THE UNIVERSITY OF MELBOURNE

Automated interpretation of imagery and lidar point clouds for intelligent transport applications

Kourosh Khoshelham
University of Melbourne

5 August, 2020

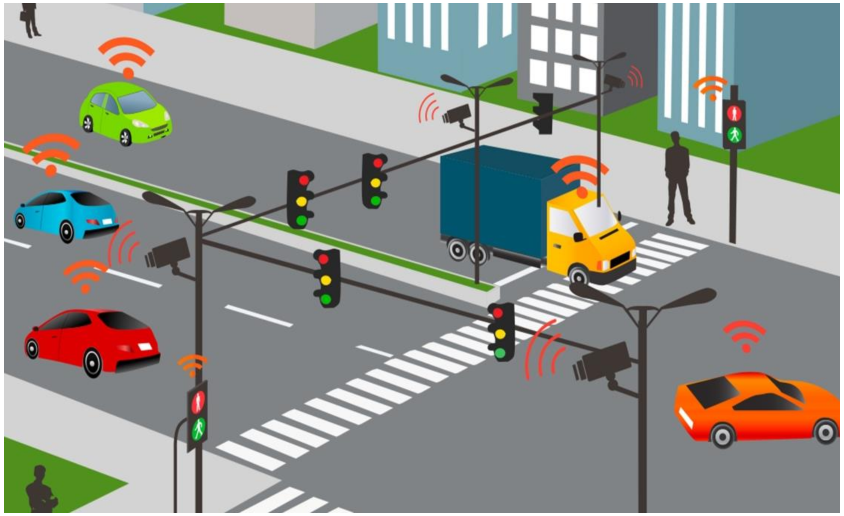
Spatial data for transport applications

- Spatial data enables
  - improving traffic flow
  - improving road safety
  - reducing emissions
  - autonomous driving
  - connected transport
  - …

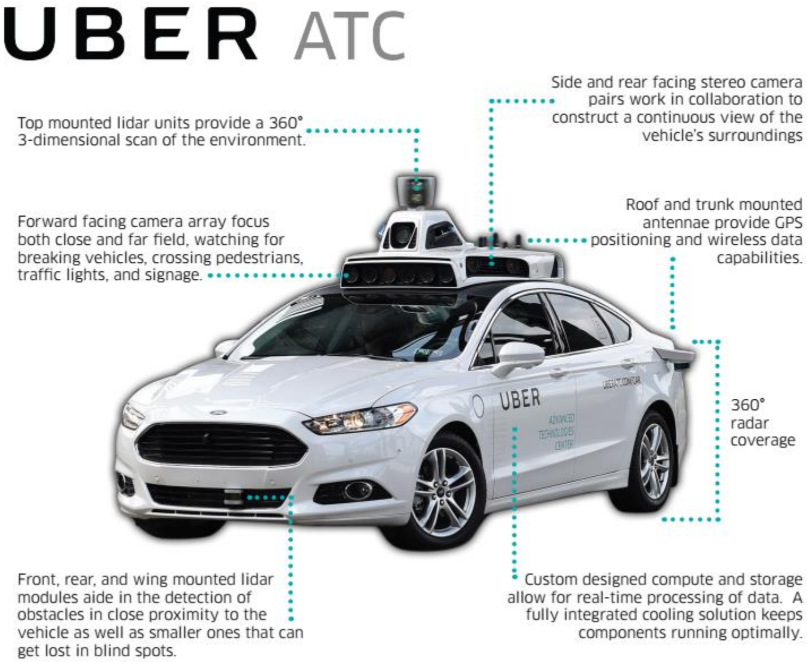
- Role of automation

Data acquisition modes

Sensors installed in the environment



Sensors installed on vehicles



Example applications

- Sensors installed in the environment
  - Smart parking
  - Congestion detection
  - Detection of anomalous driving behaviour

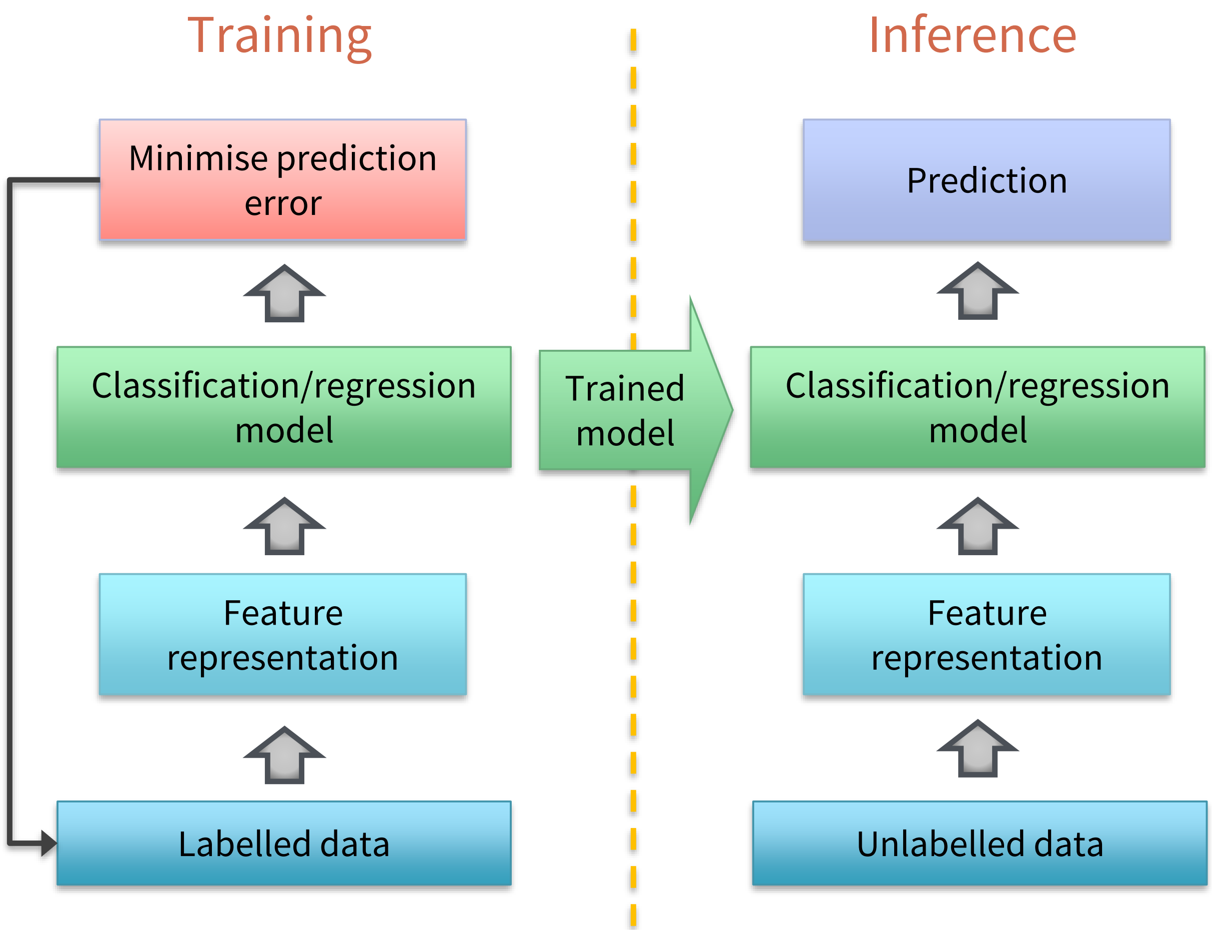
- Vehicle-borne sensors
  - Vehicle localisation
  - Efficient generation of HD Maps
  - Detection of unsigned road incidents

Automation: conventional approach

Machine learning:

→ Train a model using labelled data

→ Apply the trained model to unlabelled data



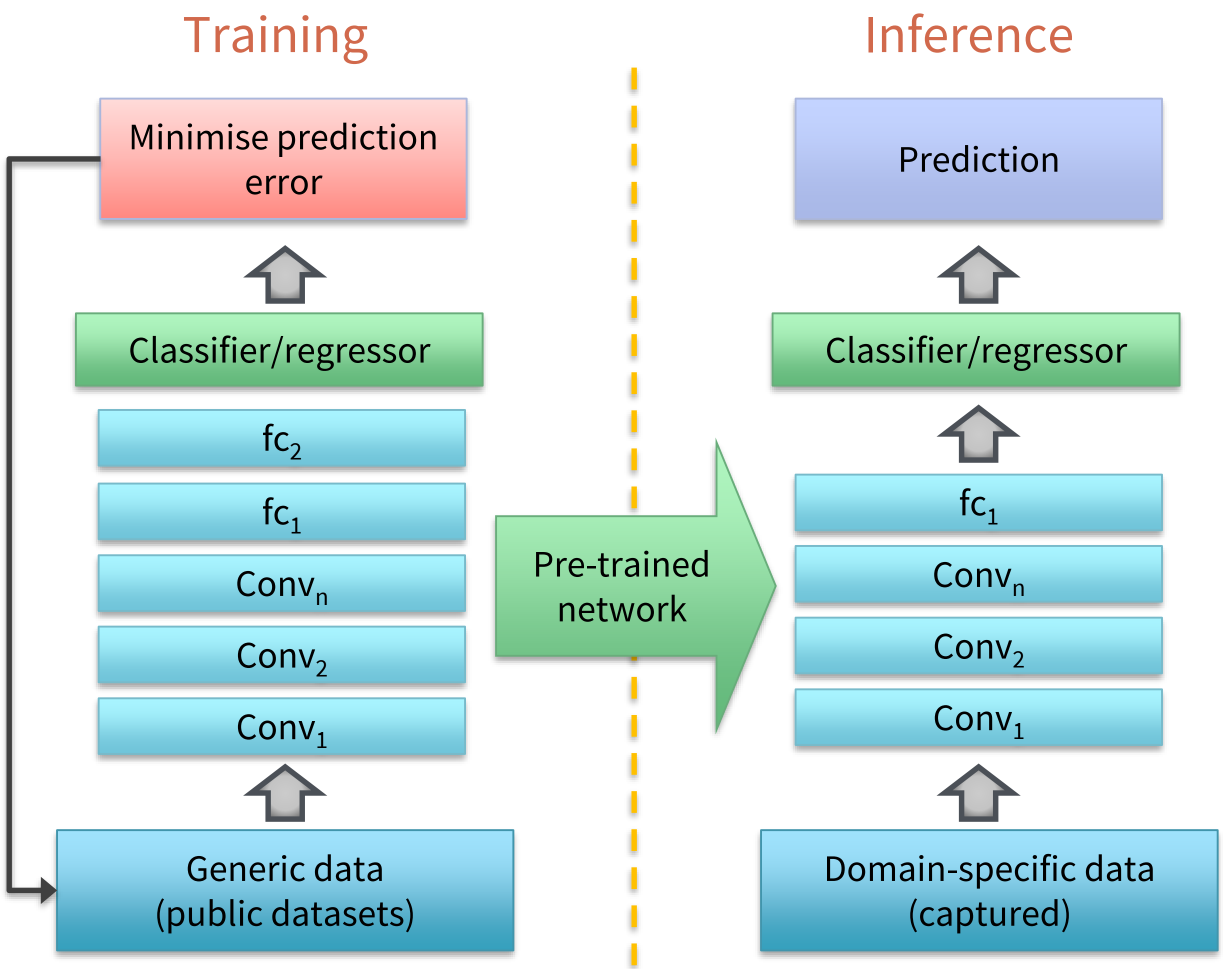
Automation: modern approach

Deep learning:

→ Learning feature representations using deep neural networks

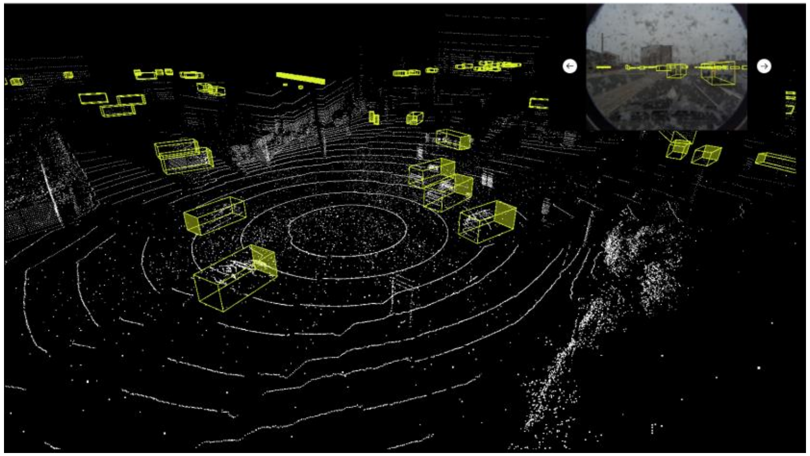
Transfer learning:

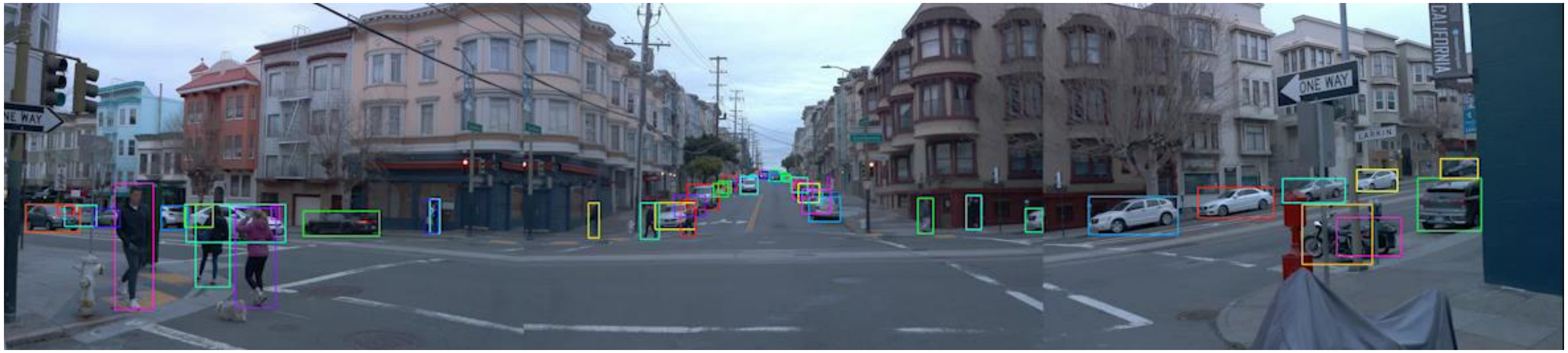
→ Training using generic public data and transferring the learned knowledge to interpret domain-specific captured data



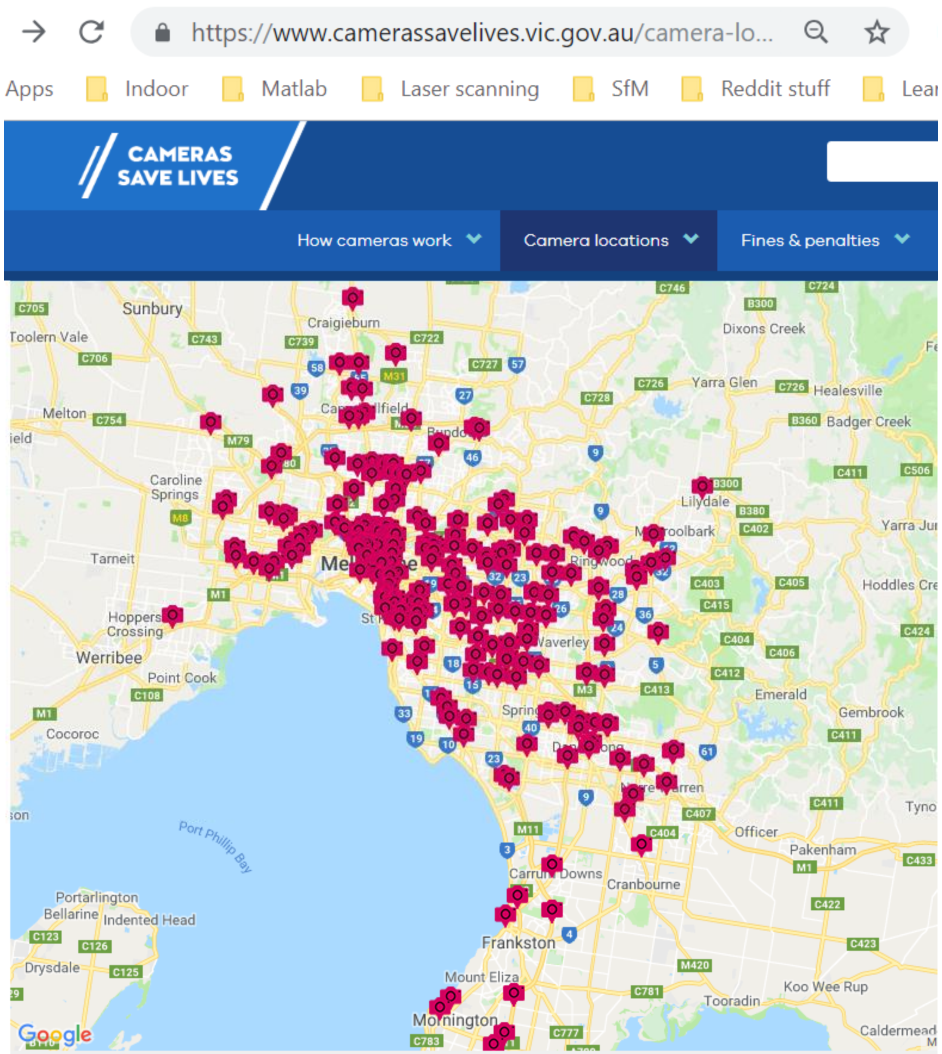
Example public datasets

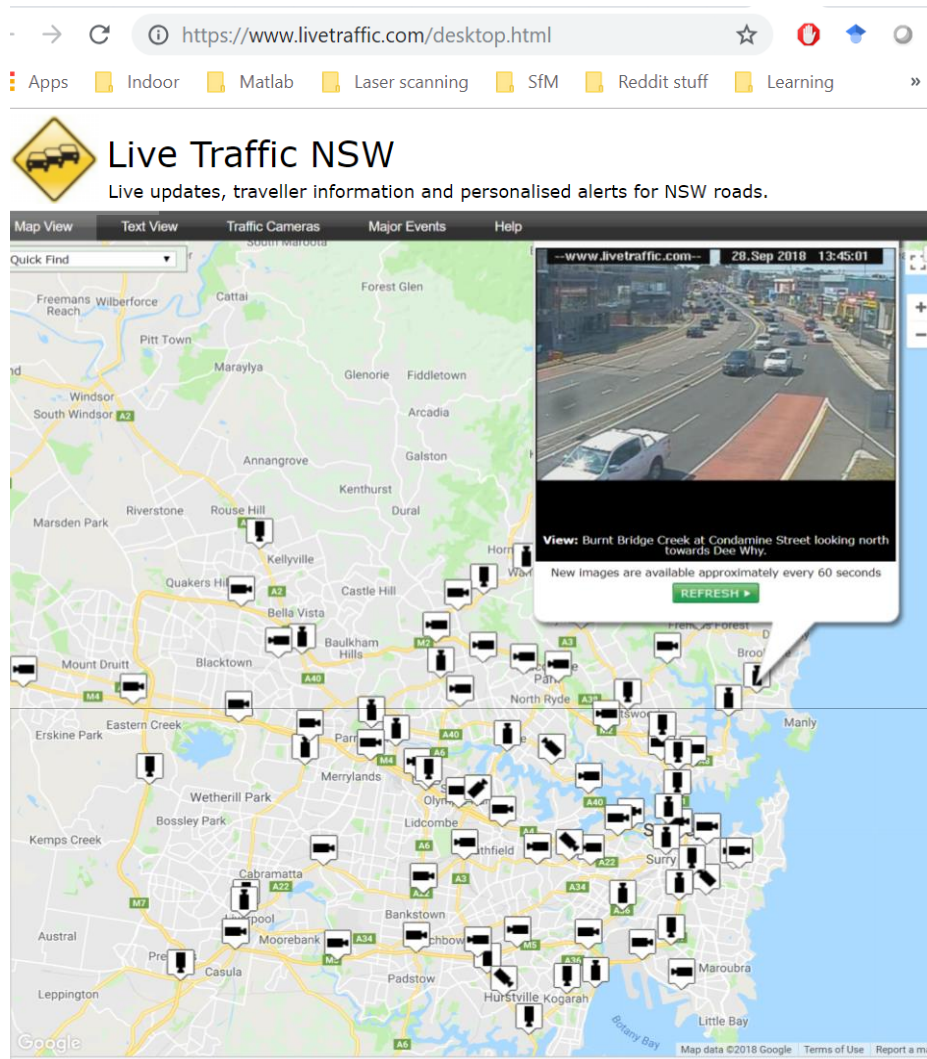
- KITTI Vision Benchmark Suite
- Oxford Robotcar Dataset
- Waymo Open Dataset
- Lyft Open Data
- University of Michigan Dataset (NCLT)
- CADC Dataset
- …





Access to live data





Smart parking

- Sensors detect vacancy/occupancy
- Cloud computing platform generates occupancy map
- Navigation app guides user to nearest vacant parking space
- Communication via short range (e.g. Zigbee) and long range (e.g. LoRaWan) protocols.

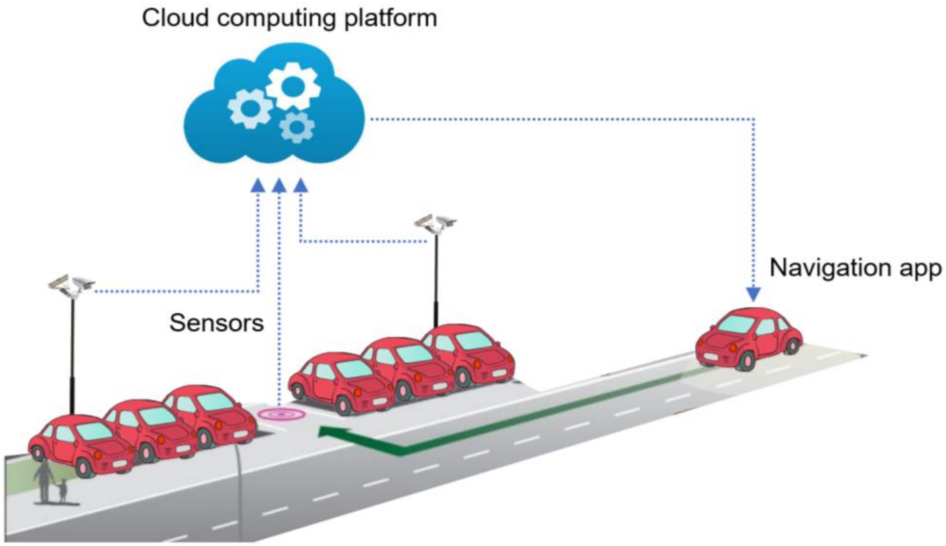
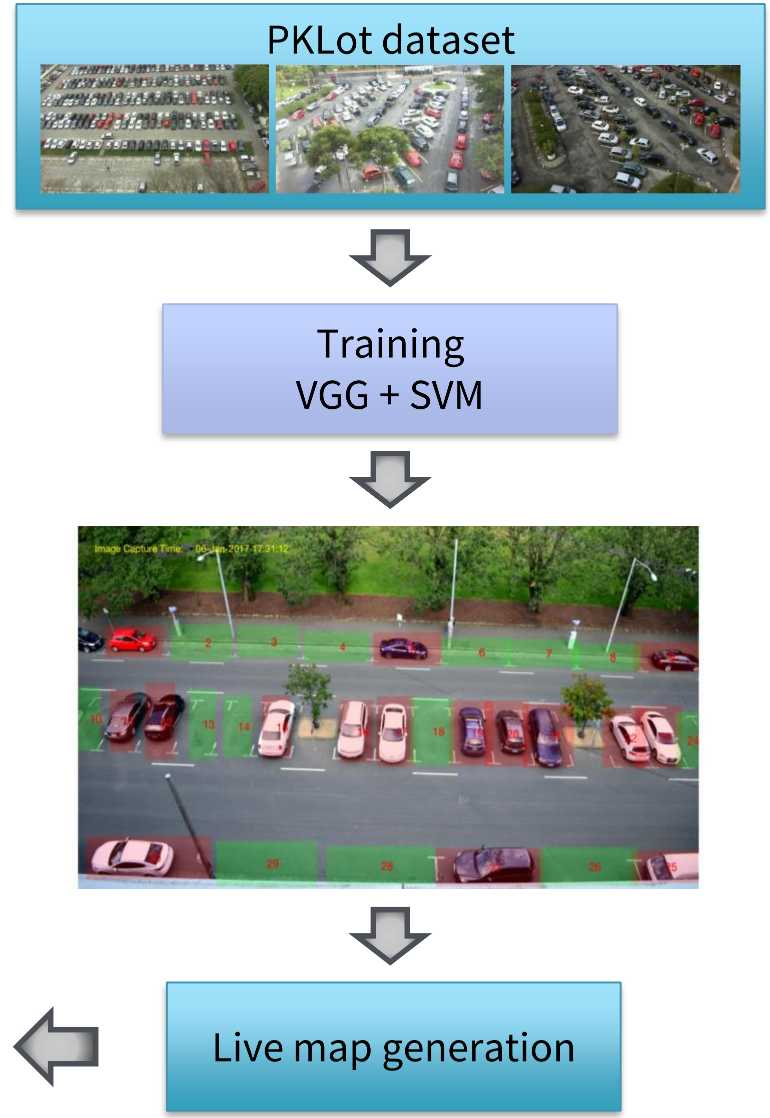


Image-based parking occupancy detection

- Training a classifier using a public dataset
- Occupancy detection on live images
- Parking occupancy live map
- Parking guidance system, booking, etc.



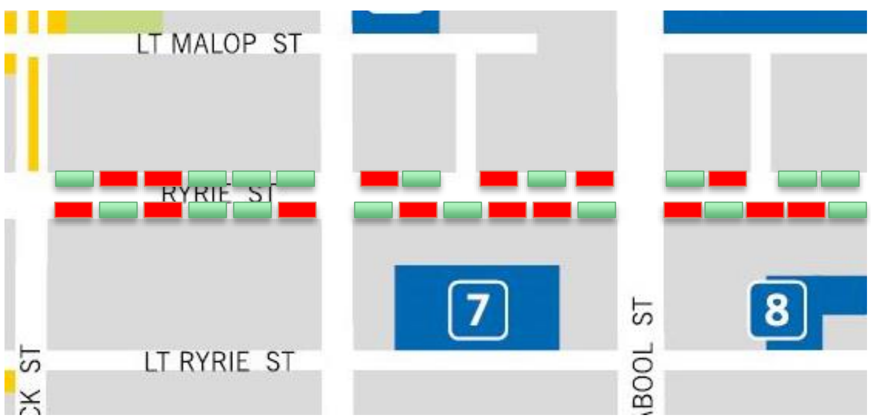
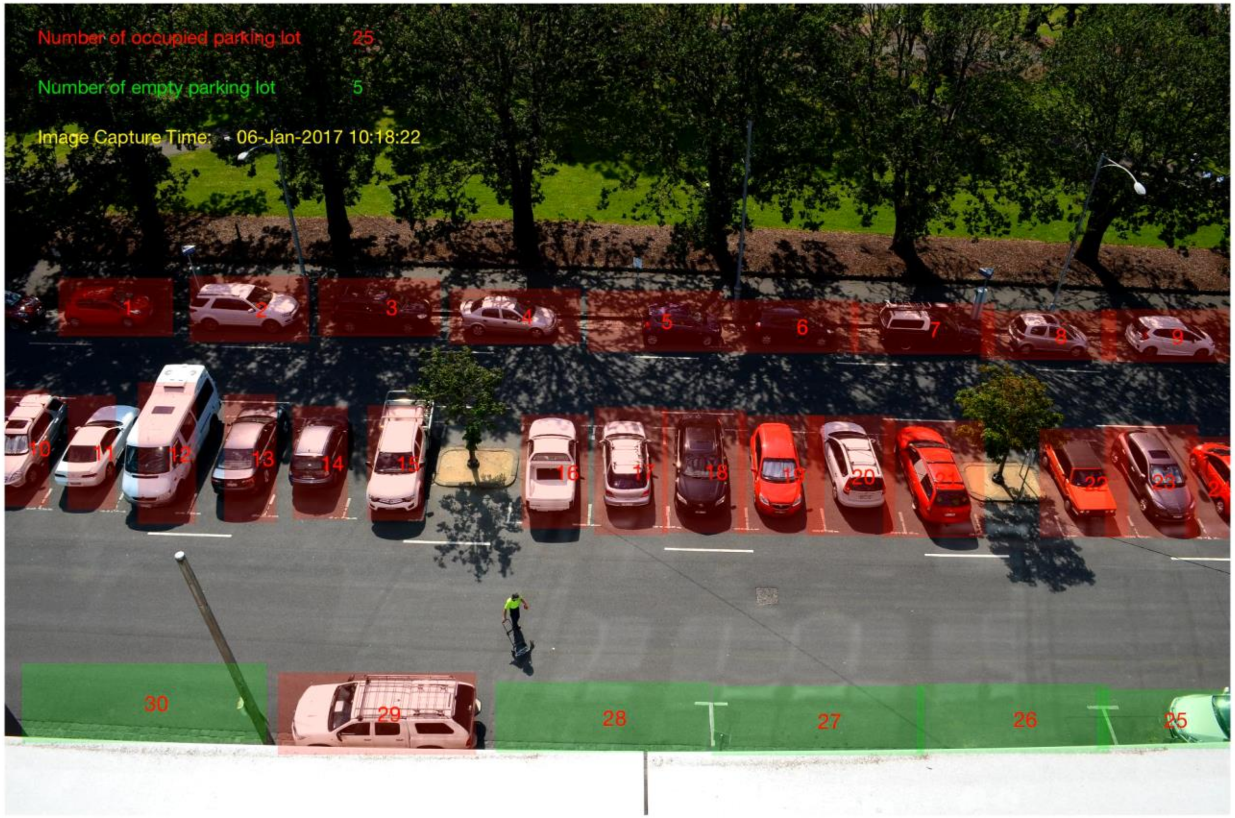


Image-based parking occupancy detection



97%

Detection accuracy

67 ms

Average processing time on a standard CPU

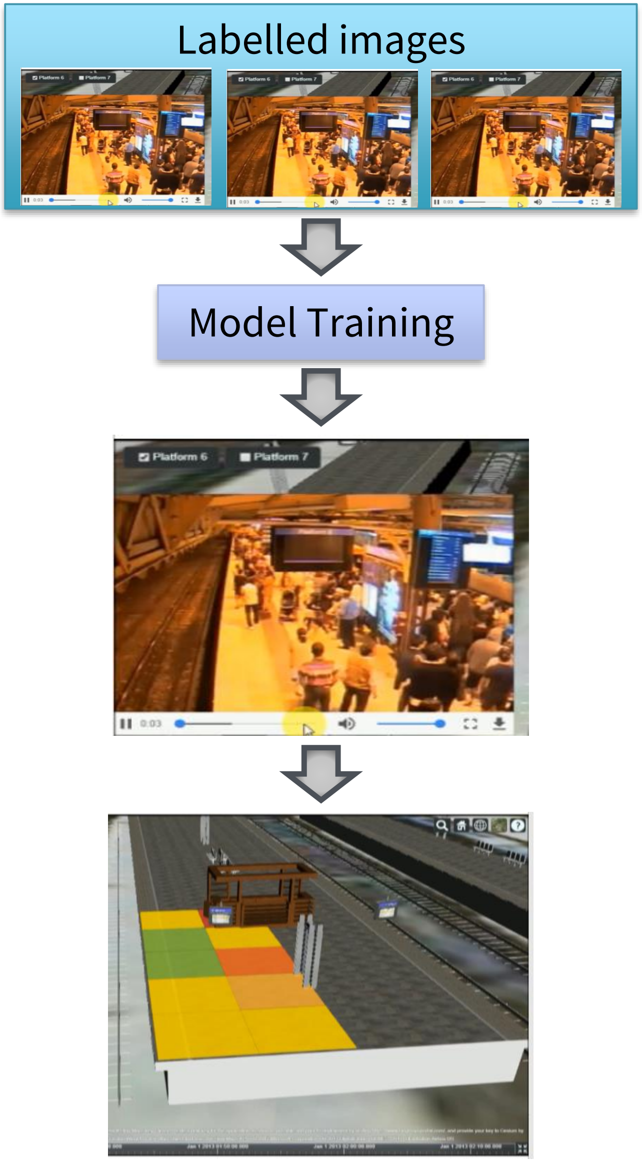
Acharya, D., Yan, W., Khoshelham, K., 2018. Real-time Image-based Parking Occupancy Detection Using Deep Learning in: Peters, S., Khoshelham, K. (Eds.), Proceedings of Research@Locate18. CEUR Workshop Proceedings, Adelaide Australia, pp. 33-40.

Congestion Detection

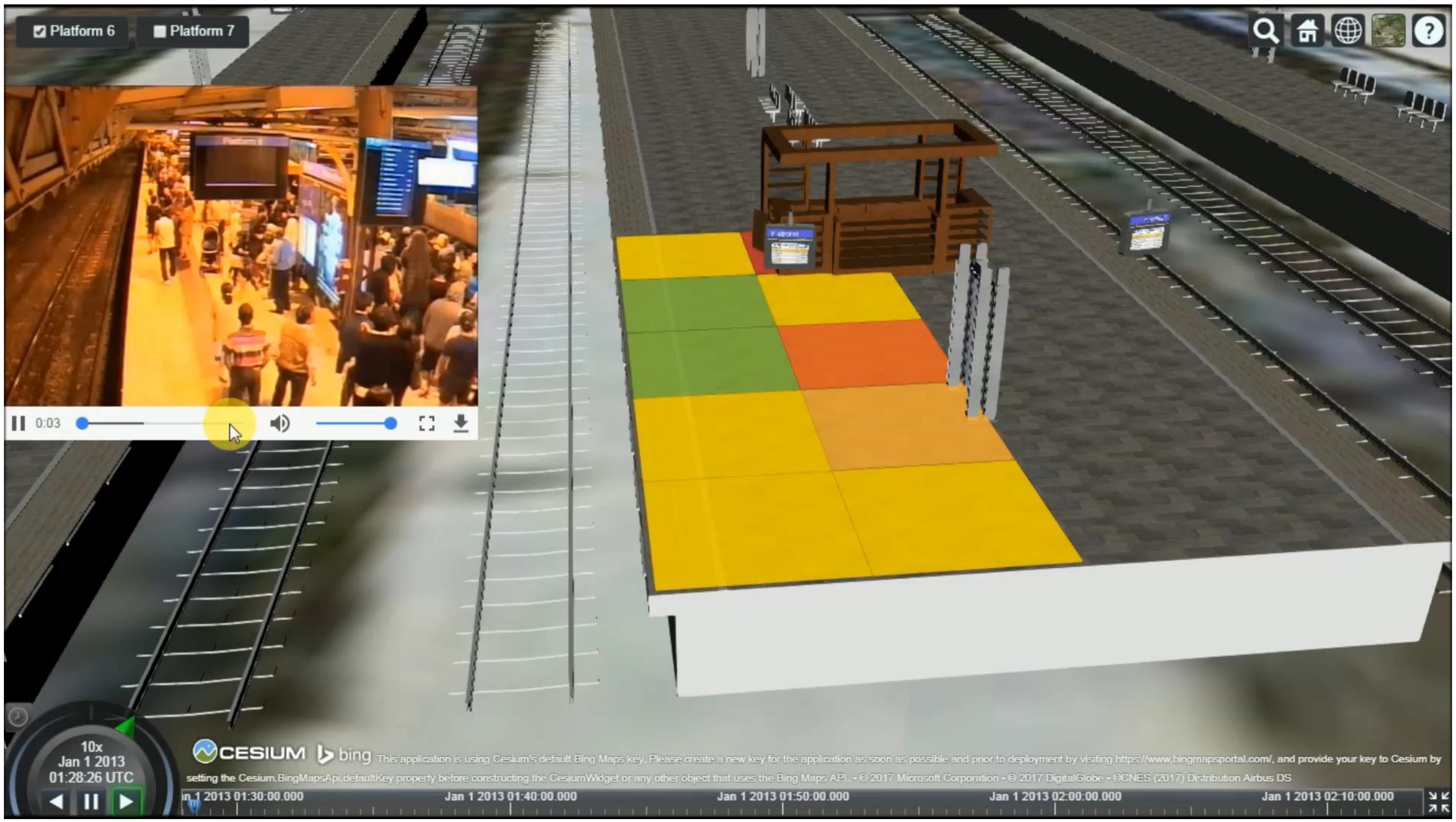
- Indirect approach:
  - Estimate count, density, flow
  - Infer congestion
  - Suitable for vehicles?

- <u>Direct approach</u>:
  - Label images according to congestion level
  - Train a model using labelled images
  - Apply the model to live images
  - Suitable for pedestrians

- Live map of congestion



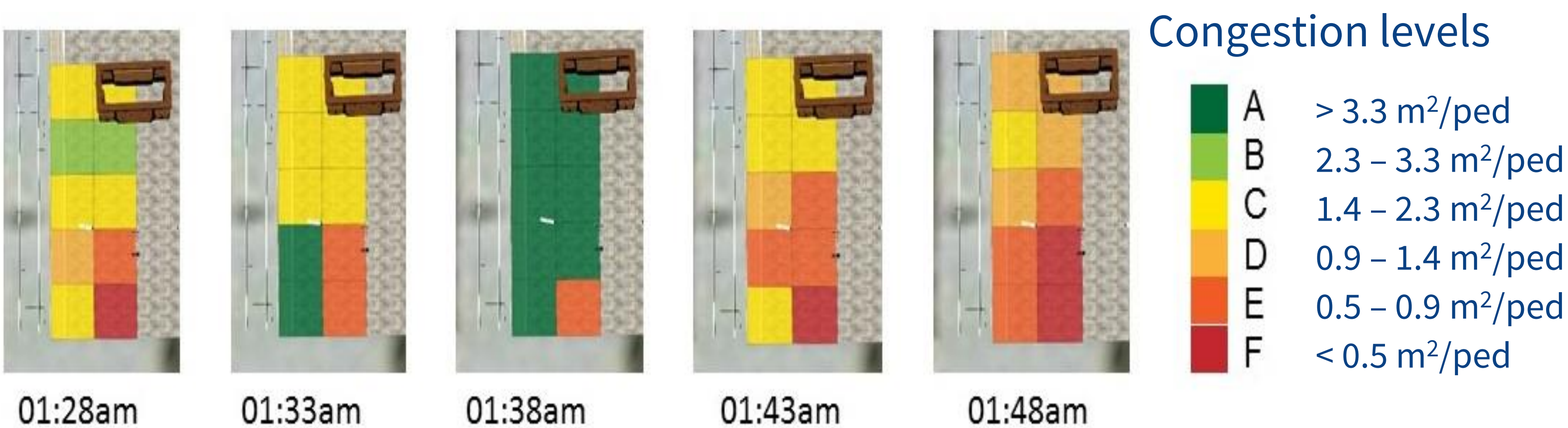
Congestion Detection



Li, Y., Khoshelham, K., Sarvi, M., Haghani, M., 2019. Direct generation of level of service maps from images using convolutional and long short-term memory networks. Journal of Intelligent Transportation Systems, 1-9.

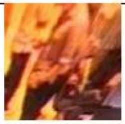
Congestion Detection

Temporal analysis of congestion



Li, Y., Khoshelham, K., Sarvi, M., Haghani, M., 2019. Direct generation of level of service maps from images using convolutional and long short-term memory networks. Journal of Intelligent Transportation Systems, 1-9.

Congestion Detection

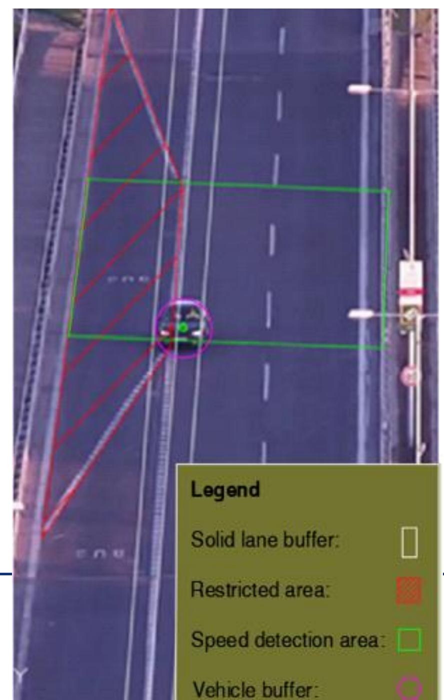
	LOS A	LOS B	LOS C	LOS D	LOS E	LOS F
Example						
No of samples	764	237	532	472	520	195
Confusion matrix						
Predicted label						
LOS A	667	76	15	3	1	2
LOS B	53	154	16	12	2	0
LOS C	13	37	450	23	7	2
LOS D	7	24	46	348	32	15
LOS E	0	7	17	32	426	38
LOS F	0	2	8	13	26	146
Total accuracy	80.6%					

	Define ROI	Gridding	Extract features	Classification	Visualisation	Total
Time(s)	0.00088	0.00136	0.37087	0.00188	0.00015	0.38

Li, Y., Khoshelham, K., Sarvi, M., Haghani, M., 2019. Direct generation of level of service maps from images using convolutional and long short-term memory networks. Journal of Intelligent Transportation Systems, 1-9.

Detection of Anomalous Driving Behaviour

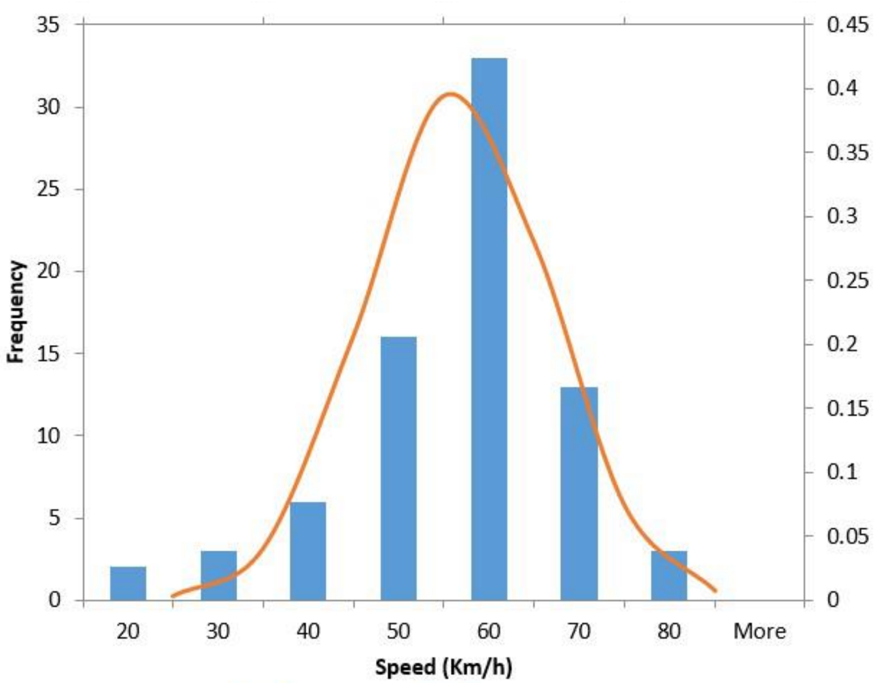
- Speed anomaly
- Solid line crossing
- Entering restricted zones

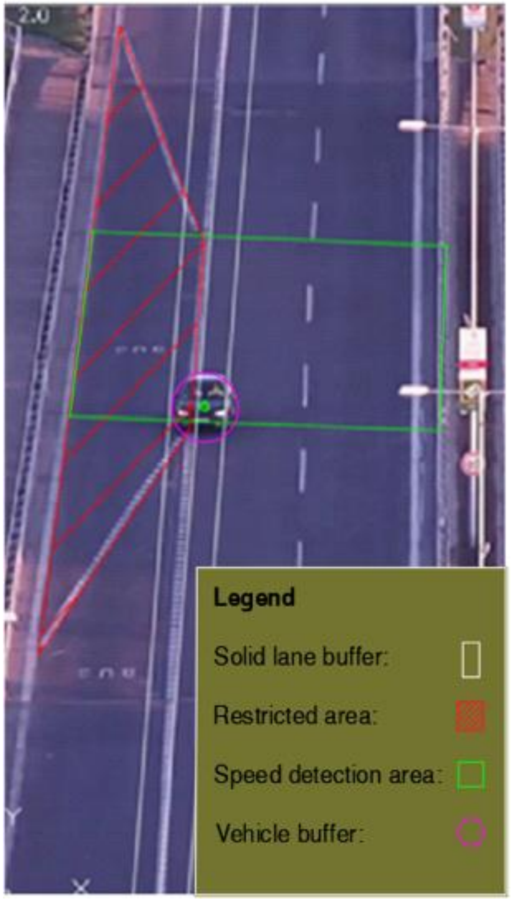


Detection of Anomalous Driving Behaviour

200 ms

Average processing time per frame on a GPU

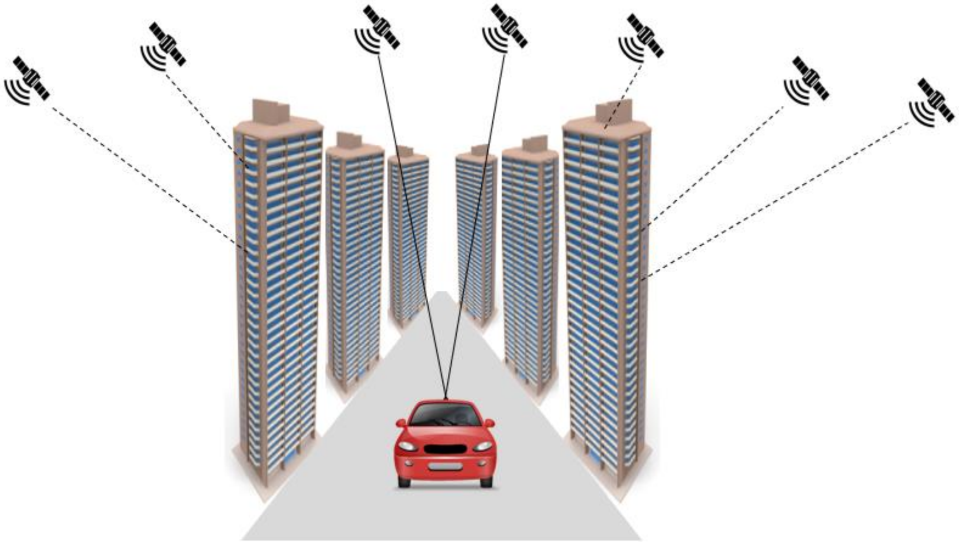




Behavior Class	Recall	Precision
solid line crossing detection	0.889	0.865
entering restricted areas detection	0.730	0.964

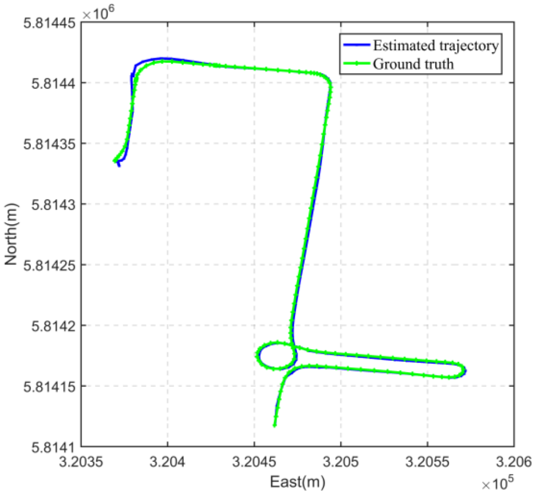
Vehicle localisation

- Generally relies on GNSS
- Not available in dense urban environments
- Alternatives:
  - Wireless (e.g. UWB)
  - Inertial (IMU)
  - Visual/lidar odometry
  - HD Maps/ landmarks



Vehicle localisation

Omnidirectional Visual Inertial Odometry (OVIO)



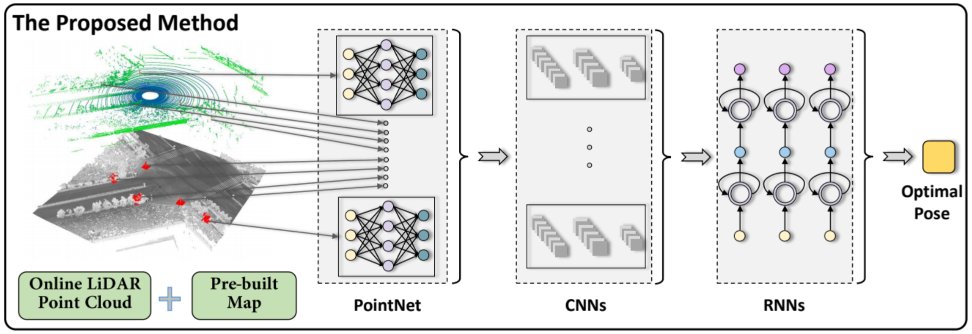
Accuracy: 0.33% @ 600 m
(drift problem)

Ramezani, M., Khoshelham, K., Fraser, C., 2018. Pose estimation by Omnidirectional Visual-Inertial Odometry. Robotics and Autonomous Systems 105, 26-37.

Vehicle localisation

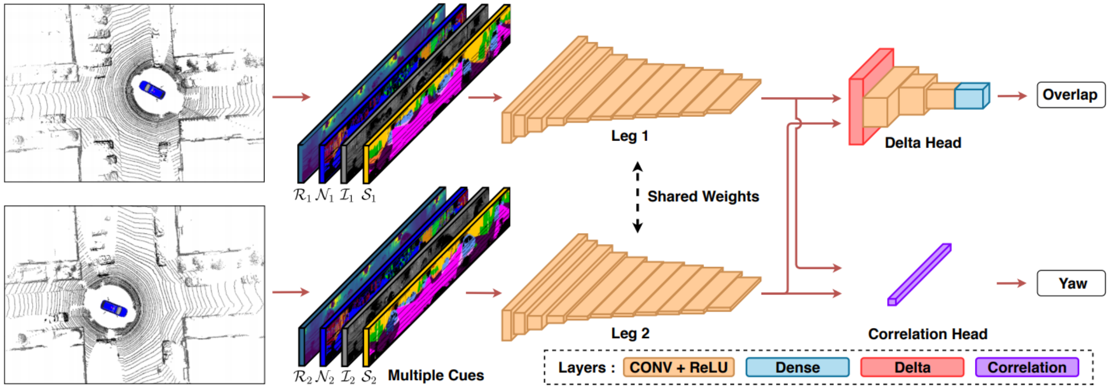
Registration with previous lidar scans

L3-Net (Lu et al., 2019)



W. Lu, Y. Zhou, G. Wan, S. Hou and S. Song, "L3-Net: Towards Learning Based LiDAR Localization for Autonomous Driving," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 6382-6391.

OverlapNet (Chen et al., 2020)



Chen, X., Läbe, T., Milioto, A., Röhling, T., Vysotska, O., Haag, A., ... & Fraunhofer, F. K. I. E. OverlapNet: Loop Closing for LiDAR-based SLAM. Robotics: Science and Systems, 2020.

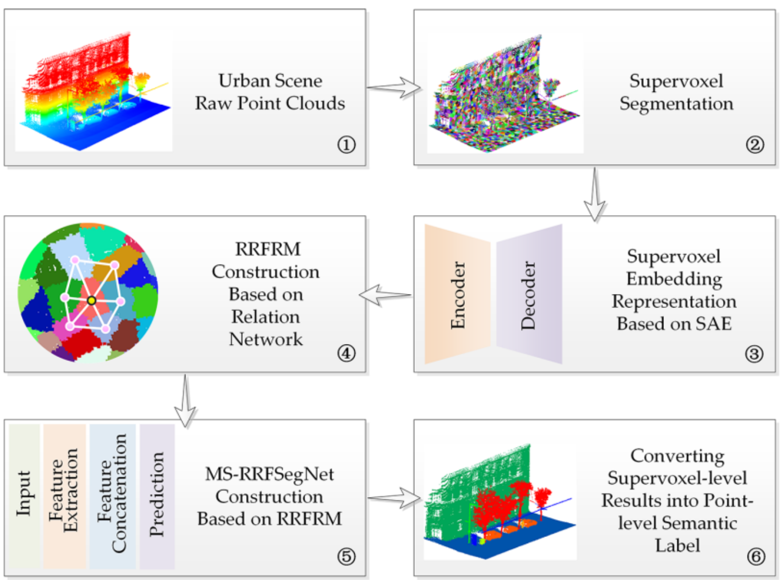
Efficient generation of HD Maps

- Training a semantic segmentation network using a public dataset
- Applying the trained network to data from a different environment
- Challenges:
  - Scarcity of training samples
  - Domain adaptation

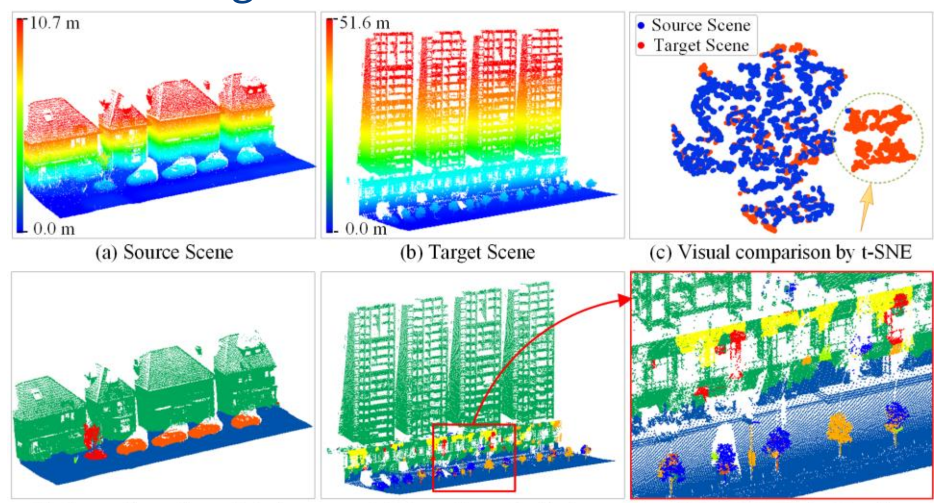
Fig. 1: An example mobile lidar system (left), the raw point cloud (middle), and the classified point cloud (right).

Efficient generation of HD Maps

Semantic segmentation using a network trained on PL3D dataset



Differences between the source and target scene



(a) Source Scene (b) Target Scene (c) Visual comparison by t-SNE

(d) Source Scene Ground Truth (e) Target Scene Prediction (f) Local Prediction Details

Luo, H., Chen, C., Fang, L., Khoshelham, K., Shen, G., 2020. MS-RRFSegNet: Multiscale Regional Relation Feature Segmentation Network for Semantic Segmentation of Urban Scene Point Clouds. IEEE Transactions on Geoscience and Remote Sensing.

Efficient generation of HD Maps

Domain adaptation using PointDAN (Qin et al., 2019)

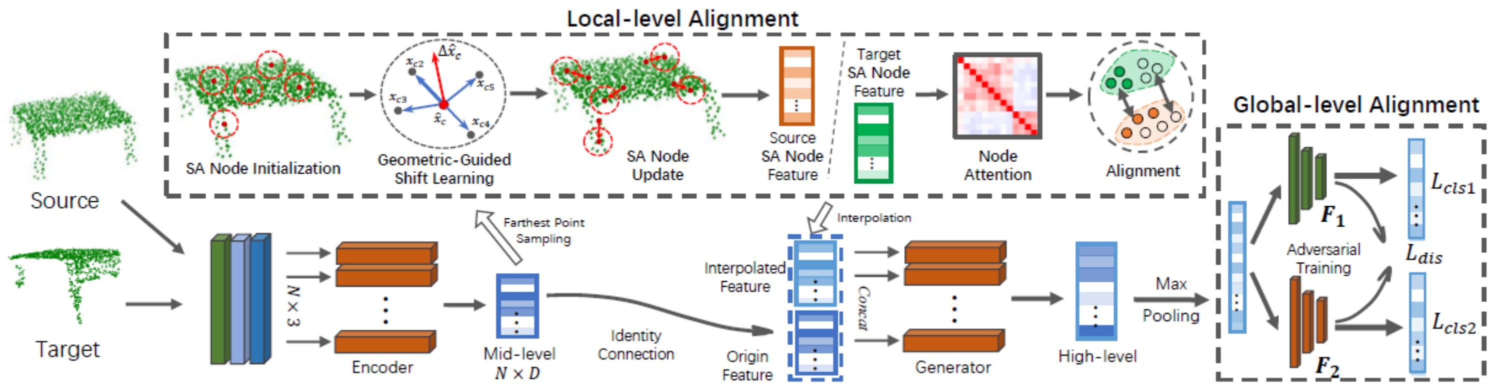


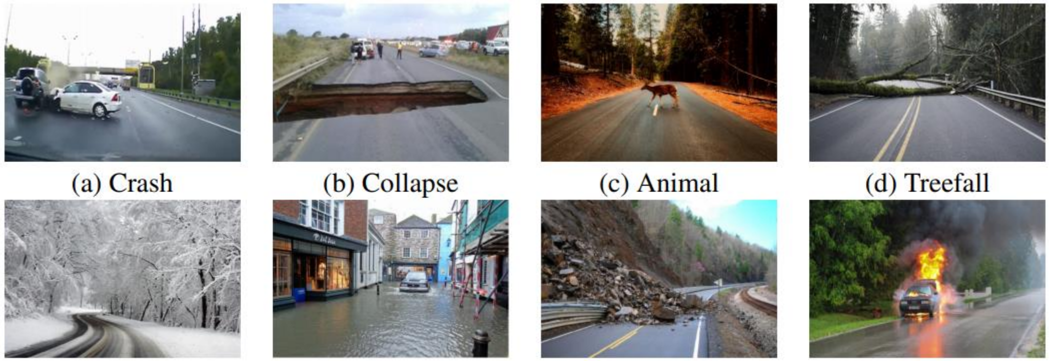
Figure 2: Illustration of PointDAN which mainly consists of local-level and global-level alignment.

Qin, C., You, H., Wang, L., Kuo, C.-C.J., Fu, Y., 2019. PointDAN: A Multi-Scale 3D Domain Adaption Network for Point Cloud Representation, Advances in Neural Information Processing Systems, pp. 7190-7201.

Detection of unsigned road incidents

- Detection by first vehicle at the scene
- Communication to other vehicles (V2V or V2I)
- Other vehicles reroute to avoid the incident
- Approach:
  - Train a CNN using images sourced from Google, Bing, Flickr
  - Apply the trained network to data captured by vehicle-borne sensors

True label	F1	Top-1
Animal on Road	0.9021	95.56%
Road Collapse	0.9174	92.59%
Vehicle Crash	0.9394	96.88%
Fire	0.9848	97.00%
Flooded Road	0.8806	88.68%
Landslide	0.9028	92.86%
Treefall	0.9241	91.78%
Snow on Road	0.9689	96.10%
Negative	0.9854	98.26%



(a) Crash (b) Collapse (c) Animal (d) Treefall
(e) Snow (f) Flood (g) Landslide (h) Fire

Levering, A., Tomko, M., Tuia, D., Khoshelham, K., 2020. Detecting Unsigned Physical Road Incidents from Driver-view Images. IEEE Transactions on Intelligent Vehicles.

Detection of unsigned road incidents

Challenges: severity and relevance

(a) Will the animal cross the road (relevance)?

(b) How severe is the flood?

(c) Will the animal cross the road and how severe is the consequence?

(a) Animal

(b) Flood

(c) Animal

Levering, A., Tomko, M., Tuia, D., Khoshelham, K., 2020. Detecting Unsigned Physical Road Incidents from Driver-view Images. IEEE Transactions on Intelligent Vehicles.

Summary

- Imagery and lidar data offer a great potential for intelligent transport applications
- Automated interpretation is key
- Practical applicability defined by:
  - Accuracy
  - Reliability
  - Computational efficiency
- Role of deep learning, transfer learning, and domain adaptation

Acknowledgements

Works presented were results of collaborations with students and colleagues:

- Debaditya Acharya
- Yan Li
- Milad Ramezani
- Haifeng Luo
- Alex Levering
- Jason Yan
- Xiao Zheng
- Fumi Wu
- Weizhang Chen
- Elham Naghizadeh
- Martin Tomko
- Majid Sarvi

Thank you

k.khoshelham@unimelb.edu.au

THE UNIVERSITY OF MELBOURNE