# **Towards Robust In-Domain and Out-of-Domain Generalization: Contrastive** Learning with Prototype Alignment and Collaborative Attention

Yuan-Jhe Kuo, Cheng-Yu Yang, and Chiou-Ting Hsu National Tsing Hua University, Taiwan

# **Robust Domain Generalization**

# **Proposed Framework**

### Goal

- Robust image classification model
  - Good out-of-domain performance ullet
  - Good in performance
  - Robust to noisy source labels

## Idea of domain generalization

- Reducing class ambiguities
- Preserving in-domain class distributions 2.
- Increasing robustness by collaborative learning





#### $C_1$ $C_2$ $D_1$ $D_2$ Proposed Method **Margin-Based Contrastive Learning** • **Prototype Alignment** Push away highly overlapped classes Maintain intra-class features Margin-Based Contrastive Learning Prototype Alignment $\mathbf{\star}$ 🗡 η $D_2$ $D_1$ $C_1$ $C_2$ Lalign L<sub>margin</sub> $max(\eta - |\theta(f(x),y) - \theta(f(x),k)|, 0)$ $\left\| Proj(g(x)) - m_c(t) \right\|_2$ $(\overline{x,y})$ $\overline{k \in K_{y}}$ $\overline{c=1} \forall (x,y), y=c$

**Collaborative Attention Combine positive and negative learning** 

**Positive feature**:  $z^+ = z \odot dil(M^+)$ 

**Negative feature**: 
$$z^{-} = z \odot (\frac{1}{2}M^{-} + \frac{1}{2}(1 - M^{+}))$$

**Cross-entropy:** 

$$L_{pos} = L_{ce}(h(z^+); y^+)$$

 $L_{neg} = L_{ce}(h(A(z^{-})); y^{-})$ 



$$L_{all} = L_{main} + \lambda_1 L_{margin} + \lambda_2 L_{align} + \lambda_3 L_{collab}$$

Method	PACS	VLCS	Avg.	Method	PACS	VLCS	Avg.
ERM [21]	$97.75 \pm 0.41$	$87.21\pm0.72$	92.48	ERM [21]	$87.56 \pm 0.33$	$78.13\pm0.16$	82.85
CORAL [20]	$97.64 \pm 0.33$	$86.88 \pm 0.87$	92.26	CORAL [20]	$87.40 \pm 0.19$	$78.20\pm0.26$	82.80
RSC [12]	$97.01 \pm 0.58$	$86.48 \pm 0.54$	91.75	RSC [12]	$84.04 \pm 0.67$	$77.99 \pm 0.27$	81.02
SagNet [19]	$97.53 \pm 0.40$	$86.86 \pm 0.83$	92.20	SagNet [19]	$86.52 \pm 0.63$	$77.82 \pm 0.22$	82.17
Mixup [23]	$97.92 \pm 0.54$	$86.89 \pm 0.93$	92.41	Mixup [23]	$85.90 \pm 0.08$	$78.61\pm0.14$	82.26
Mixstyle [26]	$97.31 \pm 0.68$	$86.89 \pm 0.92$	92.10	Mixstyle [26]	$86.15 \pm 0.41$	$78.00\pm0.39$	82.08
ARM [25]	$97.86 \pm 0.45$	$87.08 \pm 0.95$	92.10	ARM [25]	$87.31 \pm 0.16$	$78.10\pm0.25$	72.71
	$07.00 \pm 0.43$	$07.00 \pm 0.95$	02.02	SAM [7]	$86.28 \pm 0.37$	$78.19\pm0.26$	82.23
SAM [7]	$97.84 \pm 0.27$	$86.20 \pm 0.55$	92.02	EoA [1]	87.55	78.86	83.21
MIRO [4]	$97.74 \pm 0.11$	$87.57 \pm 1.00$	92.66	MIRO [4]	$87.57 \pm 0.21$	$79.08\pm0.35$	83.33
Ours	$97.80 \pm 0.54$	$87.43 \pm 0.82$	92.62	Ours	$\textbf{87.68} \pm \textbf{0.48}$	$\textbf{79.27} \pm \textbf{0.26}$	83.48

### t-SNE visualization

### In-domain Testing (w/ SWAD)

### **Out-of-domain Testing (w/ SWAD)**

Components				ResNet-50		
Lmain	L <sub>collab</sub>	Lmargin	Lalign	ID	OOD	
$\checkmark$				$97.89 \pm 0.18$	$86.90 \pm 0.30$	
<ul> <li>✓</li> </ul>	$\checkmark$			$97.36 \pm 0.86$	86.97 ± 0.79	
✓		$\checkmark$		$97.60 \pm 0.17$	$87.13 \pm 0.36$	
<ul> <li>✓</li> </ul>			$\checkmark$	97.57 $\pm$ 0.24	$86.84 \pm 0.14$	
<ul> <li>✓</li> </ul>		$\checkmark$	$\checkmark$	$97.68 \pm 0.32$	87.45 ± 0.12	
✓	$\checkmark$	$\checkmark$		$97.56 \pm 0.33$	87.66 ± 0.37	
<ul> <li>✓</li> </ul>	$\checkmark$	$\checkmark$	$\checkmark$	$97.80 \pm 0.54$	$87.68 \pm 0.48$	

	symm inc							
Method		In-Domain		Out-of-Domain				
	0	0.2	0.4	0	0.2	0.4		
ERM [21]	$97.75 \pm 0.41$	$97.35\pm0.50$	$96.15\pm0.51$	$87.56 \pm 0.33$	$\textbf{86.34} \pm \textbf{0.19}$	$84.72\pm0.81$		
RSC [12]	$97.01 \pm 0.58$	$96.40 \pm 0.82$	$95.11\pm0.76$	$84.04 \pm 0.67$	$82.91 \pm 0.92$	$78.62 \pm 1.04$		
Mixup [23]	$97.92 \pm 0.54$	$97.23 \pm 0.46$	$96.37\pm0.44$	$85.90 \pm 0.08$	$85.36 \pm 0.43$	$84.13 \pm 0.52$		
SagNet [19]	$97.53 \pm 0.40$	$97.05\pm0.66$	$96.40\pm0.82$	$86.52 \pm 0.63$	$85.50\pm0.31$	$83.60 \pm 0.51$		
CutMix [24]	$97.77 \pm 0.16$	$97.32\pm0.20$	$96.15\pm0.59$	$85.31 \pm 0.26$	$84.58 \pm 0.63$	$82.50\pm0.53$		
SAM [7]	$97.84 \pm 0.27$	$97.17 \pm 0.38$	$96.16\pm0.55$	$86.28 \pm 0.37$	$85.65 \pm 0.40$	$83.66 \pm 0.20$		
Ours	$97.80 \pm 0.54$	$97.35 \pm 0.38$	$\textbf{96.58} \pm \textbf{0.48}$	$87.68 \pm 0.48$	$86.22 \pm 0.57$	$\textbf{84.92} \pm \textbf{0.70}$		

#### **Robustness against noisy labels (w/ SWAD)**

Ablation study (w/ SWAD)