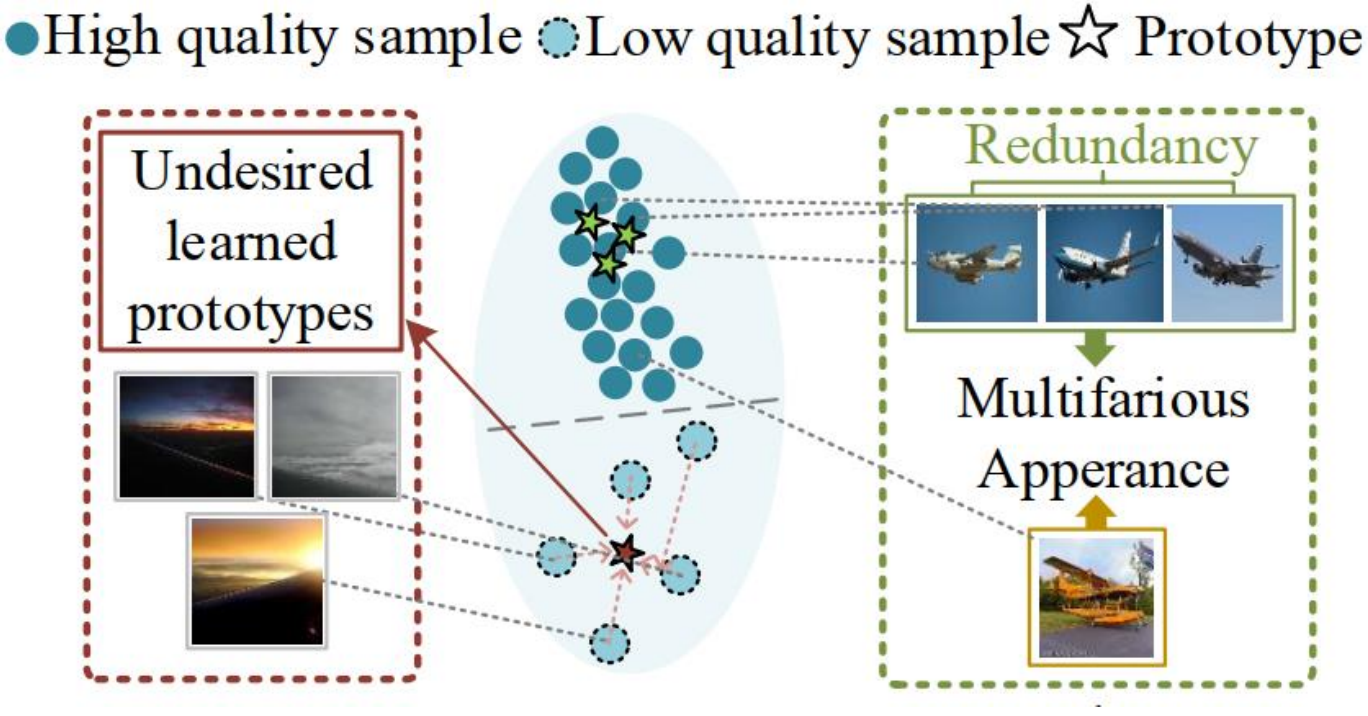
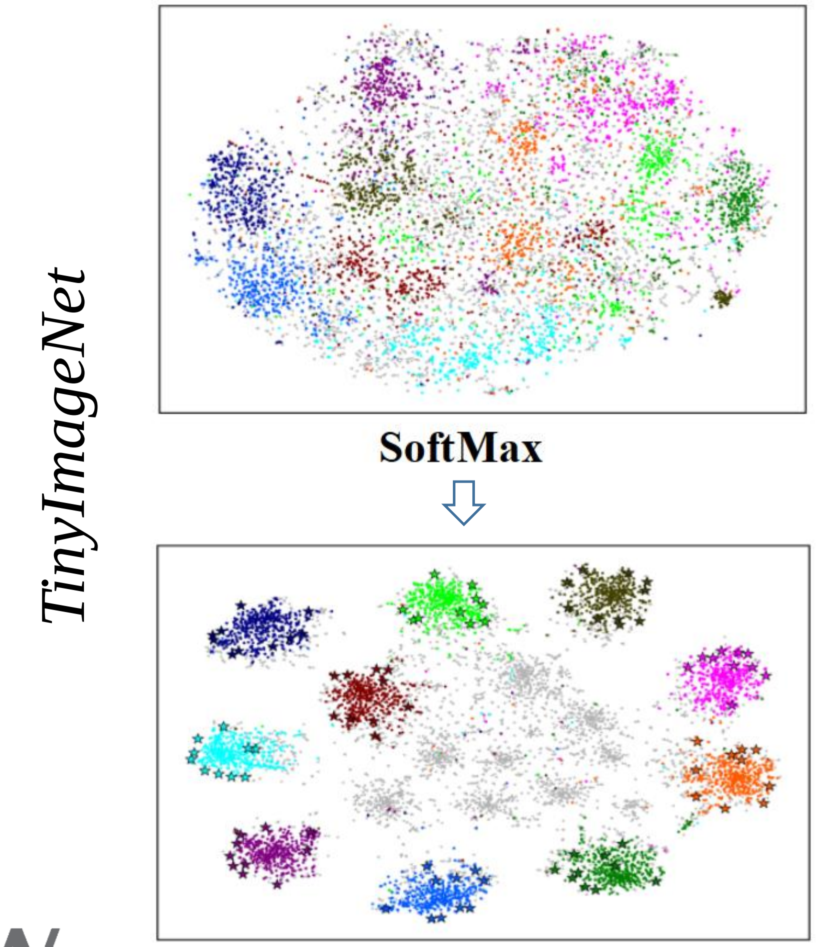
PMAL: Open Set Recognition via Robust Prototype Mining

- Problems of learning protoypes in OSR



(a) Quality (b) Diversity

- Effect Visualization



PMAL

Contact E-mail: lujing6@hikvision.com Jing Lu HIKVISION

PMAL: Open Set Recognition via Robust Prototype Mining

Jing Lu[1*], **Yunlu Xu**[1*], **Hao Li**[1], **Zhanzhan Cheng**[1,2†], **Yi Niu**[1]

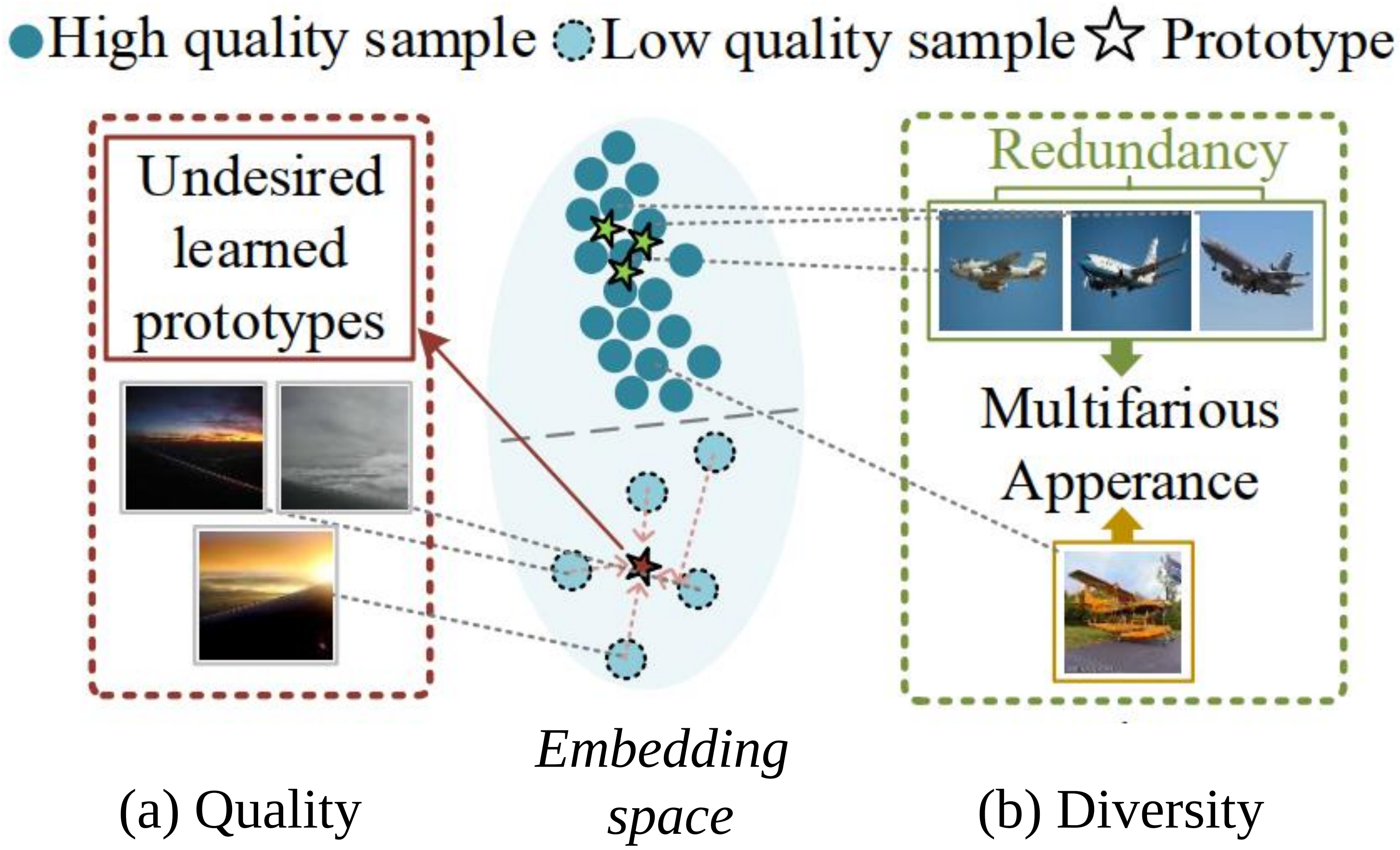
[1] Hikvision Research Institution, Hangzhou, China

[2] Zhejiang University, Hangzhou, China

{lujing6, xuyunlu, lihao50, chengzhanzhan, niuyi}@hikvision.com

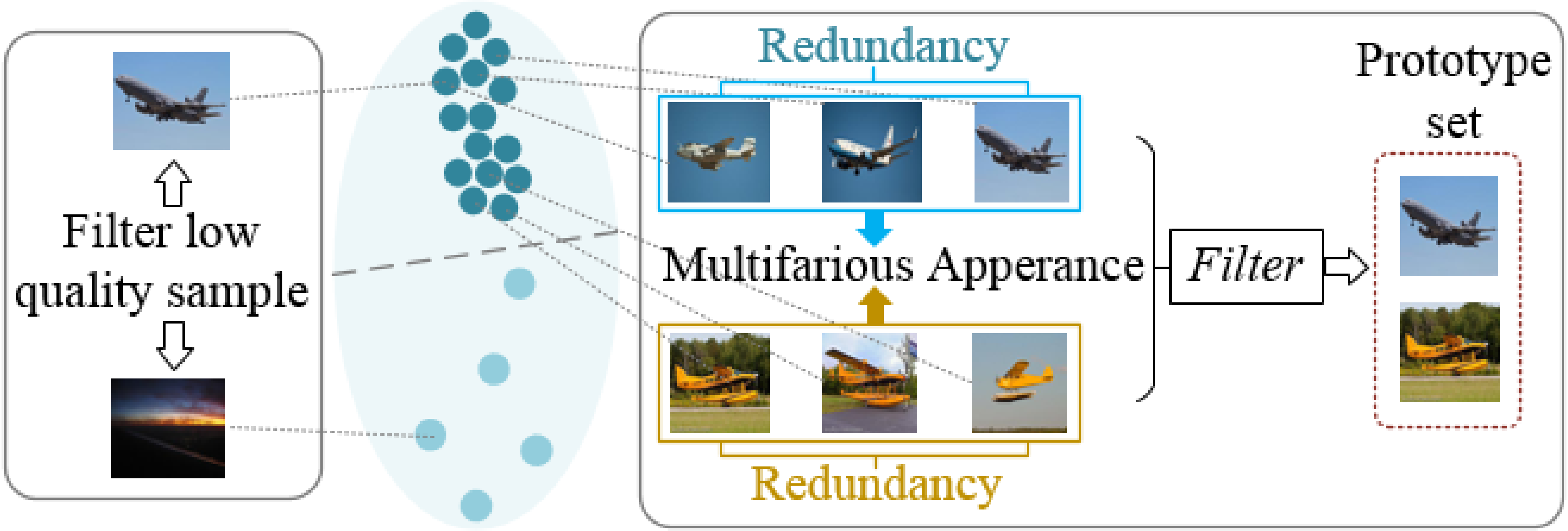
Problems

- ***Implicit*** *Learned Prototypes*



Motivation

- ***Explicit*** *Prototype Mining:* PMAL

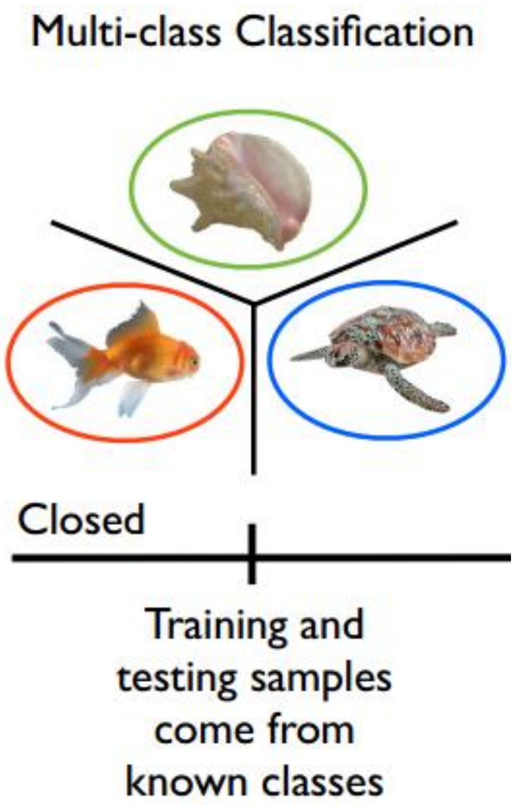


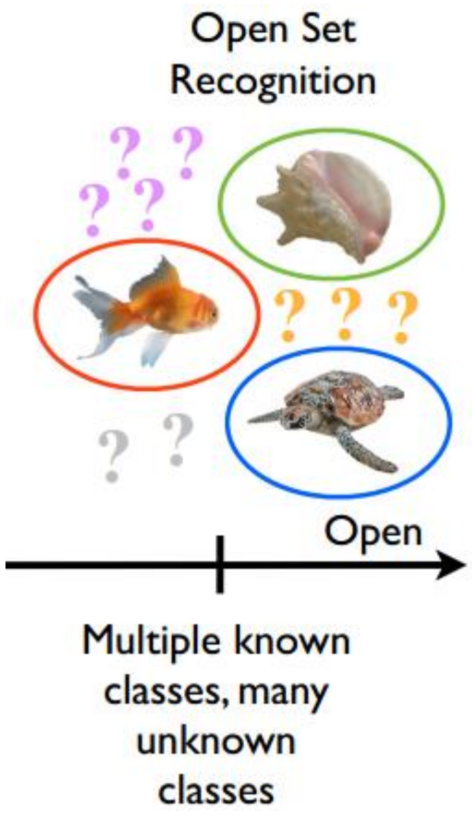
(a) *Mine High-quality Candidates*

(b) *Filter with diversity*

Preliminary

- Open Set Recognition(OSR)





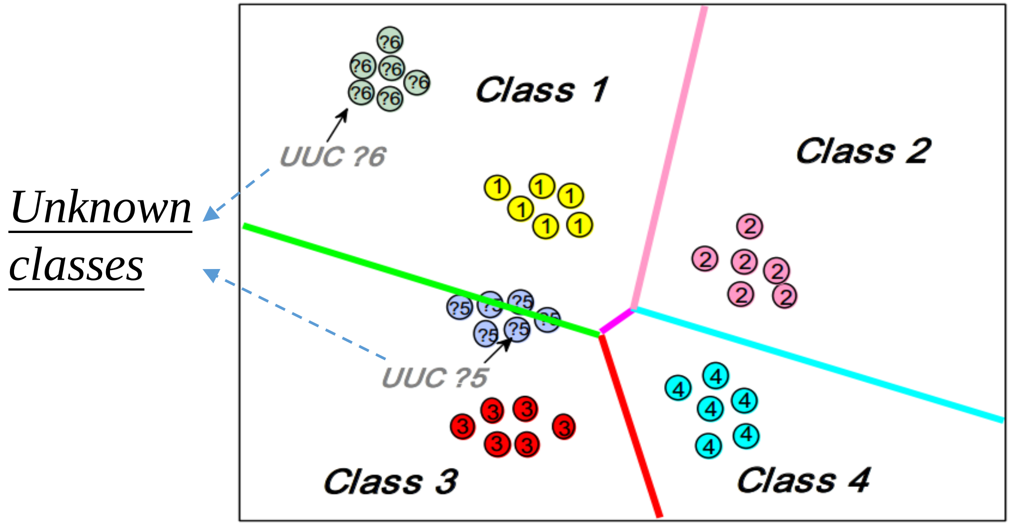
- *Ideal: Close Set Assumption*
- *Actually: Open Set Environment*

Scheirer, W. J., et al. 2013. Towards Open Set Recognition. TPAMI.

Preliminary

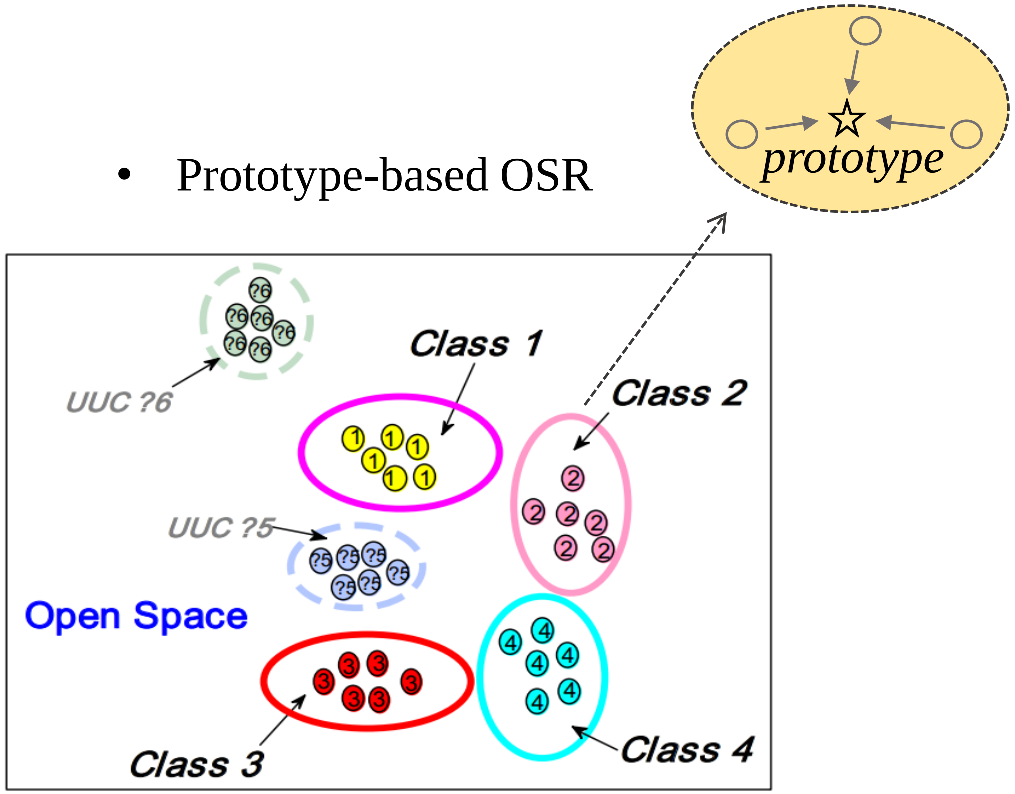
- Prototype-based OSR

    - Softmax-based close-set recognition



✓ Unable to tell UNKNOWN classes

- Prototype-based OSR



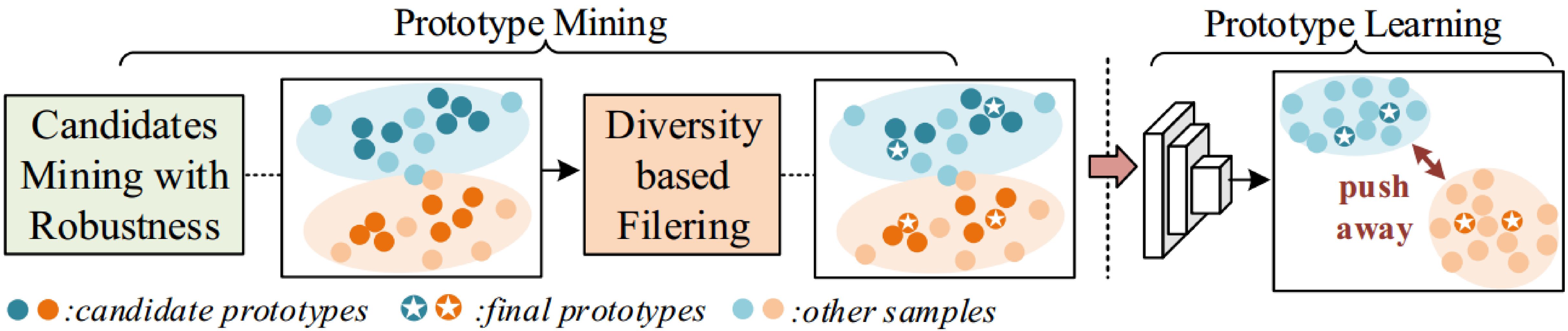
✓ Learn compact intra-class embedding

✓ Reserve more space for UNKNOWN classes

Jaderberg, M, et al. 2015. Spatial Transformer Networks. NeurIPS.

Method

- Overview of PMAL



Method

- Mine *high-quality* samples as prototype candidates

☐ Data uncertainty modelling[1]

Equa. 1: $z(x_i) = \phi(x_i) + n(x_i),\ n(x_i) \sim \mathcal{N}(0, \sigma(x_i))$

⇨ High-quality samples satisfy[2]:

$$z_i \approx \phi_i$$

☐ Key properties of high-quality samples

Equa. 2: $\boxed{d_{\mathcal{M}}(\phi_i^1, \phi_j^1) \approx d_{\mathcal{M}}(\phi_i^2, \phi_j^2)}$

Mahalanobis distance

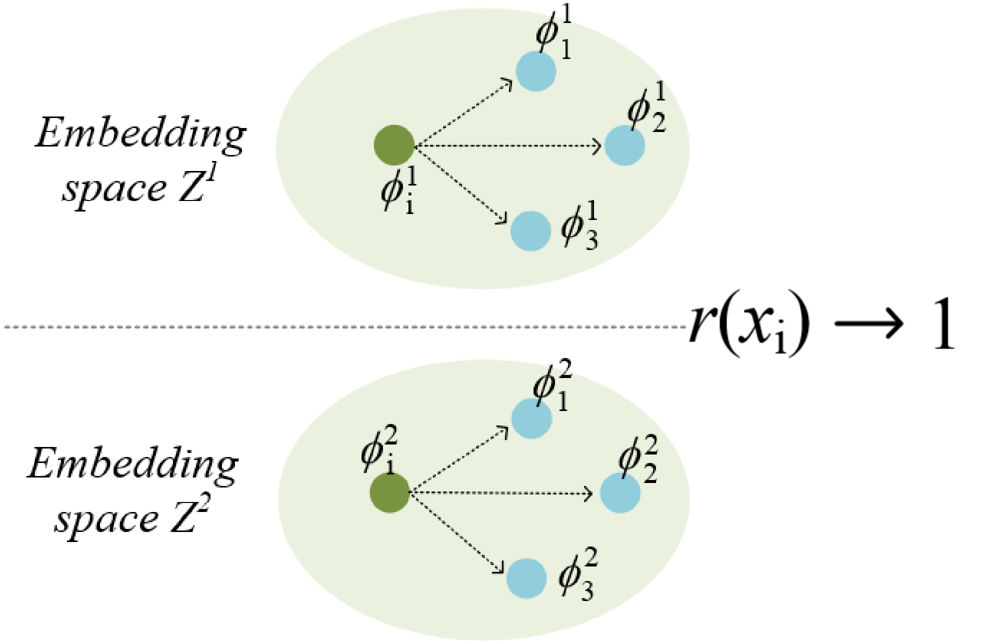
☐ *Embedding Topology*

Equa. 3: $t(z_i) \triangleq (d_{\mathcal{M}}(z_i, z_1), \ldots, d_{\mathcal{M}}(z_i, z_N))$

☐ *Embedding Topology robustness*

Equa. 4: $r(x_i) \triangleq exp(-||t(z_i^1) - t(z_i^2)||_2)$

☐ For a high-quality samples x_i



(a)

Method

- Mine *high-quality* samples as prototype candidates

✓ Data uncertainty modelling[1]

Equa. 1: $z(x_i) = \phi(x_i) + n(x_i),\ n(x_i) \sim \mathcal{N}(0, \sigma(x_i))$ ⇨

High-quality samples satisfy[2] :

$$z_i \approx \phi_i$$

✓ Key properties of high-quality samples

Equa. 2: $\boxed{d_{\mathcal{M}}(\phi_i^1, \phi_j^1) \approx d_{\mathcal{M}}(\phi_i^2, \phi_j^2)}$

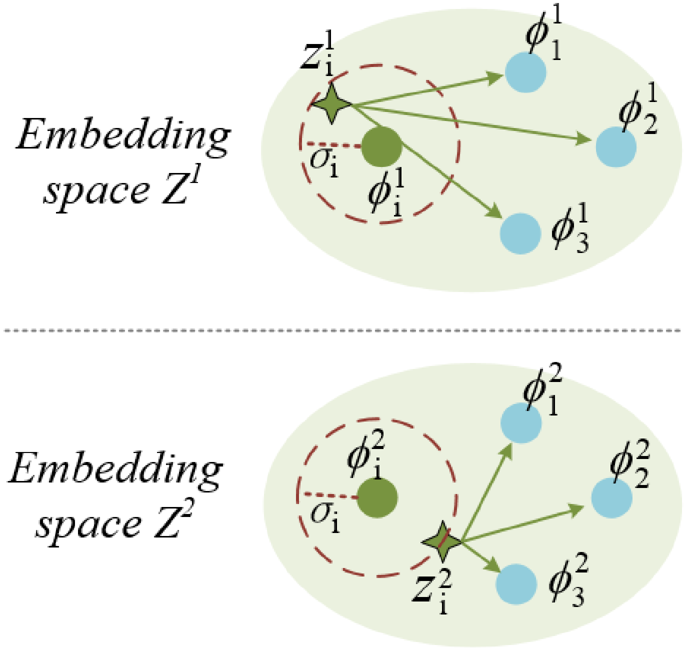
Mahalanobis distance

✓ Embedding Topology robustness

Equa. 3: $t(z_i) \triangleq (d_{\mathcal{M}}(z_i, z_1), \ldots, d_{\mathcal{M}}(z_i, z_N))$

Equa. 4: $r(x_i) \triangleq exp(-||t(z_i^1) - t(z_i^2)||_2)$

☐ For a low-quality samples x_i

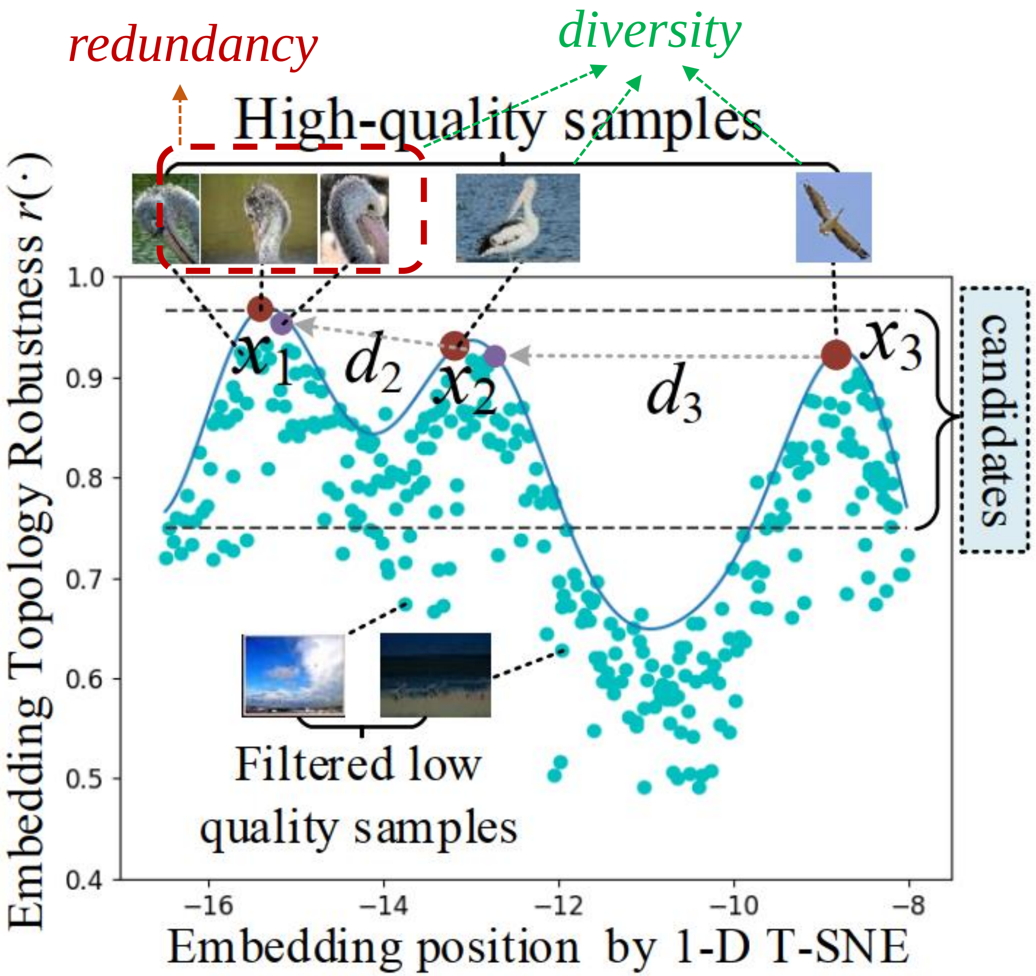


$r(x_i) < 1$

(b)

Method

- Filter with diversity



☐ Local maximum robustness

☐ Large embedding distance

- Greedy filtering algorithm

$$P_k = \bigcup_{i=1}^{T} \{x_i \mid \max_{x_i \in C_k} \{ \min_{x_j \in C_k} d_{\mathcal{M}}(z_i, z_j) \mid r(x_j) > r(x_i)\}\}$$

C_k: candidate set of class k

P_k: final prototype set of class k

Method

- Embedding optimization

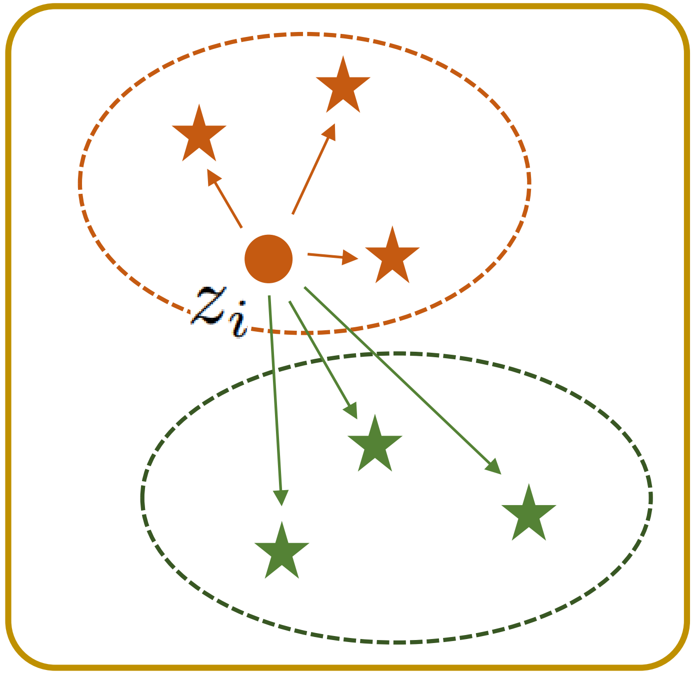
  ✓ **Point to Set distance**

$$z_i \xrightarrow{\textit{query}} z(P_k) = (z(p_{k,1}), \ldots, z(p_{k,T})) \in \mathbb{R}^{D\times T}$$

Equa.1: $$z_i^{att}(P_k) = SoftMax\left(\frac{z_i^{\mathrm{T}} z(P_k)}{\sqrt{d}}\right) z(P_k)$$

$$\Downarrow$$

Equa.2: $$d(z_i, z(P_k)) = 1 - \frac{{z_i}^{\mathrm{T}} z_i^{att}(P_k)}{|{z_i}^{\mathrm{T}}||z_i^{att}(P_k)|}$$



Method

- Embedding optimization

PMAL:
Explicitly updated by model forward pass

Existed methods:
Implicitly learned

Equa.1:

$$\mathcal{L}_p = \frac{1}{N}\sum_{i=1}^{N}[d(z_i, z(P_m)) - d(z_i, z(P_u)) + \delta]_+$$

Equa.2:

$$P_u = \underset{P_k \in P \setminus P_m}{\arg\min}\,(d(z_i, z(P_k)))$$

Ablation

- Each component

Table 3: Ablations of each module on TinyImageNet.

Components		(a)	(b)	(c)	(d)	(e)	(f)
PM	High-Quality	✓		✓		✓	✓
	Diversity		✓		✓	✓	✓
EO	Point-to-Set			✓	✓		✓
AUROC		80.3	78.1	81.6	80.2	81.9	**83.1**

- Both high quality and diversity matters for prototypes.
- Point-to-set distance helps learning better embedding space.

Ablation

- Each component

Table 3: Ablations of each module on TinyImageNet.

Components		(a)	(b)	(c)	(d)	(e)	(f)
PM	High-Quality	✓		✓		✓	✓
	Diversity		✓		✓	✓	✓
EO	Point-to-Set			✓	✓		✓
AUROC		80.3	78.1	81.6	80.2	81.9	**83.1**

- Both high quality and diversity matters for prototypes.
- Point-to-set distance helps learning better embedding space.

Ablation

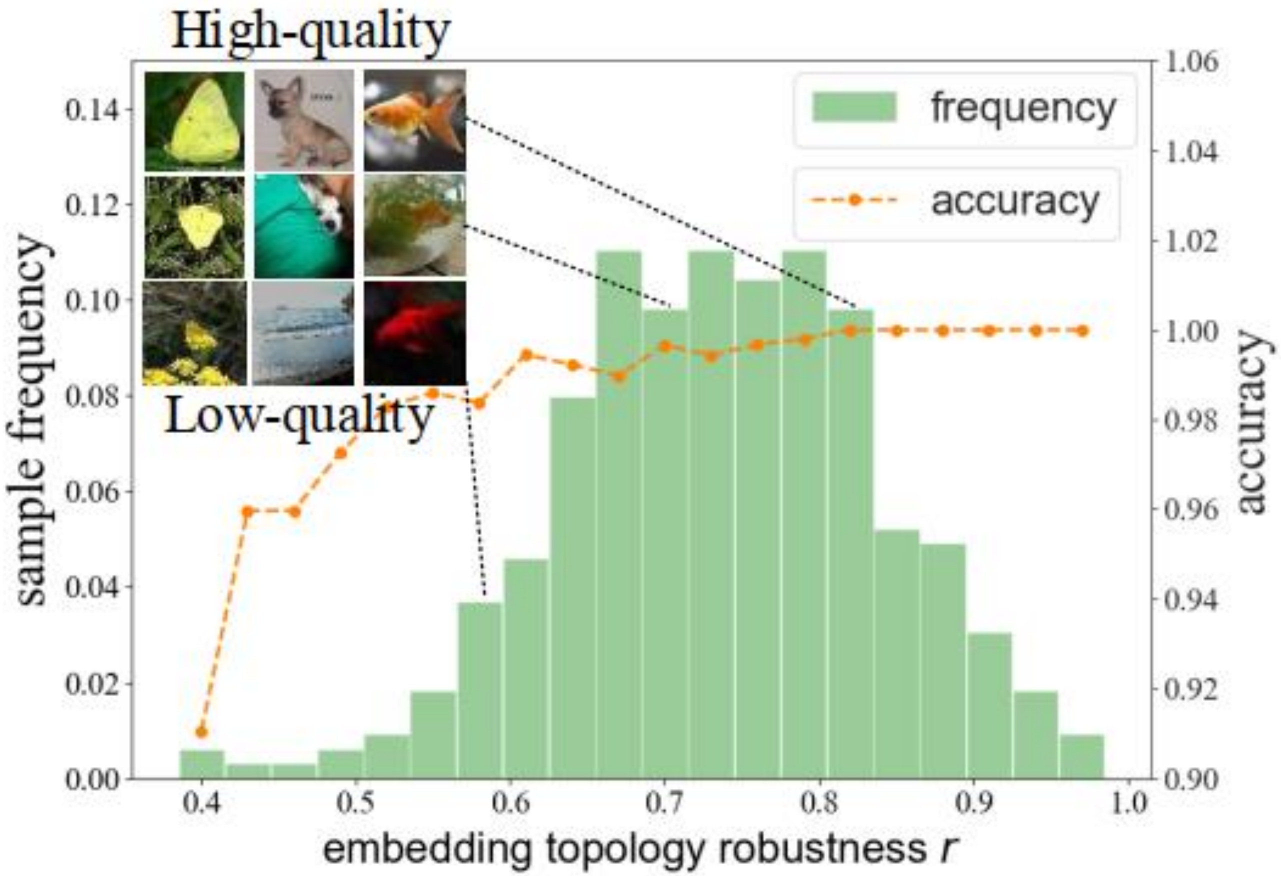
- Replace components in PM with existed strategies

Table 4: Comparisons with other methods on the *quality* and *diversity* property.

	Method	ACC	AUROC
Mine High-quality Candidates	(a)Probability	81.9	79.3
	(b)Deep Ensembles	82.3	80.5
	(c)MC-dropout	81.6	78.8
Diversity filtering	(a)Randomization	81.5	79.1
	(b)Clustering	81.8	79.6
	Ours	**84.7**	**83.1**

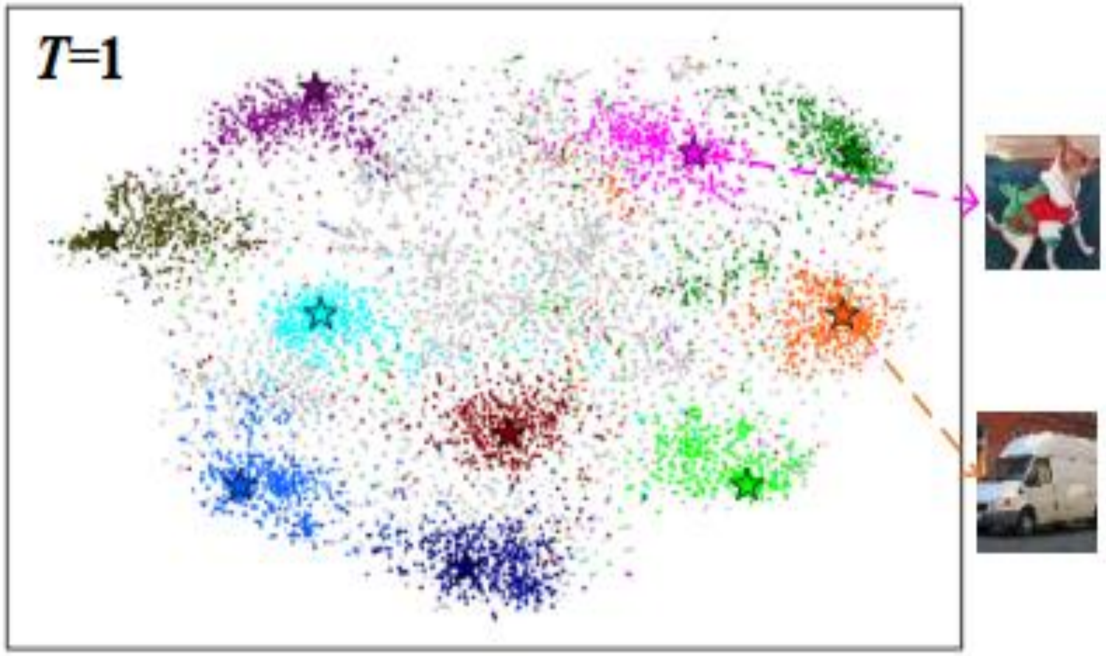
Visualization

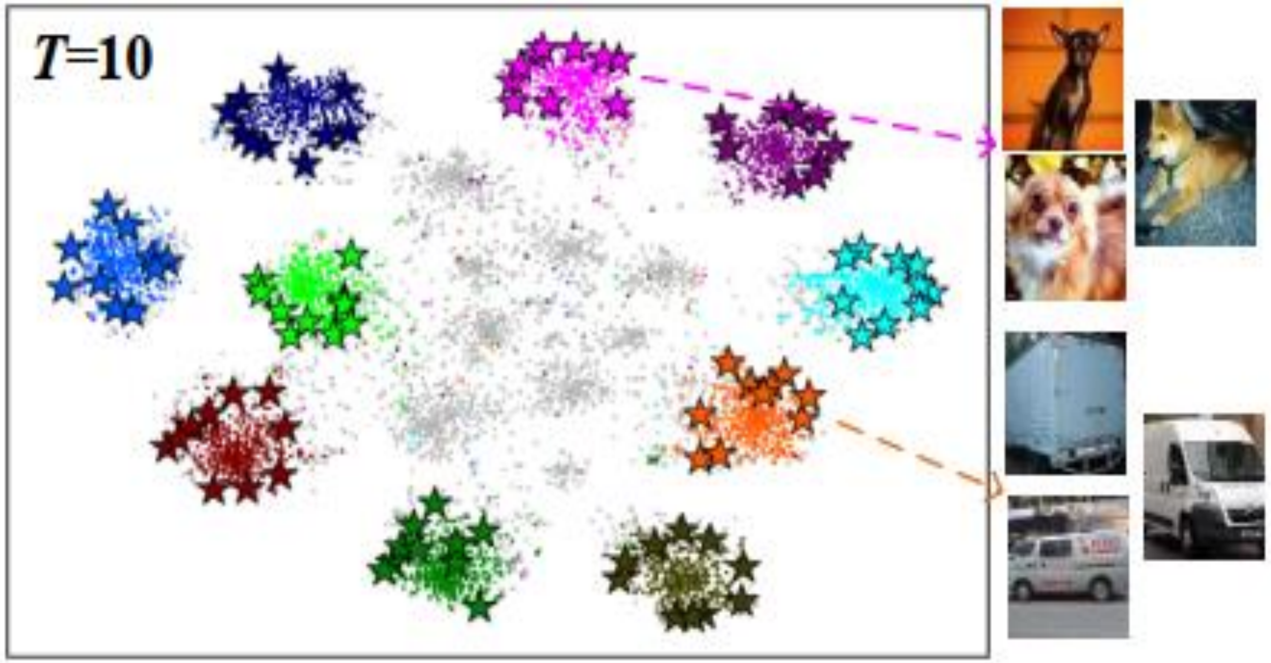
- High quality



Visualization

- Diversity



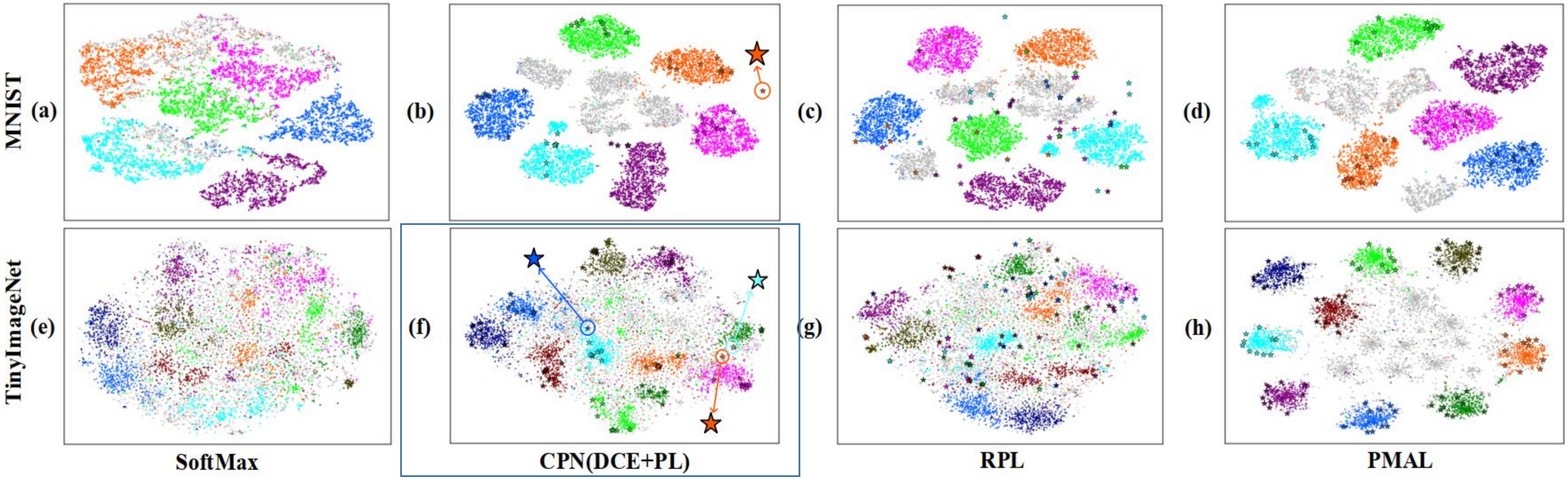


✓ Multifarious prototypes

Visualization

- Embedding space

Each color denotes different classes and ‘gray’ denotes unknowns



- On *simple* MNIST, all prototype-based methods performs satisfying.
- On more *complex* TinyImageNet, PMAL performs much better.

Performance

- Mainstream small-scale benchmarks

Table 1: Close set ACC and Open set AUROC on small datasets. '*' denotes implemented results and 'C' is short for 'CIFAR'.

Methods	Close set ACC						Open set AUROC					
	MNIST	SVHN	C10	C+10	C+50	TINY	MNIST	SVHN	C10	C+10	C+50	TINY
SoftMax	99.5	94.7	80.1	-	-	-	97.8	88.6	67.7	81.6	80.5	57.7
CPN (Yang et al.)	99.7	96.7	92.9	94.8*	95.0*	81.4*	99.0	92.6	82.8	88.1	87.9	63.9
PROSER (Zhou, Ye, and Zhan)	-	96.5	92.8	-	-	52.1	94.3	-	89.1	96.0	95.3	69.3
CGDL (Sun et al.)	99.6	94.2	91.2	-	-	-	99.4	93.5	90.3	95.9	95.0	76.2
OpenHybrid (Zhang et al.)	94.7	92.9	86.8	-	-	-	99.5	94.7	95.0	96.2	95.5	79.3
RPL-OSCRI (Chen et al.)	99.5*	95.3*	94.3*	94.6*	94.7*	81.3*	99.3	95.1	86.1	85.6	85.0	70.2
ARPL (Chen et al.)	99.5	94.3	87.9	94.7	92.9	65.9	99.7	96.7	91.0	97.1	95.1	78.2
RPL-WRN (Chen et al.)	99.6*	95.8*	95.1*	95.5*	95.9*	81.7*	99.6	96.8	90.1	97.6	96.8	80.9
PMAL-OSCRI	99.6	96.5	96.3	96.4	96.9	84.4	99.5	96.3	94.6	96.0	94.3	81.8
PMAL-WRN	**99.8**	**97.1**	**97.5**	**97.8**	**98.1**	**84.7**	**99.7**	**97.0**	**95.1**	**97.8**	**96.9**	**83.1**

Performance

- Mainstream small-scale benchmarks

Table 1: Close set ACC and Open set AUROC on small datasets. '*' denotes implemented results and 'C' is short for 'CIFAR'.

Methods	Close set ACC						Open set AUROC					
	MNIST	SVHN	C10	C+10	C+50	TINY	MNIST	SVHN	C10	C+10	C+50	TINY
SoftMax	99.5	94.7	80.1	-	-	-	97.8	88.6	67.7	81.6	80.5	57.7
CPN (Yang et al.)	99.7	96.7	92.9	94.8*	95.0*	81.4*	99.0	92.6	82.8	88.1	87.9	63.9
PROSER (Zhou, Ye, and Zhan)	-	96.5	92.8	-	-	52.1	94.3	-	89.1	96.0	95.3	69.3
CGDL (Sun et al.)	99.6	94.2	91.2	-	-	-	99.4	93.5	90.3	95.9	95.0	76.2
OpenHybrid (Zhang et al.)	94.7	92.9	86.8	-	-	-	99.5	94.7	95.0	96.2	95.5	79.3
RPL-OSCRI (Chen et al.)	99.5*	95.3*	94.3*	94.6*	94.7*	81.3*	99.3	95.1	86.1	85.6	85.0	70.2
ARPL (Chen et al.)	99.5	94.3	87.9	94.7	92.9	65.9	99.7	96.7	91.0	97.1	95.1	78.2
RPL-WRN (Chen et al.)	99.6*	95.8*	95.1*	95.5*	95.9*	81.7*	99.6	96.8	90.1	97.6	96.8	80.9
PMAL-OSCRI	99.6	96.5	96.3	96.4	96.9	84.4	99.5	96.3	94.6	96.0	94.3	81.8
PMAL-WRN	**99.8**	**97.1**	**97.5**	**97.8**	**98.1**	**84.7**	**99.7**	**97.0**	**95.1**	**97.8**	**96.9**	**83.1**

Performance

- More large-scale benchmarks

Table 2: Comparisons on 3 large-scale datasets. We denote ‘ImageNet’ as ‘IN’ for simplicity.

Method	Close Set ACC			Open Set AUROC			Additional Params		
	IN-LT	IN-100	IN-200	IN-LT	IN-100	IN-200	IN-LT	IN-100	IN-200
Softmax	37.8	81.7	79.7	53.3	79.7	78.4	0	0	0
CPN	37.1	86.1	82.1	54.5	82.3	79.5	2M	0.2M	0.4M
RPL	39.0	81.8*	80.7*	55.1	81.2*	80.2*	2M	0.2M	0.4M
RPL++	39.7	-	-	55.2	-	-	4M	-	-
PMAL	**42.9**	**86.2**	**84.1**	**71.7**	**94.9**	**93.9**	0	0	0

✓ More obvious advantages on complicated scenarios

Contact Information

PMAL: Open Set Recognition via Robust Prototype Mining

Contact E-mail: lujing6@hikvision.com Jing Lu

Document Audio Video Analysis & Recognition LAB

Our Team Homepage:

https://davar-lab.github.io/