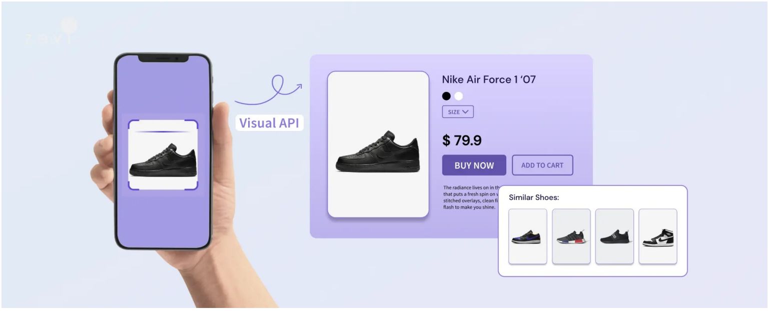
Semantic Feature Learning for Universal Unsupervised Cross-domain Retrieval

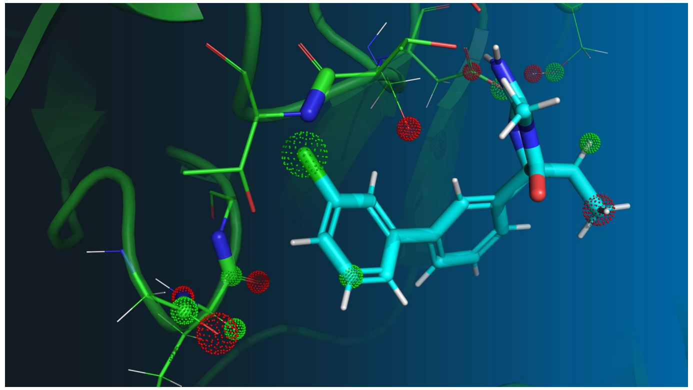
Lixu Wang[1], Xinyu Du[2], Qi Zhu[1]

1. Northwestern University, 2. General Motors Global R&D

NeurIPS 2024

Background





Challenges

- Images are unlabeled
- The semantic categories across image databases are uncertain

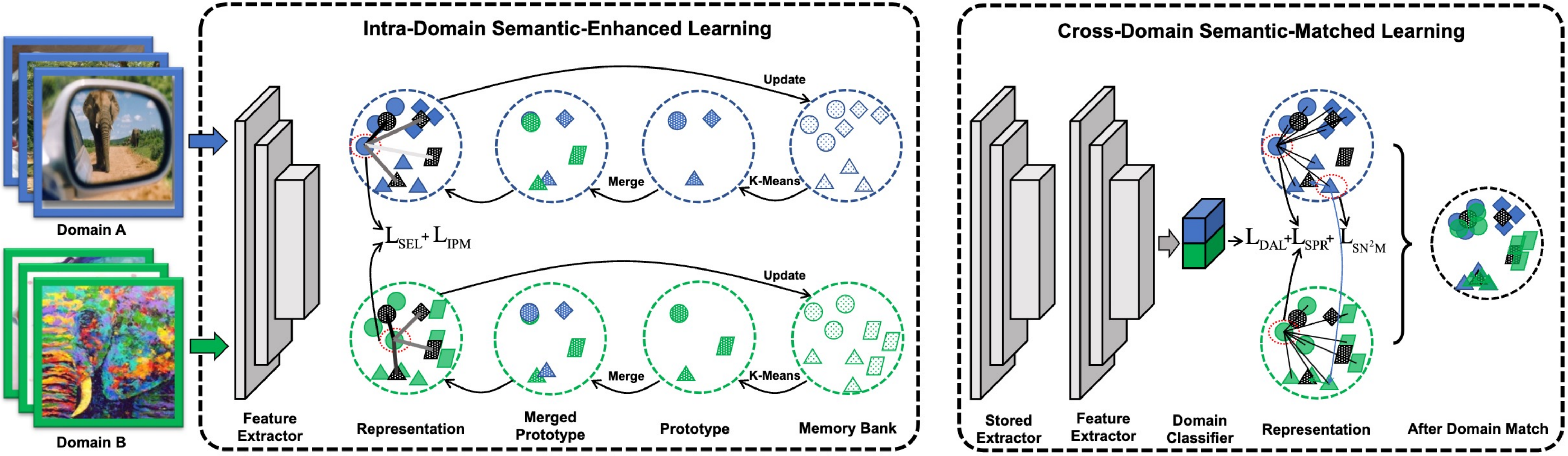
Problem Formulation

- Domains A and B are distinct but semantically relevant, and their label space may be different.

- Retrieval process

$$R(f_\theta, \boldsymbol{x}_i^{\mathrm{A}}) = \begin{cases} \text{null, if } y_i^{\mathrm{A}} \in \mathcal{Y}^{\mathrm{A}} \setminus \mathcal{Y}^{\mathrm{B}} \\ \text{sort}_{\uparrow}(\mathcal{S})[1:k], \text{ otherwise} \end{cases}$$

UEM



Domain A
Domain B
Intra-Domain Semantic-Enhanced Learning
Update
Merge
K-Means
$L_{SEL}+L_{IPM}$
Update
Merge
K-Means
Feature Extractor
Representation
Merged Prototype
Prototype
Memory Bank
Cross-Domain Semantic-Matched Learning
$L_{DAL}+L_{SPR}+L_{SN^2M}$
Stored Extractor
Feature Extractor
Domain Classifier
Representation
After Domain Match

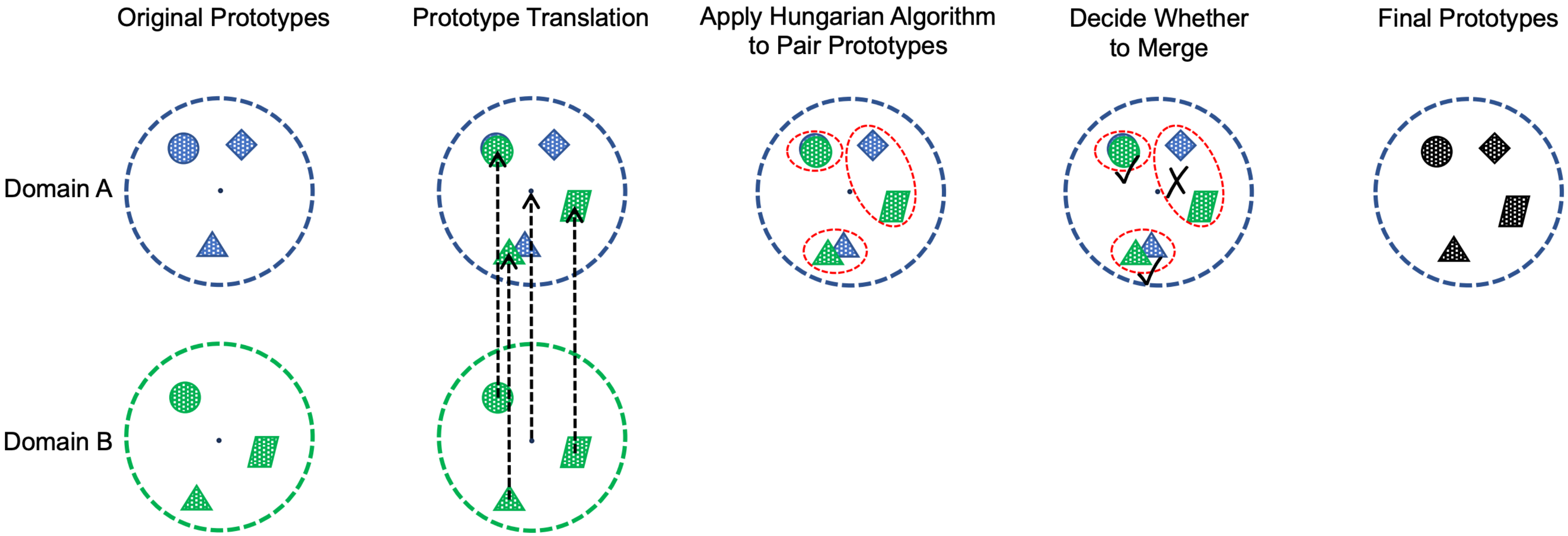
Intra-Domain Semantic-Enhanced Learning

- Challenges of applying instance discrimination (ID) to learn categorical semantics in universal unsupervised cross-domain retrieval
  - Distinct label spaces make ID learn distinct categorical structure

Theorem 1 (Geometry Distinctness): *Suppose data distributions of two domains (A and B) have mutually disjoint supports, and they are uniform over these supports. Assuming the support sets of domains A and B are not identical, the optimal feature extractors that minimize the instance discrimination loss of different domains present distinct geometric feature spaces.*

Intra-Domain Semantic-Enhanced Learning

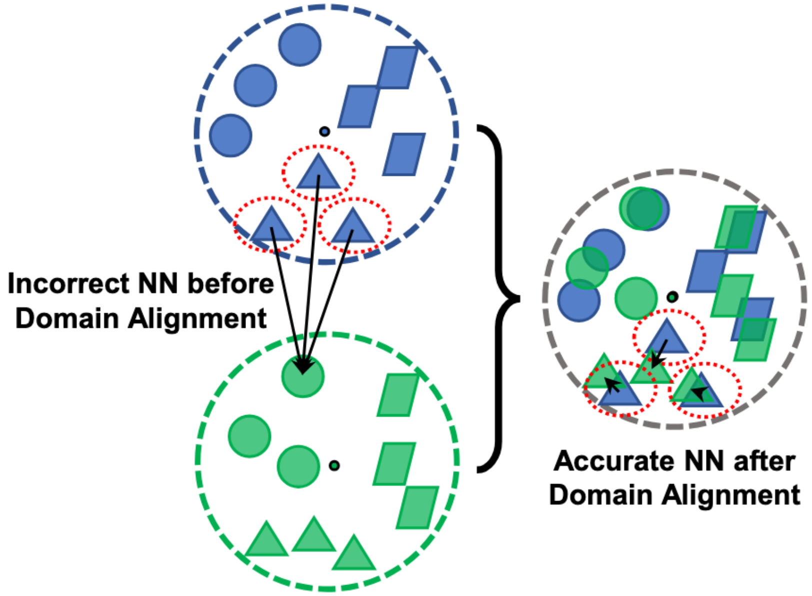
- Our solution: pre-set a unified prototypical structure across domains

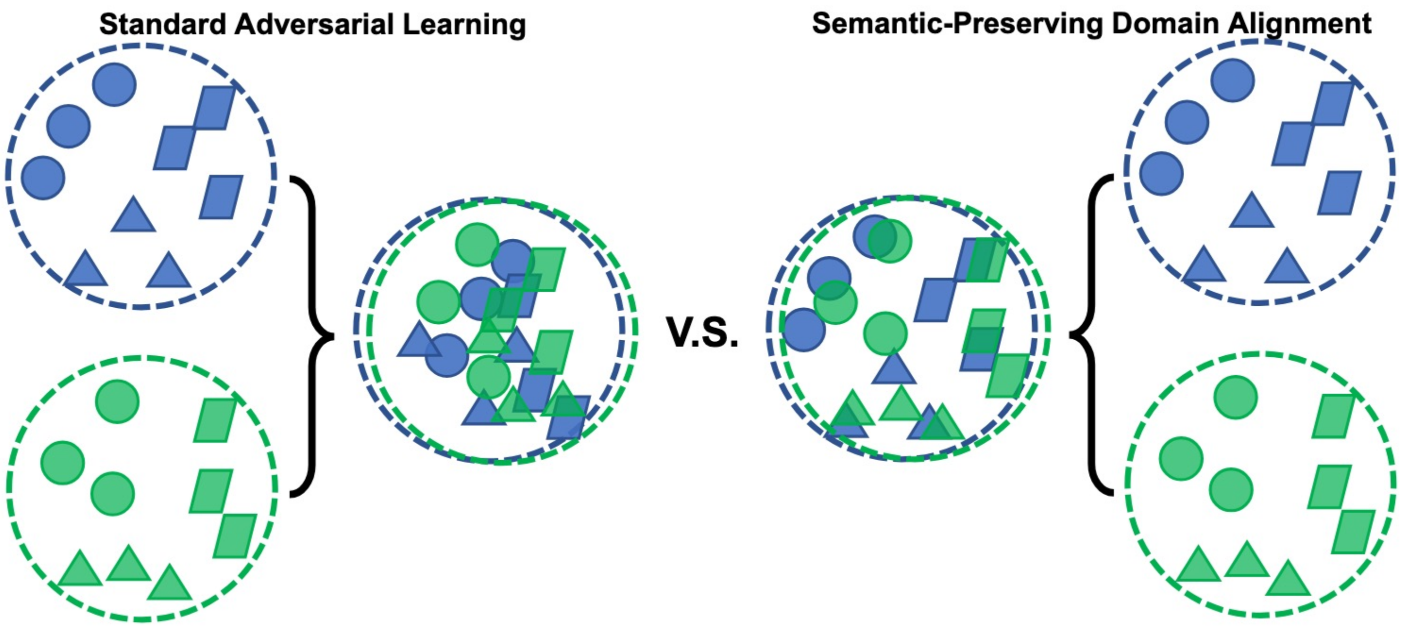


$$\mathcal{L}_{\mathrm{IPM}} = \mathcal{L}_{\mathrm{INCE}} + \alpha \mathcal{L}_{\mathrm{PNCE}}, \text{ where } \alpha = \frac{1}{1 + \exp\left(0.5E - e\right)}$$

Cross-Domain Semantic-Matched Learning

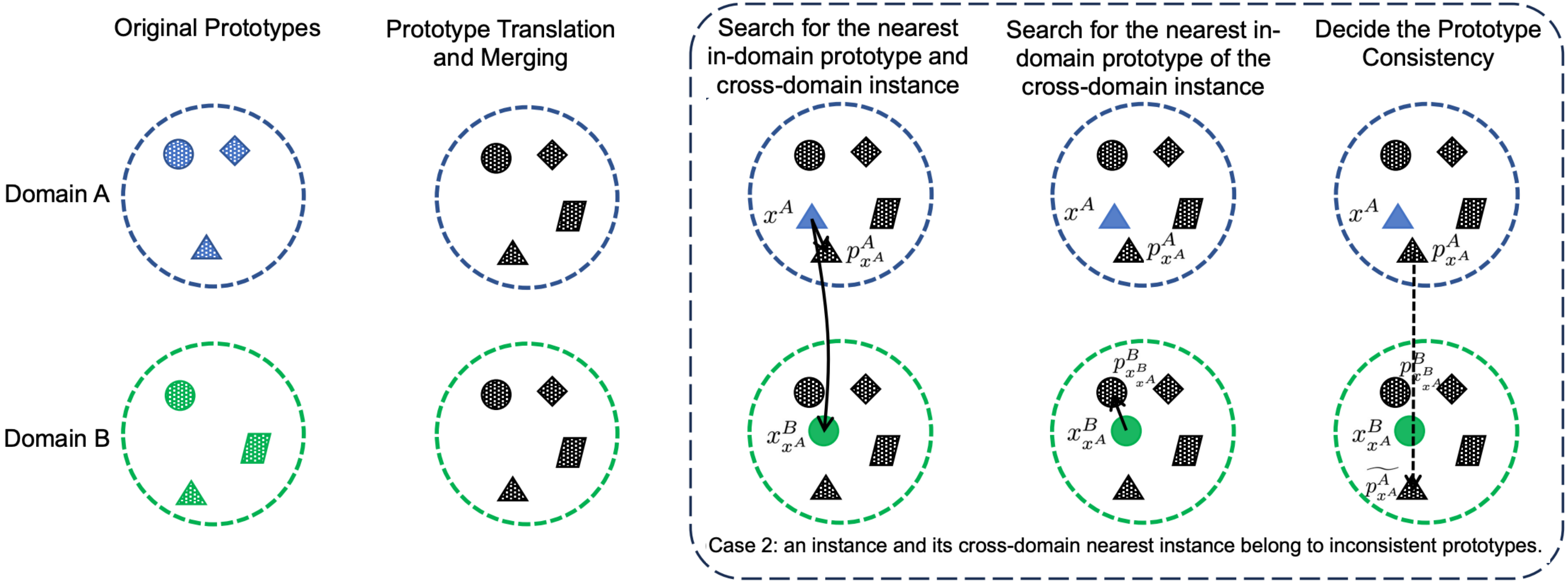
- Semantic-Preserving Domain Alignment





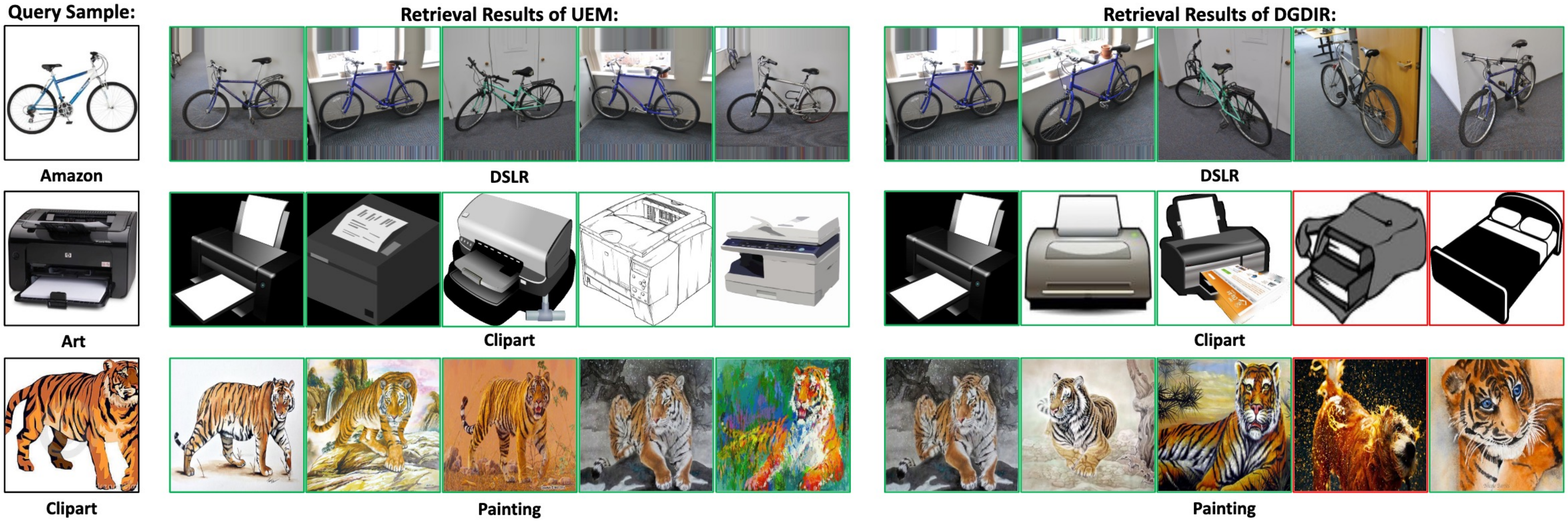
Cross-Domain Semantic-Matched Learning

- Switchable Nearest Neighboring Match



Experiments

Closed-set Cross-domain Retrieval



Experiments

Partial Cross-domain Retrieval on Office-Home

Methods	A→C	A→P	A→R	C→A	C→P	C→R	P→A	P→C	P→R	R→A	R→C	R→P	Avg.
CDS	$22.0_{\pm1.1}$	$31.1_{\pm0.7}$	$32.5_{\pm2.0}$	$26.5_{\pm1.0}$	$25.6_{\pm0.2}$	$27.9_{\pm1.5}$	$30.0_{\pm0.9}$	$31.8_{\pm1.1}$	$40.5_{\pm2.7}$	$32.3_{\pm1.8}$	$25.5_{\pm1.2}$	$37.6_{\pm3.0}$	$30.3_{\pm1.1}$
PCS	$24.5_{\pm0.4}$	$36.5_{\pm1.2}$	$38.8_{\pm2.0}$	$24.9_{\pm1.6}$	$28.8_{\pm1.1}$	$29.0_{\pm1.0}$	$28.6_{\pm2.1}$	**35.3**$_{\pm0.7}$	$41.7_{\pm1.4}$	**37.5**$_{\pm2.0}$	$26.9_{\pm1.6}$	$40.0_{\pm0.9}$	$32.7_{\pm0.8}$
DARL	$25.5_{\pm1.5}$	$34.7_{\pm2.0}$	$29.8_{\pm3.1}$	$25.0_{\pm1.9}$	$23.9_{\pm1.7}$	$27.5_{\pm1.5}$	$26.8_{\pm2.6}$	$31.9_{\pm1.1}$	$40.0_{\pm2.3}$	$35.5_{\pm1.4}$	$27.7_{\pm2.0}$	$40.0_{\pm1.5}$	$30.7_{\pm1.6}$
DN2A	**25.9**$_{\pm0.9}$	**37.0**$_{\pm1.4}$	$29.5_{\pm2.0}$	$25.2_{\pm1.0}$	$27.0_{\pm0.5}$	**30.5**$_{\pm1.1}$	$29.0_{\pm1.3}$	$31.5_{\pm0.7}$	$40.6_{\pm0.4}$	$35.7_{\pm1.7}$	$28.0_{\pm0.6}$	$41.0_{\pm1.1}$	$31.7_{\pm0.5}$
UCDIR	$23.0_{\pm1.0}$	$28.7_{\pm2.2}$	$31.0_{\pm0.9}$	$26.0_{\pm1.6}$	$22.0_{\pm1.1}$	$23.5_{\pm1.6}$	**31.1**$_{\pm1.5}$	$30.4_{\pm0.2}$	$40.2_{\pm0.6}$	$36.9_{\pm1.2}$	$27.0_{\pm2.1}$	$36.8_{\pm0.7}$	$29.7_{\pm0.7}$
CoDA	$22.5_{\pm1.2}$	$34.2_{\pm1.0}$	$35.7_{\pm2.0}$	$25.0_{\pm1.7}$	$29.5_{\pm0.8}$	$30.0_{\pm0.9}$	$30.7_{\pm1.1}$	$32.0_{\pm1.5}$	$43.2_{\pm1.3}$	$35.2_{\pm2.2}$	$28.5_{\pm1.4}$	$41.3_{\pm0.3}$	$32.3_{\pm0.7}$
DGDIR	$24.4_{\pm0.5}$	$30.9_{\pm2.0}$	**41.0**$_{\pm0.7}$	**27.2**$_{\pm1.2}$	**30.5**$_{\pm2.4}$	$29.6_{\pm1.7}$	$30.4_{\pm2.6}$	$33.2_{\pm1.0}$	**45.5**$_{\pm0.2}$	$37.1_{\pm1.1}$	**30.9**$_{\pm1.5}$	**42.0**$_{\pm1.0}$	**33.6**$_{\pm0.5}$
Ours	**40.5**$_{\pm1.4}$	**45.8**$_{\pm2.0}$	**48.0**$_{\pm2.1}$	**35.1**$_{\pm1.0}$	**39.2**$_{\pm0.5}$	**41.1**$_{\pm0.9}$	**52.4**$_{\pm3.0}$	**46.0**$_{\pm2.1}$	**55.0**$_{\pm1.7}$	**49.0**$_{\pm2.1}$	**43.1**$_{\pm1.0}$	**56.7**$_{\pm1.2}$	**46.0**$_{\pm1.1}$

Open-set Cross-domain Retrieval on DomainNet

Methods	Qu→Cl		Cl→Pa		Pa→In		In→Sk		Sk→Re		Avg.	
	Shared-set mAP@All / Open-set Acc											
CDS	22.4	58.9	34.5	65.2	25.5	60.7	25.0	59.2	33.7	64.9	28.2	61.8
PCS	23.3	57.8	34.2	67.8	24.9	60.5	**27.8**	**65.4**	34.7	66.9	29.0	63.7
DARL	21.9	54.4	32.5	60.2	22.0	53.9	26.6	60.6	32.3	62.8	27.1	58.4
DN2A	22.7	56.6	33.4	60.7	21.9	55.2	24.8	57.8	34.0	61.2	27.4	58.3
UCDIR	**24.4**	**59.0**	34.4	67.0	26.7	**62.9**	25.6	63.4	**35.5**	**68.2**	**29.3**	**64.1**
CoDA	24.2	58.8	35.6	66.6	**27.0**	61.1	24.9	58.0	34.6	62.9	**29.3**	61.5
DGDIR	23.5	57.5	**36.0**	**68.3**	25.8	60.6	25.8	59.1	35.0	64.3	29.2	62.0
Ours	**30.3**	**72.9**	**40.2**	**88.1**	**36.0**	**82.5**	**31.1**	**78.2**	**36.6**	**83.0**	**34.8**	**80.9**

Ablation Study

Variations	Office-31		Office-Home		DomainNet	
	Shared-set mAP@All / Open-set Acc					
Ours w/o P.M.	68.3	78.8	45.7	72.6	30.9	70.2
Ours w/o SEL	72.2	88.3	47.5	81.7	33.0	76.4
Ours w/o SPDA	62.2	61.9	39.0	60.8	24.4	60.5
Ours w/o SN^{2}M	75.0	89.2	48.8	82.8	33.3	80.5
Ours	77.4	92.5	50.2	86.7	34.8	80.9