Anomaly Segmentation in Driving Scene

2022 하계 세미나

Sogang University
Vision & Display Systems Lab, Dept. of Electronic Engineering

Presented by
강병우

Outline

- Background
  - Anomaly detection (AD)
  - Previous anomaly detection methods
  - Inabilities of previous anomaly segmentation methods
- Synboost
  - [1] Pixel-wise Anomaly Detection in Complex Driving Scenes (CVPR, 2021)
- PEBAL
  - [2] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)

Background

- Anomaly Detection (AD)
  - Process of locating unusual points or patterns in a set of data
  - High variance of driving scene normal data
  - No anomalous data during training
  - Various anomalous objects are critical for driving scene

Texture AD

Semantic AD

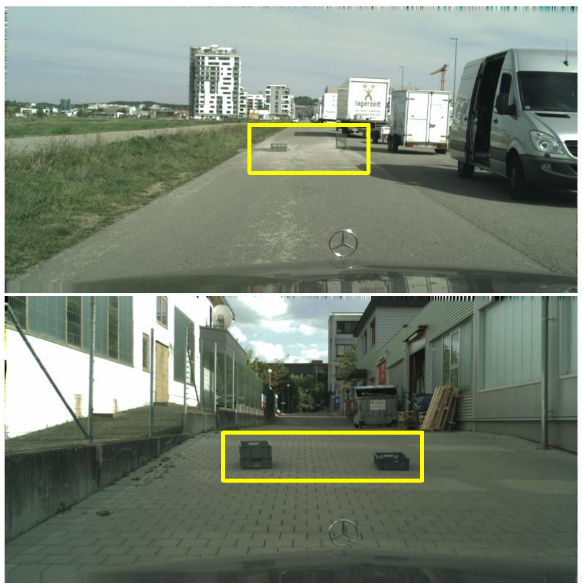
< Fishyscapes L&F >

< Fishyscapes static >

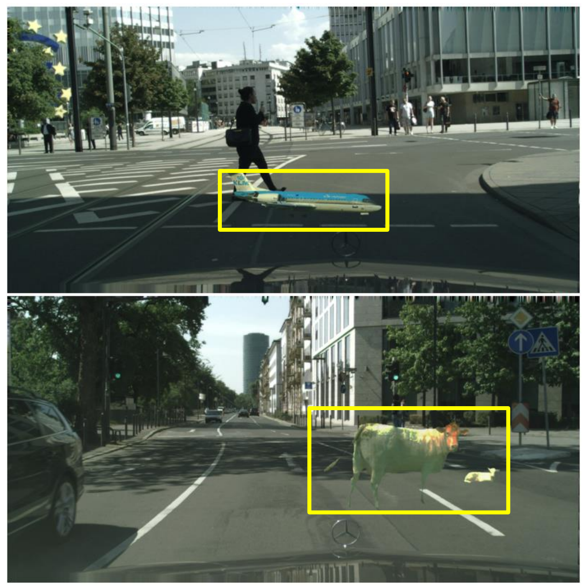
< Road anomaly >

Background

- Anomaly Detection (AD)
  - Datasets
    - Fishyscapes lost & found, Fishyscapes static, Road anomaly



< Fishyscapes lost & found >

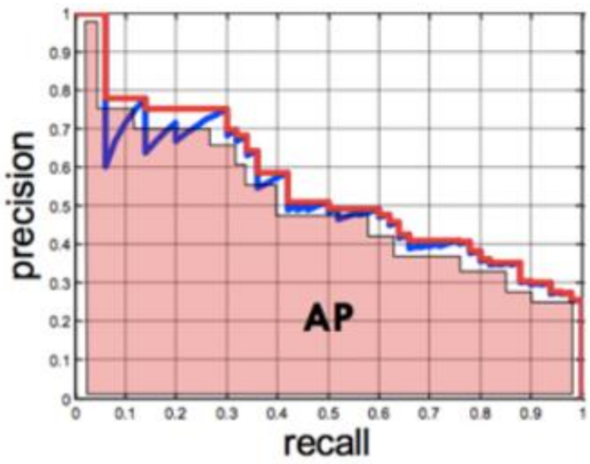


< Fishyscapes static >

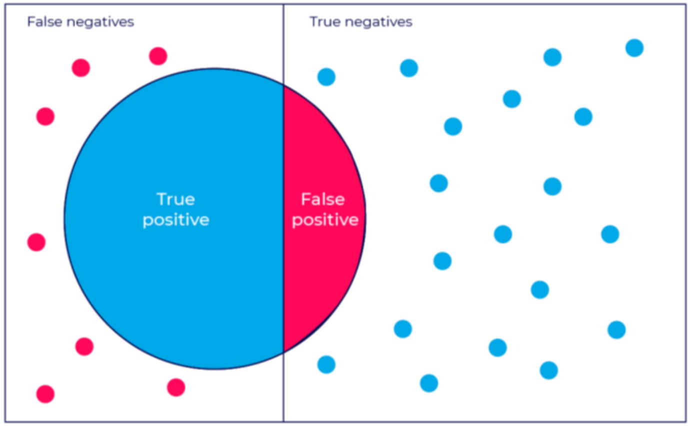
< Road anomaly >

Background

- Anomaly Detection (AD)
  - Metrics
    - Average precision (AP)
      - Area under the graph line in the precision-recall curve (AUPRC)
    - False positive rate at 95% true positive rate (FPR95)
      - Percentage of negative answers judged to be correct



< AP >

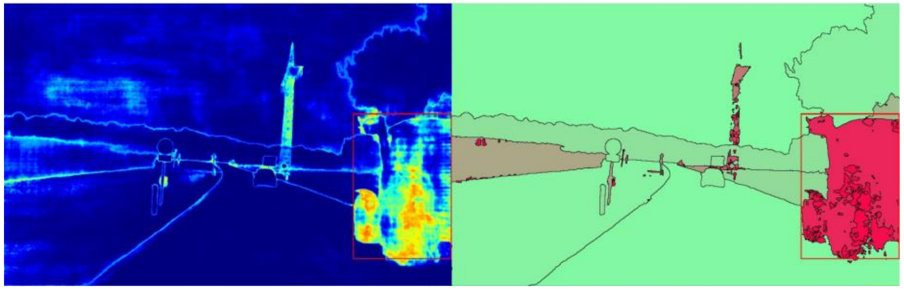


< FPR95 >

Background

- Anomaly Detection (AD)
  - Previous anomaly segmentation methods
    - Uncertainty-based anomaly segmentation
      - Estimate higher uncertainty for inputs that are anomalies than normal based on segmentation prediction map
      - Estimated uncertainty is often high at object boundaries
      - Yield many false positive predictions on the pixel level

$$H_x = -\sum_{c \in \text{classes}} p(c)\, log_2\, p(c)$$



< Softmax entropy based anomaly segmentation >

< Inability case >

Background

- Anomaly Detection (AD)
  - Previous anomaly segmentation methods
    - Anomaly segmentation via outlier exposure
      - Outliers from ImageNet or COCO void class of Cityscape
      - Requires re-training the segmentation network as a multi-task model
      - Outlier exposure can lead to a deterioration of the segmentation of inliers

< Outlier training data >

< deterioration of the segmentation of inliers >

Background

- Anomaly Detection (AD)
  - Previous anomaly segmentation methods
    - Inabilities of existing anomaly segmentation methods

Previous anomaly segmentation methods Limits

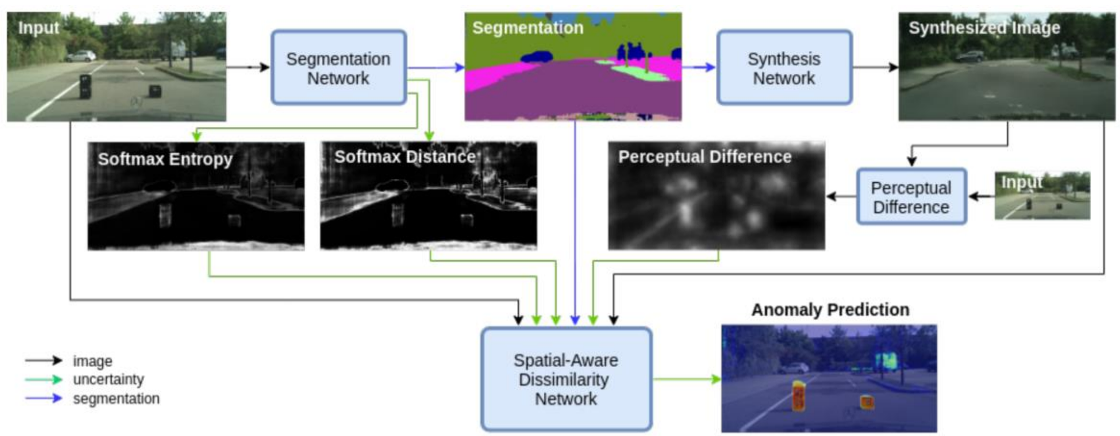
2개의 그림에서 다른한곳을 찾으세요

< Input image >

< re-synthesized the image >

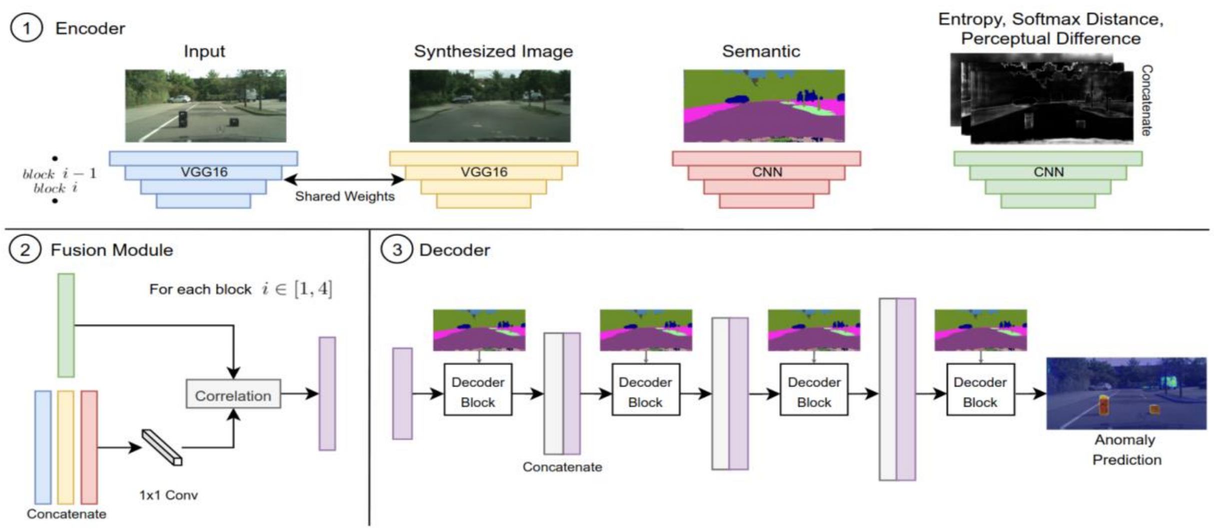
Synboost

- [1] Pixel-wise Anomaly Detection in Complex Driving Scenes (CVPR, 2021)
  - Framework
    - Extract softmax entropy, softmax distance, segmentation map
      - Softmax entropy : Segmentation prediction Entropy per each pixel
      - Softmax distance : Difference between the two largest softmax values
      - Segmentation map : Synthesize the image based on the segmentation map
    - Detect anomalies by comparing input and synthesized image



Synboost

- [1] Pixel-wise Anomaly Detection in Complex Driving Scenes (CVPR, 2021)
  - Dissimilarity module
    - Combines features to predict the anomaly segmentation map
    - Feature extractor of Input and synthesized image share same weights
    - For detect various scale of objects, concatenate 4 kind scales of correlated blocks at decoder blocks



Synboost

- [1] Pixel-wise Anomaly Detection in Complex Driving Scenes (CVPR, 2021)

Synboost

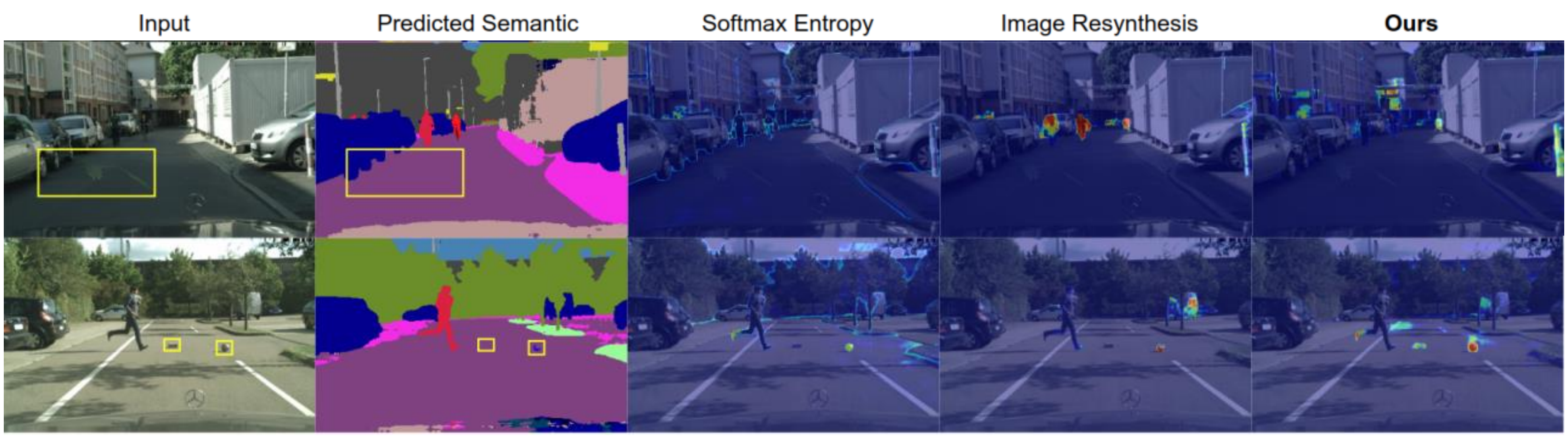
- [1] Pixel-wise Anomaly Detection in Complex Driving Scenes (CVPR, 2021)
    - Experiment

	Method	FS L&F ↑AP	FS L&F ↓FPR95	FS Static ↑AP	FS Static ↓FPR95	FS Web Oct. 2020 ↑AP	FS Web Oct. 2020 ↓FPR95	CS ↑mIOU
No Retrain	Softmax Entropy [14]	2.93	44.83	15.41	39.75	16.61	39.79	80.30
	Embedding Density [4]	4.65	24.36	62.14	**17.43**	29.16	38.80	80.30
	Image Resynthesis++ [24]	5.70	48.05	29.60	27.13	12.46	51.29	**83.50**
	Ours	***43.22***	***15.79***	***72.59***	*18.75*	***61.31***	***18.89***	***83.50***
Retrain	Bayesian DeepLab [27]	9.81	38.46	48.70	15.50	35.80	25.67	73.80
	Dirichlet DeepLab [26]	34.28	47.43	31.30	84.60	30.02	76.62	70.50
	Outlier Head [3]	30.92	22.18	**84.02**	**10.34**	**63.99**	**18.79**	77.30

Table 1. **Comparison between anomaly segmentation methods.** Our method achieves higher AP and lower FPR95 than previous methods that do not compromise segmentation performance (class mIOU on Cityscapes). It also achieves second-best performance when compared to all existing approaches.

Synboost

- [1] Pixel-wise Anomaly Detection in Complex Driving Scenes (CVPR, 2021)
  - Failure cases
    - When both softmax entropy and image resynthesis fail to detect an anomaly object, it will always display wrong final prediction
    - Not accurate to detecting small and distant anomalous objects



PEBAL

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)

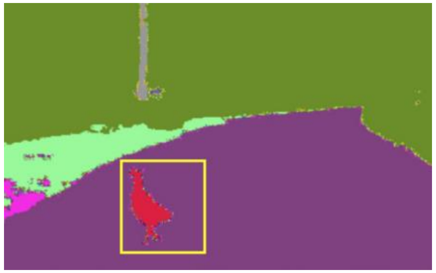
Synboost Limits

ℓ_{pal} ℓ_{ebm} ℓ_{reg}

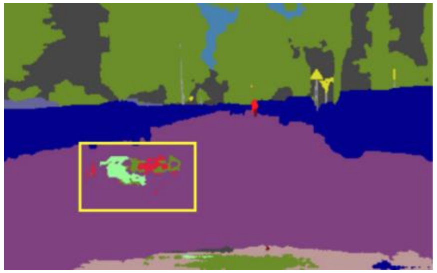
PEBAL ℓ_{pal} ℓ_{ebm} ℓ_{reg}

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)
  - Abstention learning : "Add an anomaly class"
    - Using Abstention learning to solve overconfident segmentation problem
    - Requires penalty factor to regularize the classification of anomalous pixels
    - Adaptive penalties are required for different pixels in a complex driving scene

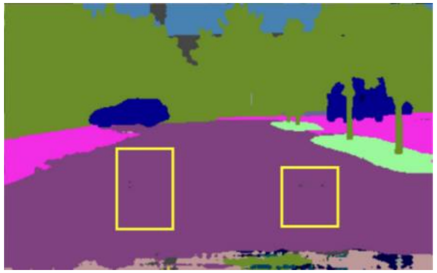
Penalty factor $a_\omega = (-E_\theta(\mathbf{x})_\omega)^2$



< Miss classified >



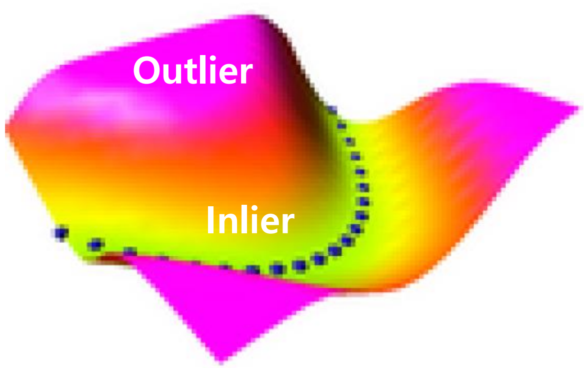
< Noisy classified >



< Non detected >

PEBAL ℓ_{pal} ℓ_{ebm} ℓ_{reg}

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)
  - Energy-biased abstention learning
    - Basically, Energy based on physics
      - The lower the energy, the more stable, the higher the energy, the more unstable
    - $E_\theta(\mathbf{x})_\omega$ computed with the logsumexp operator
    - Penalty factor a_ω using $E_\theta(\mathbf{x})_\omega$
      - Outlier : high energy and lower penalty factor
      - Inlier : low energy and higher penalty factor



$$E_\theta(\mathbf{x})_\omega = -\log \sum_{y \in \{1,\ldots,Y\}} \exp(f_\theta(y;\mathbf{x})_\omega)$$

Penalty factor $a_\omega = (-E_\theta(\mathbf{x})_\omega)^2$

PEBAL ℓ_{pal} ℓ_{ebm} ℓ_{reg}

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)
  - ℓ_{pal}
    - [2] Deep Gamblers: Learning to Abstain with Portfolio Theory

$$\max_f W(\mathbf{b}(f), \mathbf{p}) = \max_{\mathbf{w}} \sum_i^B \log \left[f_{\mathbf{w}}(x_i)_{j(i)} o + f_{\mathbf{w}}(x_i)_{m+1} \right]$$

$\Pr(j|x) = f(x)_j$

✓ Probability of x feature belongs to j class

$\max \mathbb{E}[\log p(j|x)] = \max_{\mathbf{w}} \mathbb{E}[\log f_{\mathbf{w}}(x)_j]$

✓ Learning to find the w value with max probability that the x feature belongs to j class

Classification as a horse race

✓ o value means betting money

✓Adding abstention term to consider the horse loses the game

PEBAL ℓ_{pal} ℓ_{ebm} ℓ_{reg}

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)
  - ℓ_{pal}

$$\ell_{pal}(\theta, \mathbf{y}, \mathbf{x}, E_\theta(\mathbf{x})) = -\sum_{\omega \in \Omega} \log \left(f_\theta(y_\omega; \mathbf{x})_\omega + \frac{f_\theta(Y+1; \mathbf{x})_\omega}{a_\omega} \right)$$

$f_\theta(y_\omega; \mathbf{x})_\omega$: logit of inlier class
$f_\theta(Y+1; \mathbf{x})_\omega$: logit of outlier class

- Adaptive penalties are required for different pixels in a complex driving scene
- ℓ_{pal} mitigate the overconfident problem of segmentation network
  - Inlier case
    - ✓ High penalty factor means inlier logit is dominant
  - Outlier case
    - ✓ Low penalty factor means inlier logit and outlier logit are dominant

Penalty factor $a_\omega = (-E_\theta(\mathbf{x})_\omega)^2$

$$E_\theta(\mathbf{x})_\omega = -\log \sum_{y \in \{1,\ldots,Y\}} \exp(f_\theta(y; \mathbf{x})_\omega)$$

서강대학교 SOGANG UNIVERSITY

VDS LAB

PEBAL $\boxed{\ell_{pal} \quad \ell_{ebm} \quad \ell_{reg}}$

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)
  - ℓ_{ebm}
    - EBM losses effectively create an energy gap between normal and abnormal pixels
    - Inlier case
      - If energy of inlier pixel is higher than m_{in}, loss reduce the energy of liner pixel
    - Outlier case
      - If energy of inlier pixel is lower than m_{out} , loss reduce the energy of liner pixel

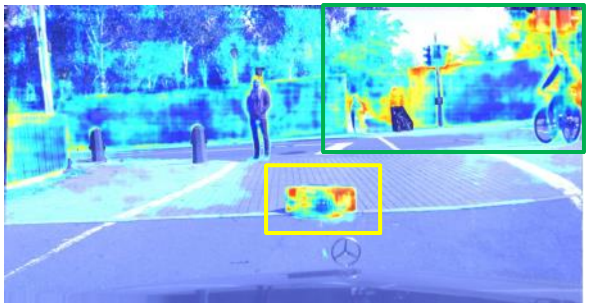
$$\ell_{ebm}^{in}(E_\theta(\mathbf{x})) = \sum_{\omega \in \Omega} \left(\max(0, E_\theta(\mathbf{x})_\omega - m_{in})\right)^2 \qquad \ell_{ebm}^{out}(E_\theta(\mathbf{x})) = \sum_{\omega \in \Omega} \left(\max(0, m_{out} - E_\theta(\mathbf{x})_\omega)\right)^2$$

[1] Tian, Yu, et al. "Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes." (ECCV 2022)

PEBAL $\boxed{\ell_{pal} \quad \ell_{ebm} \quad \ell_{reg}}$

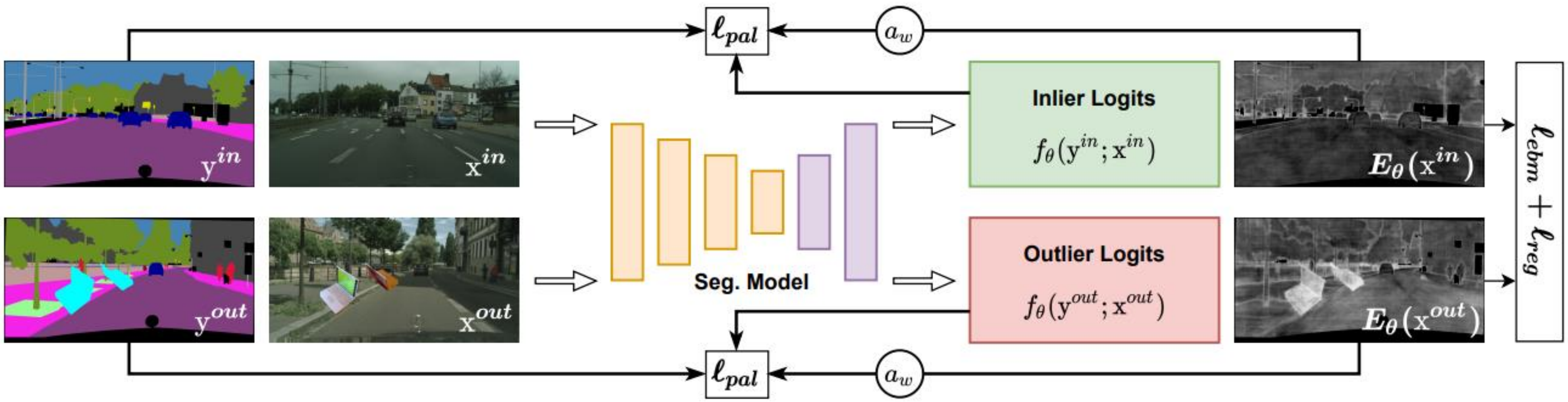
- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)
  - ℓ_{reg}
    - Smoothness regularization
      - Fail to classify neighboring anomaly pixels inside boundary of anomalous object
      - Neighboring pixels are not that abruptly different energy
    - Sparsity regularization
      - Anomalous objects are rare in driving scene
      - Most of the pixels are inlier classes, so their energy is low

$$\ell_{reg}(E_\theta(\mathbf{x})) = \sum_{\omega \in \Omega} \underbrace{\beta_1 |E_\theta(\mathbf{x})_\omega - E_\theta(\mathbf{x})_{\mathcal{N}(\omega)}|}_{\text{Smoothness}} + \underbrace{\beta_2 |E_\theta(\mathbf{x})_\omega|}_{\text{Sparsity}},$$



PEBAL ℓ_{pal} ℓ_{ebm} ℓ_{reg}

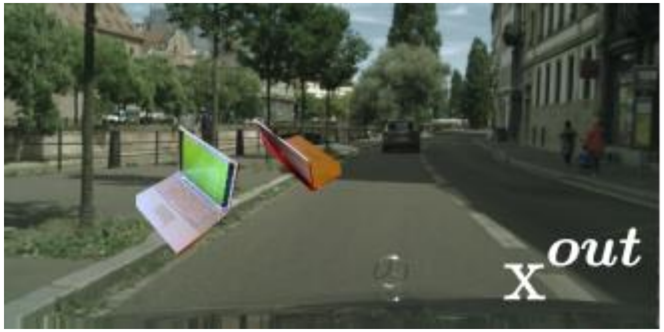
- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)



- ℓ_{pal} : Mitigate the overconfident problem of segmentation network
- ℓ_{ebm} : Effectively create an energy gap between normal and abnormal pixels
- ℓ_{reg} : Considering smoothness and sparsity of anomalous objects

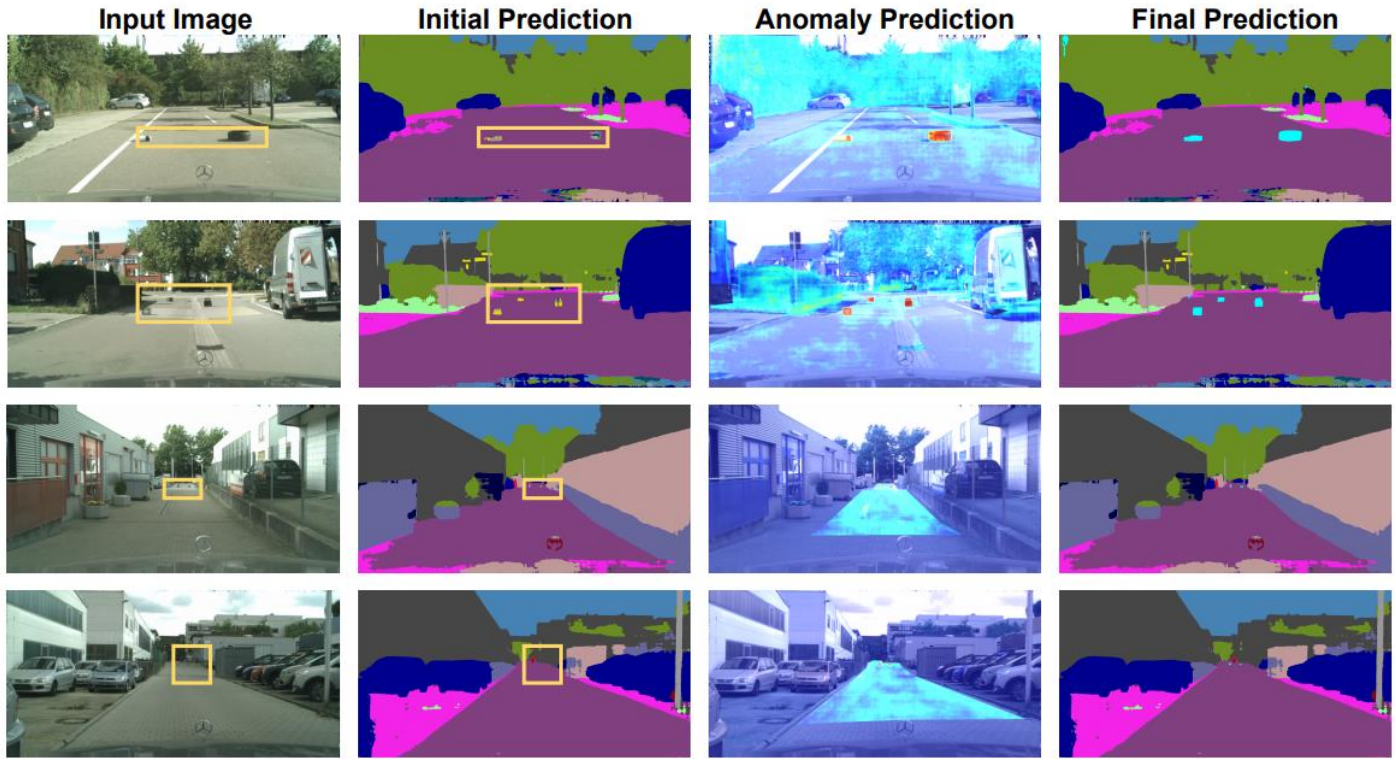
PEBAL ℓ_{pal} ℓ_{ebm} ℓ_{reg}

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)
  - AnomalyMix
    - Cuts the anomalous objects from an outlier dataset and paste them into the images of the inlier dataset
      - Containing a combination of inlier and outlier pixels
    - Balanced learning and keeping the visual context of inlier labels
    - Can form a potentially infinite number of training images for outlier dataset



PEBAL $\boxed{\ell_{pal} \quad \ell_{ebm} \quad \ell_{reg}}$

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)



PEBAL ℓ_{pal} ℓ_{ebm} ℓ_{reg}

- [1] Pixel-wise Energy-biased Abstention Learning for Anomaly Segmentation on Complex Urban Driving Scenes (ECCV, 2022)

Models	re-training	Extra Network	OoD Data	FS LostAndFound		FS Static	
				AP ↑	FPR95 ↓	AP ↑	FPR95 ↓
Discriminative Outlier Detection Head [2]	✔	✔	✔	31.31	19.02	96.76	0.29
MSP [18]	✘	✘	✘	1.77	44.85	12.88	39.83
Entropy [19]	✘	✘	✘	2.93	44.83	15.41	39.75
SML [21]	✘	✘	✘	31.05	21.52	53.11	19.64
kNN Embedding - density [4]	✘	✘	✘	3.55	30.02	44.03	20.25
Bayesian Deeplab [35]	✔	✘	✘	9.81	38.46	48.70	15.05
Density - Single-layer NLL [4]	✘	✔	✘	3.01	32.9	40.86	21.29
Density - Minimum NLL [4]	✘	✔	✘	4.25	47.15	62.14	17.43
Image Resynthesis [31]	✘	✔	✘	5.70	48.05	29.6	27.13
OoD Training - Void Class	✔	✘	✔	10.29	22.11	45.00	19.40
Dirichlet Deeplab [34]	✔	✘	✔	34.28	47.43	31.30	84.60
Density - Logistic Regression [4]	✘	✔	✔	4.65	24.36	57.16	13.39
SynBoost [11]	✘	✔	✔	43.22	15.79	72.59	18.75
Ours	✘	✘	✔	**44.17**	**7.58**	**92.38**	**1.73**

Thank you for listening