# Supplementary Material for "Dual Bipartite Graph Learning: A General Approach for Domain Adaptive Object Detection"

Chaoqi Chen[1], Jiongcheng Li[2], Zebiao Zheng[2], Yue Huang[2], Xinghao Ding[2], Yizhou Yu[1*]
[1]The University of Hong Kong [2]Xiamen University
cqchen1994@gmail.com, jiongchengli@stu.xmu.edu.cn, zbzheng@stu.xmu.edu.cn
huangyue05@gmail.com, dxh@xmu.edu.cn, yizhouy@acm.org

**Definition 1.** ***Source and Target Risks.*** *The source risk $R_s(h)$ and target risk $R_t(h)$ of $h$ w.r.t. $\mathcal{L}$ under source distribution $P$ and target distribution $Q$ are defined as:*

$$\begin{aligned} R_s(h) &\triangleq \mathbb{E}_{(x,y)\sim P}\mathcal{L}(h(x),y) = \sum_{i=1}^{K+1}\pi_i^s R_{s,i}(h) \\ R_t(h) &\triangleq \mathbb{E}_{(x,y)\sim Q}\mathcal{L}(h(x),y) = \sum_{j=1}^{K+1}\pi_j^t R_{t,j}(h) \end{aligned} \tag{1}$$

*where $\pi_i^s = P(y=i)$ and $\pi_j^t = Q(y=j)$ are class-prior probabilities of $P$ and $Q$. Then, we have*

$$\begin{aligned} R_s(h) &= \sum_{i=1}^{K}\pi_i^s R_{s,i}(h) + \pi_{K+1}^s R_{s,K+1}(h) = R_s^*(h) + \Delta_s \\ R_t(h) &= \sum_{j=1}^{K}\pi_j^t R_{t,j}(h) + \pi_{K+1}^t R_{t,K+1}(h) = R_t^*(h) + \Delta_t \end{aligned} \tag{2}$$

**Definition 2.** ***$\mathcal{H}\Delta\mathcal{H}$-Distance.*** *For any $h, h' \in \mathcal{H}$, the domain discrepancy between source and target distributions can be defined as:*

$$d_{\mathcal{H}\Delta\mathcal{H}}(P,Q) \triangleq \sup_{h,h'\in\mathcal{H}} |\mathbb{E}_P\mathcal{L}(h,h') - \mathbb{E}_Q\mathcal{L}(h,h')| \tag{3}$$

**Definition 3.** ***Ideal Joint Hypothesis.*** *Let $\mathcal{H}$ be the hypothesis class, the ideal joint hypothesis is the hypothesis which minimizes the combined error,*

$$h^* = \arg\min_{h\in\mathcal{H}} R_s^*(h) + R_t^*(h) \tag{4}$$

Given the hypothesis space $\mathcal{H}$ with a condition that constant function $K+1 \in \mathcal{H}$, for $\forall h \in \mathcal{H}$, the expected error on target samples $R_t(h)$ is bounded as,

$$\begin{aligned} \frac{R_t(h)}{1-\pi_{K+1}^t} \le\, & R_s^*(h) + d_{\mathcal{H}\Delta\mathcal{H}}(P_{X|Y\le K}, Q_{X|Y\le K}) + \lambda \\ & + \frac{\Delta_t}{1-\pi_{K+1}^t} \end{aligned} \tag{5}$$

where the shared error $\lambda = \min_{h\in\mathcal{H}} \frac{R_t^*(h)}{1-\pi_{K+1}^t} + R_s^*(h)$, $R_s^*(h) = \sum_{i=1}^{K}\pi_i^s R_{s,i}(h)$, and $\Delta_t = \pi_{K+1}^t R_{t,K+1}(h)$. We show the derivation of Inequality (5) as follows.

*Proof.* Given the above definitions, we have

$$\begin{aligned} R_t(h) &= R_t^*(h) + \Delta_t \\ &\le R_t^*(h^*) + (1-\pi_{K+1}^t)\mathbb{E}_{Q_{X|Y\le K}}\mathcal{L}(h,h^*) + \Delta_t \\ &\le R_t^*(h^*) + (1-\pi_{K+1}^t)(\mathbb{E}_{P_{X|Y\le K}}\mathcal{L}(h,h^*) \\ &+ d_{\mathcal{H}\Delta\mathcal{H}}(P_{X|Y\le K}, Q_{X|Y\le K})) + \Delta_t \end{aligned} \tag{6}$$

According to the triangle inequality, we have

$$\begin{aligned} \mathbb{E}_{P_{X|Y\le K}}\mathcal{L}(h,h^*) &\le \mathbb{E}_{P_{X|Y\le K}}\mathcal{L}(h,i) + \mathbb{E}_{P_{X|Y\le K}}\mathcal{L}(h^*,i) \\ &\le R_s^*(h) + R_s^*(h^*) \end{aligned}$$

Then, the Eq. (6) can be formulated as,

$$\begin{aligned} R_t(h) \le\, & R_t^*(h^*) + (1-\pi_{K+1}^t)(R_s^*(h) + R_s^*(h^*) \\ & + d_{\mathcal{H}\Delta\mathcal{H}}(P_{X|Y\le K}, Q_{X|Y\le K})) + \Delta_t \end{aligned}$$

Finally, we have

$$\begin{aligned} \frac{R_t(h)}{1-\pi_{K+1}^t} &\le \frac{R_t^*(h^*)}{1-\pi_{K+1}^t} + R_s^*(h) + R_s^*(h^*) \\ &+ d_{\mathcal{H}\Delta\mathcal{H}}(P_{X|Y\le K}, Q_{X|Y\le K}) + \frac{\Delta_t}{1-\pi_{K+1}^t} \\ &= R_s^*(h) + d_{\mathcal{H}\Delta\mathcal{H}}(P_{X|Y\le K}, Q_{X|Y\le K}) + \lambda + \frac{\Delta_t}{1-\pi_{K+1}^t} \end{aligned}$$

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

*Corresponding author