

Large scale labelled video data augmentation for semantic segmentation in driving scenarios

July 26, 2017

1 Introduction

This document lists and describes three videos submitted as a supplementary material for the main paper.

2 Supplementary Video Information

The details of the aforementioned videos are listed below.

- **VID-01-CamVid-Label-Augmentation-Results.mp4**

This video shows a sample of our augmented data for the CamVid [1] dataset. It consists of 134-frames from subsequence 001tp. Each frame contains (clockwise from top-left) an original image, corresponding augmented frame and hand labelled frame if it exists. No manual filtering was performed. Note that all propagated training frames were used for training.

- **VID-02-CitySapes-Label-Augmentation-Results.mp4**

This video shows a sample of our augmented data for the CityScapes [2] dataset. It consists of five 21-frame subsequences taken from different cities in the training dataset. As in the previous video, each frame contains (clockwise from top-left) an original image, corresponding augmented frame and hand labelled frame if it exists.

- **VID-03-CamVid-Instance-Label-Augmentation-Results.mp4**

This video shows a sample of our augmented data for an extension of the CamVid [1] dataset that includes car instances (referred to as the CamVid-Instance dataset). It consists of 134 frames from sequence 006r0.

As in the previous videos, each frame contains (clockwise from top-left) an original image, corresponding augmented frame and hand labelled frame if it exists.

References

- [1] G. J. Brostow, J. Fauqueur, and R. Cipolla. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, 30(2):88–97, 2009.
- [2] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.