Supplementary Material for
Semantics-Preserving Locality Embedding for Zero-Shot Learning

We now provide technical details on the derivations of our proposed zero-shot learning (ZSL) algorithms. Recall that the inductive ZSL problem can be expressed as follows:

$$\begin{aligned}\min_{\mathbf{A}_S,\mathbf{A}_F} \; & E_C(\mathbf{A}_S,\mathbf{A}_F)+\lambda_1 E_S(\mathbf{A}_F)+\lambda_2\Omega(\mathbf{A}_F,\mathbf{A}_S)\\ \text{s.t.} \;\; & \mathbf{Z}\mathbf{H}\mathbf{Z}^\top=\mathbf{I}.\end{aligned} \tag{i}$$

where $\mathbf{Z} = [\mathbf{A}_F^\top\mathbf{X}, \mathbf{A}_S^\top\mathbf{S}] \in \mathbb{R}^{d_k\times(N+C)}$ indicates the concentrated projected data matrix, $\mathbf{S} = [\mathbf{s}_1,\mathbf{s}_2,\ldots,\mathbf{s}_C] \in \mathbb{R}^{d_s\times C}$, and $\mathbf{I}$ is the identity matrix. To solve the minimization problem of (i), we define $\widetilde{\mathbf{X}} = \begin{pmatrix}\mathbf{X} & \mathbf{0}_{d_F\times C}\\ \mathbf{0}_{d_S\times N} & \mathbf{S}\end{pmatrix}$ and an augment transformation matrix $\mathbf{A} = [\mathbf{A}_F;\mathbf{A}_S]$ with $\mathbf{Z} = \mathbf{A}^\top\widetilde{\mathbf{X}}$. In order to derive the formula more aesthetically, we now treat $\mathbf{s}_i$ belongs to class i. As suggested in [1],

$$E_C(\mathbf{A}_S,\mathbf{A}_F)=\sum_{j=1}^{C}\|\mathbf{A}_S^\top\mathbf{s}_j-\frac{1}{N_j}\sum_{i=1}^{N_j}\mathbf{A}_F^\top\mathbf{x}_i^j\|^2 \tag{ii}$$

can be rewritten in the following concise form:

$$E_C(\mathbf{A}_S,\mathbf{A}_F)=tr(\mathbf{A}^\top\widetilde{\mathbf{X}}\mathbf{M}\widetilde{\mathbf{X}}^\top\mathbf{A}), \tag{iii}$$

where $tr(\cdot)$ denotes the trace sum, and the each entry in the matrix $\mathbf{M}$ is defined as follows:

$$\mathbf{M}_{ij}=\begin{cases}\frac{1}{N_cN_c} & \text{if } i,j\le N \text{ and } y_i=y_j=c\\ 1 & \text{if } i,j>N \text{ and } i=j\\ \frac{-1}{N_c} & \text{if } \begin{cases} i\le N, j>N\\ i>N, j\le N\end{cases} \text{ and } y_i=y_j=c\\ 0 & \text{otherwise.}\end{cases} \tag{iv}$$

Similarly,

$$E_S(\mathbf{A}_F)=\frac{1}{2}\sum_{j=1}^{C}\{\frac{1}{N_j^2}\sum_{i=1}^{N_j}\sum_{k=1}^{N_j}\|\mathbf{A}_F^\top\mathbf{x}_i^j-\mathbf{A}_F^\top\mathbf{x}_k^j\|^2\} \tag{v}$$

can also be written in a trace norm form as follows:

$$E_S(\mathbf{A}_S,\mathbf{A}_F)=tr(\mathbf{A}^\top\widetilde{\mathbf{X}}\mathbf{L}\widetilde{\mathbf{X}}^\top\mathbf{A}), \tag{vi}$$

where $\frac{1}{N_j^2}$ is for normalization, $\mathbf{L} = \mathbf{D} - \mathbf{W}$, and $\mathbf{D}$ is a diagonal matrix with $(\mathbf{D})_{ii} = \sum_j \mathbf{W}_{ij}$ and $\mathbf{W}$ is the affinity matrix defined as:

$$\mathbf{W}_{ij}=\begin{cases}1 & \text{if } i,j\le N \text{ and } y_i=y_j\\ 0 & \text{otherwise.}\end{cases} \tag{vii}$$

By combining the above terms, we convert the whole formula (i) as:

$$\begin{aligned}\min_{\mathbf{A}=[\mathbf{A}_S;\mathbf{A}_F]} \; & tr(\mathbf{A}^\top\widetilde{\mathbf{X}}(\mathbf{M}+\lambda_1\mathbf{L})\widetilde{\mathbf{X}}^\top\mathbf{A})+\lambda_2\left(\|\mathbf{A}_S\|^2+\|\mathbf{A}_F\|^2\right)\\ \text{s.t.} \;\; & \mathbf{A}^\top\widetilde{\mathbf{X}}\mathbf{H}\widetilde{\mathbf{X}}^\top\mathbf{A}=\mathbf{I},\end{aligned} \tag{viii}$$

where we use the Frobenius norm as the regularizer. Now, the transformation $\mathbf{A}$ can be derived by solving the d_k smallest eigenvectors of the following generalized eigenvalue decomposition problem:

$$(\widetilde{\mathbf{X}}(\mathbf{M}+\lambda_1\mathbf{L})\widetilde{\mathbf{X}}^\top+\lambda_2\mathbf{I})\mathbf{A}=\Phi\widetilde{\mathbf{X}}\mathbf{H}\widetilde{\mathbf{X}}^\top\mathbf{A}. \tag{ix}$$

Similarly, the transductive version of our ZSL can also be solved by:

$$(\hat{\mathbf{X}}(\hat{\mathbf{M}}+\lambda_1\hat{\mathbf{L}})\hat{\mathbf{X}}^\top+\lambda_2\mathbf{I})\mathbf{A}=\Phi\hat{\mathbf{X}}\hat{\mathbf{H}}\hat{\mathbf{X}}^\top\mathbf{A}, \tag{x}$$

where $\hat{\mathbf{X}} = \begin{pmatrix}\mathbf{X} & \mathbf{X}^U & \mathbf{0}_{d_F\times C} & \mathbf{0}_{d_F\times C^U}\\ \mathbf{0}_{d_S\times N} & \mathbf{0}_{d_S\times N^U} & \mathbf{S} & \mathbf{S}^U\end{pmatrix}$. Note that $\hat{\mathbf{M}}$ and $\hat{\mathbf{L}}$ can be computed with pseudo labels as in (iv) and (vii), respectively.

References

[1] Mingsheng Long, Jianmin Wang, Guiguang Ding, Jiaguang Sun, and Philip S Yu. Transfer feature learning with joint distribution adaptation. In *ICCV*, 2013. 1