Paper

ATTA: Anomaly-aware Test-Time Adaptation for Out-of-Distribution Detection in Segmentation

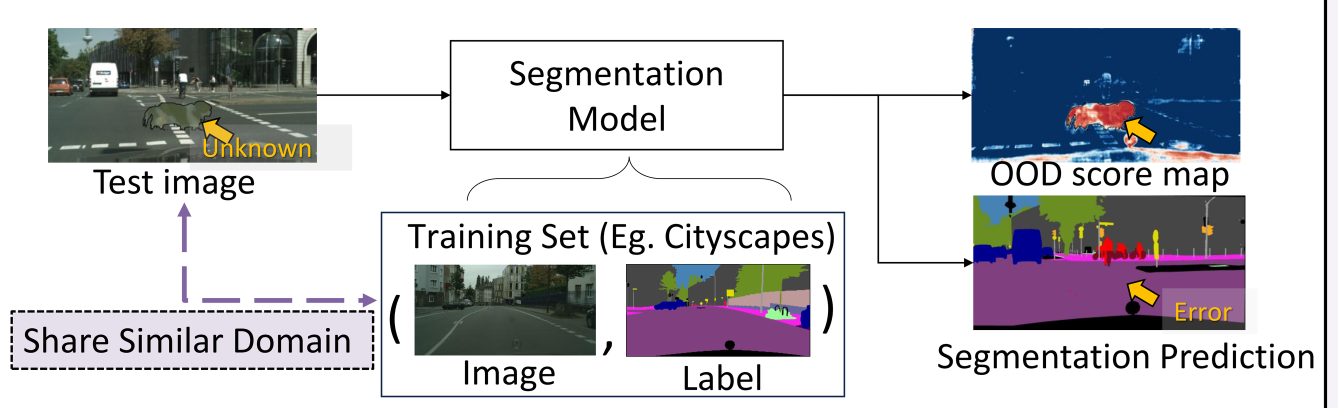
Zhitong Gao, Shipeng Yan, Xuming He   ShanghaiTech University

Code

INTRODUCTION

Dense Out-of-Distribution (OOD) Detection:

- **Goal:** Generate pixel-wise identification of unknown objects.
- **Assumption:** Training and testing data share similar domain.



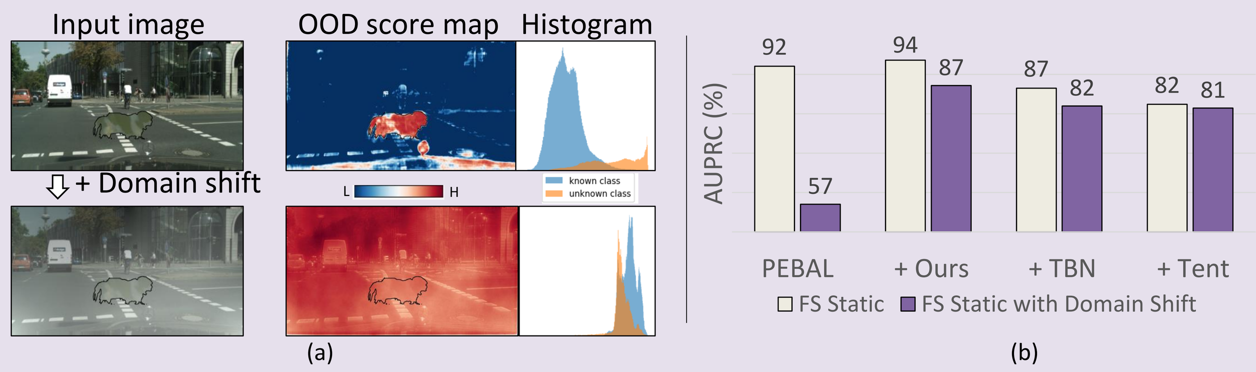
In real-world situations, **domain shift** often exists and significantly affects the performance of existing OOD detection models.

Domain-Shift Often Exists

E.g. various road, weather and lighting conditions.

Example images from existing OOD Detection Benchmarks

Domain-Shift Significantly Affects OOD Detection Performance



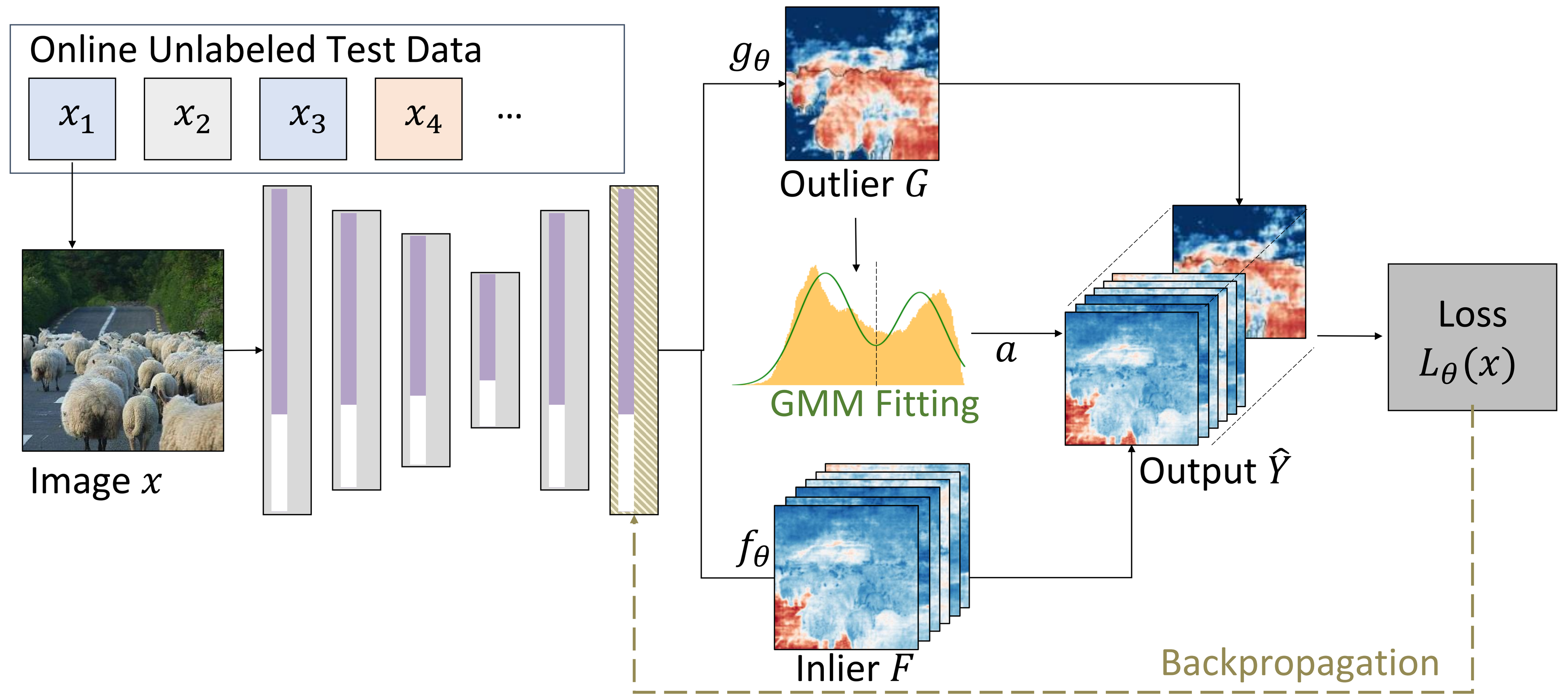
We Study Dense OOD Detection with Domain Shift

- **Strategy**: Adapt OOD detection models during test.
- **Challenges:** Apply existing test-time adaptation (TTA) techniques could:
  - Impair OOD detection performance on images from *seen domains*.
  - Indiscriminately reduce uncertainty of *unknown classes*.
- **Main Idea:** A dual-level test-time adaptation framework that
  - Simultaneously detects *two types* of distribution shift;
  - Performs online model adaptation in a *selective* manner.

METHOD

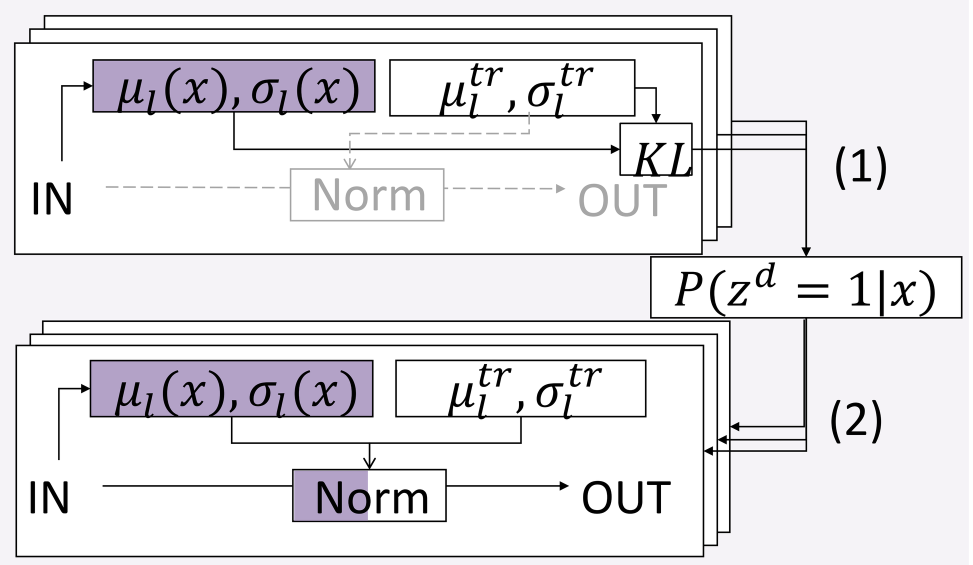
Anomaly-aware Test-Time Adaptation (ATTA)

- **Setting:** During testing, an image (or a batch) sequentially arrives, and our goal is to update the model parameters for each batch and produce anomaly-aware semantic segmentation outputs.



A. Selective Batch Normalization (SBN)

- **Goal**: Determines the existence of domain shift in the image and compensate for the input distribution deviation.



IN: BN input; OUT: BN output.

(1) Estimate the **probability of domain-shift** by considering image-level stat.

$$= h_{(a,b)}\left(\sum_{l=1}^{L} KL\left(N(\mu_l(x), \sigma_l(x)) \,||\, N(\mu_l^{tr}, \sigma_l^{tr})\right)\right)$$

Test feature Stat.; $h_{(a,b)}$: sigmoid((x + a)/b); Training feature Stat.

(2) Update **BN statistics:**

$$\hat{\mu}_l = P(z^d = 1|x) * \mu_l(x) + P(z^d = 0|x) * \mu_l^{tr}$$

$$\hat{\sigma}_l^2 = P(z^d = 1|x) * \sigma_l^2(x) + P(z^d = 0|x) * (\sigma_l^{tr})^2$$

B. Anomaly-aware Self-Training (AST)

- **Goal**: Further enhance OOD detection and closed-set prediction.
- **Overall Objective**: Promote the confidence of both the inlier and outlier predictions.

$$L_\theta(x) = -\sum_i \sum_{c=1}^{C+1} w_c \hat{Y}_{c,i} \log(\hat{Y}_{c,i})$$

$\hat{Y}_{c,i} \rightarrow P_\theta(y_i = c\,|x)$

Anomaly-aware Output Representation

$$\hat{Y}_{c,i} = \begin{cases} F_{c,i} * (1 - \hat{G}_i), & if\ i \in [1, C] \\ \hat{G}_i, & if\ i = C + 1 \end{cases}$$

Outlier Probability

$$\hat{G}_i = \text{Sigmoid}\left(\frac{G - a(x)}{b(x)}\right)$$

$b(x) \rightarrow$ Std.

$a(x)$: $\max\{a: \pi_1 N(a|\mu_1, \sigma_1) = \pi_2 N(a|, \mu_2, \sigma_2)\}$

- **Training:** update the final classification block in an episodic manner.

$$\min_\theta -\sum_i \left(\sum_{c=1}^{C} F_{c,i}(1 - \hat{G}_i)\log\left(F_{c,i}(1 - \hat{G}_i)\right) + \lambda \cdot \hat{G}_i \log \hat{G}_i\right)$$

$$\approx -\sum_{i \in D} \left(\sum_{c=1}^{C} F_{c,i}(1 - t_i)\log\left(F_{c,i}(1 - \hat{G}_i)\right) + \lambda \cdot t_i \log \hat{G}_i\right)$$

$D = \{i: \hat{G}_i < \tau_1\ or\ \hat{G}_i > \tau_2\}$; $t_i = 0 \cdot [\hat{G}_i < \tau_1] + 1 \cdot [\hat{G}_i > \tau_1]$

RESULTS

Simulated Domain-Shift Dataset

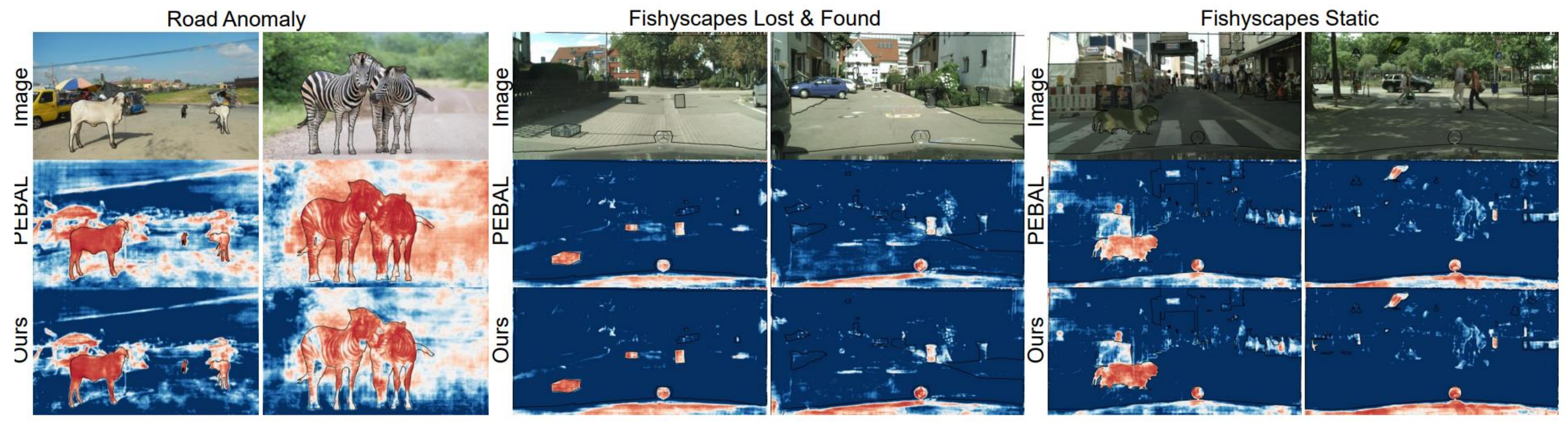
- Original FS Static dataset (white rows); FS Static -C dataset (gray rows).

	MSP [17]	Entropy [24]	Max logit [15]	Energy [30]	Meta-OOD [4]	PEBAL [44]	+ Ours	+ TBN [36]	+ Tent [46]
AUC ↑	92.36	93.14	95.66	95.90	97.56	99.61	**99.66**	99.25	99.04
	70.85	71.23	74.13	74.02	78.34	67.63	**99.21**	98.96	98.93
AP ↑	19.09	26.77	38.64	41.68	72.91	92.08	**93.61**	86.51	82.38
	10.52	14.32	23.60	22.36	52.31	57.02	**87.14**	81.97	81.42
FPR_{95} ↓	23.99	23.31	18.26	17.78	13.57	1.52	**1.15**	2.33	4.09
	100.0	100.00	89.94	89.94	100.0	97.17	**2.94**	4.26	4.43

Existing Dense OOD Detection Benchmarks

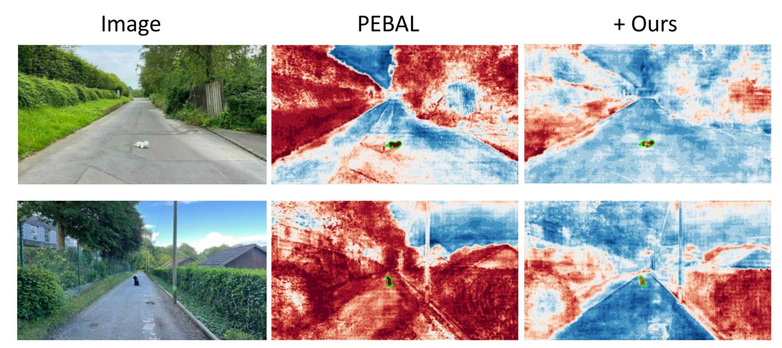
- Road Anomaly (high domain shift), FS LostAndFound, FS Static .

Methods	OoD Data	Road Anomaly AUC ↑	AP ↑	FPR_{95} ↓	FS LostAndFound AUC ↑	AP ↑	FPR_{95} ↓	FS Static AUC ↑	AP ↑	FPR_{95} ↓
MSP [17]	✗	67.53	15.72	71.38	89.29	4.59	40.59	92.36	19.09	23.99
Entropy [24]	✗	68.80	16.97	71.10	90.82	10.36	40.34	93.14	26.77	23.31
Mahalanobis [26]	✗	62.85	14.37	81.09	96.75	56.57	11.24	96.76	27.37	11.7
Meta-OoD [4]	✓	-	-	-	93.06	41.31	37.69	97.56	72.91	13.57
Synboost [10]	✓	81.91	38.21	64.75	96.21	60.58	31.02	95.87	66.44	25.59
DenseHybrid [14]	✓	-	-	-	99.01	69.79	5.09	99.07	76.23	4.17
Max Logit [15]	✗	72.78	18.98	70.48	93.41	14.59	42.21	95.66	38.64	18.26
+ ATTA (Ours)	-	**76.60**	**23.96**	**63.49**	**93.53**	**17.39**	**40.69**	95.48	**41.23**	20.89
Energy [30]	✗	73.35	19.54	70.17	93.72	16.05	41.78	95.90	41.68	17.78
+ ATTA (Ours)	-	**77.41**	**25.27**	**62.57**	93.30	**17.47**	43.32	**96.0**	**41.84**	**17.63**
PEBAL [44]	✓	87.63	45.10	44.58	98.96	58.81	4.76	99.61	92.08	1.52
+ ATTA (Ours)	-	**92.11**	**59.05**	**33.59**	**99.05**	**65.58**	**4.48**	**99.66**	**93.61**	**1.15**



- SMIYC Benchmark (high domain shift)

RoadAnomaly21	AP↑	FPR_{95} ↓	sIoU↑	PPV↑
PEBAL [44]	49.1	40.8	38.9	27.2
+ ATTA (Ours)	**67.0**	**31.6**	**44.6**	**29.6**
RoadObstacle21	**AP↑**	**FPR_{95} ↓**	**sIoU↑**	**PPV↑**
PEBAL [44]	5.0	12.7	29.9	7.6
+ ATTA (Ours)	**76.5**	**2.8**	**43.9**	**37.7**



Ablation Study

- Ablation of Two main modules.

SBN	AST	AUC ↑	AP ↑	FPR_{95} ↓
✗	✗	87.63	45.10	44.58
✗	✓	88.72	48.11	43.66
✓	✗	90.84	55.81	37.48
✓	✓	**92.11**	**59.05**	**33.59**

- Ablation of internal design of each module.

Train	Batch	Entropy	Norm	AUC ↑	AP ↑	FPR_{95} ↓
✗	✓	anomaly-aware	GMM	86.29	48.65	57.03
✓	✗	anomaly-aware	GMM	88.72	48.11	43.66
✓	✓	seen-class only	-	90.46	54.64	39.28
✓	✓	anomaly-aware	z-score	91.25	56.65	36.33
✓	✓	anomaly-aware	GMM	**92.11**	**59.05**	**33.59**

- Inference Time Analysis

Methods	Direct Inference	ATTA (Ours)	ATTA (Ours) w/o SBN	Tent	ODIN	SynBoost	Mahalanobis
Time (s)	1.2	2.7	1.5	5.1	9.2	3.0	224.2