

A Details about Neighborhood Batch Sampling

In this section, we cover more details in regard to Neighborhood Batch Sampling (NBS). We have considered two instantiations of the translation mapping R and similarity scores α , based on hard k -nearest neighbor search and soft selection, respectively. Given a novel class y_n , we want to select the base classes $\{y_b\}$ that are semantically similar to the y_n query.

Hard assignments (NBS-H) This sampling method retrieves k uniformly weighted nearest base classes. NBS-H can be formulated as follows,

$$R(y_n) = \arg \min_{\mathcal{Y}'_b \subset \mathcal{Y}_b, |\mathcal{Y}'_b|=k} \sum_{y_b \in \mathcal{Y}'_b} \|\mathbf{l}_{y_b} - \mathbf{l}_{y_n}\|_2^2, \quad \alpha(y_b, y_n) = \frac{1}{k} \forall y_b \in R(y_n). \quad (11)$$

Similar heuristics are used in previous works [11, 12] as well by introducing a new hyper-parameter k . Though NBS-H may save computational resources, in practice, we find it too sensitive to the selection of k . In addition to that, it treats all selected base classes as equally related to the target novel class y_n , which slows the convergence and hurts the performance.

Soft assignments (NBS-S) In this case, all base classes are considered, and weighted by the softmax score over the learned metrics,

$$R(y_n) = \mathcal{Y}_b, \quad \alpha(y_b, y_n) = \frac{\exp(-\|\mathbf{l}_{y_b} - \mathbf{l}_{y_n}\|_2^2)}{\sum_{y'_b \in \mathcal{Y}_b} \exp(-\|\mathbf{l}_{y'_b} - \mathbf{l}_{y_n}\|_2^2)}. \quad (12)$$

Through the ablation study, we showed that this batch sampling technique is more effective than NBS-H given enough computational resources.

B Details about Intermediate GAN Objectives

In this section, we formulate our full objectives for intermediate variants derived for the imbalanced set-to-set translation.

c-GAN Its full objective could be defined as a basic minimax game,

$$G_n^* = \arg \min_{G_n} \max_{D_n} \mathcal{L}_{\text{adv}}(G_n, D_n, \mathcal{B}, \mathcal{N}). \quad (13)$$

cCyc-GAN Accordingly, its full objective can be directly derived from cycle-consistency,

$$G_n^* = \arg \min_{G_n, G_b} \max_{D_n, D_b} \mathcal{L}_{\text{adv}}(G_n, D_n, \mathcal{B}, \mathcal{N}) + \mathcal{L}_{\text{adv}}(G_b, D_b, \mathcal{N}, \mathcal{B}) + \lambda_{\text{cyc}} \mathcal{L}_{\text{cyc}}(G_n, G_b). \quad (14)$$

C Details about Computing Subgradient of Ky Fan m -norm

Theorem 1 Given a matrix \mathbf{X} and its Ky Fan m -norm $\|[\mathbf{X}]_m\|_* = \sum_i \sigma_i(\tilde{\mathbf{X}})$ where $\tilde{\mathbf{X}} = \mathbf{U}\Sigma\mathbf{V}^T$ is the m -truncated SVD and $\sigma_i(\cdot)$ is the i -th largest singular value, we have,

$$\frac{d\|[\mathbf{X}]_m\|_*}{d\mathbf{X}} = \mathbf{U}\mathbf{V}^T \quad (15)$$

Proof Rewrite Ky Fan m -norm by its sub-differential set,

$$\|[\mathbf{X}]\|_* = \text{tr}(\Sigma) = \text{tr}(\Sigma\Sigma^{-1}\Sigma) \quad (16)$$

Then,

$$d\|[\mathbf{X}]_m\|_* = \text{tr}(\Sigma\Sigma^{-1}d\Sigma) \quad (17)$$

Since we have,

$$d\mathbf{X} = d\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T + \mathbf{U}d\mathbf{\Sigma}\mathbf{V}^T + \mathbf{U}\mathbf{\Sigma}d\mathbf{V}^T \quad (18)$$

Therefore,

$$\begin{aligned} \mathbf{U}d\mathbf{\Sigma}\mathbf{V}^T &= d\mathbf{X} - d\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T - \mathbf{U}\mathbf{\Sigma}d\mathbf{V}^T \\ \Rightarrow d\mathbf{\Sigma} &= \mathbf{U}^T d\mathbf{X}\mathbf{V} - \mathbf{U}^T d\mathbf{U}\mathbf{\Sigma} - \mathbf{\Sigma}d\mathbf{V}^T\mathbf{V} \end{aligned} \quad (19)$$

By the diagonality of $\mathbf{\Sigma}$ and anti-symmetricity of \mathbf{U} , \mathbf{V} ,

$$\begin{aligned} \mathbf{U}^T d\mathbf{U}\mathbf{\Sigma} + \mathbf{\Sigma}d\mathbf{V}^T\mathbf{V} &= 0 \\ \Rightarrow d\mathbf{\Sigma} &= \mathbf{U}^T d\mathbf{X}\mathbf{V} \end{aligned} \quad (20)$$

Substitute it into Equation 17,

$$\begin{aligned} d\|[\mathbf{X}]_m\|_* &= \text{tr}(\mathbf{\Sigma}\mathbf{\Sigma}^{-1}d\mathbf{\Sigma}) = \text{tr}(\mathbf{U}^T d\mathbf{X}\mathbf{V}) = \text{tr}(\mathbf{U}^T\mathbf{V}d\mathbf{X}) \\ \Rightarrow \frac{d\|[\mathbf{X}]_m\|_*}{d\mathbf{X}} &= \mathbf{U}\mathbf{V}^T \end{aligned} \quad (21)$$

□