Appendix for "Generative Inference Network for Imbalanced Domain Generalization"

I. Experimental Setting

Normal Domain Generalization: We randomly select one of four domains as target domain from the original PACS [1] or VLCS [2] while the remaining three as multiple source domains without any changes. The specific statistics of these two datasets are summarized in Table I and Table II, respectively. With respect to them, these benchmarks both contain somewhat imbalanced data scale across various domains and categories. For example, there are more instances in "**S**ketch" domain than that of "**P**hoto" domain, and the sample size of "Dog" is larger than that of "House". In addition, the original Office-Home includes four domains, i.e., Realworld (**Rw**), Clipart (**Cl**), Art (**Ar**), Product (**Pr**) with each domain from 65 categories. The specific sample size for each domain is **Ar** (2,427), **Cl** (4,365), **Pr** (4,439) and **Rw** (4,357), respectively.

Imbalanced Domain Generalization: To further clearly illustrate the effect of imbalanced data scale on learning high-generalization classification model, we develop a more practical and challenging scenario named Imbalanced Domain Generalization (IDG) with the corresponding experimental settings. Specifically, three of four domains from PACS or VLCS are selected as multiple source domains, where we randomly remove instances from several domains or categories. And there is no additional operation on target domain. The statistics of multi-source domains (PACS and VLCS) under IDG scenario are shown in Figures 1 and 2. For Office-Home, we adopt the similar manner to design more significant difference of data scale from domain-level and category-level. The specific statistics of each domain is shown in Figure 3.

TABLE I: The specific statistics of the original PACS benchmark.

Class \ Domain	**A**rt	**C**artoon	**P**hoto	**S**ketch	Total
Dog	379	389	189	772	1729
Elephant	255	457	202	740	1654
Giraffe	285	346	182	753	1566
Guitar	184	135	186	608	1113
Horse	201	324	199	816	1540
House	295	288	280	80	943
Person	449	405	432	160	1446
Total	2048	2344	1670	3929	9991

TABLE II: The specific statistics of VLCS benchmark where each domain is used as source domain.

Class \ Domain	**C**altech	**L**abelme	Pascal **V**oc	**S**un	Total
Bird	166	56	231	14	467
Car	86	846	489	652	2073
Chair	83	62	300	725	1170
Dog	47	29	294	21	391
Person	609	866	1049	885	3409
Total	991	1859	2363	2297	7510

II. Theoretical Analysis

Lemma. In probability theory, suppose random variable x comes from the sample space $\Omega = \{1, 2, \cdots, n\}$ with the corresponding probabilities $P_1 \leq P_2 \leq \cdots \leq P_n$, where $P_i = P(x = i) \geq 0, \sum_{i=1}^{n} P_i = 1$. Under this condition, we have the conclusion:

$$1 - P_n \leq 2(\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1} P_i P_j + \sum_{i=1}^{n-1} P_i P_n). \tag{1}$$

Proof.

$$\begin{aligned} 1 - P_n &= \sum_{i=1}^{n-1} P_i = \sum_{i=1}^{n-1}\sum_{j=1}^{n} P_i P_j \\ &= 2\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1} P_i P_j + \sum_{i=1}^{n-1} P_i P_n + \sum_{i=1}^{n-1} P_i^2 \\ &\leq 2(\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1} P_i P_j + \sum_{i=1}^{n-1} P_i P_n). \end{aligned} \tag{2}$$

Theorem 1. Given the prior probabilities of multiple source domains $\{P_s^1, P_s^2, \cdots, P_s^K\}$ and the corresponding label probabilities within the k-th source domain $\{Q_k^1, Q_k^2, \cdots, Q_k^C\}$, and the probability densities of the latent variable $\{q_z^1, q_z^2, \cdots, q_z^C\}$ where $q_z^c(\mathbf{z}) = q(\mathbf{z}|\mathbf{y} = c)$, the error bound of the generative annotation is formulated as the following with the generalization error ϵ:

$$\begin{aligned} &|E(\mathbf{y}) - E(\hat{\mathbf{y}})| \\ &= 1 - \int \max\Big\{\sum_{k=1}^{K} P_s^k Q_k^1 q_z^1, \cdots, \sum_{k=1}^{K} P_s^k Q_k^C q_z^C\Big\} d\mathbf{z} \leq \epsilon, \end{aligned} \tag{3}$$

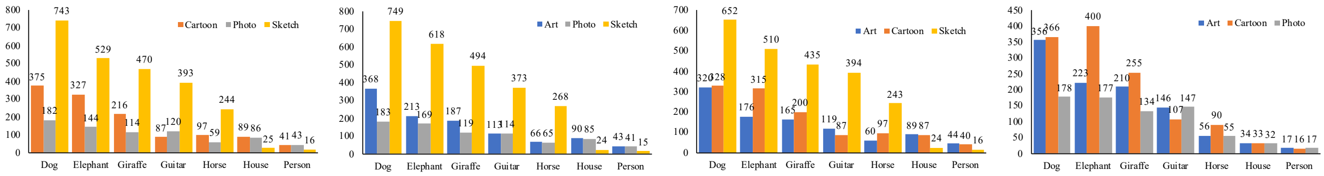
where $\hat{\mathbf{y}}$ is the generative annotation from classifier.

Proof. With respect the above definitions, we can easily obtain the marginal distribution of the variable $\mathbf{z}$ as:

$$\begin{aligned} &\widetilde{q}_z(\mathbf{z}) = \sum_{c=1}^{C}\sum_{k=1}^{K} P_s^k Q_k^c q_z^c, \\ &\mathbb{E}_z[P_y^u P_y^v] = \int \frac{(\sum_{k=1}^{K} P_s^k Q_k^u q_z^u(\mathbf{z}))(\sum_{k=1}^{K} P_s^k Q_k^v q_z^v(\mathbf{z}))}{\widetilde{q}_z(\mathbf{z})} d\mathbf{z}, \end{aligned} \tag{4}$$

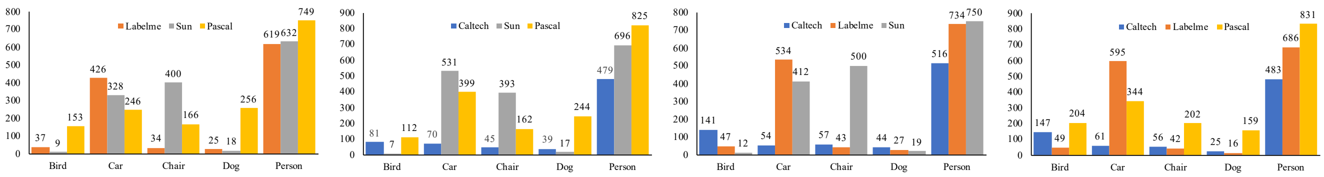
where $P_y^c = P(\mathbf{y} = c|\mathbf{z})$ is the probability conditioned on $\mathbf{z}$. From another perspective, we can formulate Eq. (4) as:

$$\begin{aligned} \mathbb{E}_z[P_y^u P_y^v] &= 2\sum_{u=1}^{C-1}\sum_{v=u+1}^{C} P_y^u P_y^v \\ &= 2\sum_{u=1}^{C-2}\sum_{v=u+1}^{C-1} P_y^u P_y^v + 2\sum_{u=1}^{C-1} P_y^u P_y^C. \end{aligned} \tag{5}$$



(a) Target: Art-painting (b) Target: Cartoon (c) Target: Photo (d) Target: Sketch

Fig. 1: Experimental Settings corresponding to IDG scenario over PACS benchmark.



(a) Target: Caltech (b) Target: LabelMe (c) Target: Pascal-Voc (d) Target: Sun

Fig. 2: Experimental Setting corresponding to IDG scenario over VLCS benchmark.

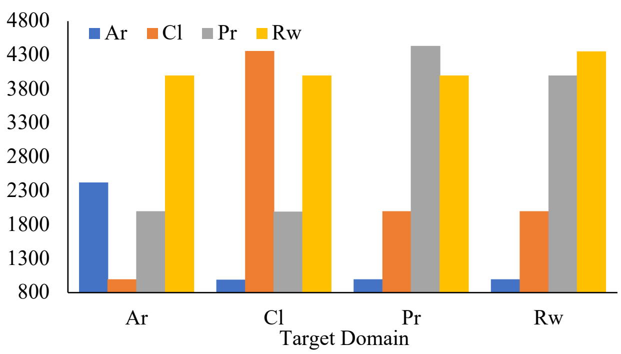


Fig. 3: IDG setting on Office-Home benchmark.

In addition, the error defined in **Theorem 1** is also rewritten as follows:

$$\begin{aligned} &|E(\mathbf{y}) - E(\hat{\mathbf{y}})| \\ &= 1 - \int \max\left\{\sum_{k=1}^{K} P_s^k Q_k^1 q_z^1, \cdots, \sum_{k=1}^{K} P_s^k Q_k^C q_z^C\right\} d\mathbf{z} \qquad (6) \\ &= \mathbb{E}_z[1 - \max_{c=1,2,\cdots,C} P(\mathbf{y} = c|\mathbf{z})]. \end{aligned}$$

And we sort $\{P_y^1, P_y^2, \cdots, P_y^C\}$ with ascending order, i.e. $P_y^1 \le P_y^2 \le \cdots \le P_y^C$. According to the aforementioned lemma, we achieve the following conclusion:

$$\mathbb{E}_z[1 - P(\mathbf{y} = C|\mathbf{z})] \le \mathbb{E}_z[P_y^u P_y^v]. \qquad (7)$$

When considering the error bound ϵ as $\mathbb{E}_z[P_y^u P_y^v]$, the final conclusion of theorem holds.

References

[1] D. Li, Y. Yang, Y.-Z. Song, and T. M. Hospedales, "Deeper, broader and artier domain generalization," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5542–5550.

[2] C. Fang, Y. Xu, and D. N. Rockmore, "Unbiased metric learning: On the utilization of multiple datasets and web images for softening bias," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1657–1664.