Supplementary Materials

Traffic Sign Detection and Classification on the Austrian Highway Traffic Sign Data Set

1 Schema of Metadata Tables

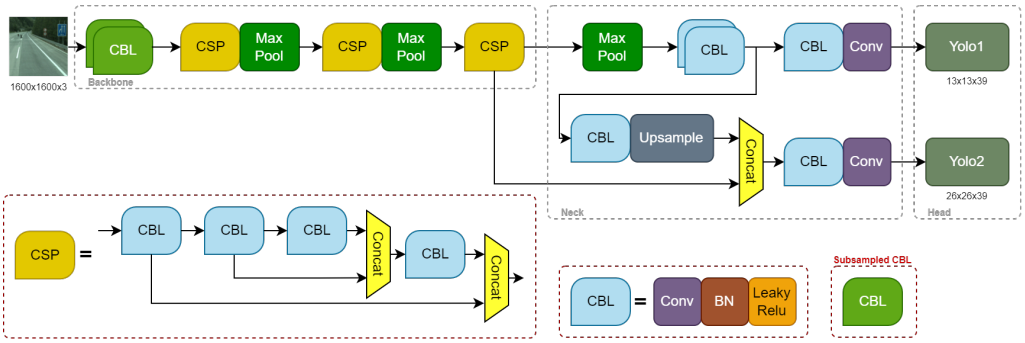
Column	Type	Description
image_id	Integer	Identifier of scene image containing the annotation; equals file name of corresponding JPEG file
annotation_id	Integer	Unique identifier of each annotation; linked to column **id** in metadata table of ATSD-Signs (Table 2)
xtl	Float	x-coordinate of top-left corner of bounding box, in pixels
ytl	Float	y-coordinate of top-left corner of bounding box, in pixels
xbr	Float	x-coordinate of bottom-right corner of bounding box, in pixels
ybr	Float	y-coordinate of bottom-right corner of bounding box, in pixels
group_id	Integer	Identifier for grouping signs; rows with same non-negative group_id semantically belong together, e. g., as main sign and additional panel
type	Category	Type (material) of annotated traffic sign; one of 'plate', 'led', 'prismatic', 'backlit'
not_normal_to_roadway	Bool	Whether traffic sign is not oriented normal to roadway; cf. Figure 6a in main text
unusual_sign	Bool	Whether traffic sign is in some sense 'unusual'; cf. Figure 6j in main text
crossed_out	Bool	Whether traffic sign is crossed out; cf. Figure 6c in main text
weather	Category	Weather conditions when image was acquired; either 'normal' or 'rain'
lighting	Category	Lighting conditions under which image was acquired; one of 'normal', 'underexposed', 'overexposed', 'shades'
fog	Bool	Whether there was fog when image was acquired; cf. Figure 6e in main text
tunnel	Bool	Whether image was acquired in a tunnel; cf. Figure 6f in main text
damaged	Bool	Whether traffic sign is damaged; cf. Figure 6g in main text
trimmed	Bool	Whether traffic sign is not visible entirely on scene image; cf. Figure 6h in main text
covered	Bool	Whether traffic sign is partly occluded, e. g., by vehicles or plants; cf. Figure 6i in main text
multiple_signs_visible	Bool	Whether multiple signs are visible in annotated bounding box; cf. Figure 6b in main text
caption	Bool	Whether image caption is visible in annotated bounding box; cf. Figure 6k in main text
class_id	String	Identifier of class annotated traffic sign belongs to; two-digit category identifier followed by '_' followed by two-digit class identifier; same as in ATSD-Signs

Supplementary Table S1: Schema of ATSD-Scenes metadata tables ('meta train.csv', 'meta test.csv'). Rows correspond to individual traffic sign annotations in traffic scene images.

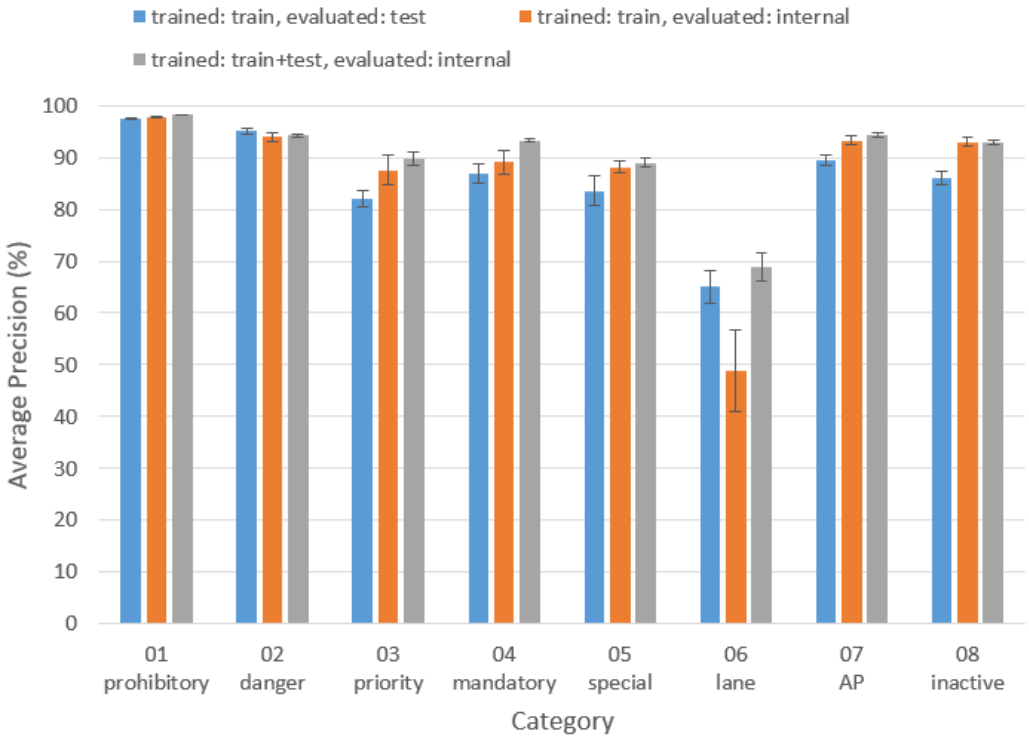
Column	Type	Description
id	Integer	Unique identifier of each patch; equals file name of corresponding JPEG file
type	Category	Type (material) of traffic sign visible on each patch; one of ‘plate’, ‘led’, ‘prismatic’, ‘backlit’
not_normal_to_roadway	Bool	Whether traffic sign is not oriented normal to roadway; cf. Figure 6a in main text
fog	Bool	Whether there was fog when image was acquired; cf. Figure 6e in main text
tunnel	Bool	Whether image was acquired in a tunnel; cf. Figure 6f in main text
damaged	Bool	Whether traffic sign is damaged; cf. Figure 6g in main text
trimmed	Bool	Whether traffic sign is not visible entirely on the patch; cf. Figure 6h in main text
covered	Bool	Whether traffic sign is partly occluded, e. g., by vehicles or plants; cf. Figure 6i in main text
rain	Bool	Whether it was raining when image was acquired; cf. Figure 6d in main text
overexposed	Bool	Whether image is overexposed
underexposed	Bool	Whether image is underexposed
shades	Bool	Whether shadow is cast on parts of traffic sign

Supplementary Table S2: Schema of ATSD-Signs metadata tables (‘meta train.csv’, ‘meta test.csv’). Rows correspond to individual traffic sign patches.

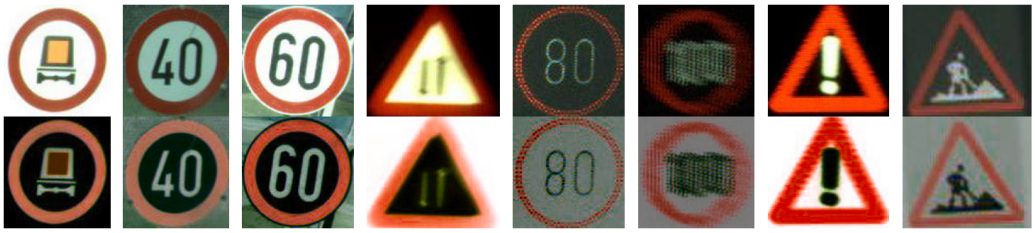
2 Details on Baseline Traffic Sign Recognition Models



Supplementary Figure S1: Architecture of the Yolov4-tiny detection model. This is a one-stage detector consisting of three main parts: backbone, neck and head. The feature formation is done by the backbone using CSPDarknet53-tiny.The neck mixes and aggregates the features formed in the backbone using a feature pyramid network, which aggregates parameters from different backbone scales to a single feature map. The Yolo head uses this feature map to locate the bounding boxes and classify the detected objects.



Supplementary Figure S2: Average precision of each category including the standard deviation averaging three indepen-dent models. Category 'lane' scores the lowest average precision among all categories, which may be caused by the small number of samples in the data set. In contrast, 'prohibitory' achieves an average precision of up to 98.20%.



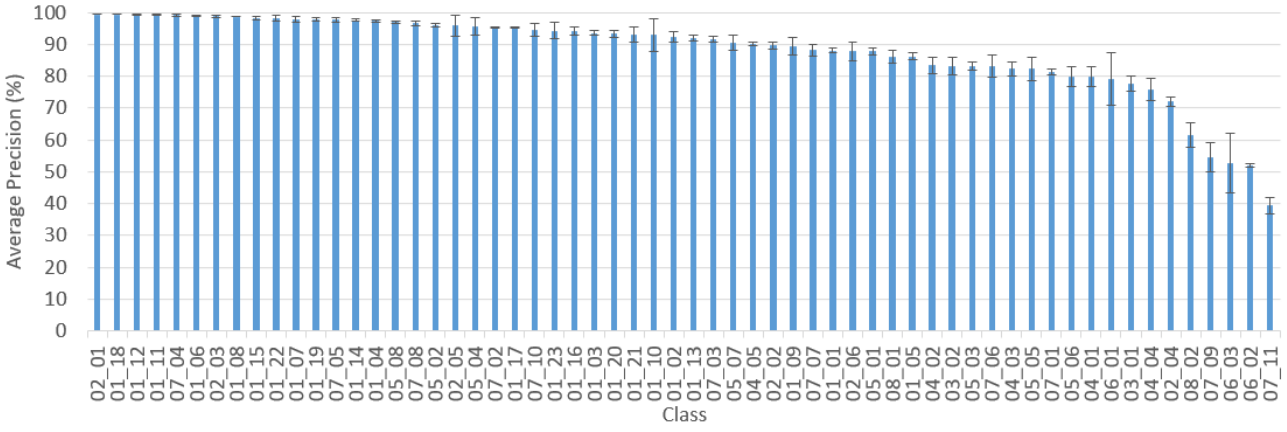
Supplementary Figure S3: Exemplary results of our LED augmentation. Top: original image; bottom: transformed image (non-LED → LED, and vice versa).

Augmentation	Test Set		Internal Set	
	accuracy (%)	balanced acc. (%)	accuracy (%)	balanced acc. (%)
none	97.15±0.19	94.97±0.69	96.63±0.11	94.38±0.51
geometric	97.39±0.40	95.28±0.59	96.90±0.48	94.18±0.64
geometric+LED	97.61±0.24	95.75±0.53	97.20±0.21	94.94±0.06
none			97.82±0.07	96.02±0.18
geometric			98.18±0.11	97.14±0.20
geometric+LED			98.27±0.25	97.65±0.28

Supplementary Table S3: Classification results on the public test set and the internal set, averaged over three independent runs (models trained with same hyperparameters but different random seeds). Accuracy (Acc.) and balanced accuracy (BAcc.) are displayed as *mean*±*SD*. Rows without results on the test set refer to models trained on all public data, including both training- and test set.

Augmentation	Test Set mAP (%)		Internal Set mAP (%)	
	keep	suppress	keep	suppress
none	87.74±2.29	87.40±1.87	89.03±2.34	89.49±2.24
geometric	87.56±2.28	87.43±1.69	89.04±2.10	89.47±2.04
geometric+LED	87.87±2.29	87.74±1.71	89.30±2.22	90.07±1.96
none			91.68±0.67	91.97±0.71
geometric			91.69±0.72	92.01±0.71
geometric+LED			91.99±0.66	92.46±0.68

Supplementary Table S4: Results of the full detection- and classification pipeline on the public test set and the internal set, averaged over nine independent runs (three detectors combined with three classifiers in round-robin fashion). Mean average precision (mAP) is displayed as *mean*±*SD*. Rows without results on the test set refer to models trained on all public data, including both training- and test set. 'suppress' means that detections with disagreement between detector and classifier were discarded, 'keep' means that disagreements were not taken into account.



Supplementary Figure S4: Average precision of each class, for full pipelines evaluated on the public test set without suppression of detections based on detection- and classification disagreement. Classifiers were trained with geometric and LED augmentation. The two classes that can be recognized best are 'Wild animals_crossing' (02 01; 99.50±0.00%AP) and 'No trucks' (01 18; 99.50±0.00% AP), and the two classes that cause most problems are 'Additional info: except with yellow-red rotating light' (07 11; 39.40±5.28% AP) and 'Lane open' (06 02; 52.01±1.34% AP).