IDD: A Dataset for exploring problems in Autonomous Navigation in Unconstrained Environments

Girish Varma[1], Anbumani Subramanian[2], Anoop Namboodiri[1], Manmohan Chandraker[3], C V Jawahar[1]
1. IIIT Hyderabad 2. Intel Bangalore 3. University of California, San Diego

http://idd.insaan.iiit.ac.in/

IEEE 2019 Winter Conference on Applications of Computer Vision

Autonomous Navigation Dataset

- Images from road scenes.
- Pixel level/Bounding box Annotations.
- Semantic/instance segmentation, Detection.
- A basic primitive for Autonomous Navigation.

Cityscapes: Structured Traffic

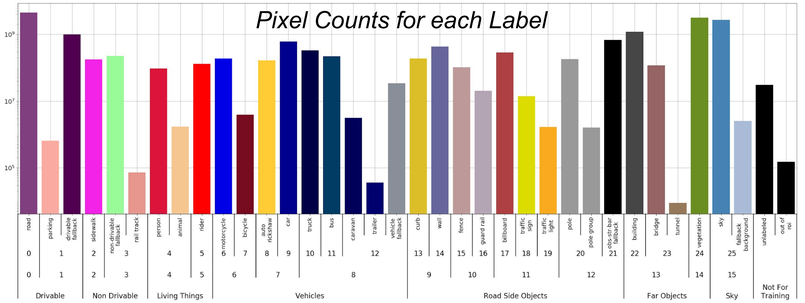
Unstructured Driving Conditions: A Challenge for Cityscapes Trained Model

Our Dataset

Largest dataset with calibrated camera setup.

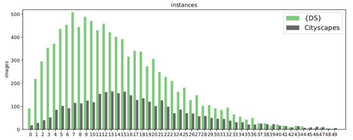
Dataset	Calibration	Nearby frames / Video	Distortion /Night	#Images/ #Sequences	#Labels Train/Total	Average Resolution
Cityscapes [5]	✓	✓		5K / 50	19/34	2048x1024
IDD	✓	✓		10K / 180	30/34	1678x968
BDD100K [26]		✓	✓	10K / 10K	19/30	1280x720
MVD [16]				25K / -	65/66	>1920x1080

Label Hierarchy & Statistics



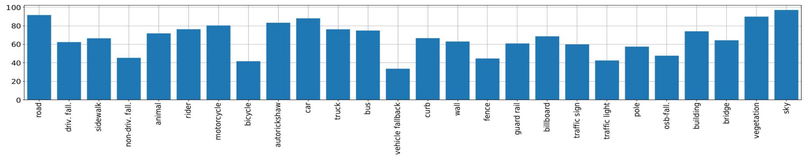
Instance Counts Comparison with Cityscapes

We have more pixels of truck, bus, motorcycle, guard rail, bridge and rider. The pixel counts for new labels are also high. For traffic participants, we have almost double the counts compared to CS.



Benchmarking

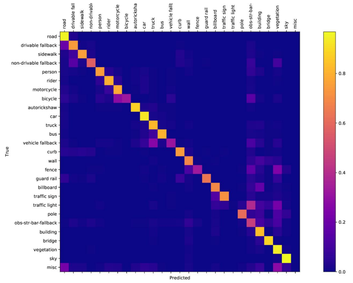
Trained DRN-D38 Model : Analysis



Mean Iou: 66%

Some most ambigous classes

- motorcycle and bicycle.
- billboard and traffic sign.
- obs-str-bar-fallback, vegetation & traffic-light.
- building and billboard.
- vegetation and wall, pole, fence.
- drivable, non-drivable, vegetation.



Domain Discrepancy

Train	Test	road	sidewalk	person	motorcycle	bicycle	car	truck	bus	wall	fence	traffic sign	traffic light	pole	building	vegetation	sky	mIoU of common labels
CS	DS	72	22	30	47	10	58	30	19	17	13	19	8	23	32	76	68	34
DS	CS	81	26	74	34	55	85	16	17	21	24	25	21	47	77	90	88	49
BD	ID	83	0	38	44	2	52	21	13	0	0	0	0	36	42	83	94	32
ID	BD	84	16	57	34	44	77	14	24	10	33	18	13	41	68	82	87	44
CS	CS	98	84	81	60	76	94	56	78	49	58	77	67	62	92	92	94	76
MV	MV	85	58	73	55	61	90	61	65	45	58	72	67	50	86	90	98	70
ID	ID	92	68	73	80	42	89	79	78	64	45	60	38	58	75	90	97	70
BD	BD	95	62	61	32	22	90	52	57	25	45	52	58	49	85	87	97	60

Workshop & Challenge

AutoNUE, ECCV '18. Munich, Germany.
http://cvit.iiit.ac.in/scene-understanding-challenge-2018/
http://cvit.iiit.ac.in/autonue2018/

Method	AP	AP@50
*MaskRCNN [8] with ResNet101	0.268	0.499
*PANet [14]	0.376	0.661

Method	% mIoU at Levels		
	L1	L2	L3
ERFNet	-	-	55.4
DRN-D-38	85.9	72.6	66.6
*DeeplabV3+ [4]	89.8	78.0	74.0
*PSPNet [27]	89.9	78.0	74.1
*Wider Resnet-38, DeeplabV3 Decoder, Inplace ABN [20], Ensemble of 4	89.7	77.9	74.3

Qualitative Results