Unsupervised Contrastive Representation Learning for 3D Mesh Segmentation

Ayaan Haque[1,2], Hankyu Moon[2], Heng Hao[2], Sima Didari[2], Jae Oh Woo[2], Patrick Bangert[2]
[1]UC Berkeley, [2]Samsung SDS Research America

Overview

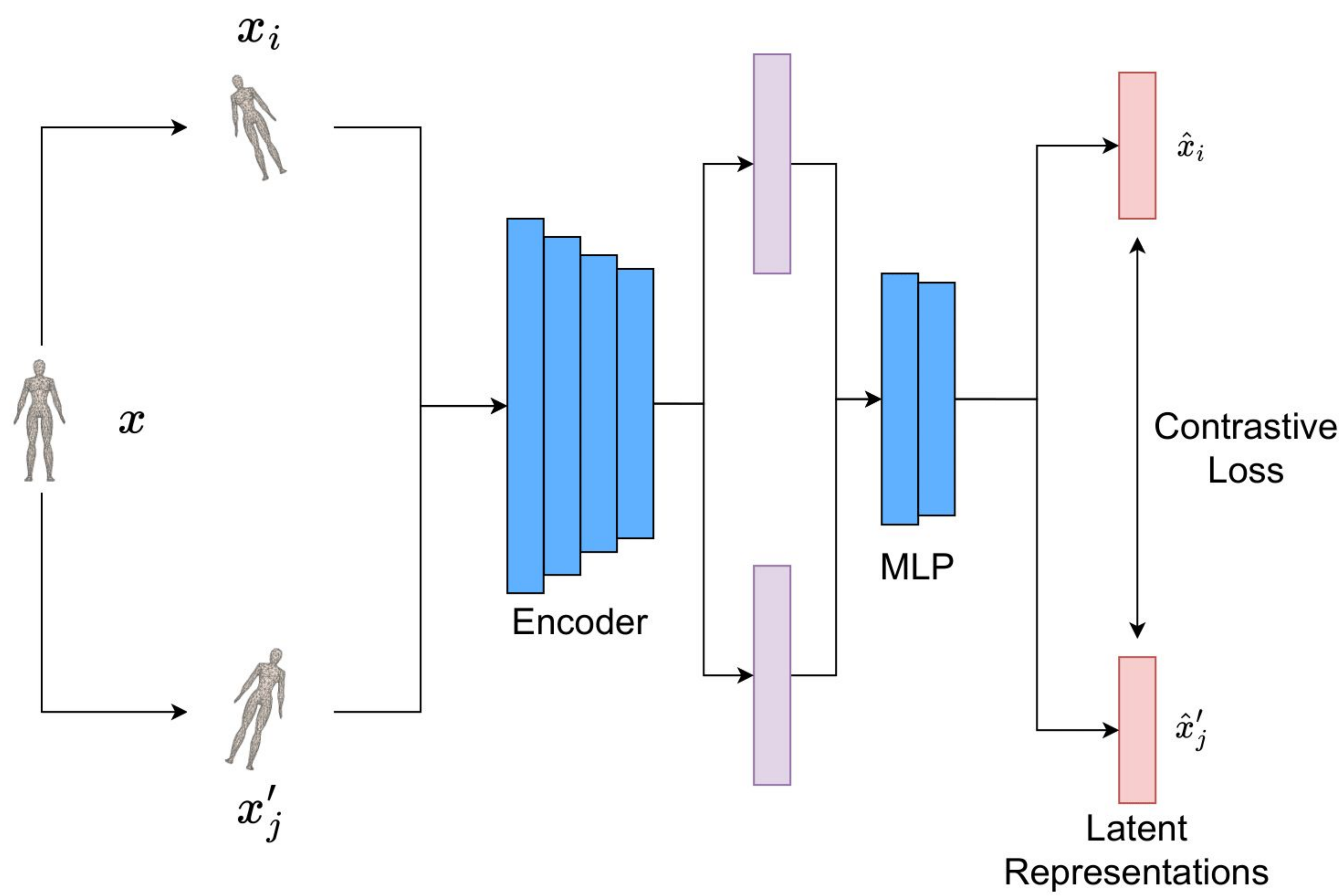
Problem

- Training deep learning networks requires high supervision, but acquiring 3D mesh labels is challenging
- Contrastive pre-training reduces burden of supervision, but difficult to apply for 3D representations

Solution: SSL-MeshCNN

- Contrastive learning algorithm for MeshCNN to improve downstream segmentation performance
- Tailored augmentation policy specialized for meshes, model learns efficient representations
- Pre-training method **reduces amount of labels** needed for mesh segmentation by **33%**

SSL-MeshCNN



Pre-Training

- Mesh-specific contrastive pre-training algorithm inspired by popular method SimCLR
  - Apply stochastic augmentation policy on all meshes to create positive and negative pairs
  - Augmentations include anisotropic scaling, vertex shifting, edge flipping
- Pre-train network using a contrastive loss on all pairs of meshes in unsupervised fashion

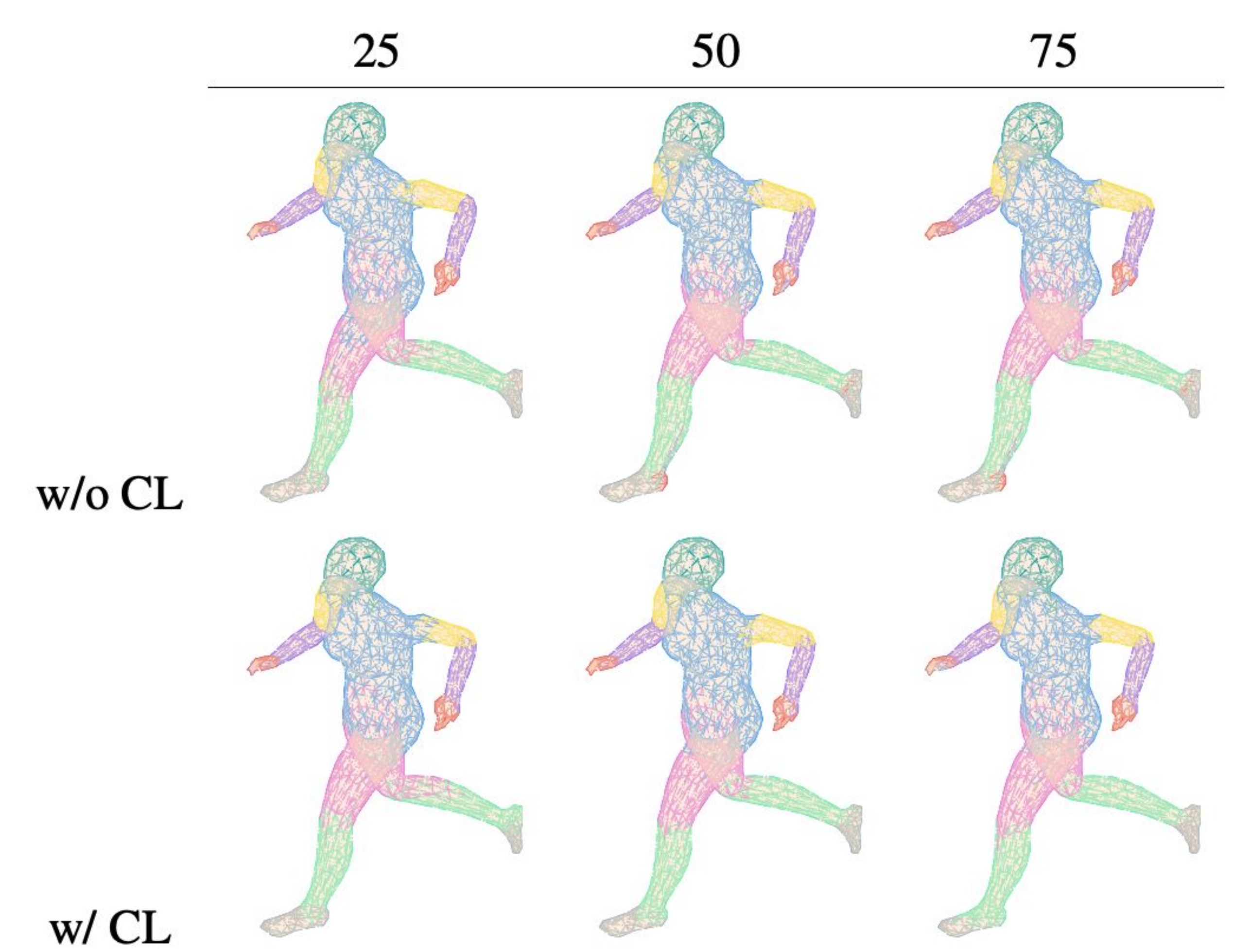
Downstream Segmentation Training

- Save pre-trained encoder, attach Mesh-Decoder to encoder to form MeshU-Net
- Perform supervised semantic segmentation of meshes using MeshU-Net (encoder-decoder framework) using only meshes with corresponding labels

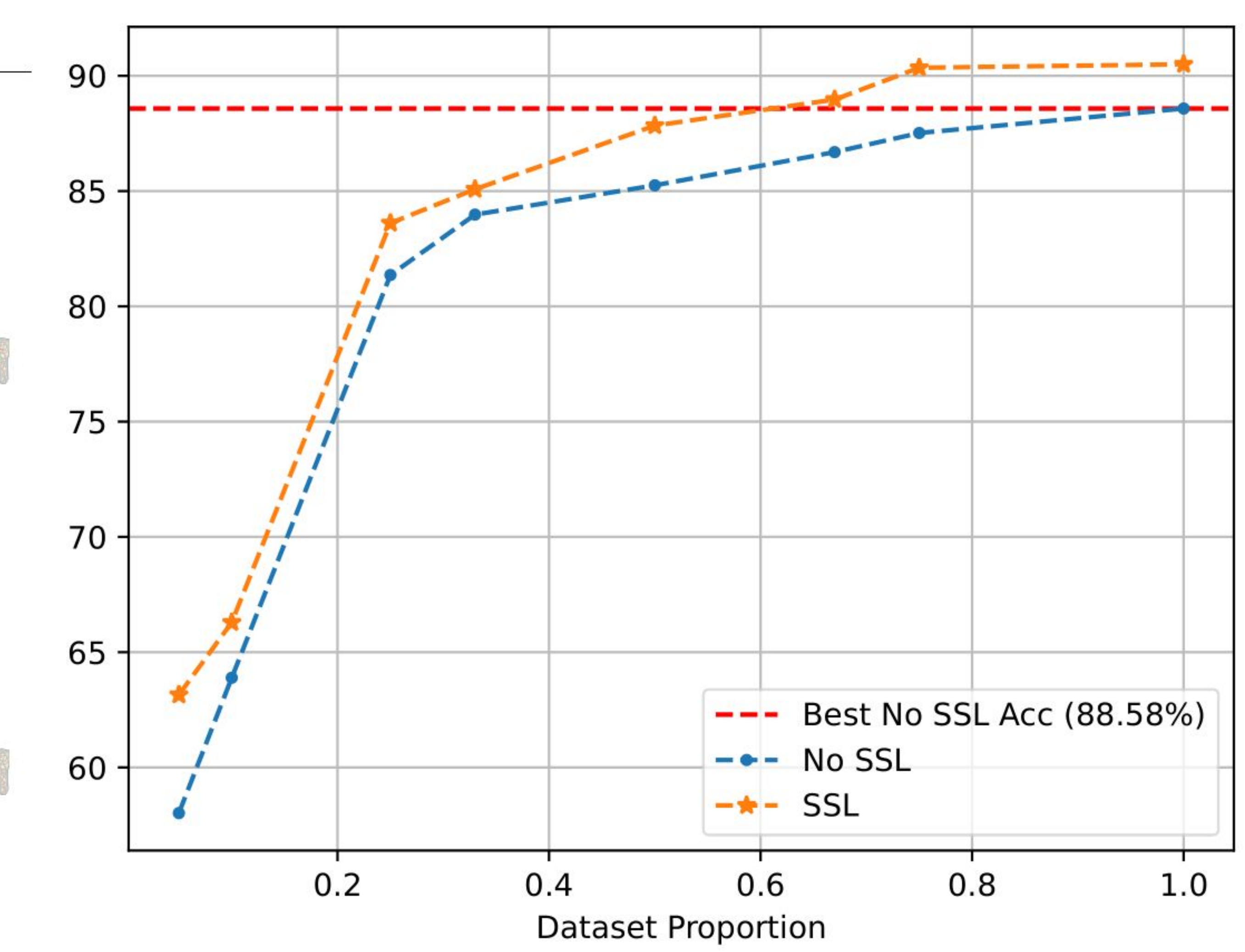
Loss Function

$$\mathcal{L}(\hat{x}_i, \hat{x}'_j) = -\log \frac{\exp\left(\text{sim}\left(\hat{x}_i, \hat{x}'_j\right)/\tau\right)}{\sum_{k=1}^{2M_1} \mathbb{1}_{[k \neq i]} \exp\left(\text{sim}\left(\hat{x}_i, \hat{x}'_k\right)/\tau\right)}$$

Results

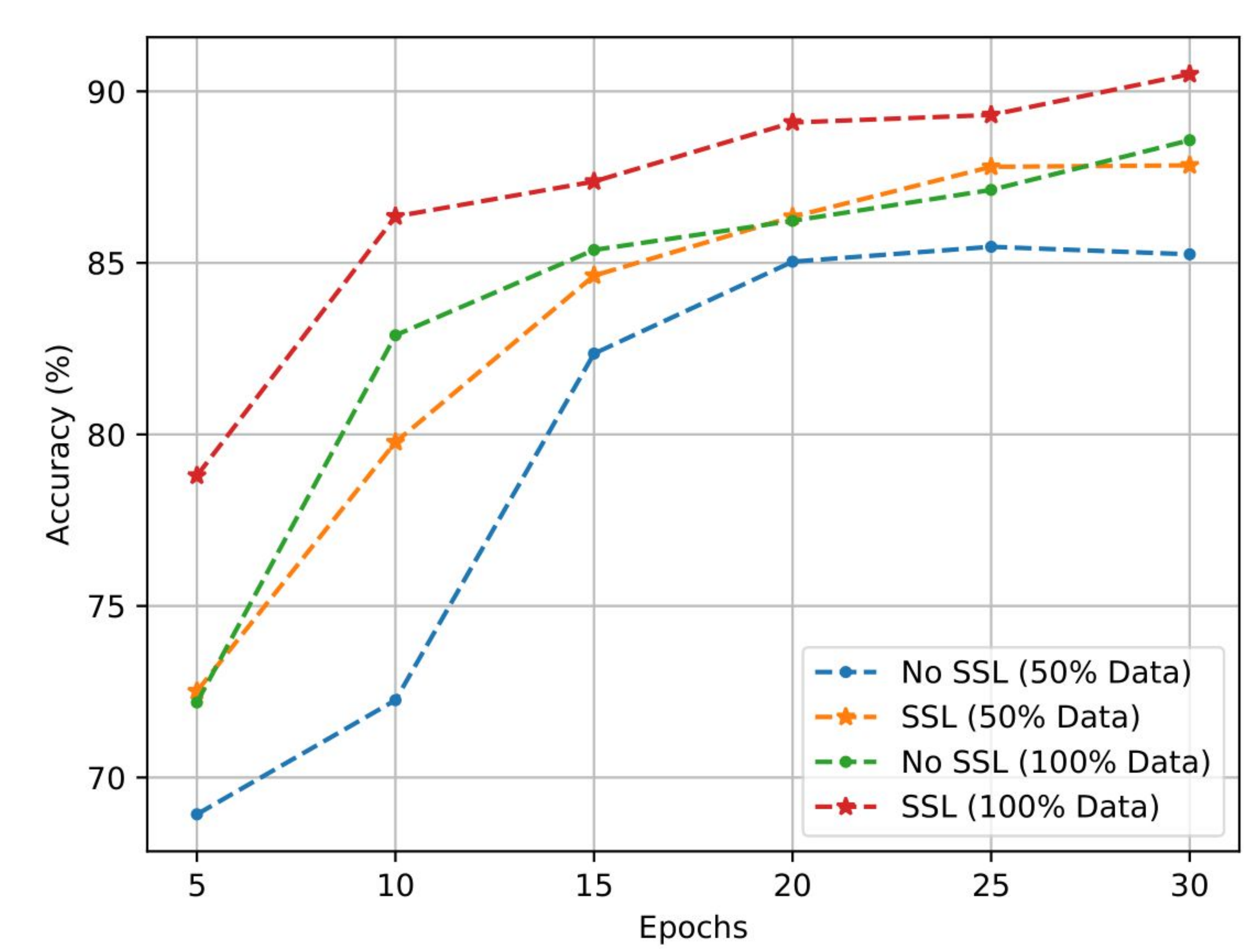


Predicted segmentation visualizations at varying proportions of dataset (%) w/ & w/o pre-training

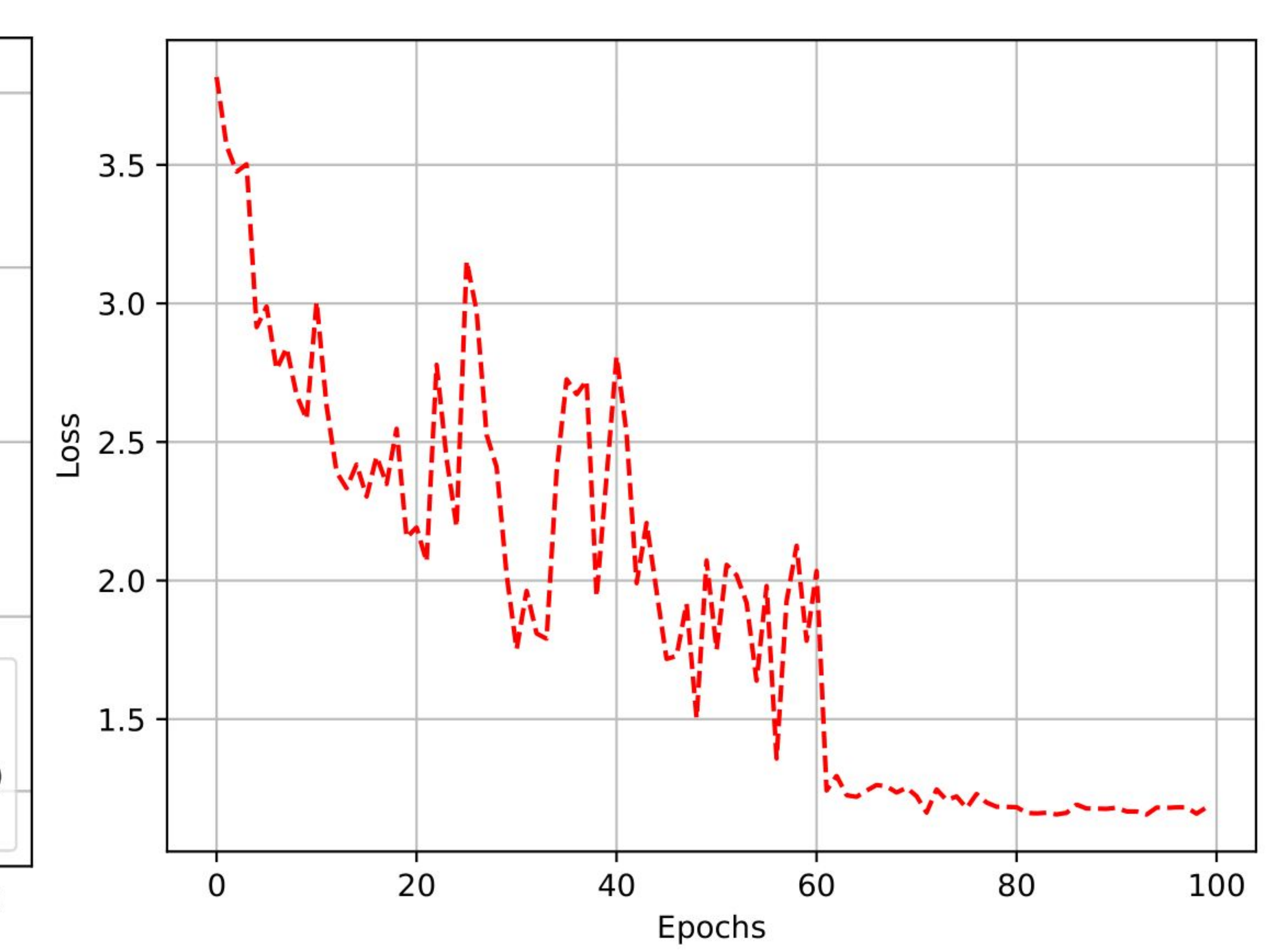


Accuracy at varying quantities of labeled data with and without contrastive learning

Training Metrics



Convergence rates for training at 50% and 100% data proportions w/ and w/o CL



Loss (NT-Xent) values at different epochs for contrastive pre-training