Algorithm 1: Self-paced contrastive learning in pretraining stage.

Input: Unlabeled dataset $\mathcal{U}$ and their respective meta-label set D_{meta}; Encoder of the segmentation network $E(\cdot)$; Temperature τ; Learning pace γ scheduler ;

Output: Pre-trained model parameters $\{\boldsymbol{\theta}\}$ for $E(\cdot)$;

Initialize network parameters $\boldsymbol{\theta}$;

Initialize hyper-parameters: learning pace: $\gamma \leftarrow \gamma_0 =$ Scheduler (0) ;

for epoch = 1, ..., n_{epochs} **do**

- **for** n = 1, ..., n_{iter} **do**
  - Sample unlabeled training batch $\{\mathcal{U}_n\}$;
  - For all $\mathbf{x}_u \in \mathcal{U}_n$, do random transformation and get $\mathbf{x}_u^T$;
  - Compute z using non-linear project $g(\cdot)$ for the features $E(\mathbf{x}_u)$;
  - Compute the sample-wise contrastive loss using Eq. (2):
  - $\ell_{ij} = -\log \frac{\exp\left(z_i^{\intercal} z_j / \tau\right)}{\sum_{a \in \mathcal{A}(i)} \exp\left(z_i^{\intercal} z_a / \tau\right)}$;
  - Compute self-paced importance weight ω_{ij} using Eq. (6):
  - $w_{ij}^{*} = \underset{w_{ij} \in [0,1]}{\arg\min}\ w_{ij}\, \ell_{ij} + R_{\gamma}(w_{ij})$;
  - Compute self-paced contrastive loss using Eq. (3):
  - $\mathcal{L}_{\text{SP-con}}^{k} = \frac{1}{2N} \sum_{i=1}^{2N} \frac{1}{|\mathcal{P}^k(i)|} \sum_{j \in \mathcal{P}^k(i)} w_{ij}\, \ell_{ij} + R_{\gamma}(w_{ij})$;
  - According to Eq. (14), do a batch gradient descent step on the model' parameters $\boldsymbol{\theta}$;
  - Update the model' parameters $\boldsymbol{\theta}$;
- Adjust the SGD learning rate;
- Update learning pace according to the scheduler: $\gamma \leftarrow$ Scheduler (epoch)

return $\{\boldsymbol{\theta}\}$;

Your Paper

You

August 10, 2021

Abstract

Your abstract.

1 Introduction