CSC 2515 Projects

Raquel Urtasun

Select ONE and only ONE project from this list. Make sure that you do not miss the two deadlines: project proposal submission and final project submission. Make sure to read all the instructions about what to submit for each deadline.

1 Possible Projects

1.1 Autonomous Driving: Road-Estimation

The goal of this project is to create a classifier that given an image it is able to create a pixel wise segmentation of the image in terms of what is road and what is non-road (i.e., binary classification). For this purpose download the base kit of KITTI [3] with: left color images, calibration and training labels data from `http://www.cvlibs.net/datasets/kitti/eval_road.php`. Split the training data into training, validation and testing. The splits should be 60%/10% and 30% of the data. Do not submit to the test set benchmark. Best performing methods will be invited to do so. For more information about the road dataset see [2].

Note that the task is to label each pixel, and thus there is a "classification" problem to be solved per pixel. To easy the computation, for each image first compute super-pixels using for example SLIC `http://ivrg.epfl.ch/research/superpixels#SLICO`. Note that you can use other super pixel code/algorithms if you want to. Classification can then be performed by treating each super pixel as an example. See Fig. 1 for a illustration of the problem, the ground truth as well as the super pixels.

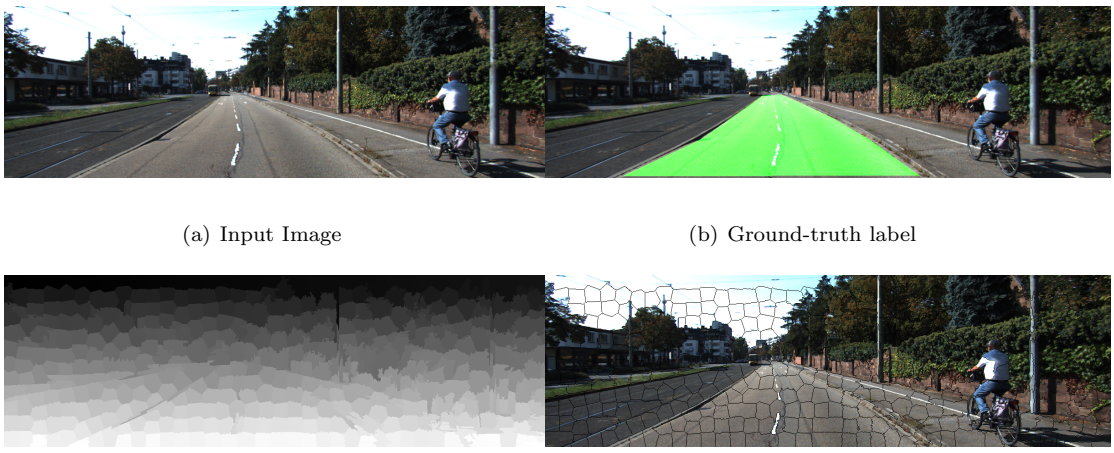
The purpose of this project is to investigate machine learning techniques to solve this task. Try things that we have seen in class and feel free to also try other techniques. Write in your report what you have done, what you observe, what you have tried, why you did what you did, etc.

Make a section of your report about what would you do to scale what you have done to work with pixels, as in that case you have millions of examples. You don't need to implement anything but add a discussion of your ideas.

BONUS: The use of "fancy" techniques such as Markov Random Fields should only be attempted if you want an extra bonus on the project (up to 30% extra). These type of techniques are only cover in class as a review of exciting things in machine learning and thus I do not expect students to try such things for the purpose of the project.

1.2 Street View House Numbers

This project consists on classifying digits from street view images. Towards this goal, download the Format 2 images from the Street View House numbers dataset [1] with `train_32x32.mat`,



(a) Input Image (b) Ground-truth label

(c) SLIC superpixel (d) SLIC super-pixel Boundary

Figure 1: An example from KITTI road segmentation benchmark.

test_32x32.mat data, which you can find at `http://ufldl.stanford.edu/housenumbers/`. In this task, all the images have a fixed 32×32 resolution with character-level ground truth labels. For each example, the labeled character is centered at the image. See Fig. 2 for an illustration. There are ten classes in total (one for each digit). Divide the training into train and validation (e.g., 80% and 20%).

Note that the data is collected from street-view images, thus there exist vast intra-class variations. To generate competitive performance, you may want to consider exploiting good feature representations that are robust to those variations, whether they should be hand-crafted features or learned features. In order to boost performance, you may also want to consider augmenting the training data with extra_32x32.mat.

The purpose of this project is to investigate machine learning techniques to solve this task. Try things that we have seen in class and feel free to use other techniques if you feel like it. Write in your report what you have done, what you observe, what you have tried, why you did what you did, etc.

Make a section of your report about what would you do to scale what you have done to work with pixels, as in that case you have millions of examples. You don't need to implement anything but add a discussion of your ideas.

BONUS: Bonus will be given to attempt to solve both text detection and recognition task, i.e. Format 1 in SVHN dataset (with up to 30% extra). In this task, images are not well cropped with also multiple labeled digits appear in one image. Your task is to first detect the digits with a bounding box then conduct classification on each cropped digit.

1.3 Own Project

You can work on a different project if you want to. Towards this goal, you should submit a one page proposal describing in detail the goal of the project, the data as well as which techniques you

Figure 2: Examples from street view house numbers benchmark.

employ to use. You are also encouraged to apply machine learning techniques on the open problems in your own research areas.

BONUS: Bonus will be given according to the importance and difficulty of your problem, the techniques you use and the performance you achieved compared with the state-of-the-art (with up to 30% extra).

2 Deadlines and Submission Instructions

Make sure that you don't miss any of the following deadlines:

- The **project proposal** should be submitted electronically via CDF website. Choose assignment name `"Proposal"` and rename your submission to `proposal.pdf`. The submission deadline is noon on **March 5th**. Indicate in the project proposal whether your graduation date is June 2015 or not. This is important to schedule your presentations
- The **project report** should also be submitted via CDF by noon on **April 10th**. Choose assignment name `"Report"` and rename your submission to `report.pdf`. A penalty of 25% would be applied per day that the submission is late.
- **Project presentations** will be scheduled starting the week of April 13th. More details to come.
- You are free to use any third-party ideas or code that you wish. You must properly provide references to any work that is not your own in the write-up. The project is not intended to be a stressful exercise; instead it is a chance for you to experiment, to think, to play and to

hopefully have fun! Start with simple methods that work more or less out of the box and go from there.

References

[1] Yuval Netzer et al. "Reading digits in natural images with unsupervised feature learning." *NIPS workshop on deep learning and unsupervised feature learning*, 2011.

[2] Jannik Fritsch , Tobias Kuhnl, and Andreas Geiger. "A New Performance Measure and Evaluation Benchmark for Road Detection Algorithms." *ISTC* 2013.

[3] Andreas Geiger, Philip Lenz, and Raquel Urtasun. "Are we ready for autonomous driving? the kitti vision benchmark suite." *CVPR* 2012.