KI WISSEN
Automotive AI Powered by Knowledge

1.4 Static Scene Knowledge for Pedestrian Detection

Vera Stehr | Valeo

Detection

Knowledge Integration into Architecture

Idea

- Static cameras on walls or ceiling

→ Static field of view

→ Use static image (without pedestrians, cars, mobile objects) as prior knowledge

Figure 1: example image pair - left the live camera image, right the corresponding static image

Background Aware Shared Resnet (BASR)

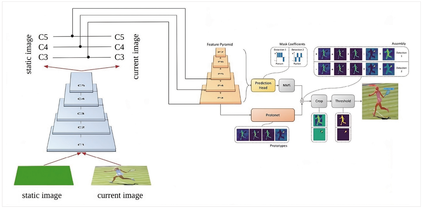

Figure 2: Architecture of BASR, based on YOLACT [1] (adapted from [1])

- Both images pass the same backbone with the same weights
- Merge of C3, C4, C5 feature maps of static and current image

Background Aware Individual Resnet (BAIR)

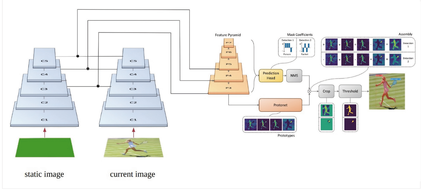

Figure 3: Architecture of BAIR, based on YOLACT [1] (adapted from [1])

- Separate backbones (weights) for static and current image

Results

	bounding box		mask	
	recall	av. precision	recall	av. precision
YOLACT [1]	69.2 %	63.5 %	59.8 %	53.4 %
BASR (C4, C5)	70.7 %	64.5 %	61.0 %	54.2 %
BAIR (C3, C4, C5)	71.5 %	66.1 %	60.6 %	54.2 %
improv. BASR	1.5 %	1.0 %	1.2 %	0.8 %
improv. BAIR	2.3 %	2.6 %	0.8 %	0.8 %

Table 1: Best results for the BASR and BAIR architecture. The layers in parentheses are those which have been merged.

Conformity Check with Knowledge

Idea

- Filtering out false positive detections
- Compare content of predicted bounding box with same cutout in empty image

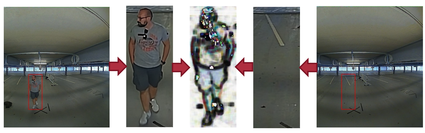

Figure 4: Detection in the camera image (left), the same corresponding empty image (right) and the difference of the cutout calculated with the structural similarity (SSIM).

Normalization

- Diminishes photometric disturbances between live and static image

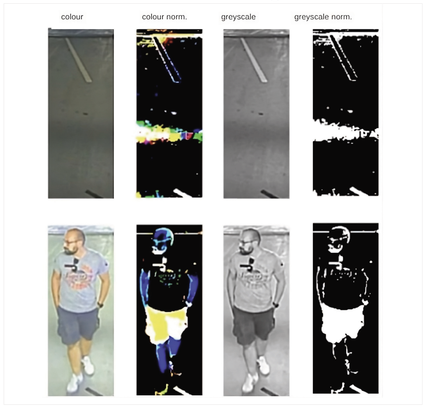

Figure 5: Detected object (photometrically disturbed) and the corresponding cutout in the empty image (first column), normalized images (second column), grey converted images (third column) and grey converted normalized images (fourth column).

Results

empty – pedestrian photometrically disturbed				
	average	std. dev.	min	max
color	0.53	0.05	0.42	0.63
color norm.	0.47	0.01	0.46	0.49
grey	0.54	0.04	0.43	0.64
grey norm.	0.47	0.01	0.45	0.49
empty – empty photometrically disturbed				
	average	std. dev.	min	max
color	0.95	0.05	0.73	1.00
color norm.	0.95	0.02	0.89	1.00
grey	0.95	0.05	0.74	1.00
grey norm.	0.95	0.03	0.88	1.00

Table 2: Statistical analysis with 1000 image pairs each, where the respective second image is randomly photometrically disturbed (brightness, contrast, saturation, hue). Average, standard deviation, minimum and maximum of the SSIM is calculated.

Recall of the filter	99 %
Precision of the filter	96 %
Specifity of the filter	88 %

Table 3: Recall (probability that a true positive detection passes the filter, precision (probability that an object that passes the filter is actually an object) and specifity (probability that a detection that is filtered out is actually false positive).

References:

[1] Bolya et al., IEEE/CVF Int. Conf. on Computer Vision, 9156 (2019)
[2] Wang et al., IEEE Transact. on Image Processing, 600 (2004)

Partners

Continental, Bosch, Valeo, [at], AVL, BTC embedded systems, Capgemini engineering, e:fs TechHub GmbH, DFKI, DLR, fortiss, Fraunhofer IAIS, Fraunhofer FOKUS, FZI, Universität des Saarlandes, bast

External partners

eject

For more information contact:
vera.stehr@valeo.com

KI Wissen is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

KI FAMILIE

VDA LEITINITIATIVE

Funded by the European Union NextGenerationEU

Supported by: Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Bundestag