Supplementary Material:
Progressive Graph Learning for Open-Set Domain Adaptation

1. Proof for Theorem 2.1

Proof. Let h^* be the idea joint classifier which minimizes the combined error,

$$h^* = \arg\min_{h\in\mathcal{H}} R_s(h) + R_t^*(h). \tag{1}$$

Given the definition of target risk, we have,

$$R_t(h) = \sum_{i=1}^{C} \pi_i^t R_{t,i}(h) + \pi_{C+1}^t R_{t,C+1}(h). \tag{2}$$

For brevity, we denote the first term as $R_t^*(h)$ and the second term as Δ_o. Then,

$$\begin{aligned} R_t(h) &= R_t^*(h) + \Delta_o, \\ &\leq R_t^*(h^*) + (1-\pi_{C+1}^t)\mathbb{E}_{\mathbb{Q}^t_{X|Y\leq C}}\mathcal{L}(h,h^*) + \Delta_o, \\ &\leq R_t^*(h^*) + (1-\pi_{C+1}^t)(\mathbb{E}_{\mathbb{P}^s_X}\mathcal{L}(h,h^*) \\ &+ disc(\mathbb{Q}^t_{X|Y\leq C}, \mathbb{P}^s)) + \Delta_o. \end{aligned} \tag{3}$$

Based on the triangle inequality we have,

$$\begin{aligned} \mathbb{P}^s_X\mathcal{L}(h,h^*) &\leq \mathbb{P}^s_X\mathcal{L}(h,i) + \mathbb{P}^s_X\mathcal{L}(h^*,i), \\ &\leq R_s(h) + R_s(h^*). \end{aligned} \tag{4}$$

Thus, the Equation 3 can be rewritten as,

$$\begin{aligned} R_t(h) \leq R_t^*(h^*) &+ (1-\pi_{C+1}^t)\big(R_s(h) + R_s(h^*) \\ &+ disc(\mathbb{Q}^t_{X|Y\leq C}, \mathbb{P}^s_X)\big) + \Delta_o. \end{aligned} \tag{5}$$

Lastly, we can easily obtain,

$$\begin{aligned} \frac{R_t(h)}{1-\pi_{C+1}^t} &\leq R_s(h) + disc(\mathbb{Q}^t_{X|Y\leq C}, \mathbb{P}^s_X) + \lambda \\ &+ \frac{\pi_{c+1}^t}{1-\pi_{c+1}^t} R_{t,C+1}(h), \end{aligned} \tag{6}$$

where the shared error $\lambda = \min_{h\in\mathcal{H}} \frac{R_t^*(h)}{1-\pi_{c+1}^t} + R_s(h)$. ■

2. Proof for Theorem 3.1

Proof. By definition, the target risk with the shared classifier $\tilde{h} \in \mathcal{H}_1$ and the pseudo labeling function $h_b \in \mathcal{H}_2$ can be represented as,

$$\begin{aligned} R_t(\tilde{h}, h_b) &= (1-\pi_\alpha)R_t^*(\tilde{h}) + \pi_\alpha R_t^*(h_b) \\ &+ \pi_\alpha \pi_{C+1}^t R_{t,C+1}(h_b) + \delta, \end{aligned} \tag{7}$$

where $\delta = (1-\pi_\alpha)\cdot const$ denotes the constant since the loss function $\mathcal{L}(\cdot,\cdot)$ is bounded. Let h^* be the idea shared classifier which minimizes the combined error,

$$h^* = \arg\min_{\tilde{h}\in\mathcal{H}_1} R_s(h) + R_t^*(h). \tag{8}$$

Now we consider,

$$\begin{aligned} R_t(\tilde{h}, h_b) &\leq (1-\pi_\alpha)\big[R_t^*(h^*) + (1-\pi_{C+1}^t)\mathbb{E}_{\mathbb{Q}^t_{XY|Y\leq C}}\mathcal{L}(\tilde{h}, h^*)\big] \\ &+ \pi_\alpha\big[R_t^*(h_b) + \pi_{C+1}^t R_{t,C+1}(h_b)\big] + \delta, \\ &\leq (1-\pi_\alpha)\big[R_t^*(h^*) + (1-\pi_{C+1}^t)\big(\mathbb{E}_{\mathbb{P}^s_X}\mathcal{L}(\tilde{h}, h^*) \\ &+ disc(\mathbb{Q}^t_{X|Y\leq C}, \mathbb{P}^s_X)\big)\big] + \pi_\alpha\big[\pi_{C+1}^t R_{t,C+1}(h_b) \\ &+ R_t^*(h_b)\big] + \delta. \end{aligned} \tag{9}$$

Based on the triangle inequality we have,

$$\begin{aligned} \mathbb{P}^s_X\mathcal{L}(\tilde{h},h^*) &\leq \mathbb{P}^s_X\mathcal{L}(\tilde{h},i) + \mathbb{P}^s_X\mathcal{L}(h^*,i), \\ &\leq R_s(\tilde{h}) + R_s(h^*). \end{aligned} \tag{10}$$

Therefore, we can reformulate the original inequality in the following,

$$\begin{aligned} R_t(\tilde{h}, h_b) &\leq (1-\pi_\alpha)\big[R_t^*(h^*) + (1-\pi_{C+1}^t)\big(R_s(\tilde{h}) + R_s(h^*) \\ &+ disc(\mathbb{Q}^t_{X|Y\leq C}, \mathbb{P}^s_X)\big)\big] \\ &+ \pi_\alpha\big[\pi_{C+1}^t R_{t,C+1}(h_b) + R_t^*(h_b)\big] + \delta. \end{aligned} \tag{11}$$

As the index-based thresholds $\alpha_u^{(m)}, \alpha_k^{(m)} \rightarrow \alpha^*$, we have $R_t^*(h_b) \rightarrow R_t^*(h)$. The equation (11) can be transformed into,

$$\begin{aligned} \frac{R_t(\tilde{h}, h_b)}{1-\pi_{C+1}^t} &\leq (1-\pi_\alpha)(R_s(\tilde{h}) + disc(\mathbb{Q}^t_{X|Y\leq C}, \mathbb{P}^s_X)) + \lambda \\ &+ \frac{\pi_\alpha \pi_{C+1}^t}{1-\pi_{C+1}^t} R_{t,C+1}(h_b) + \delta, \end{aligned} \tag{12}$$

where $\lambda = \min_{\tilde{h}\in\mathcal{H}_1} (1-\pi_\alpha)R_s(\tilde{h}) + \frac{R_t^*(\tilde{h})}{1-\pi_{C+1}^t}$. ■