

DATA FROM SKY

ultimate aerial analysis



Data From Sky



Satellites, 50 cm/pixel



Autonomous drones, 20 cm/pixel



Fixed cameras – tall buildings



Advanced traffic analysiseffective and safe traffic



Applications in traffic

Monitoring: traffic intensities

Capacity: analysis of traffic flow

Safety: audit of critical regions, trajectory interactions, conflicts

Meta-Analysis: gap-time and time to follow, emission of solids a NOx







How it works







Tracking





Previous Approach (2014)

Cascade Classifier - AdaBoost

- 80 k training samples
- Prefiltration of detection candidates (by road surface and movement cues)

Accuracy: ≈80 % on validation data







We have data... Let's use them!

Annotated high-resolution video data

- 280+ hours of total footage length
- 500 000+ tracked vehicles:
- 50 000+ km traveled by vehicles

Prospective HUGE training/validation set



Deep Neural Networks



Our tools





Dedicated Learning Server with nVidia GeForce **GTX 1080**

- 8 GB GDDR5X RAM
- 2560 CUDA cores
- 9 TFLOPS







Different Data







Classification – proof of concept

Can DNN work with such tiny objects?



Tried: ConvNet, GoogLeNet, Custom Architecture

Problems: Training from scratch, too aggressive data reduction, low abstraction.

Final Architecture: ResNet-50

- residual, 50 layers
- 7.4 M training samples
- Accuracy: > 98 % on validation data



Sliding Window

Shortest path from Classifier to Detector.

- WAY too slow!
 - 400 k+ detection candidates per frame (without constraints)
 - few minutes per FullHD frame



- Prefiltration of detection candidates (by road surface and movement cues)
 - Noticeable speed-up (in order of 100x) ☺
 - Introduces unwanted localization bias \otimes



Faster R-CNN

Complex End-To-End Solution.

- Human Attention
- Input = FullHD image
- Output = List of Detections

Modifications:

- VGG-CNN1024
- Custom RPN
- Wrapper for our datasets





Output

ResNet-50 + Sliding Window



mAP = 0.647 Detection Time: 77 s/frame mAP = 0.986 Detection Time: 0.110 s/frame

Faster R-CNN









Localisation Offset

Oblique capture of the scene







Accurate Localisation = Spatial Regressor

Idea from J. Prokaj: *Persistent Tracking for Wide Area Aerial Surveillance* (CVPR2014) Principle: "Teach network to show where the vehicle is in the sample."



(X, Y) = displacement to vehicle position W = certainty of estimation

- Consensus from nearby samples
- Robust to bias from candidate filtration
- Can be iteratively applied



Be Careful When Diving Into DNN!

- Good Classifier ≠ Good Detector
- Find balance between abstraction and data reduction
- Tiny input for tiny data
- Even small DNN do really need huge memory
- We ran out of file-system **inodes** due to huge amount of training data.
 - Change configuration of Ext4, or use XFS, ReiserFS
- Moving datasets with millions of samples is time-consuming
 - Archives / virtual hard drives



Looking forward to tell you more...

Adam Babinec +420 775 345 465 info@rcesystems.cz www.datafromsky.com

