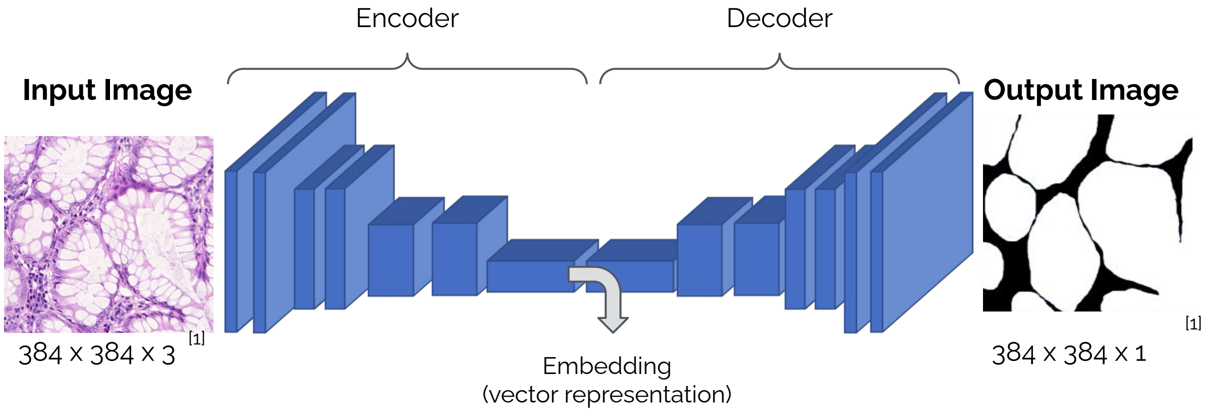
Active Learning

3

Self-Supervised Pre-training

Image Segmentation

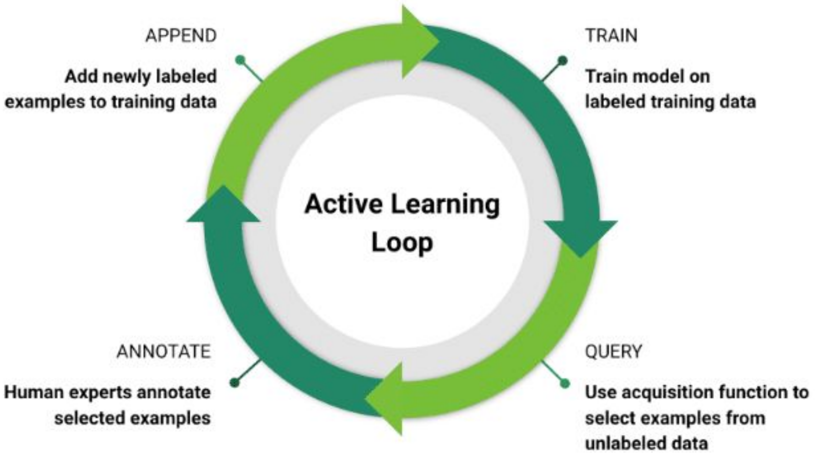
- Computer Vision task
- aim is to **train a neural network** to **classify all pixels of an input image** into object classes
- Results in a **segmentation mask**



[1] K. Sirinukunwattana, J. P. W. Pluim, H. Chen, X. Qi, P. Heng,Y. B. Guo, L. Y. Wang, B. J. Matuszewski, E. Bruni, U. Sanchez,A. Böhm, O. Ronneberger, B. B. Cheikh, D. Racoceanu, P. Kainz, REFERENCES |45M. Pfeiffer, M. Urschler, D. R. J. Snead, and N. M. Rajpoot,"Gland segmentation in colon histology images: The glas challengecontest,"CoRR, vol. abs/1603.00275, 2016.

Active Learning

- aim is to **select most informative training samples** and reduce need to annotate large datasets
- Approaches are categorized by:
  - Sampling strategy
  - Informativeness metric



[2]

[2] https://blogs.nvidia.com/blog/2020/01/16/what-is-active-learning/, Accessed 10.06.2020

Self-Supervised Pre-training

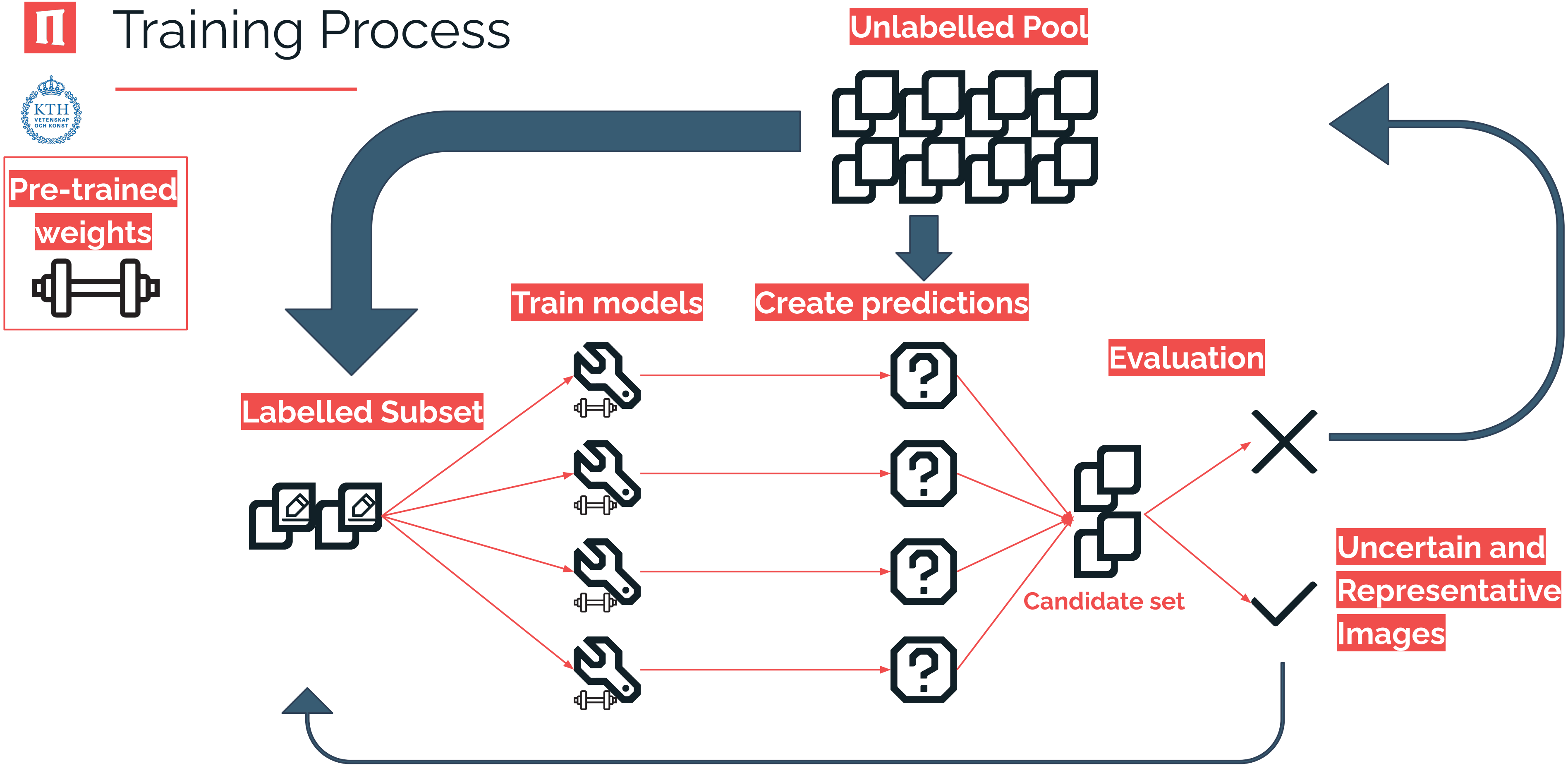
Self-Supervision

- Learning technique that **doesn't require human-annotated labels to learn**
- Supervisory signal is obtained from the data itself

Pre-training

- refers to **training a model with one task** to form parameters that are useful for **downstream tasks** (such as image segmentation)
- **network pre-conditioner**: sets the parameters in the appropriate range for further supervised training
- Transfer weights for supervised training

Training Process

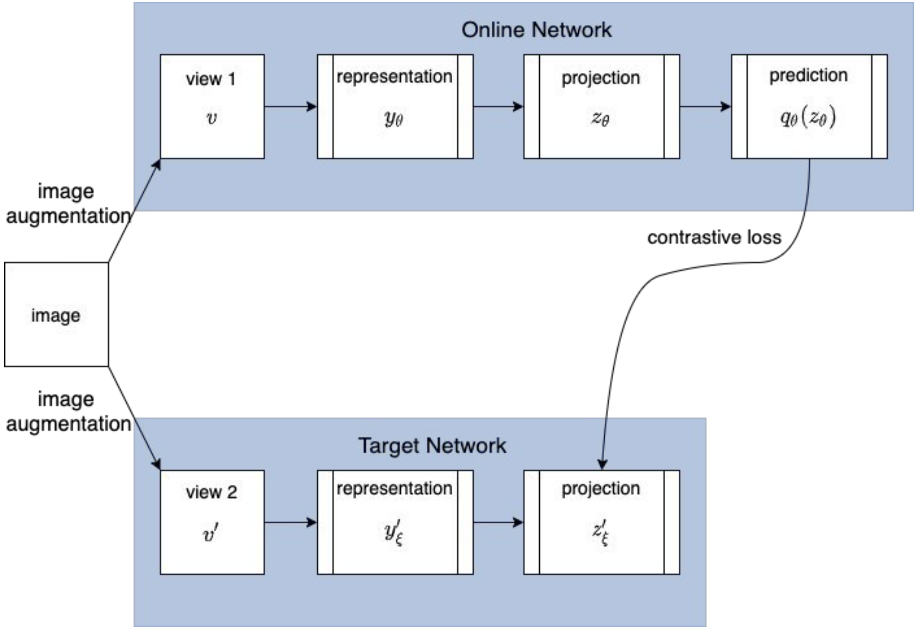


Self-Supervised Pre-Training

Bootstrap your own latent: A new approach to self-supervised Learning (2020)

Pre-training with BYOL (Bootstrap your own latent)

- Pretrain on same training dataset that is used for downstream task
- **Generate two views of the same image** by applying random augmentations
- Use online predictor to predict target representation
  - “Similar samples have similar representations”
- **Contrastive loss**: minimizes the distance between representations of view 1 and view 2

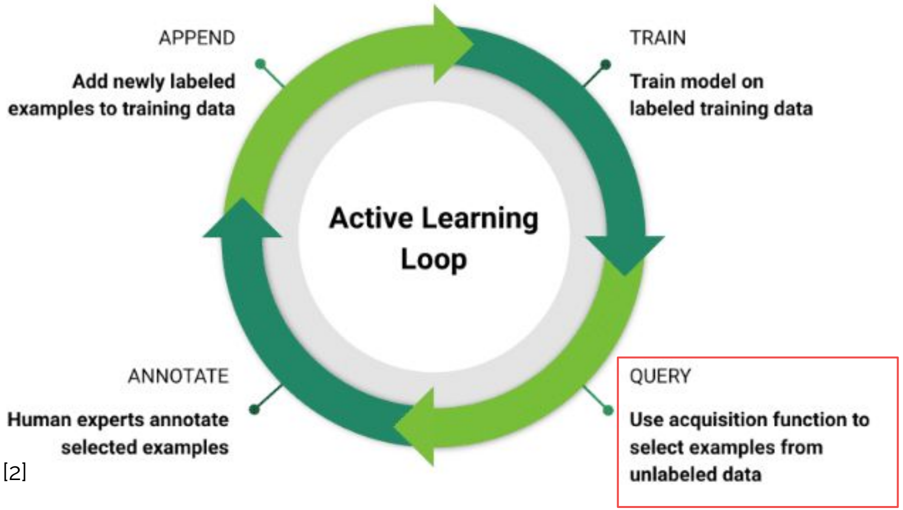


Active Learning

Suggestive Annotation: A Deep Active Learning Framework for Biomedical Image Segmentation (2017)

Peltarion
Holländargatan 17
SE-111 60 Stockholm
Sweden

Active Learning - Query Function



[2]

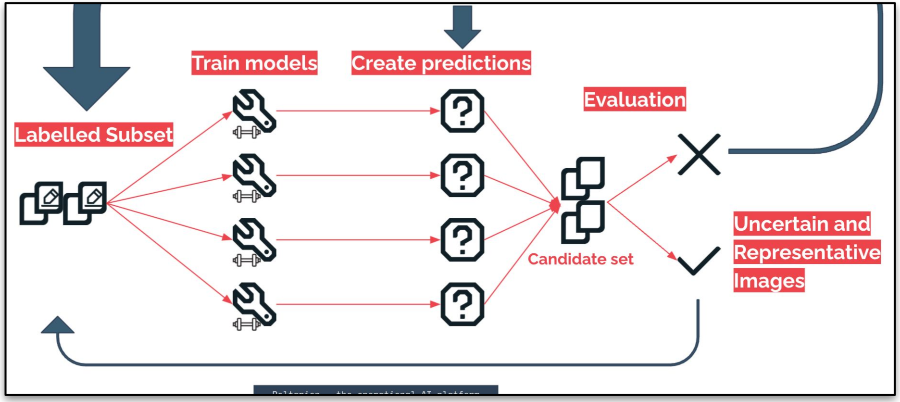
1. Uncertainty
2. Maximum Set Cover

Active Learning - Query Function

1. Uncertainty

- Calculating the **prediction uncertainty** between four different bootstrapped models:
  - Predict **segmentation mask** for all unlabelled images
  - Calculate **mean and variance** between the four different models
- Select images with highest variance (**models disagree most**)

- Result: **candidate set**



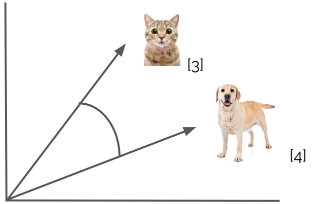
Peltarion - the operational AI platform

Active Learning - Query Function

2. Maximum Set Cover

Aim: select images that cover most diverse cases in pool of unlabelled image

1. Predict **image embeddings** for all images
2. **Calculate cosine similarity** between embeddings of each image in the candidate set and each image in the unlabelled pool



3. Iteratively select an image from the candidate set that is similar to the largest number of other image that have previously not been covered before

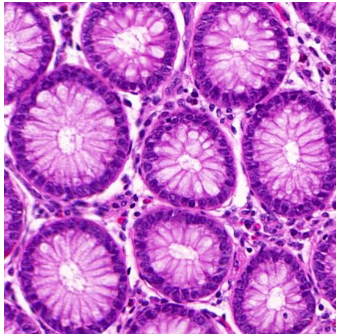
[3] https://timesofindia.indiatimes.com/life-style/relationships/pets/5-things-that-scare-and-stress-your-cat/articleshow/67586673.cms, Accessed: 16.06.2021

[4] https://www.akc.org/most-popular-breeds/, Accessed: 16.06.2021

Data

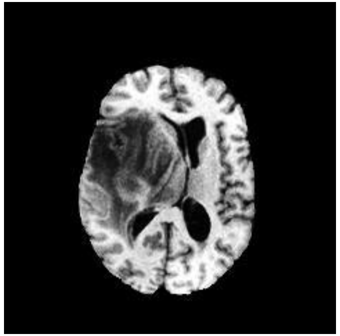
Gland Dataset

- Gland morphology is used as a key criterion for cancer grading
- 85 training and 60 test images



BraTS Dataset

- Multimodal brain Magnetic Resonance Images (MRI)
- ~58000 images from 369 patients



1. Experiment

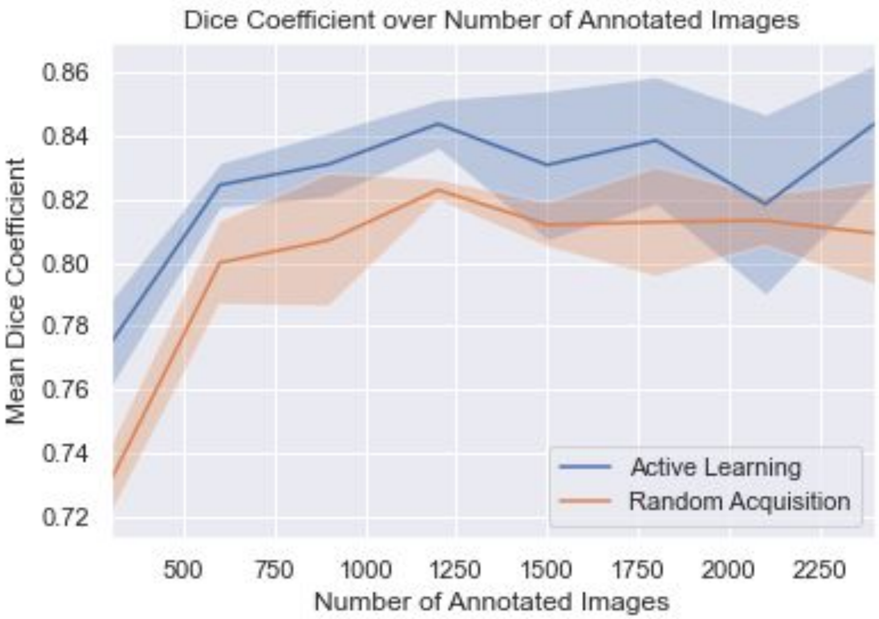
Active Learning: Batches of training data are added with an active learning acquisition function.

VS.

Random Acquisition: In each iteration, new training data is selected randomly.

Results - **BraTS Dataset**

Active Learning



Experiments

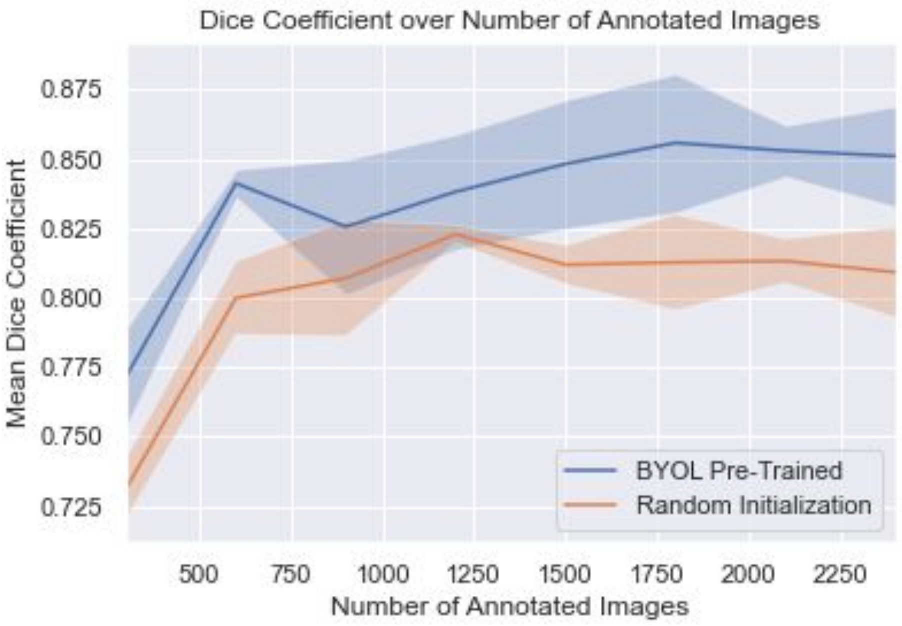
Random Initialization: The model weights are randomly initialized.

VS.

BYOL Pre-Training: The model is finetuned after transferring pre-trained weights.

Results - **BraTS Dataset**

Pre-training



Experiments

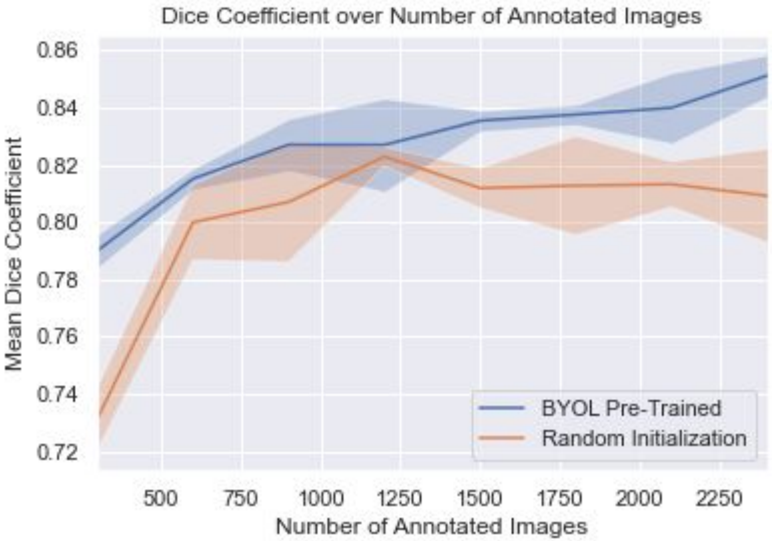
Random Acquisition without Pre-Training: The model weights are randomly initialized. In each following iteration, new training data is selected randomly.

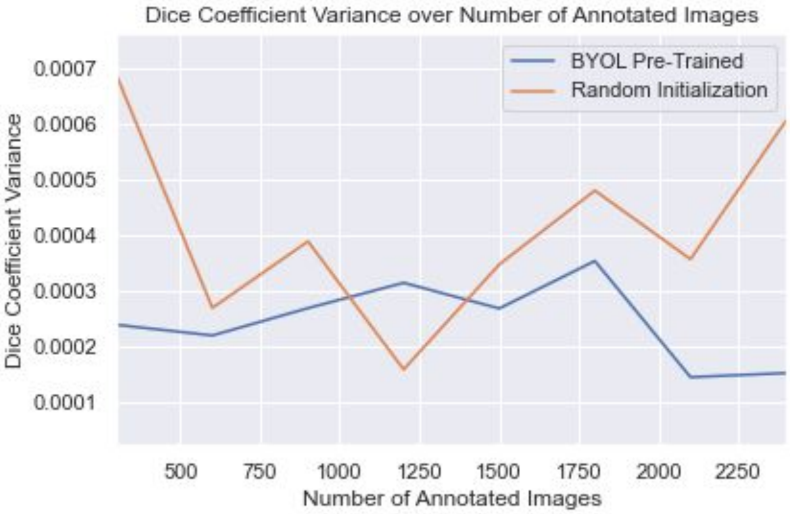
VS.

Active Learning with BYOL Pre-Training: The model is finetuned after transferring pre-trained weights. Subsequent batches of training data are added with an active learning acquisition function.

Results - **BraTS Dataset**

Combined Approach - leads to higher Model Robustness





Conclusions

- An active learning strategy based on **uncertainty** and **representativeness** improves model performance in comparison to random acquisition
- BYOL pre-training **improves model performance** in comparison to random initialization in the case of BraTS data (complex and imbalance)
- BYOL pre-training in combination with active learning **increases model robustness** in comparison to random initialization for both datasets

Thank you!

Peltarion
Holländargatan 17
SE-111 60 Stockholm
Sweden