

Supplementary Material for Improving Representation Learning for Histopathologic Images with Cluster Constraints

Weiye Wu¹ Chongyang Gao² Joseph DiPalma¹ Soroush Vosoughi¹ Saeed Hassanpour¹
¹Dartmouth College ²Northwestern University

1. Implementation Details

This section provides details on the implementation of various self-supervised learning methods. The hyper-parameters and data augmentations for these methods are summarized below.

1.1. SimCLR

We used settings similar to those in the commonly used SimCLR repository (<https://github.com/sthalles/SimCLR>).

We configured the SimCLR model using the following hyper-parameters and augmentations:

- Optimizer: Adam
- Scheduler: Cosine Annealing
- Learning Rate: 0.1
- Weight Decay: 1e-4
- Batch Size: 512
- Crop: Scale = (0.08, 1.0)
- ColorJitter: Scale = (0.8,0.8,0.8,0.2), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (0.1,2), Probability = 0.5
- Horizontal Flip: Probability = 0.5

1.2. SwAV

We used settings similar to those in the official SwAV repository (<https://github.com/facebookresearch/swav/>).

- Optimizer: SGD
- Scheduler: Cosine Annealing
- Learning Rate: 0.1

- Weight Decay: 1e-4
- Batch Size: 512
- Momentum: 0.9
- Crop: Scale = (0.2, 1.0)
- ColorJitter: Scale = (0.4,0.4,0.4,0.1), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (0.1,2), Probability = 0.5
- Horizontal Flip: Probability = 0.5

1.3. PCL

We used settings similar to those in the official PCL repository (<https://github.com/salesforce/PCL/>).

- Optimizer: SGD
- Scheduler: Cosine Annealing
- Learning Rate: 0.03
- Weight Decay: 1e-4
- Batch Size: 512
- Momentum: 0.9
- Crop: Scale = (0.2, 1.0)
- ColorJitter: Scale = (0.4,0.4,0.4,0.1), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (0.1,2), Probability = 0.5
- Horizontal Flip: Probability = 0.5

1.4. Barlow Twins

We used settings similar to those in the official Barlow Twins repository (<https://github.com/facebookresearch/barlowtwins>). We present two different augmentation settings here, as Barlow Twins uses asymmetric augmentation techniques for its two views.

- Optimizer: LARS
- Scheduler: Cosine Annealing
- Base Learning Rate: 2
- Weight Decay: 1e-6
- Batch Size: 512

Augmentation 1:

- Crop: Scale = (0.08, 1.0)
- ColorJitter: Scale = (0.4,0.4,0.2,0.1), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (0.1, 2.0), Probability = 1.0
- Solarization: Probability = 0.0
- Horizontal Flip: Probability = 0.5

Augmentation 2:

- Crop: Scale = (0.08, 1.0)
- ColorJitter: Scale = (0.4,0.4,0.2,0.1), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (0.1, 2.0), Probability = 0.1
- Solarization: Probability = 0.2
- Horizontal Flip: Probability = 0.5

1.5. BYOL

We used settings similar to those in the widely-used BYOL repository (<https://github.com/lucidrains/byol-pytorch>).

- Optimizer: Adam
- Learning Rate: 3e-4
- Batch Size: 512
- Crop: Scale = (0.08, 1.0)
- ColorJitter: Scale = (0.8,0.8,0.8,0.2), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (1,2), Probability = 0.2
- Horizontal Flip: Probability = 0.5

1.6. SimSiam & CluSiam

In the SimSiam and corresponding CluSiam training, we used settings similar to those in the official SimSiam repository (<https://github.com/facebookresearch/simsiam>).

- Optimizer: SGD
- Scheduler: Cosine Annealing
- Learning Rate: 0.1
- Weight Decay: 1e-4
- Batch Size: 512
- Momentum: 0.9
- Fix Prediction Learning Rate: True
- RandomCrop: Scale = (0.2, 1.0)
- ColorJitter: Scale = (0.4,0.4,0.4,0.1), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (0.1,2), Probability = 0.5
- Horizontal Flip: Probability = 0.5

1.7. Supervised Learning

We followed settings similar to those in SimSiam/CluSiam to train our supervised model. A weighted loss function was used to account for class imbalance in the patch dataset.

- Optimizer: Adam
- Scheduler: Cosine Annealing
- Learning Rate: 0.1
- Weight Decay: 1e-4
- Batch Size: 512
- Weighted Loss: True
- Crop: Scale = (0.2, 1.0)
- ColorJitter: Scale = (0.4,0.4,0.4,0.1), Probability = 0.8
- Grayscale: Probability = 0.2
- GaussianBlur: Scale = (0.1,2), Probability = 0.5
- Horizontal Flip: Probability = 0.5