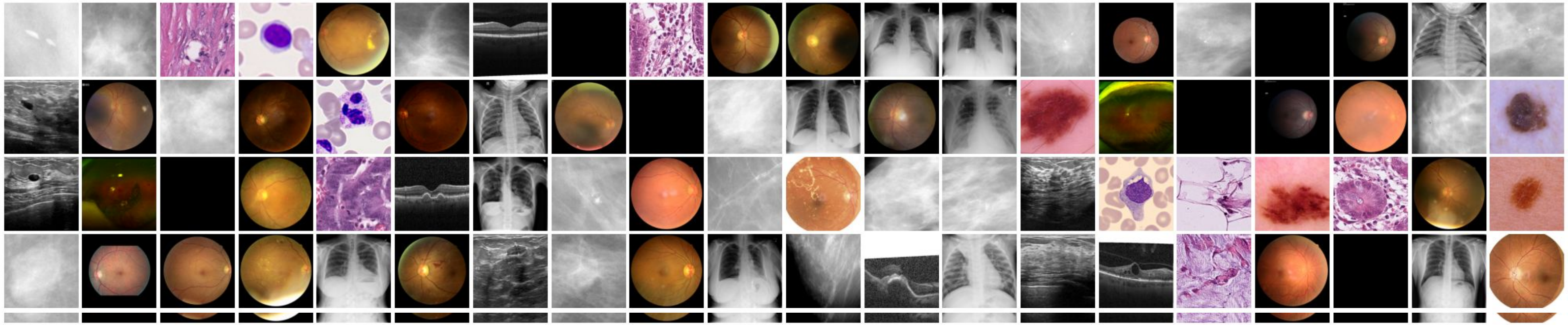
Practical Course: Machine Learning in Graphics, Vision and Language

Final presentation

Daniela Kemp, Simon Frank, Mohamed Elsherif, Gwent Krause & Tim Rebig

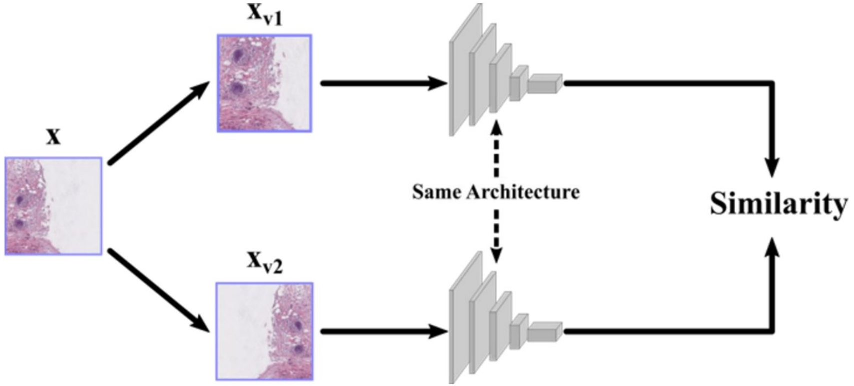
Self-Supervised Learning for Medical Image Analysis



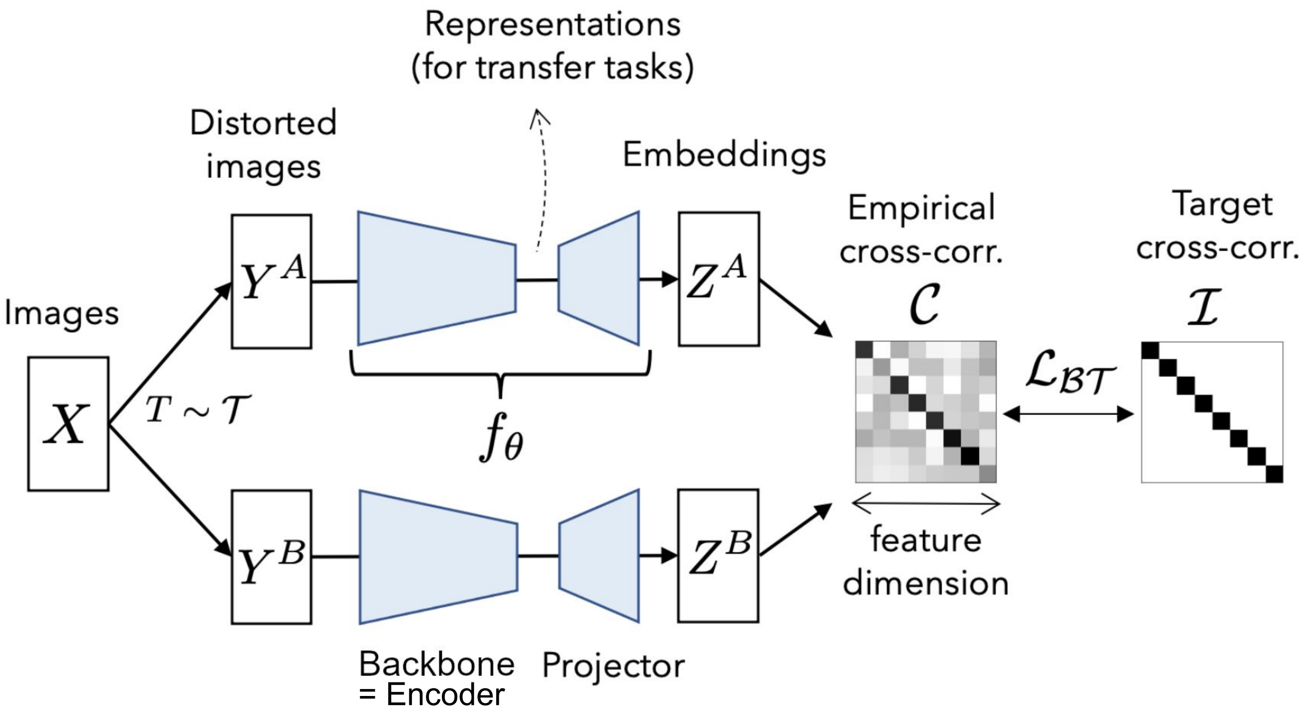
https://www.l2l-challenge.org/data.html

Self-Supervised Learning for Medical Image Analysis

- Classical data labeling requires expert knowledge making it expensive
- Need to reduce the necessity for annotated data
- Contrastive Learning:
  - SimCLR and MoCo dominant approaches, but hardware intensive
- Barlow Twins SOTA-method underexplored in the medical domain



Pre-training Method: Barlow-Twins



Barlow Twins: Self-Supervised Learning via Redundancy Reduction Jure Zbontar et. al.

Barlow-Twins Architectures

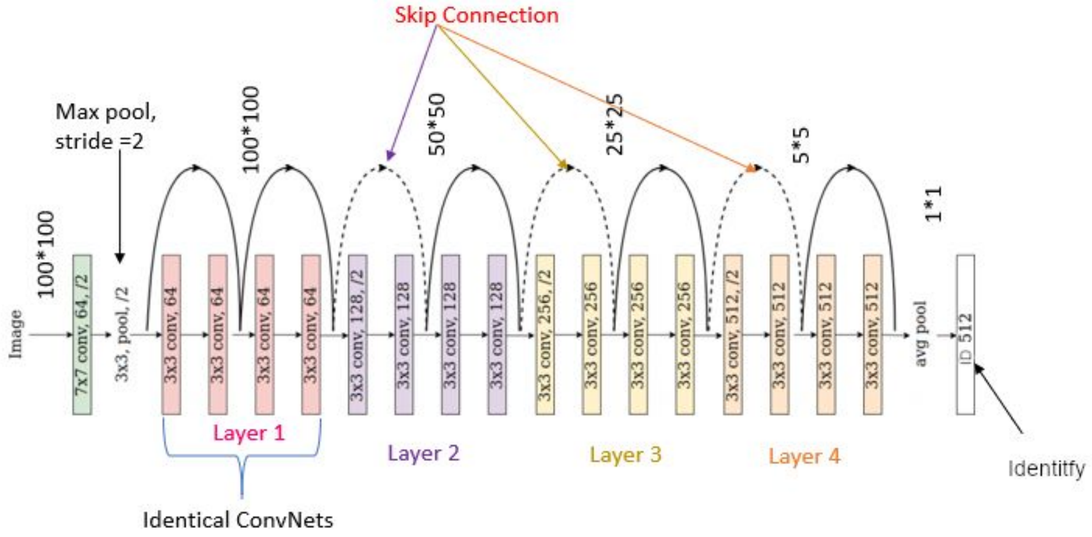
Trainable Parameters: 20.6 mio

- Resnet18 11.2 mio
- Projection Head 9.4 mio

Projection Heads:

- FC layer, batchnorm, ReLU
- FC layer, batchnorm, ReLU
- FC layer

Read-out head: MLP

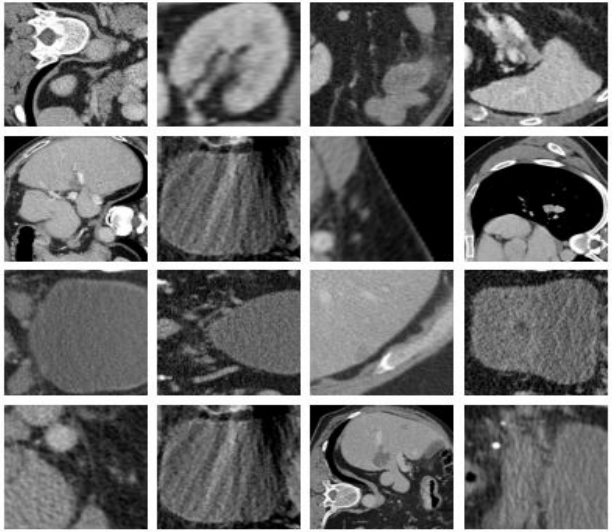


ResNet-18 Architecture

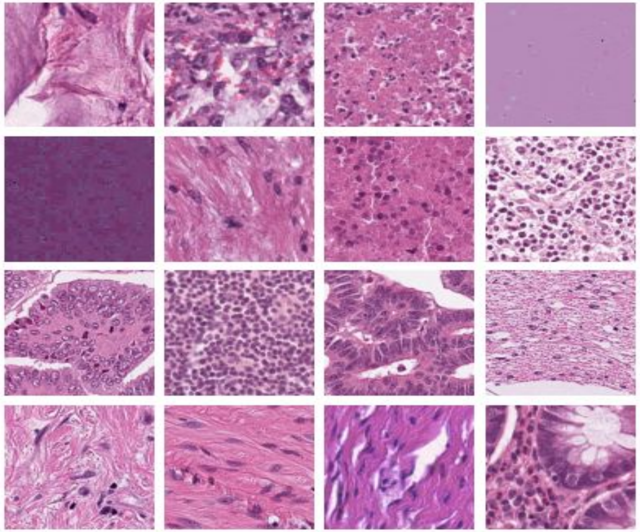
Curating medical datasets

MIMeta datasets and toolbox

- Multiple medical domains
- Provides a unified interface for loading data sets
- Images standardized to 224 x 224 pixels
- Standardised data splits



Axial organ slices



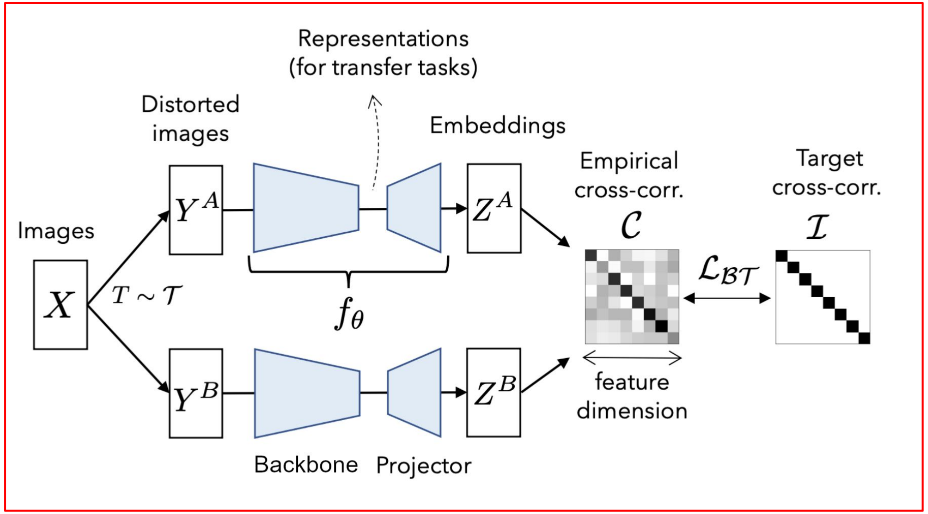
Colorectal cancer

Chosen datasets for experiments

https://www.l2l-challenge.org/data.html

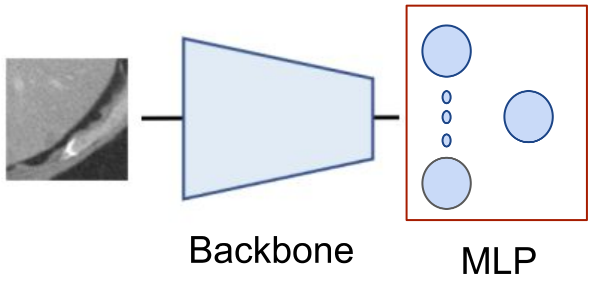
Barlow-Twin Pretraining

- Datasets: Axial Organ Slices, Coronal Organ Slices, Sagittal Organ Slices (similar domains) (3.948 images in the training set)
- Wrapper for the MIMeta Data loader to load multiple datasets
- Hyperparameters:
  - epochs: 1000
  - batch size: 256
  - learning rate: 0.03
  - optimizer: Adam



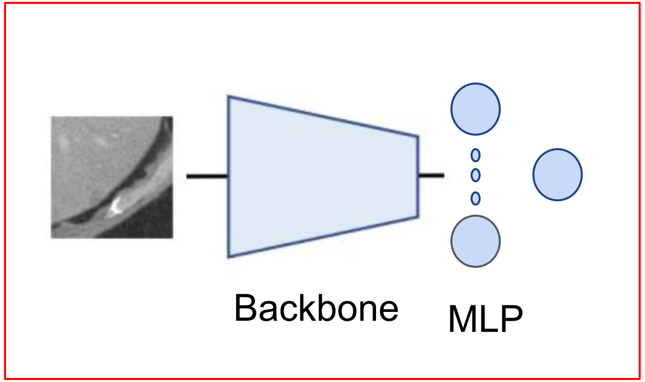
Fine-tuning Setup

- Imagenet pre-training as comparison
- Two different training method
  - Freeze Backbone train only linear Read-out head with 5.6k parameter



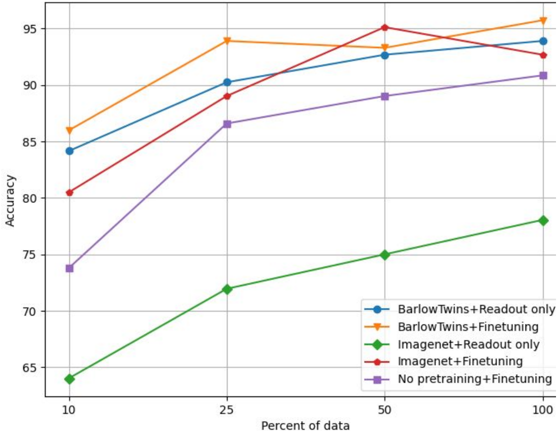
Fine-tuning Setup

- Imagenet pre-training as comparison
- Two different training method
  - Freeze Backbone train only linear Read-out head with 5.6k parameter
  - Train the whole model

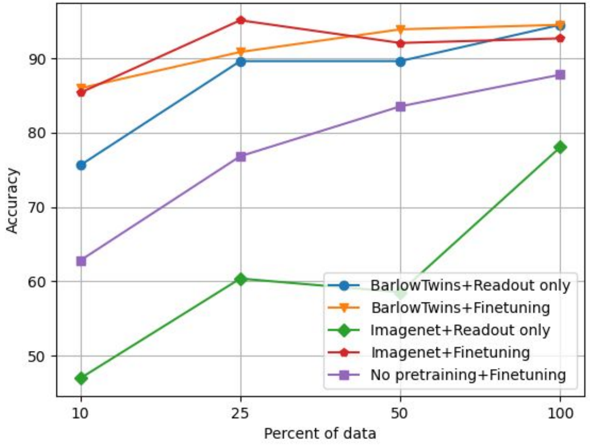


Accuracy: Comparison

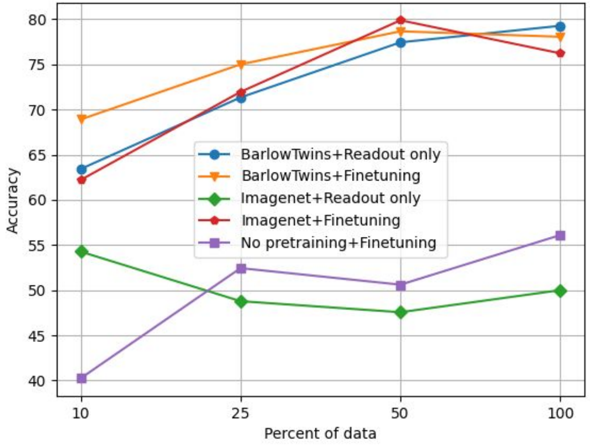
Axial

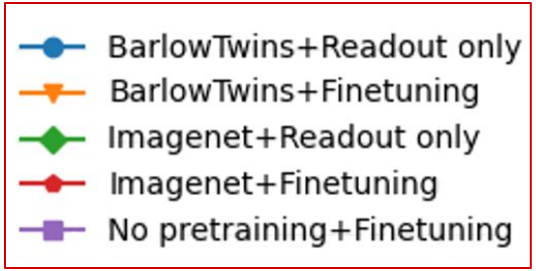


Coronal

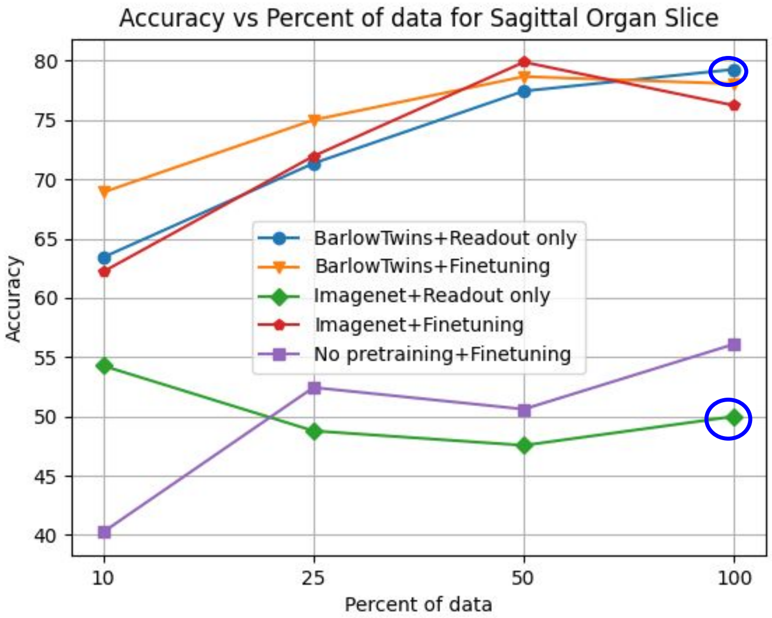


Sagittal

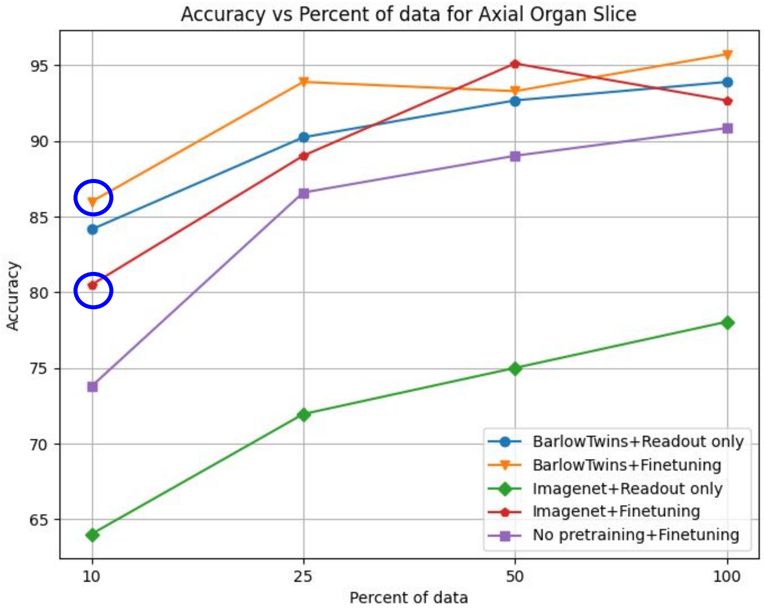


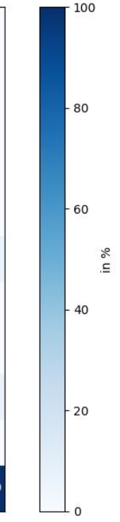


Accuracy: Comparison

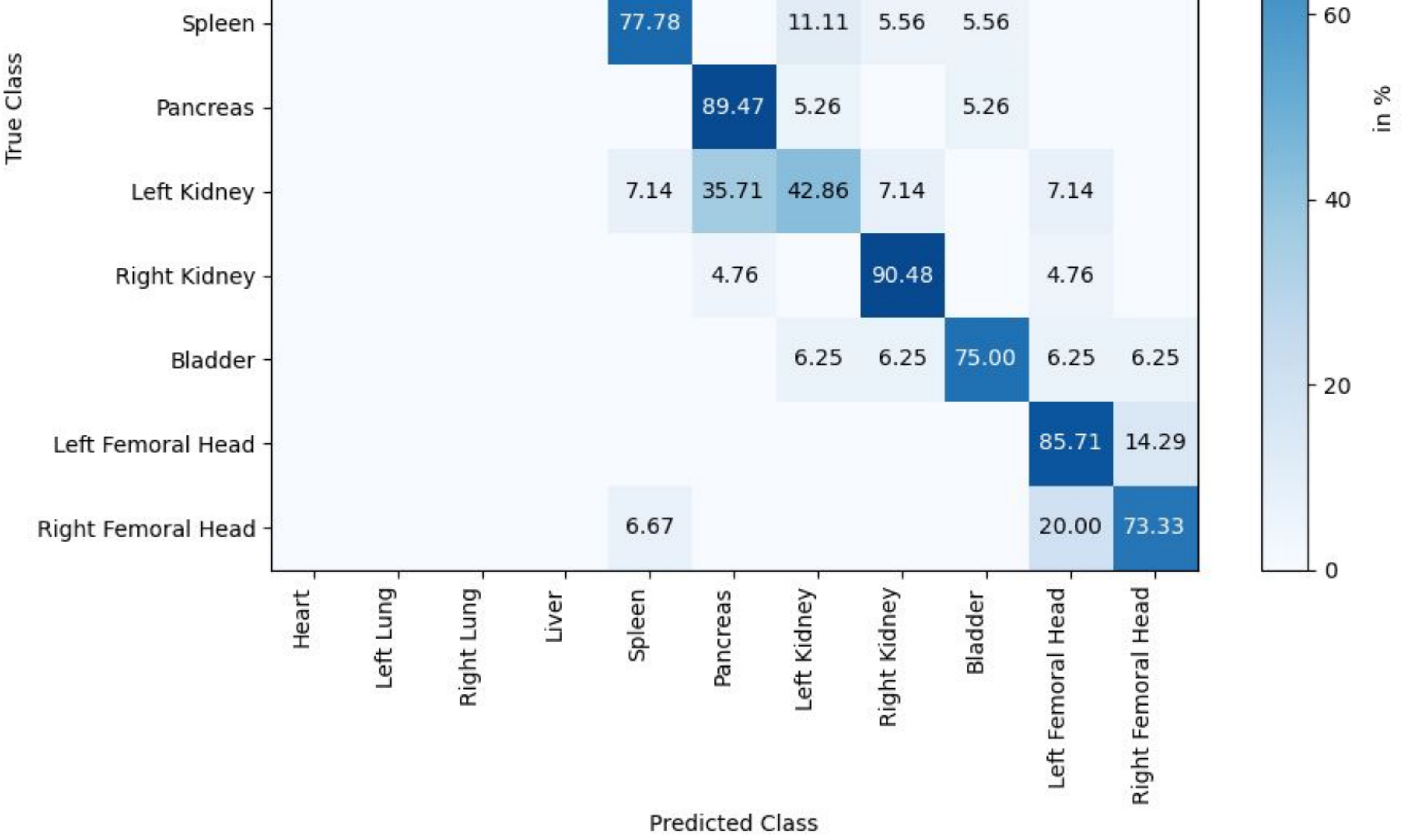


Accuracy: Comparison





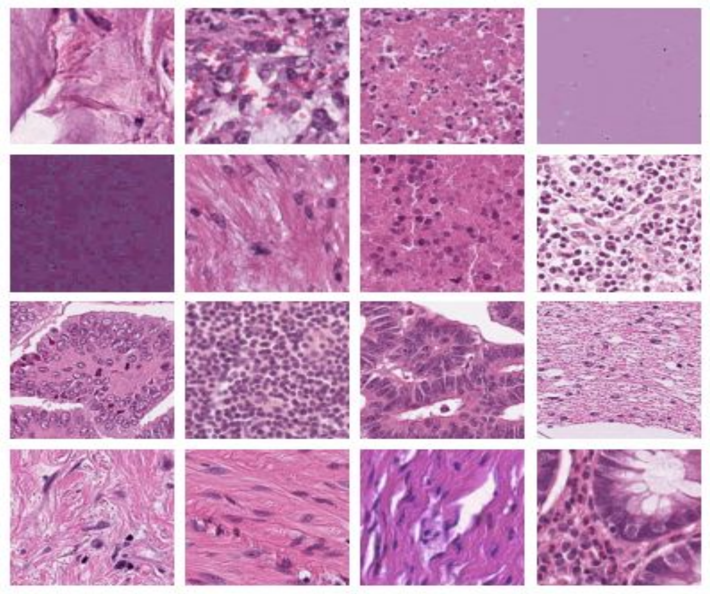
100
80
60
40
20
0
in %



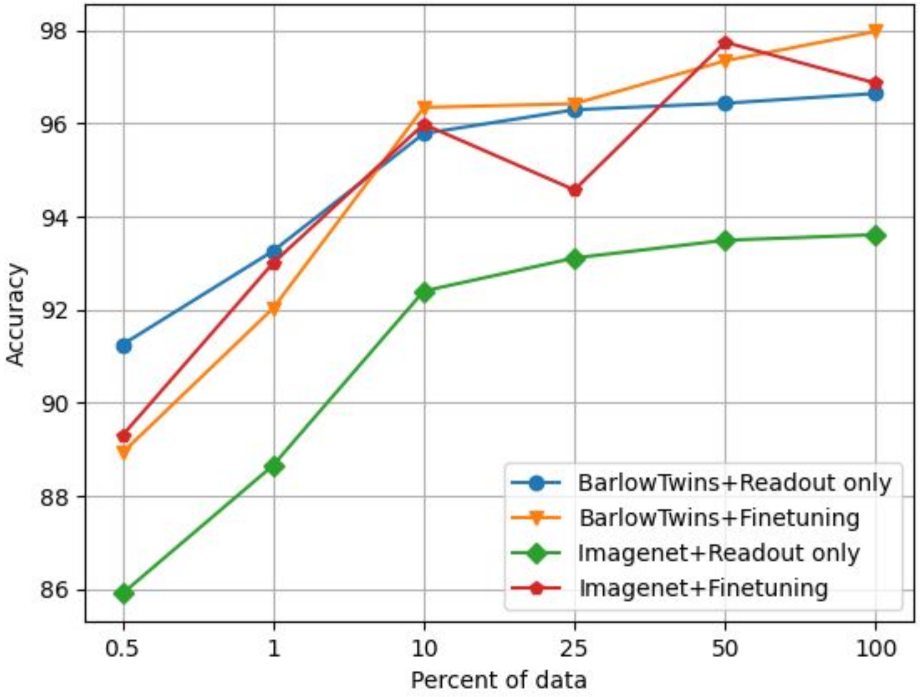
True Class
Spleen
Pancreas
Left Kidney
Right Kidney
Bladder
Left Femoral Head
Right Femoral Head
77.78
11.11
5.56
5.56
89.47
5.26
5.26
7.14
35.71
42.86
7.14
7.14
4.76
90.48
4.76
6.25
6.25
75.00
6.25
6.25
85.71
14.29
6.67
20.00
73.33
Heart
Left Lung
Right Lung
Liver
Spleen
Pancreas
Left Kidney
Right Kidney
Bladder
Left Femoral Head
Right Femoral Head
Predicted Class
60
40
20
0
in %

Colorectal Cancer Histopathology Dataset

- 107180 Total Images
- Also pretraining using Barlow Twins

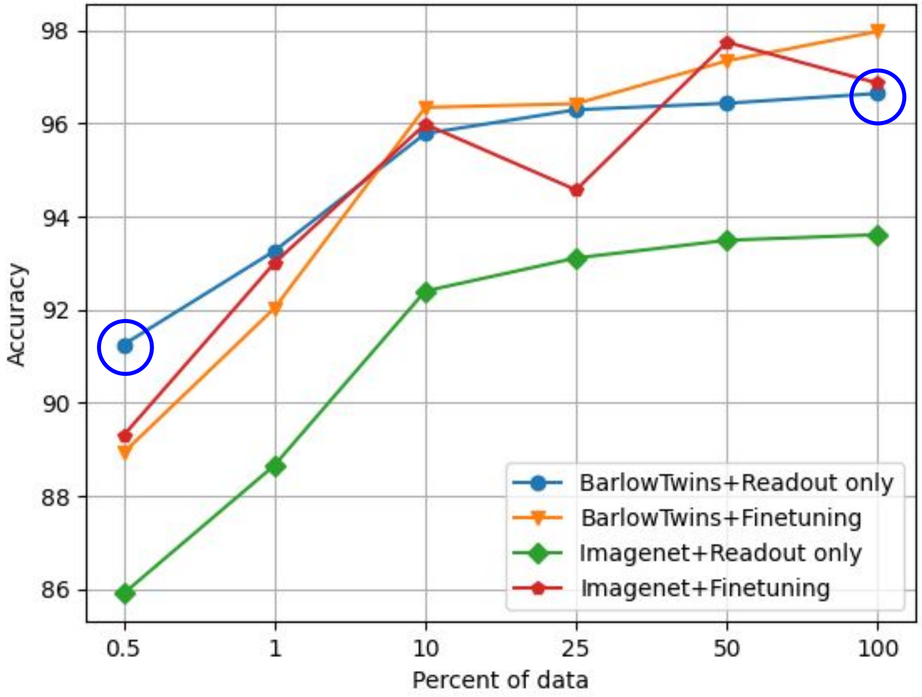


Accuracies



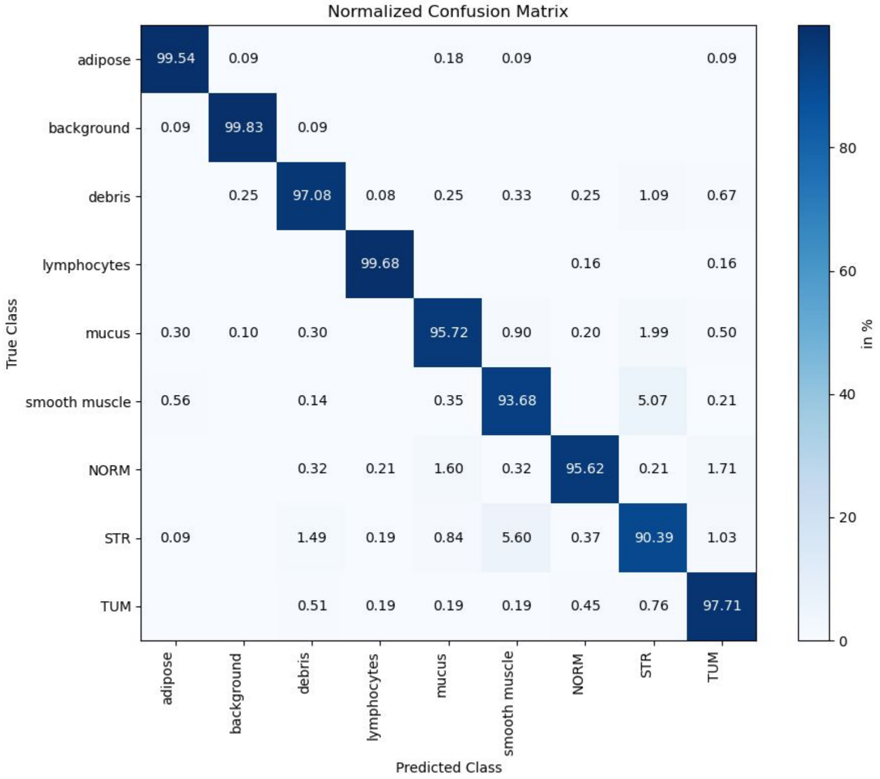
98
96
94
92
90
88
86
Accuracy
0.5
1
10
25
50
100
Percent of data
BarlowTwins+Readout only
BarlowTwins+Finetuning
Imagenet+Readout only
Imagenet+Finetuning

Accuracies



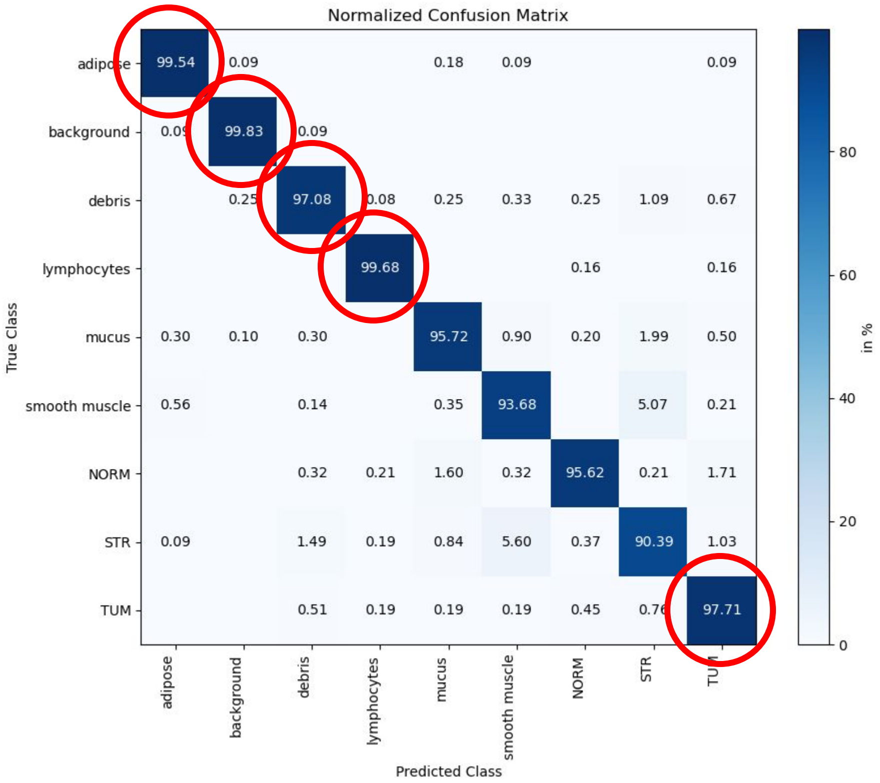
Accuracy
98
96
94
92
90
88
86
0.5
1
10
25
50
100
Percent of data
BarlowTwins+Readout only
BarlowTwins+Finetuning
Imagenet+Readout only
Imagenet+Finetuning

Barlow-Twins + only Linear Read-out-head, 100% Data



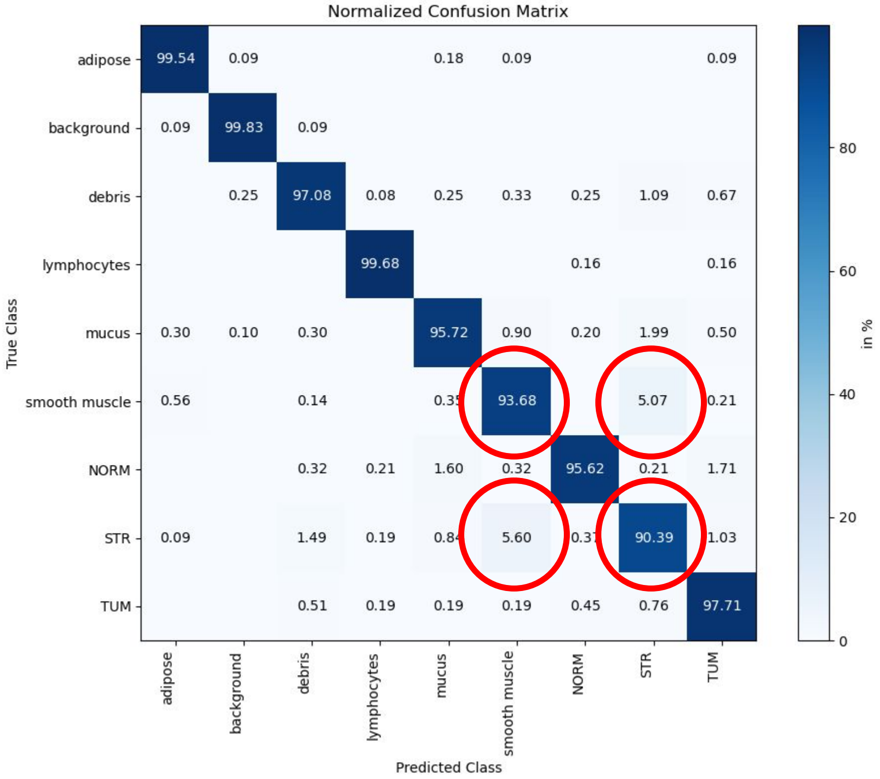
(96.64% accuracy)

Barlow-Twins + only Linear Read-out head, 100% Data



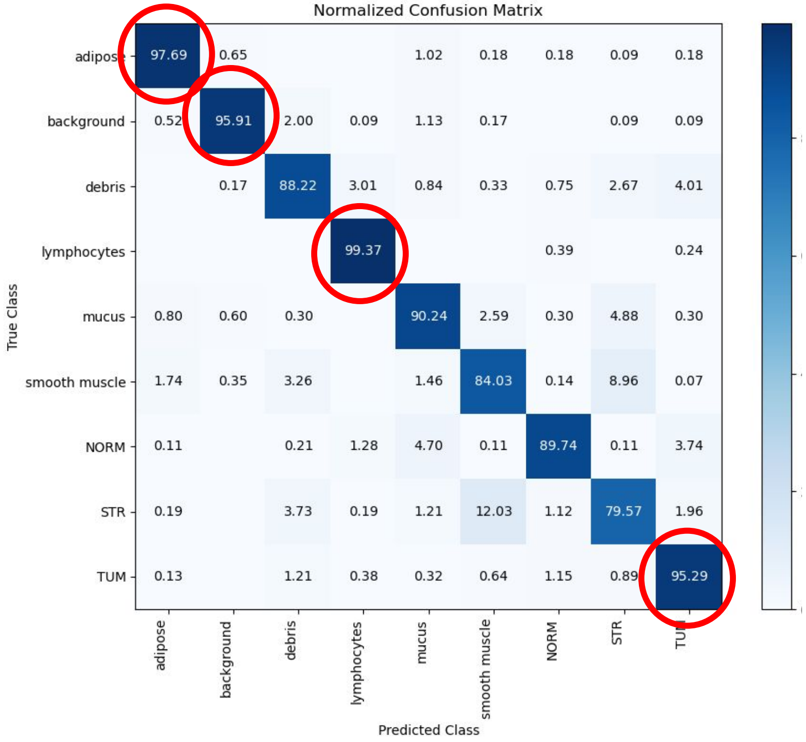
(96.64% accuracy)

Barlow-Twins + only Linear Read-out head, 100% Data

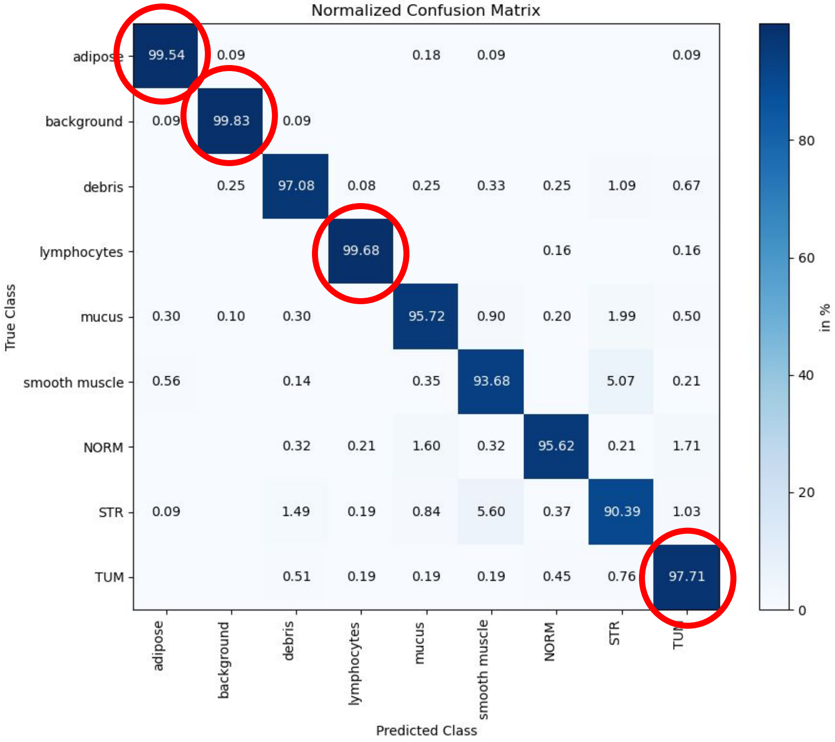


(96.64% accuracy)

Comparison - Barlow-Twins + only Linear Read-out head

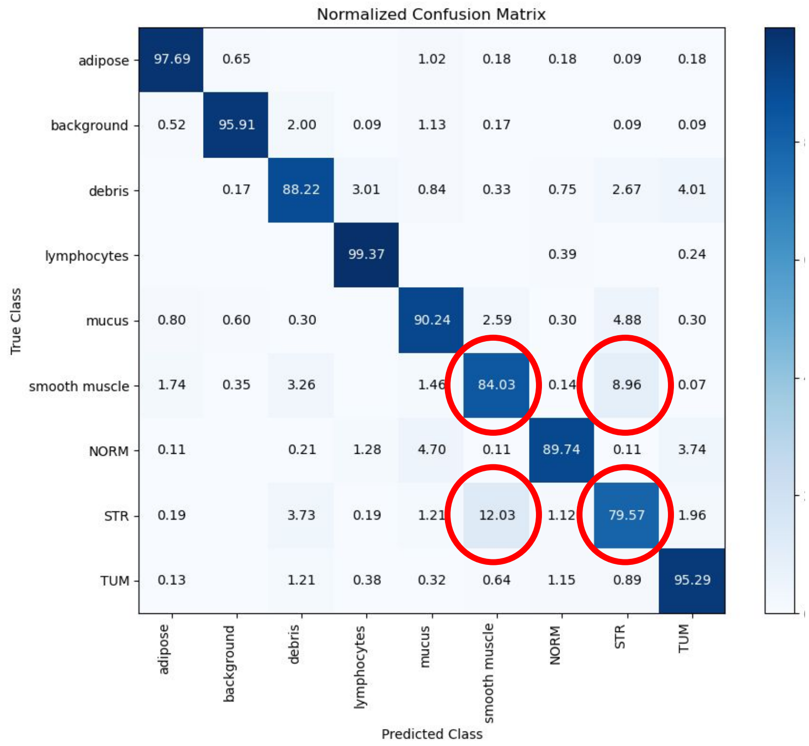


0.5% Training data

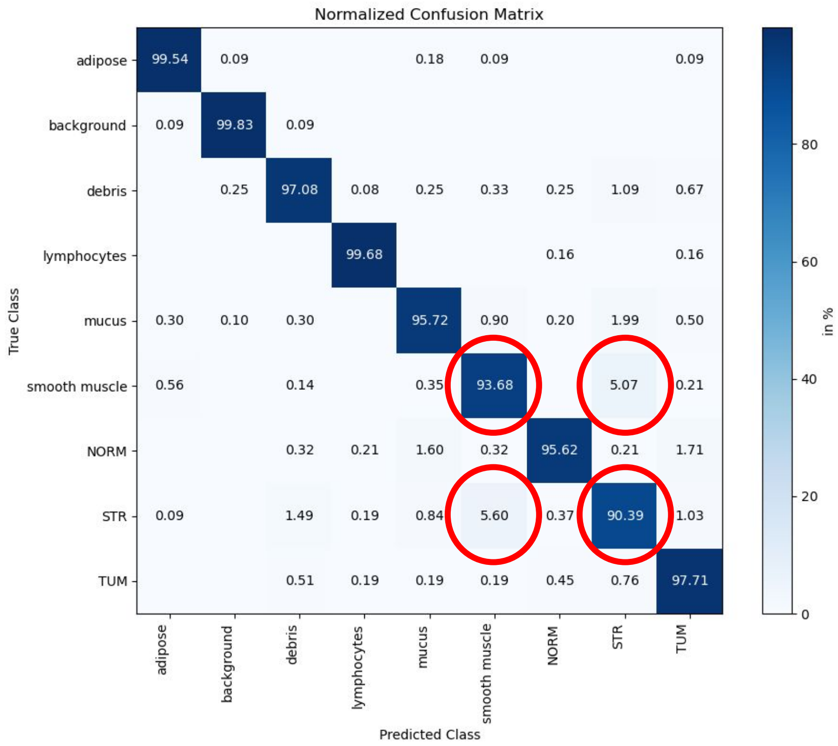


100% Training data

Comparison - Barlow-Twins + only Linear Read-out head

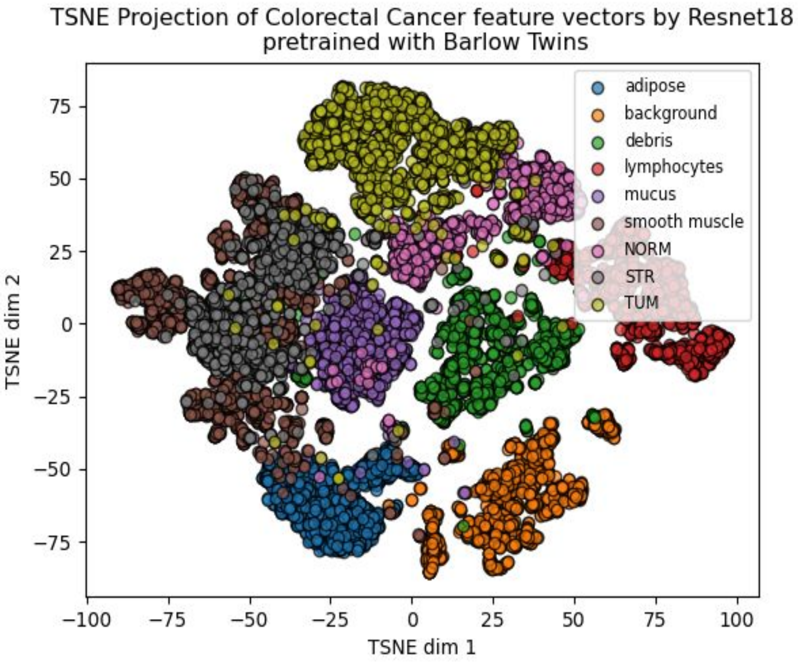


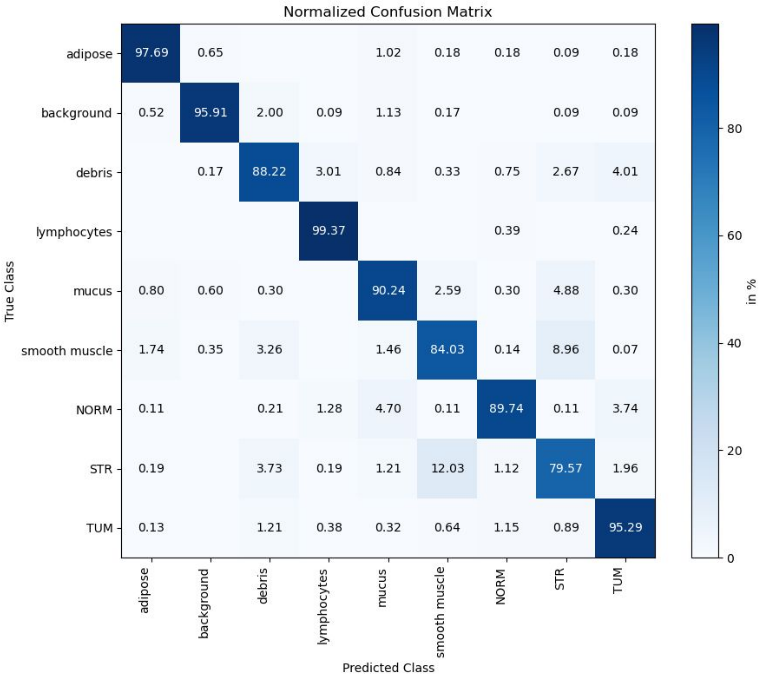
0.5% Training data



100% Training data

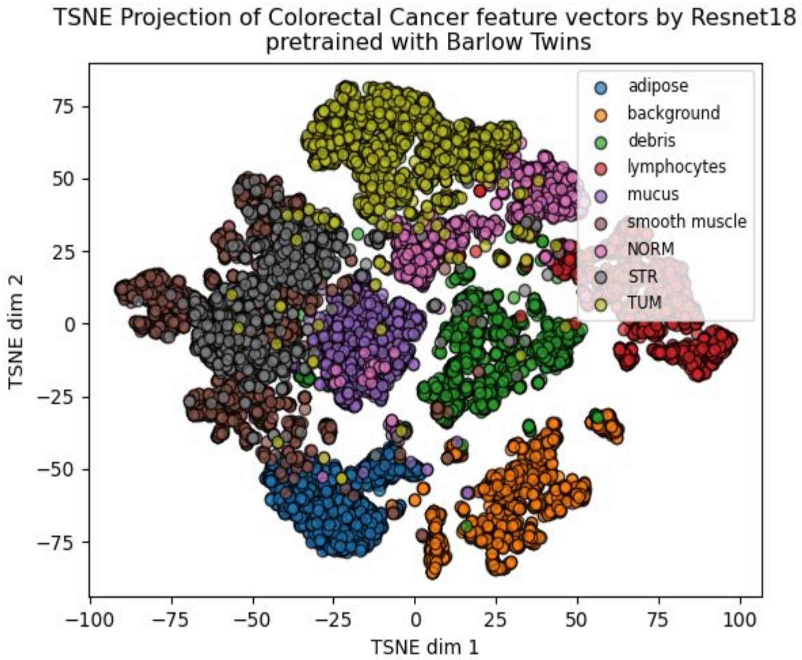
Feature analysis with t-SNE

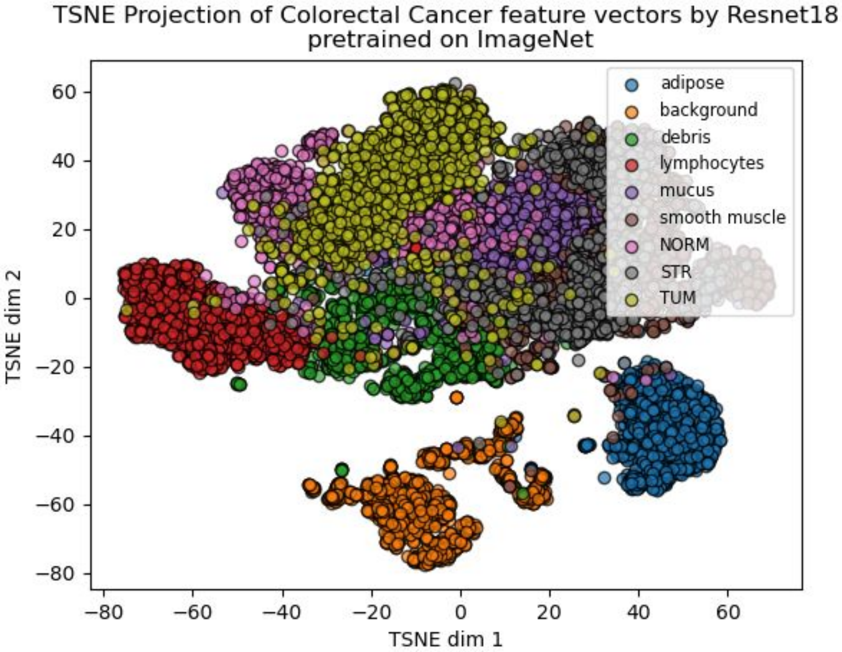




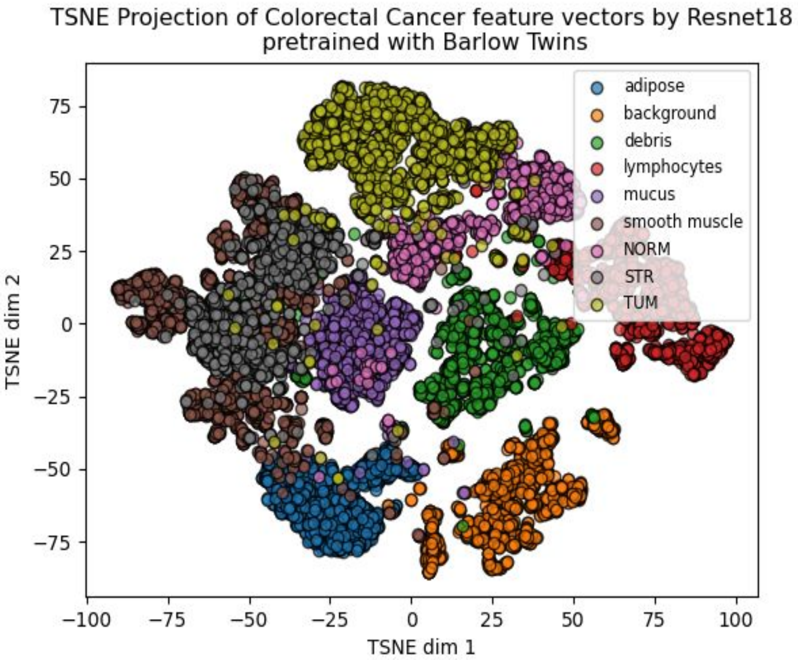
0.5% Training data

Feature analysis with t-SNE

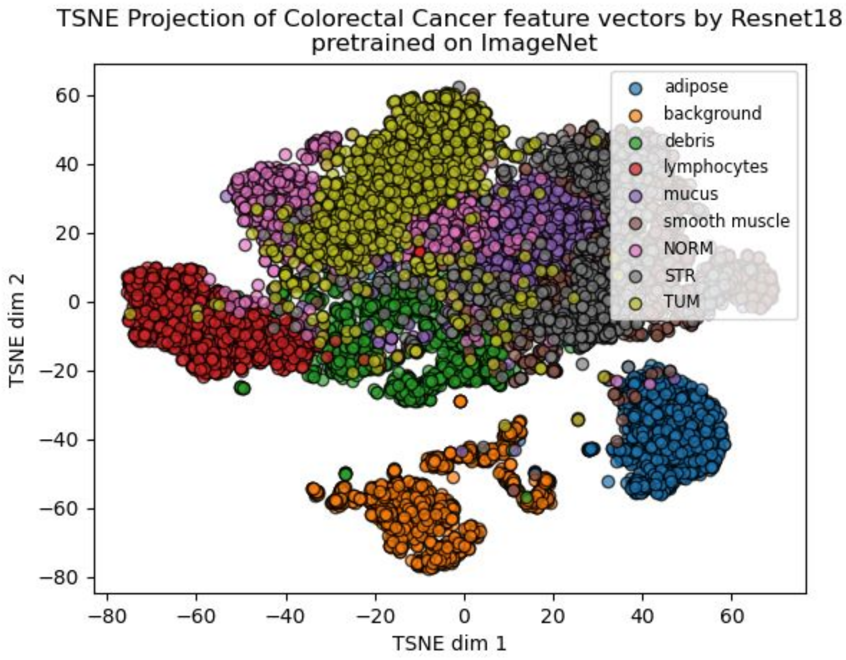




Feature analysis with t-SNE



Feature analysis with t-SNE



Barlow-Twins Conclusion

- On large dataset:
  - Marginal improvements observed
- On small dataset:
  - Pretrained features exhibit superior generalization
  - Limited fine-tuning necessary
  - Overfitting was observed during fine-tuning of the whole model
- Pre-training facilitates effective natural clustering