Supplementary Material for: Road Detection using Convolutional Neural Networks

Aparajit Narayan[1], Elio Tuci[2], Frédéric Labrosse[1], Muhanad H. Mohammed Alkilabi[1]

[1]Aberystwyth University, [2]Middlesex University London

[1]apn3,ffl,mhm1@aber.ac.uk, [2]E.Tuci@mdx.ac.uk

1 Datasets

1.1 K56

This dataset contains 154 images in an urban environment originally obtained from the KITTI dataset (see [1]). The images show well a demarcated (white lines) two lane highway road. The detection algorithm/method is requried to only consider the lane the recording platform was driving on (i.e the right lane). Apart from this other challenges include, shadows, variations in lane-markings and presence of cars in both lanes.

Figure 1: Frame from dataset K56.

1.2 Track

This dataset contains 1802 frames captured from an omni-directional camera mounted on a Pioneer 3-AT mobile robot available with the Computer Science Department at Aberystwyth University. The running track (Figure 5) is a highly visible manufactured surface accompanied by well-defined edge markings. This blue surface contrasts well against the white edges and central line as well as the grassy surroundings. However, the two adjacent running tracks are both some variation of blue, the robot driving on the lighter of the two tracks.

The challenges of the dataset include changes in surface color caused by leaves, shadows cast by nearby trees, occasional crossing white and yellow lines, and an intersection.

Figure 2: Frame from dataset Track.

1.3 Llan

This dataset contains 541 frames captured from an omni-directional camera mounted on the IRDIS rover robot (see [1]) available with the Computer Science Department at Aberystwyth University. The road consists of a mostly well-marked tarmac surface with mostly grass/shrubs on the sides. The challenges of this dataset include noise (a moving diagonal pattern in the original omnidirectional images resulting in hyperbolic lines forming on the unwrapped panoramic images due to noise in the camera), general lack of color contrast (due to an overcast sky), raised pedestrian crossings, a T-junction, and a sharp turn involving a brief change in the road surface.

Figure 3: Frame from dataset Llan.

1.4 Farm

This dataset contains 2998 frames captured from an omni-directional camera mounted on the IRDIS rover robot available with the Computer Science Department at Aberystwyth University. The challenges of this dataset include low contrast between the road and the surroundings with no crisp edge, surface discoloration due to wet patches, and changing viewpoint due to changes in the on-road position.

1.5 Rugged

This dataset contains 2760 frames captured from a GOPRO Hero4 camera (http://gopro.com). The road consists of a hilly path made of primarily dirt

Figure 4: Frame from dataset Farm.

and gravel (filmed after a spell of rain). Challenges include low contrast (due to overcast conditions), puddles, extremely delineated edges at some sequences and occasional sharp turns. There are also a few sudden changes in colour properties with the appearance of fences and barricades on the sides.

Figure 5: Frame from dataset Rugged.

1.6 Rain

This dataset contains 838 frames captured from a GOPRO Hero4 camera (http://gopro.com). The smooth tarmac road in this dataset was filmed immediately after a spell of rain had cleared up resulting in bright spots of reflection on the surface. The non-road surfaces on the sides are quite varied and include wooden barricades, grassy patches, shrubs, benches, thrash bins and even a some frames where a pedestrian crosses the robot while walking along the road edges.

Figure 6: Frame from dataset Rain.

1.7 Footpath

The road in this dataset is a marked tarmac surface that contrasts well against the mostly green (grass) and blue (running track) surroundings. This dataset comprises 1,551 images captured from an omni-directional camera mounted on a Pioneer 3-AT mobile robot available with the Computer Science Department at Aberystwyth University. The challenges of this dataset include a crossroad,

a widening of the road, and an obstacle. Also, the tarmac is covered in moss on the left-hand side at the beginning of the dataset.

Figure 7: Frame from dataset Footpath.

1.8 Lakeside

This dataset contains 8024 frames captured from an omni-directional camera mounted on the IRDIS rover robot (see [2]) available with the Computer Science Department at Aberystwyth University. The road in this dataset is made of various materials ranging from loose gray gravel to brown mud, and it presents dry and wet patches with puddles in places. The road is de-limited by grass, but the boundary road grass is not always obvious.

Figure 8: Frame from dataset Lakeside.

1.9 K59

This dataset contains 100 images in an urban environment originally obtained from the KITTI dataset. The dataset consists involves travelling along well demarcated road with kerbs, parked cars, pavements and other road-lanes on the sides. Challenges involve presence of traffic, shadows, an intersection and very low contrast between the road-lane and pavement.

Figure 9: Frame from dataset K59.

Table 1: Median and Standard Deviation of width error (pixels) of best LCNN for each colour model across all datasets. Negative values indicate width prediction being narrower than that in the ground truth.

Dataset	Track		Llan		Farm		Rugged		Rain		Footpath		Lakeside		Steep		K59		K56	
	med	std	med	std	med	std	med	std	med	std	med	std	med	std	med	std	med	std	med	std
RGB	30.3	8.2	49.4	14.0	-7.0	10.2	-4.4	15.7	-18.5	8.8	11.7	7.3	-9.0	3.5	24.4	11.2	1.3	21.4	28.0	22.1
HSV	**4.6**	**4.0**	23.7	10.9	**-1.0**	**7.2**	1.5	12.0	-23.6	7.9	**1.7**	**13.1**	3.6	3.5	23.9	11.5	7.5	6.2	26.3	11.0
YUV	26.4	4.8	29.7	15.2	-3.8	8.8	**1.1**	**11.3**	**-15.5**	**8.8**	4.3	6.8	-8.0	2.7	26.3	12.3	16.0	2.2	15.7	3.3
YCbCr	17.7	4.8	27.7	14.9	0.9	10.0	-4.7	11.9	-22.1	9.3	-5.4	9.1	**2.3**	**6.0**	19.2	15.5	8.5	2.8	**8.4**	**2.9**
lab	8.7	0.7	-3.3	10.8	13.0	3.1	-42.7	10.3	-65.8	6.6	-3.5	10.1	16.2	2.1	-16.6	11.4	14.1	2.5	9.6	2.8
CbCra	15.9	1.9	**-2.2**	**11.7**	2.6	6.1	-38.3	12.3	-58.7	9.7	-3.2	11.6	7.6	4.0	**-11.2**	**11.2**	**7.4**	**4.0**	9.5	3.1

1.10 Steep

This dataset contains 2033 frames captured from a GOPRO Hero4 camera (http://gopro.com). The road is a steep tarmac path with leaves and mud covering sections of it. Indeed one of the most challenging aspects of this dataset is the leaves on either side of the road that make it extremely delineated and make it difficult to ascertain where the actual road boundaries are. Other challenges include changes include changes in non-road surfaces with appearance of fences on either side and occasional widening of the road.

Figure 10: Frame from dataset Steep.

1.11 Width Detection

The median width errors of the best networks for each colour model with the LCNN and AlexNet architectures is presented in tables 1 and 2 respectively. It can be observed that across all colour models (and both architectures), networks tend to predict parameter better than the width. While annotating the images to generate the teaching input for training, the width of the trapezoid has been arbitrarily set to accurately capture the road width. However, in case of very ill-defined road, the definitions or road edges are fuzzy. For example, the overestimation of the width by LCNN for most colour-spaces in the dataset Steep is due to the inclusion of leaves on either side of the path within the road shape. Similarly in the case of Lakeside and Farm (which have very narrow roads), AlexNet includes extra pixels on either side of the annotated ground-truth boundaries. It should be noted that networks have been selected purely on their ability to minimize position error and that similar degrees of accuracy across datasets and colour-models can also be observed with the ASC's width prediction. Refer to the supplementary videos for a better understanding of the detection behaviour across different colour models for both architectures.

Table 2: Median and Standard Deviation of width error (pixels) of best modified AlexNet for each colour model across all datasets. Negative values indicate width prediction being narrower than that in the ground truth.

Dataset	Track		Llan		Farm		Rugged		Rain		Footpath		Lakeside		Steep		K59		K56	
	med	std	med	std	med	std	med	std	med	std	med	std	med	std	med	std	med	std	med	std
RGB	**3.2**	**2.7**	-7.1	11.0	**10.1**	**3.6**	-13.6	11.2	-49.7	8.4	**-0.3**	**8.8**	4.9	3.7	10.5	11.4	**3.3**	**6.4**	**-1.8**	**6.1**
HSV	17.8	2.6	5.4	11.4	12.8	3.9	**-6.6**	**12.0**	**-34.0**	**9.0**	3.3	11.2	**5.0**	**2.8**	14.0	10.5	8.9	2.9	16.3	5.0
YUV	18.4	1.1	**1.4**	**10.7**	15.7	2.9	-25.6	10.5	-54.9	6.6	6.7	9.3	16.6	2.5	-0.4	11.7	10.2	3.0	10.2	3.5
YCbCr	19.2	1.3	2.1	10.8	16.4	3.0	-25.3	10.6	-54.7	6.5	5.9	9.1	18.0	2.6	**-0.2**	**11.5**	9.1	3.1	8.7	2.8
lab	17.0	0.7	3.3	10.5	25.4	2.6	-33.7	10.5	-60.6	6.5	4.7	9.6	25.5	2.1	-8.0	11.3	19.5	2.1	17.9	2.0
CbCra	18.4	0.6	4.5	10.4	25.4	2.7	-33.3	10.4	-59.9	6.4	5.0	9.7	26.2	2.2	-7.8	11.1	21.0	2.1	19.0	2.4

1.12 Videos

Figure 11: To play the video, click on the image or use the following URL `https://www.youtube.com/embed/qO3Iop-SwrY`. This contains image sequences in various datasets and colour models with the detected and ground-truth road-shape (trapezoid) drawn on each frame.The yellow trapezoid shows the road-shape detected by the network and the white trapezoid shows the ground-truth shape.

Figure 12: To play the video, click on the image or use the following URL `https://www.youtube.com/embed/umx6Fr9Qe5o`. /This contains clips from two successful robot trials, using the LCNN architecture (refer to main paper) and HSV colour model.

References

[1] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)*, 2013.

[2] M. Ososinski and F. Labrosse. Automatic driving on ill-defined roads: An adaptive, shape-constrained, color-based methods. *Journal of Field Robotics*, In Press, 2014.