Maskomaly: Zero-shot Mask Anomaly Segmentation

SCAN ME

Jan Ackermann, Christos Sakaridis, Fisher Yu

https://github.com/jan-ackermann/maskomaly

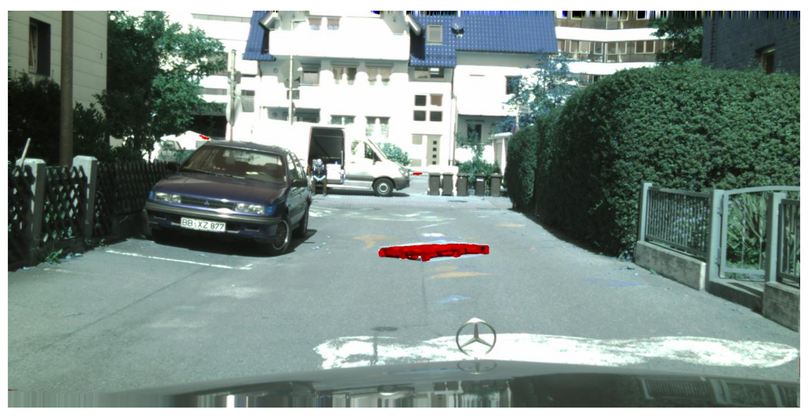
ETH zürich | CVL Computer Vision Lab

Anomaly Segmentation

Given set of **inlier classes** and **image,** segment regions that semantically not belong to any of the inlier classes

Datasets for autonomous driving consider anomalous what is not a class in **CityScapes**

Only the wooden pallet is marked red because it is not a class in CityScapes



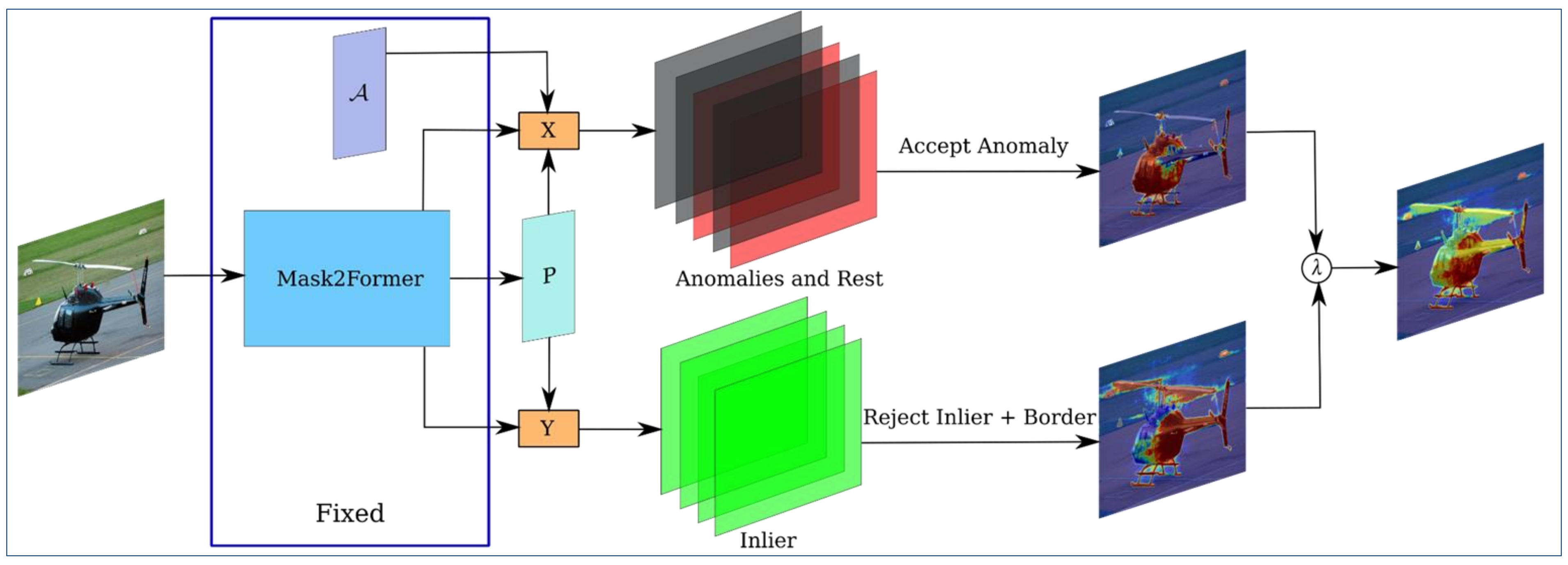
MDM – Maximal Detection Margin

- Designing method around AP is not practical (only requires existence of a good threshold)
- Need for metric that captures robustness regarding threshold choice

Columns: dense prediction, thresholded at 0.3, 0.5, 0.8
Lower column is more desirable, but AP is close (95.5% vs. 98.9%)

$$\mathrm{MDM}_d^{T_{\mathrm{margin}}}(A,B) = \max_{0\le x\le y\le 1} (y-x)\cdot \mathbb{1}\{\forall z\in[x,y]: d(A>z,B) > T_{\mathrm{margin}}\}.$$

Maskomaly



- Pure **Post-processing: no training**
- Builds up on **Mask-based** Segmentation Networks
- Combines idea of **rejecting** inlier areas and **accepting** anomalous regions
- Deals with **border regions** explicitly

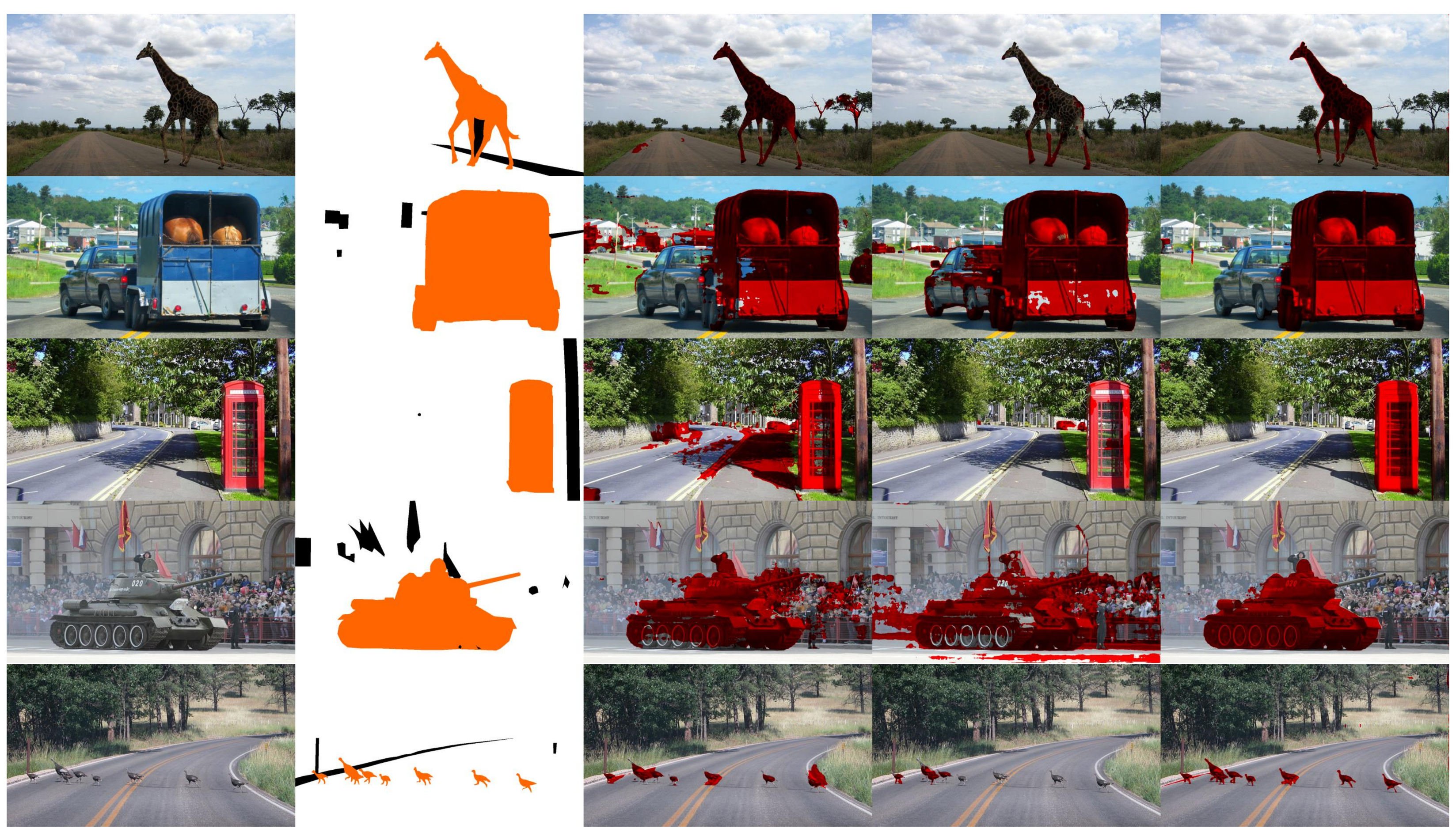
Algorithm 1 Maskomaly, Hyperparameters $=\{T_{\text{mask}}, T_{\text{b}}, \varepsilon_{\text{b}}, \lambda\}$

1: **function** MASKOMALY($M, P, \mathcal{A}, \mathcal{I}$)
2:   **for** $1 \le n \le N$ **do** ▷ Add inlier masks to $\mathcal{I}$
3:     **if** $(\operatorname{argmax}_{1\le l\le C+1} p_n[l]) \ne C+1 \wedge (\max_{1\le l\le C+1} p_n[l]) \ge T_{\text{mask}}$ **then**
4:       $\mathcal{I} \leftarrow \mathcal{I} \cup \{n\}$
5:   **for** $i,j \in H \times W$ **do** ▷ Reject inlier pixels
6:     $o_{\text{reject}}[i,j] \leftarrow \min_{n\in\mathcal{I}}(1 - m_n[i,j]\cdot \max_{1\le l\le C} p_n[l])$
7:   **for** $\{k,n\} \subseteq \mathcal{I}$ **do** ▷ Reject borders
8:     $b \leftarrow (m_k > T_{\text{b}}) \odot (m_n > T_{\text{b}})$
9:     $b \leftarrow \min(1 - b + \varepsilon_{\text{b}}, 1)$
10:     $o_{\text{reject}} = \min(o_{\text{reject}}, b)$
11:   **for** $i,j \in H \times W$ **do** ▷ Accept anomalous predictions
12:     $o_{\text{accept}}[i,j] \leftarrow \max_{n\in\mathcal{A}}(m_n[i,j]\cdot p_n[C+1])$
13:   **return** $\lambda \cdot o_{\text{reject}} + (1-\lambda)\cdot o_{\text{accept}}$ ▷ Interpolate reject and accept scores

- **Key Insight**: Mask-based Segmentation Networks learn **fixed queries** that predict masks for anomalies which are discarded by the segmentation algorithm

Used Discarded

Results



Input Ground Truth Max. Entropy ObsNet Maskomaly

- No missing patches inside segmented anomaly
- High accuracy around borders
- No predictions in non-anomalous areas
- Visualization threshold by optimizing F1

		SMIYC				
Method	Aux. data	AP↑	FPR95↓	sIoU gt↑	PPV↑	Mean F1↑
DenseHybrid	Yes	78.0	9.8	54.2	24.1	31.1
Max. Entropy	Yes	85.5	15.0	49.2	39.5	28.7
EAM	Yes	93.8	**4.1**	**67.1**	**53.8**	**60.9**
RbA	Yes	**94.5**	4.6	64.9	47.5	51.9
DenseHybrid	No	51.5	33.2	-	-	-
ObsNet	No	75.4	26.7	44.2	52.6	45.1
EAM	No	76.3	93.9	-	-	-
RbA	No	86.1	15.9	**56.3**	41.4	42.0
Maskomaly	No	**93.4**	**6.9**	55.4	**51.5**	**49.9**

		RoadAnomaly				FishyScapes Static	
Method	Aux. data	AP↑	FPR95↓	MDM↑	MDM↑	AP↑	FPR95↓
SynBoost	Yes	38.2	64.8	0.0	0.0	66.4	25.6
PEBAL	Yes	45.1	44.6	-	-	**92.1**	**1.5**
DenseHybrid	Yes	63.9	43.2	-	-	60.0	4.9
Max. Entropy	Yes	**79.7**	**19.3**	**25.2**	**9.2**	76.3	7.1
ObsNet	No	54.7	60.0	5.1	0.0	9.4	47.7
GMMSeg	No	57.7	44.3	-	-	82.6	-
EAM	Yes	66.7	13.4	-	-	**87.3**	**2.1**
Maskomaly	Yes	**70.9**	**11.9**	**67.1**	**35.9**	69.5	14.4

Ablations:

				RoadAnomaly			SMIYC	
Acc.	Rej.	Bord.	Init.	AP↑	FPR95↓	AUC↑	AP↑	FPR95↓
Yes	No	No	No	46.0	26.3	89.2	-	-
No	Yes	No	No	45.3	19.3	91.3	-	-
No	Yes	Yes	No	45.5	15.7	92.0	58.4	23.4
Yes	Yes	Yes	No	63.2	14.8	94.2	-	-
Yes	Yes	Yes	Yes	70.9	11.9	95.5	93.4	6.9

Experimental configuration:

- Backbone: **Mask2Former/Swin-L** trained on **Cityscapes** for segmentation
- Validation set: **SMIYC Validation**
- Evaluated on: SMIYC, FishyScapes, RoadAnomaly, StreetHazards

Quantitative Results:

- Maskomaly achieves **state-of-the-art** performance
- MDM is more informative than AP.

- Incrementally adding components always leads to an improvement in AP