英語論文#6

2023/12/11

M1

建元　了

論文の概要

- タイトル
  - **Cross Contrasting Feature Perturbation for Domain Generalization**
- 執筆者
  - Chenming Li, Daoan Zhang, Wenjian Huang, Jianguo Zhang
- 掲載
  - ICCV2023
- 選択理由
  - データの取得条件にずれがある学習に関する最新の論文

背景

- Out-of-distribution(OOD)問題
  - 通常の教師あり学習では、訓練データとテストデータに属するサンプルが独立同分布（IID）である
  - テストデータが訓練データと同じ確率分布に従う仮定となり、パフォーマンスの低下

  - 解決には実際と則した多様な訓練データで学習したモデルが必要

導入

- データ摂動ベースではGAN、VAEを用いる手法が採用されるが、未知のドメインを生成できないため適していない

- データ摂動ベースでは多様性を制限し、意味的一貫性を維持できない

- 潜在空間でドメインを意識した適応特徴の摂動を強制し、クラス情報を保存して意味的一貫性を制約する方法を提案

導入

- ３つの貢献
  - 最悪領域での汎化問題に対する新しい1段階オンライン学習である cross contrasting feature perturbation framework(CCFP)の提案
  - モジュールとドメインの不一致を測定するドメイン認識 Grammatrices-based metric であるlearnable domain perturbation(LDP)の開発
  - 生成モデル（GANなど）を使わず、多様なDGで最新のパフォーマンスの実現

ドメイン適応(domain Adaptation)

- **ドメイン適応（domain adaptation）**
  - 学習時の訓練データ（**Source domain**）と推論時のデータ(**Target domain**)で**ドメイン**（取得条件）が異なる場合、一般的に悪くなる
  - このドメインが変化する場合でも推論をうまく行えることを目指す

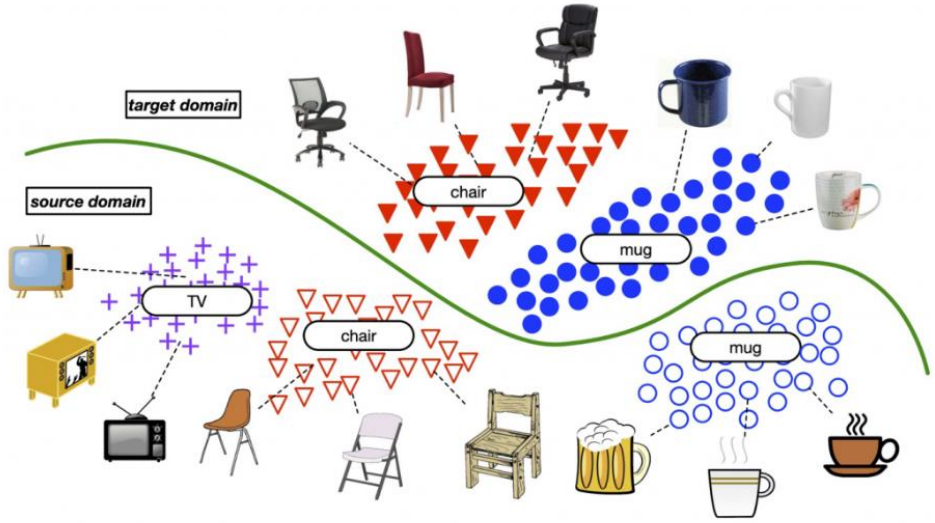

図 4　ドメイン適応のシナリオの概要（[Cao 18] より引用）。

ドメイン一般化（DG）

- 目に見えないターゲットドメインをうまく一般化できる複数のソースドメインからの表現に対応する

- 方法
  - ドメイン不変表現の学習
  - メタ学習
  - データ摂動ベース

データの摂動

- 入力空間におけるデータの摂動は、間違った相関を軽減してモデルの一般化を改善するための多様な画像を生成

例

- ２つの特徴量のインスタンスを線形補間して合成したモデルの一般化を改善
- Manifold Mixupは、画像レベルから特徴レベルまで線形補間

提案手法

- DGは以下の定式を解決することになる

$$\min_{\theta \in \Theta} \mathbb{E}_{(x,y)\sim P_{tar}} [\ell(\theta; (x, y))] \qquad (1)$$

- xは入力特徴、yは予測ラベル、 θ はモデル
- Eは期待値、lは損失関数
- Ptarはターゲットドメインの確率分布

提案手法

- DGにとっての課題はターゲットドメイン分布Pが利用できない

$$\hat{\theta}_{ERM} := \min_{\theta \in \Theta} \mathbb{E}_{(x,y)\sim P_{src}}[\ell(\theta; (x, y))], \qquad (2)$$

- ソースドメインからすべてのデータをマージして、ロスを最小限に抑える経験的リスク最小化(ERM)からアプローチ

- Psrcは訓練データ全体の経験的分布

提案手法

- ERMベースの手法はOODに対するロバスト性に欠けている
- 最悪の場合でのDGを定式化

$$\hat{\theta}_{worst-case} := \min_{\theta \in \Theta} \sup_{P: D(P, P_{src}) \leq \rho} \mathbb{E}_P[\ell(\theta; (x, y))] \quad (3)$$

- Dは確率分布空間上の距離
- 架空のターゲット分布Pから領域から距離p離れている場合、領域シフトに対して良好なパフォーマンスを得る

提案手法

- 生成モデルを使うオフラインの2段階のトレーニング手順ではコストが高い

- オンラインの1段階フレームワークであるCross Contrasting Feature Perturbation Framework（CCFP）を提案

CCFP

- 概要図

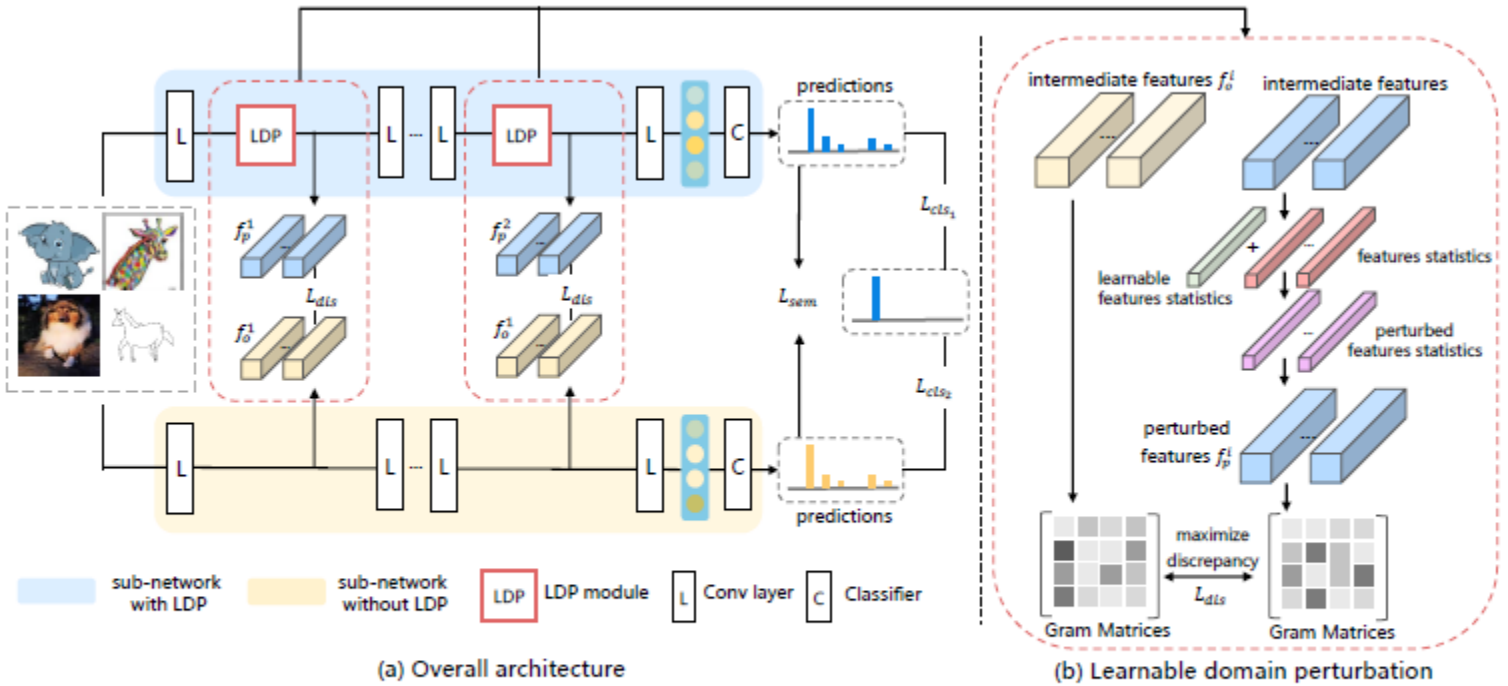

Figure 1. **An overview of our proposed CCFP.** Our framework consists of two sub-networks. (a) The bottom network is a pre-trained backbone, and the top network is the same pre-trained backbone equipped with LDP modules (red boxes). The two sub-networks have similar architecture (except for the LDP modules) but do not share parameters. The steps of feature perturbation and the calculation of L_{dis} are shown in (b).

CCFP

- 学習可能な摂動領域を利用して、架空のターゲット分布を表す摂動特徴を計算

$$\hat{\theta} := \min_{\theta \in \Theta} \sup_{p^l} \{\mathbb{E}_{p^l}[\ell(\theta; (x, y))] - \gamma D(P^l, P^l_{src})\} \quad (4)$$

- Psrcは潜在的なソースドメイン分布，　Pはソースドメインの潜在的な空間

LDP

- AdaIN
  - 入力特徴xの特徴量をスタイルイメージの特徴量に置き換える

$$AdaIN(x) = \sigma(x_s)\frac{x - \mu(x)}{\sigma(x)} + \mu(x_s) \qquad (5)$$

LDP

- LDP
  - 線形補間や不確実性モデリングではドメインの転送が制限
  - 学習可能なドメイン摂動（LDP）を設計

$$LDP(x) = (\sigma(x) + \gamma)\frac{x - \mu(x)}{\sigma(x)} + \mu(x) + \beta \qquad (6)$$

Gram-based Domain Disscrepany Metric

- グラム行列ベースの指標

$$c(x) = g \circ f^n \circ f^{n-1} \circ \cdots \circ f^1(x) \qquad (7)$$

- ドメインの不一致による損失

$$\mathcal{L}_{dis} = -\sum_{i=1}^{K} \|G(f_o^i(\mathbf{x})) - G(f_p^i(\mathbf{x}))\|_F \qquad (8)$$

意味的一貫性制約

- 最終的な分類子の予測間のL2ノルムを最小化

$$L_{sem} = ||g_o(f_o(\mathbf{x})) - g_p(f_p(\mathbf{x}))||_2^2 \qquad (9)$$

- 最終的な損失

$$\mathcal{L}_{final} = \mathcal{L}_{cls_1} + \mathcal{L}_{cls_2} + \lambda_{dis}\mathcal{L}_{dis} + \lambda_{sem}\mathcal{L}_{sem} \qquad (10)$$

提案手法（アルゴリズム）

Algorithm 1 : Cross Contrasting Feature Perturbation

Input: $\mathcal{S}_{train} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}$, batch size B, learning rate η, Adam optimizer, initial $\lambda_{dis}, \lambda_{sem}$

Initial: Parameters of CCFP *i.e.* parameters θ_0, θ_1, ϕ_0, ϕ_1, $(\gamma_k, \beta_k; k = 1 \cdots K)$ for feature extractor f_o, f_p, classifier g_o, g_p and LDP modules $P^1, P^2 \cdots P^K$ (K is defined in Eq.8).

repeat

Minimization Stage:

for $i = 1, \cdots, B$ **do**

$\mathcal{L}^i_{cls_1} = \ell(g_o(f_o(\mathbf{x}_i)), y_i)$

$\mathcal{L}^i_{cls_2} = \ell(g_p(f_p(\mathbf{x}_i)), y_i)$

$\mathcal{L}^i_{sem} = \lambda_{sem} ||f_o(\mathbf{x}_i) - f_p(\mathbf{x}_i)||_2^2$

end for

$\theta_0, \phi_0 \leftarrow \text{Adam}(\frac{1}{B}\sum_{i=1}^{B} \mathcal{L}^i_{cls_1} + \mathcal{L}^i_{sem}, \theta_0, \phi_0, \eta)$

$\theta_1, \phi_1, \gamma_k, \beta_k \leftarrow \text{Adam}(\frac{1}{B}\sum_{i=1}^{B} \mathcal{L}^i_{cls_2} + \mathcal{L}^i_{sem}, \theta_1, \phi_1, \gamma_k, \beta_k, \eta)$

Maximization Stage:

for $i = 1, \cdots, B$ **do**

$\mathcal{L}^i_{dis} = \lambda_{dis} \sum_{k=1}^{K} ||G(f_o^k(\mathbf{x}_i)) - G(f_p^k(P^k(\mathbf{x}_i)))||_F$

end for

$\gamma_k, \beta_k \leftarrow \text{Adam}(\frac{1}{B}\sum_{i=1}^{B} \mathcal{L}^j_{spe}, \gamma_k, \beta_k, \eta)$

until $\theta_0, \theta_1, \phi_0, \phi_1$ are converged

評価

- DomainBedベンチマークの7つのマルチドメイン画像分類タスク
  - Colored MNIST
  - Rotated MNIST
  - PACS
  - VLCS
  - Office-Home
  - Terra Incognita
  - DomainNet

評価

- データセットはトレーニングサブセットと検証サブセットで8:2
-

1. Training-domain validation set
2. Leave-one-out cross-validation
3. Test-domain model set (oracle)

評価

- ドメインの相関関係はd=90%, 80%, 10%

Algorithm	CMNIST	RMNIST	VLCS	PACS	OfficeHome	TerraInc	DomainNet	Avg
ERM[52]	51.5 ± 0.1	98.0 ± 0.0	77.5 ± 0.4	85.5 ± 0.2	66.5 ± 0.3	46.1 ± 1.8	40.9 ± 0.1	66.6
IRM[1]	52.0 ± 0.1	97.7 ± 0.1	78.5 ± 0.5	83.5 ± 0.8	64.3 ± 2.2	47.6 ± 0.8	33.9 ± 2.8	65.4
GroupDRO[46]	52.1 ± 0.0	98.0 ± 0.0	76.7 ± 0.6	84.4 ± 0.8	66.0 ± 0.7	43.2 ± 1.1	33.3 ± 0.2	64.8
Mixup[62]	52.1 ± 0.2	98.0 ± 0.1	77.4 ± 0.6	84.6 ± 0.6	68.1 ± 0.3	47.9 ± 0.8	39.2 ± 0.1	66.7
MLDG[29]	51.5 ± 0.1	97.9 ± 0.0	77.2 ± 0.4	84.9 ± 1.0	66.8 ± 0.6	47.7 ± 0.9	41.2 ± 0.1	66.7
CORAL[50]	51.5 ± 0.1	98.0 ± 0.1	78.8 ± 0.6	86.2 ± 0.3	68.7 ± 0.3	47.6 ± 1.0	41.5 ± 0.1	67.5
MMD[33]	51.5 ± 0.2	97.9 ± 0.0	77.5 ± 0.9	84.6 ± 0.5	66.3 ± 0.1	42.2 ± 1.6	23.4 ± 9.5	63.3
DANN[14]	51.5 ± 0.2	97.8 ± 0.1	78.6 ± 0.4	83.6 ± 0.4	65.9 ± 0.6	46.7 ± 0.5	38.3 ± 0.1	66.1
CDANN[33]	51.7 ± 0.1	97.9 ± 0.1	77.5 ± 0.1	82.6 ± 0.9	65.8 ± 1.3	45.8 ± 1.6	38.3 ± 0.3	65.6
MTL[6]	51.4 ± 0.1	97.9 ± 0.0	77.2 ± 0.4	84.6 ± 0.5	66.4 ± 0.5	45.6 ± 1.2	40.6 ± 0.1	66.2
SagNet[40]	51.7 ± 0.0	98.0 ± 0.0	77.8 ± 0.5	86.3 ± 0.2	68.1 ± 0.1	**48.6** ± 1.0	40.3 ± 0.1	67.2
ARM[67]	**56.2** ± 0.2	**98.2** ± 0.1	77.6 ± 0.3	85.1 ± 0.4	64.8 ± 0.3	45.5 ± 0.3	35.5 ± 0.2	66.1
V-REx[26]	51.8 ± 0.1	97.9 ± 0.1	78.3 ± 0.2	84.9 ± 0.6	66.4 ± 0.6	46.4 ± 0.6	33.6 ± 2.9	65.6
RSC[23]	51.7 ± 0.2	97.6 ± 0.1	77.1 ± 0.5	85.2 ± 0.9	65.5 ± 0.9	46.6 ± 1.0	38.9 ± 0.5	66.1
AND-mask[24]	51.3 ± 0.2	97.6 ± 0.1	78.1 ± 0.9	84.4 ± 0.9	65.6 ± 0.4	44.6 ± 0.3	37.2 ± 0.6	65.5
SAND-mask[24]	51.8 ± 0.2	97.4 ± 0.1	77.4 ± 0.2	84.6 ± 0.9	65.8 ± 0.4	42.9 ± 1.7	32.1 ± 0.6	64.6
Fish[48]	51.6 ± 0.1	98.0 ± 0.0	77.8 ± 0.3	85.5 ± 0.3	68.6 ± 0.4	45.1 ± 1.3	**42.7** ± 0.2	67.1
Fishr[44]	52.0 ± 0.2	97.8 ± 0.0	77.8 ± 0.1	85.5 ± 0.4	67.8 ± 0.1	47.4 ± 1.6	41.7 ± 0.0	67.1
CCFP (ours)	51.9 ± 0.1	97.8 ± 0.1	**78.9** ± 0.3	**86.6** ± 0.2	**68.9** ± 0.1	**48.6** ± 0.4	41.2 ± 0.0	**67.7**

Table 1. DomainBed with Training-domain model selection. We highlighted the best results using **bold** font.

評価

- ハイパーパラメータ（HP）検索
  - 20回の試行でDomainBedからハイパーパラメータ分布をランダム検索
  - CCFPは２つのハイパーパラメータに依存しており、0.1~10の範囲で探す
  - ImageNetで事前学習し、ResNet-50でFT

結果

- ドメイン汎化手法との比較

Algorithm	A	C	P	R	Avg.
SWAD[7]	66.1	57.7	78.4	80.2	70.6
PCL[63]	67.3	**59.9**	78.7	80.7	71.6
CCFP (ours)	**68.0**	58.6	**79.7**	**81.9**	**72.1**

Table 2. Comparison with SWAD-based state-of-the-art methods on OfficeHome benchmark. A: art, C: clipart, P: product, R: real, Avg.: average.

Algorithm	C	L	S	V	Avg.
SWAD[7]	98.8	63.3	**75.3**	79.2	79.1
PCL[63]	**99.0**	63.6	73.8	75.6	78.0
CCFP (ours)	98.9	**64.1**	74.9	**79.9**	**79.4**

Table 3. Comparison with SWAD-based state-of-the-art methods on VLCS benchmark. C: Caltech101, L: LabelMe, S: SUN09, V: VOC2007, Avg.: average.

Algorithm	L100	L38	L43	L46	Avg.
SWAD[7]	55.4	44.9	59.7	39.9	50.0
PCL[63]	58.7	46.3	60.0	43.6	52.1
CCFP (ours)	**59.9**	**47.6**	**60.8**	**43.8**	**53.0**

Table 4. Comparison with SWAD-based state-of-the-art methods on TerraIncognita benchmark. L100: Location 100, L38: Location 38, L43: Location 43, L46: Location 46, Avg.: average.

結果

- 従来の特徴摂動手法との比較

Algorithm	A	C	P	S	Avg.
ERM	81.6	78.7	95.5	78.7	83.6
Mixstyle[70]	84.0	79.9	94.3	**81.6**	84.9
DSU[32]	81.9	79.6	95.0	79.6	84.1
CCFP (ours)	**87.5**	**81.3**	**96.4**	81.4	**86.6**

Table 5. Comparison with previous feature perturbation methods on PACS benchmark. Comparison with SWAD-based state-of-the-art methods on PACS benchmark.

アブレーション研究

- 式10の損失を使わずに実行

Algorithm	A	C	P	S	Avg.
ERM	81.6	78.7	95.5	78.7	83.6
CCFP(w/o) L_{sem}	83.6	**83.9**	**96.4**	80.3	86.0
CCFP (ours)	**87.5**	81.3	**96.4**	**81.4**	**86.6**

Table 6. Comparison with result without using L_{sem} on PACS benchmark.

アブレーション研究

- 式10のLdisを削除

Algorithm	A	C	P	S	Avg.
Mixstyle[70]	84.0	79.9	94.3	81.6	84.9
Mixstyle (dual)	84.6	80.3	**96.5**	79.5	85.2
DSU[32]	81.9	79.6	95.0	79.6	84.1
DSU (dual)	86.3	79.4	94.6	**81.7**	85.5
CCFP (ours)	**87.5**	**81.3**	96.4	81.4	**86.6**

Table 7. Validation of the additional semantic consistency for previous feature perturbation methods on PACS benchmark.

アブレーション研究

- 中間層の特徴量の変化

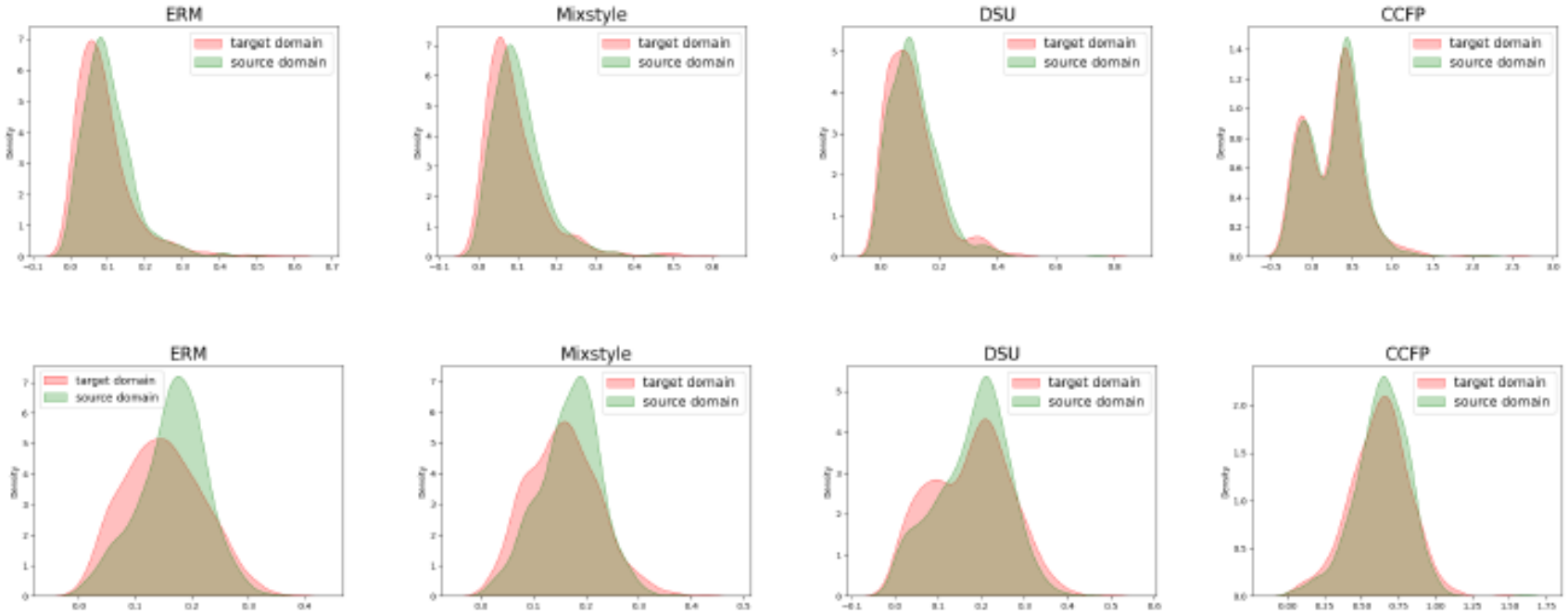

Figure 2. The visualization of feature statistics at the position 3. The top raw is the mean statistics and the bottom raw is the std statistics. We conduct the experiments on the PACS dataset with ERM, Mixstyle, DSU and our CCFP.

アブレーション研究

- LDPの挿入位置による影響

Positions	1-3	2-4	3-5	1-5	ERM
PACS	85.3	84.8	85.4	**86.6**	83.6
OfficeHome	68.4	68.5	68.3	**68.9**	64.5

Table 8. Effects of different inserted positions on PACS and OfficeHome benchmark.