Out-of-distribution Detection with Boundary Aware Learning

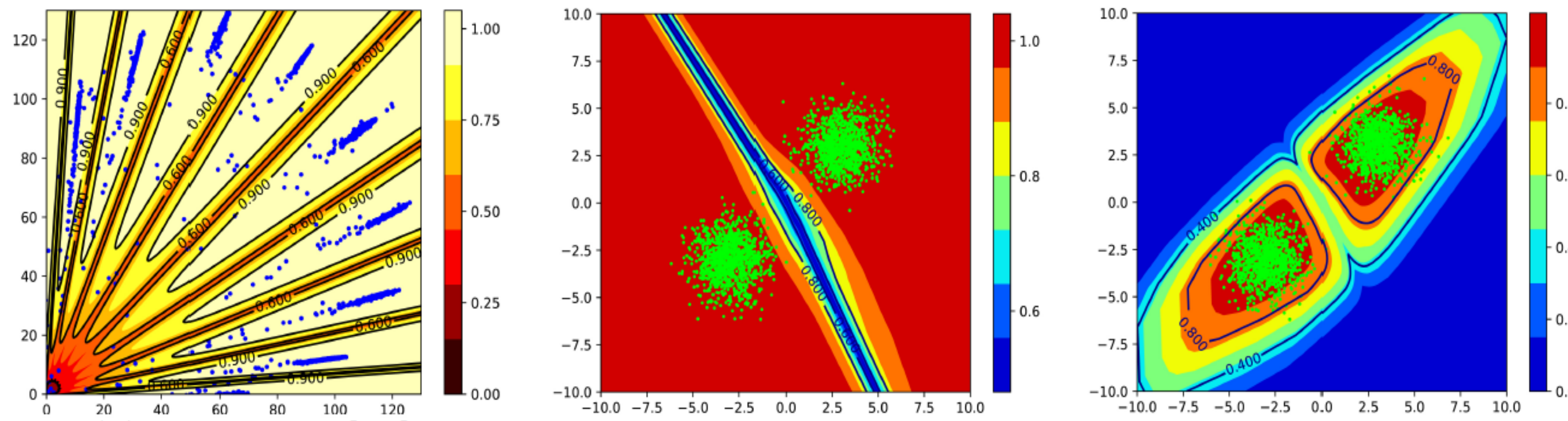
Sen Pei^{1,2}*, Xin Zhang*^{1,2}*, Bin Fan*^{4}*, and Gaofeng Meng*^{1,2,3}

^{1}*NLPR,Institute of Automation,CAS,* ^{2}*UCAS,* ^{3}*CAIR,HKISI,CAS,* ^{4}*University of Science and Technology Beijing*

TEL AVIV 2022

Introduction

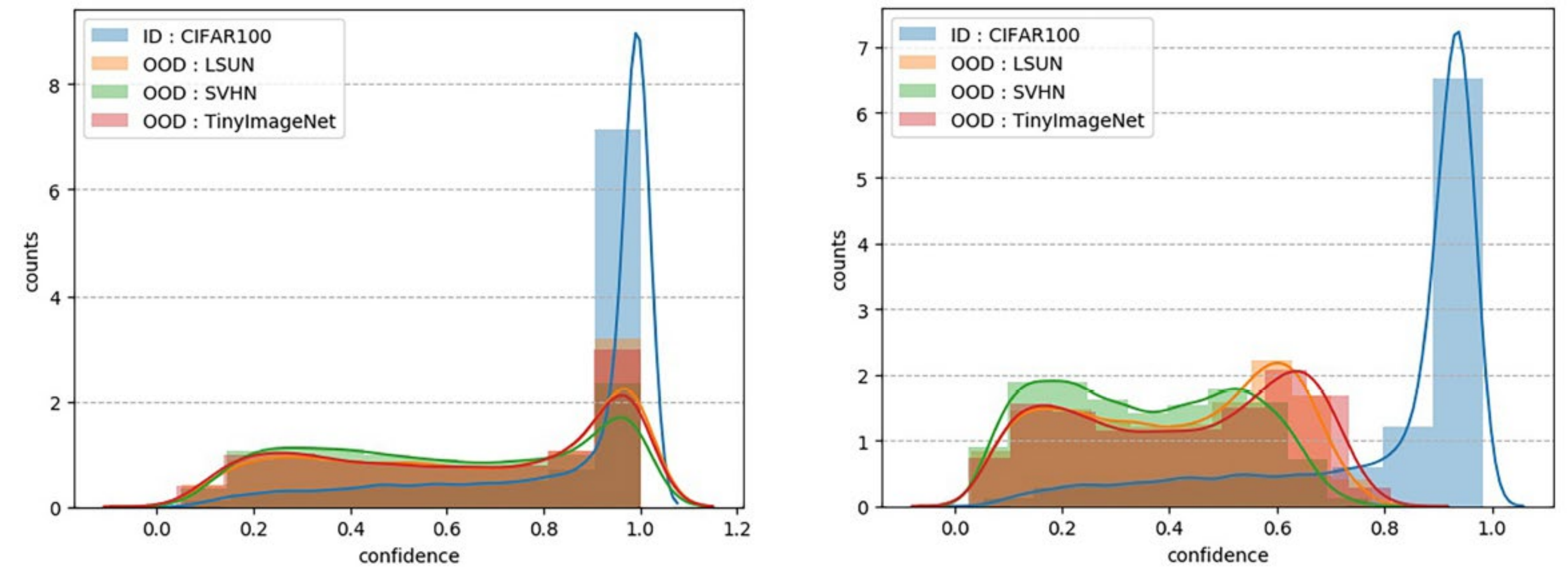
CNNs suffer from over-confidence issue



(a) ResNet18 [11] (b) MLP (c) BAL (ours)

- ➢ In-distribution features concentrate **densely** in narrow regions.
- ➢ Almost the whole **feature space** is assigned with high confidence.
- ➢ The classifier is expected to detect the **boundary** of in-distribution.
- ➢The synthetic hard OOD features lead to **compact** decision boundary.

Problem Statement: Out-of-distribution detection

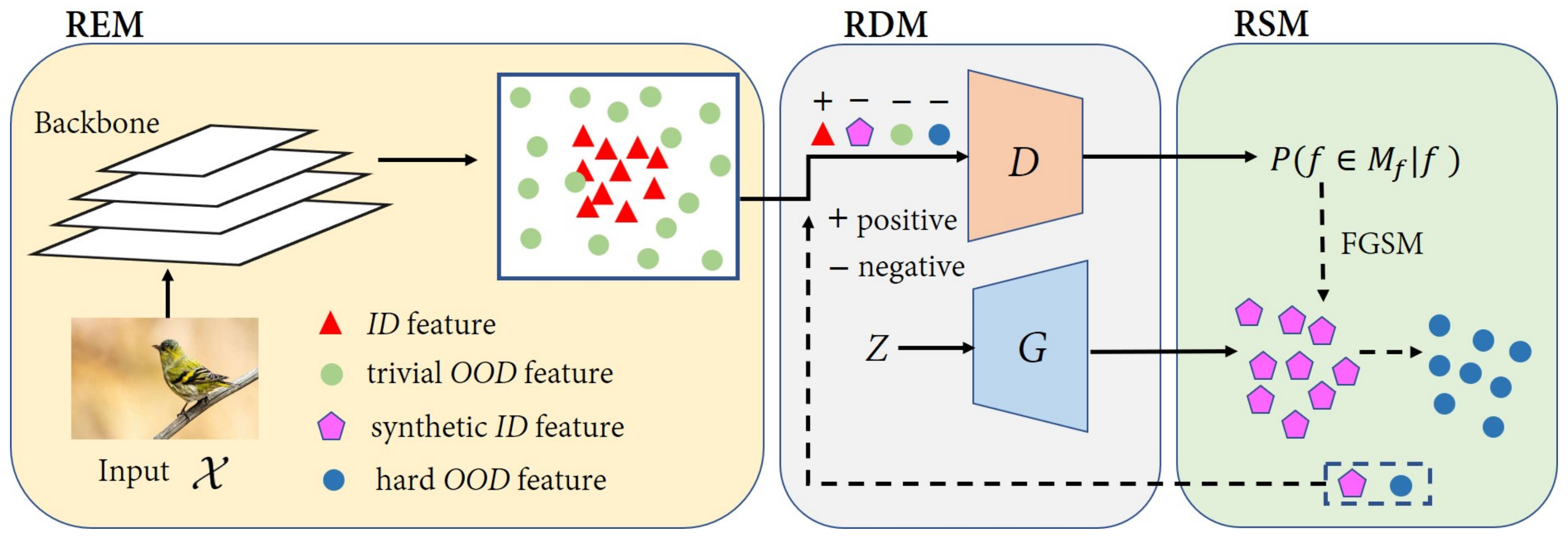


(a) Classification with ResNet18 [11] (b) Classification with BAL (ours)

- ➢ The in-distribution data is expected to get high confidence.
- ➢ The out-of-distribution data obtains much lower confidence score.
- ➢ No out-of-distribution data is accessible.

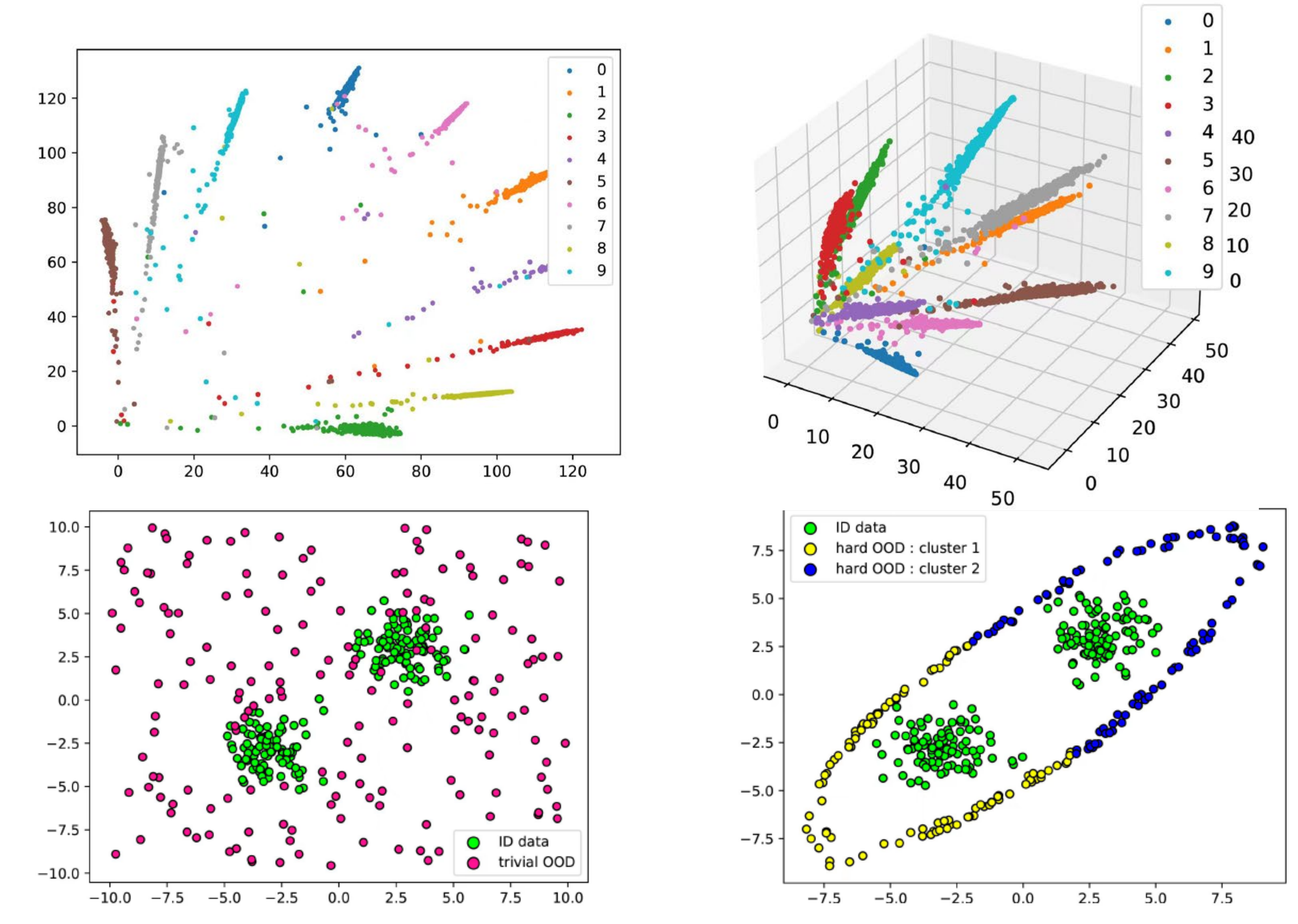
Method

Overview: Boundary Aware Learning



- ➢ RSM: Representation Sampling Module.
- ➢ REM: Representation Extraction Module.
- ➢ RDM: Representation Discrimination Module.

Insights: Focus on the learning of hard OOD



(a) *trivial OOD* via uniform sampling (b) *hard OOD* via FGSM

- ➢ trivial OOD: uniform sampling, since the ID features distribute densely.
- ➢ hard OOD: Fast Gradient Sign Method, the boundary of discriminator.

Experiments

Main Results

➢BAL achieves the state-of-the-art on common benchmarks.

ID	OOD	↓ FPR at 95% TPR					↑ AUPR in					↑ AUPR out				
		Softmax baseline[12] / AEC[6] / ODIN[25] / Generalized ODIN[15] / **BAL(ours)**														
C-10 D-BC	SVHN	59.8	57.2	63.6	44.2	**32.6**	91.9	92.3	89.1	94.6	**99.7**	87.0	92.5	83.9	88.7	**99.7**
	LSUN	33.4	27.6	5.6	5.2	**4.7**	96.4	97.3	98.9	99.0	**99.5**	94.0	96.3	98.7	**98.9**	**98.9**
	TIN	41.1	35.1	10.5	9.3	**5.0**	95.3	96.2	98.1	97.9	**98.2**	92.2	94.0	97.8	97.4	**98.0**
C-10 R-34	SVHN	67.5	57.2	64.4	12.7	**11.3**	92.2	93.4	85.8	94.5	**95.5**	84.9	84.5	81.8	93.4	**97.4**
	LSUN	54.6	34.6	26.2	21.3	**15.8**	92.3	91.8	93.7	**94.0**	93.9	88.5	92.1	93.8	93.9	**94.1**
	TIN	55.3	28.7	28.0	27.4	**21.6**	92.4	93.1	94.0	**94.3**	93.9	88.3	90.1	92.9	92.7	**93.8**
C-100 D-BC	SVHN	73.3	63.2	60.9	31.9	**21.5**	85.9	89.3	90.2	90.7	**91.5**	78.5	86.7	85.2	89.5	**92.8**
	LSUN	83.3	66.0	58.4	23.9	**11.3**	72.4	87.4	85.0	88.1	**89.3**	65.4	84.9	82.0	87.6	**88.7**
	TIN	82.4	59.7	56.9	22.7	**12.0**	73.0	83.7	84.7	86.5	**91.5**	67.4	82.9	83.0	84.3	**90.6**
C-100 R-34	SVHN	79.7	76.5	76.5	31.2	**17.3**	81.5	82.5	73.8	85.3	**87.1**	74.5	79.6	74.2	85.1	**89.3**
	LSUN	81.2	52.1	54.6	27.1	**18.7**	76.0	80.0	82.4	89.0	**91.5**	70.1	78.4	84.1	**89.0**	88.7
	TIN	79.6	55.3	50.6	29.7	**22.5**	79.2	87.1	86.8	89.3	**91.6**	72.3	85.6	87.0	88.0	**89.8**
SVHN D-BC	LSUN	22.9	22.7	22.1	18.7	**16.4**	96.7	95.4	95.3	97.2	**98.5**	88.0	88.7	**89.3**	86.3	**89.3**
	C-10	30.7	20.1	24.7	20.3	**12.1**	95.4	93.2	92.5	96.0	**97.3**	88.5	84.7	81.7	84.2	**89.9**
	TIN	21.2	18.6	19.9	15.2	**11.7**	97.0	96.1	95.5	97.3	**98.5**	88.9	90.7	90.1	**91.6**	90.6
SVHN R-34	LSUN	25.7	21.0	22.2	18.1	**13.5**	93.8	91.3	91.3	96.4	**97.8**	84.6	86.5	85.9	89.4	**92.1**
	C-10	21.7	19.5	20.0	16.7	**14.8**	94.8	92.0	91.9	97.0	**97.6**	86.4	87.3	87.1	88.2	**89.0**
	TIN	21.0	19.3	18.0	15.4	**14.3**	95.4	93.4	93.5	96.8	**98.2**	86.9	88.5	88.6	**89.4**	**89.4**

Ablation of Key Components

	↑ AUPR$_{in}$	↑ AUPR$_{out}$	↑ AUROC	↓ FPR 95
Softmax baseline	95.3	92.2	94.1	41.1
BAL (L_t)	97.0	96.0	96.6	17.9
BAL ($L_t + L_s$)	97.1	96.2	96.6	9.3
BAL ($L_t + L_u$)	97.2	96.3	96.7	8.1
BAL ($L_t + L_s + L_u$)	**98.2**	**98.0**	**97.0**	**5.0**

Visualization Results

- ➢ Classification between dogs and cats.
- ➢ Pink: Boundary Aware Learning (BAL).
- ➢ Lime: conventional max-softmax baseline.

Code

https://github.com/ForeverPs/BAL

Visualization

https://github.com/ForeverPs/BAL/tree/main/decision_boundary/results