

Recent Deep Learning Research

What does it mean?

Training Data

Can simulation give us good training data?







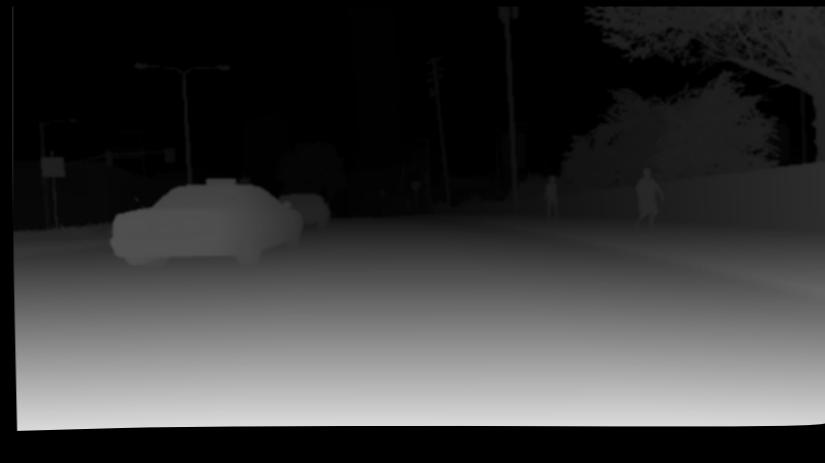




Discount
BEAUTY
STORE



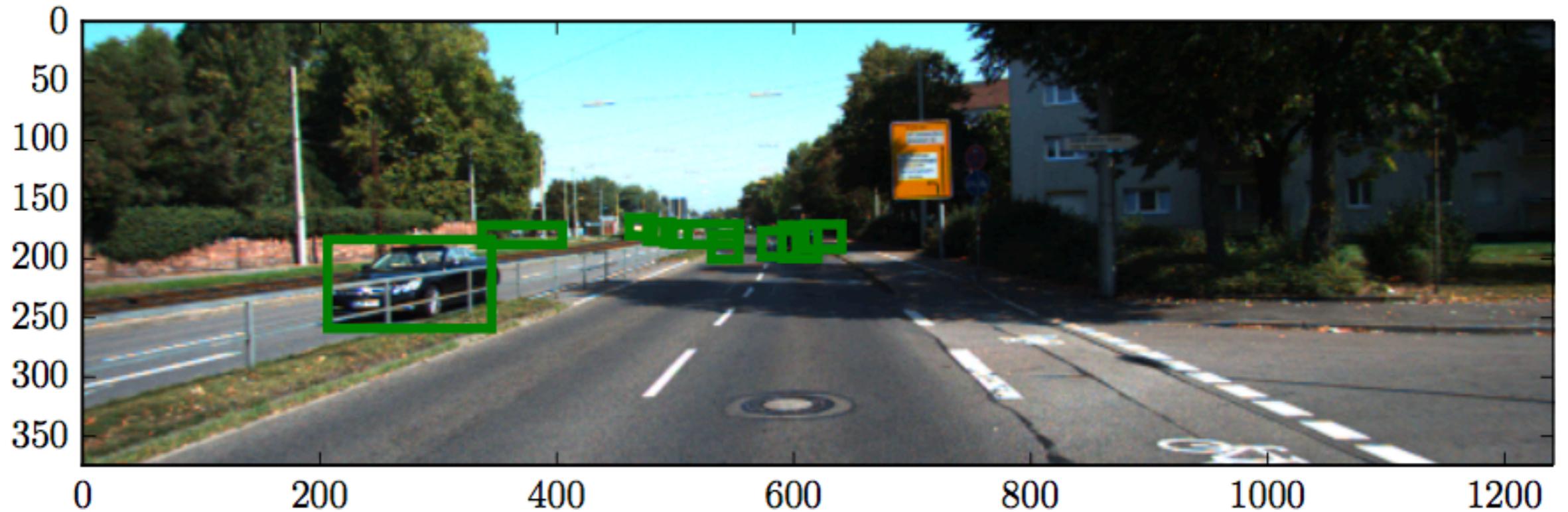




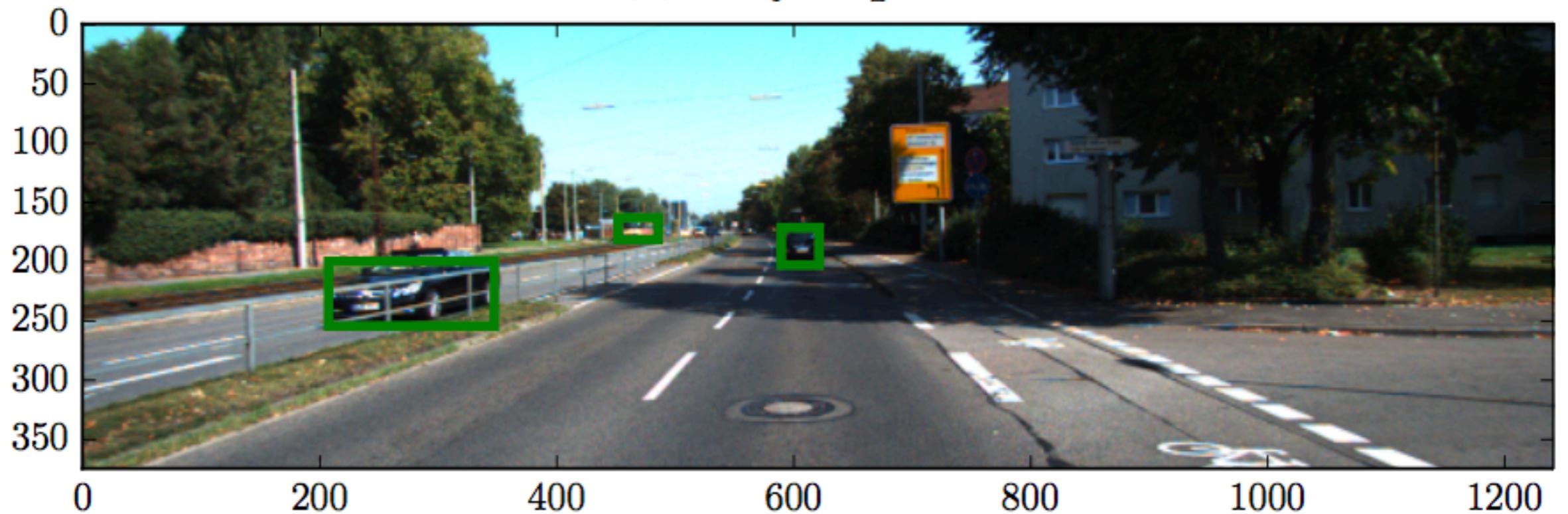


Dataset Sizes

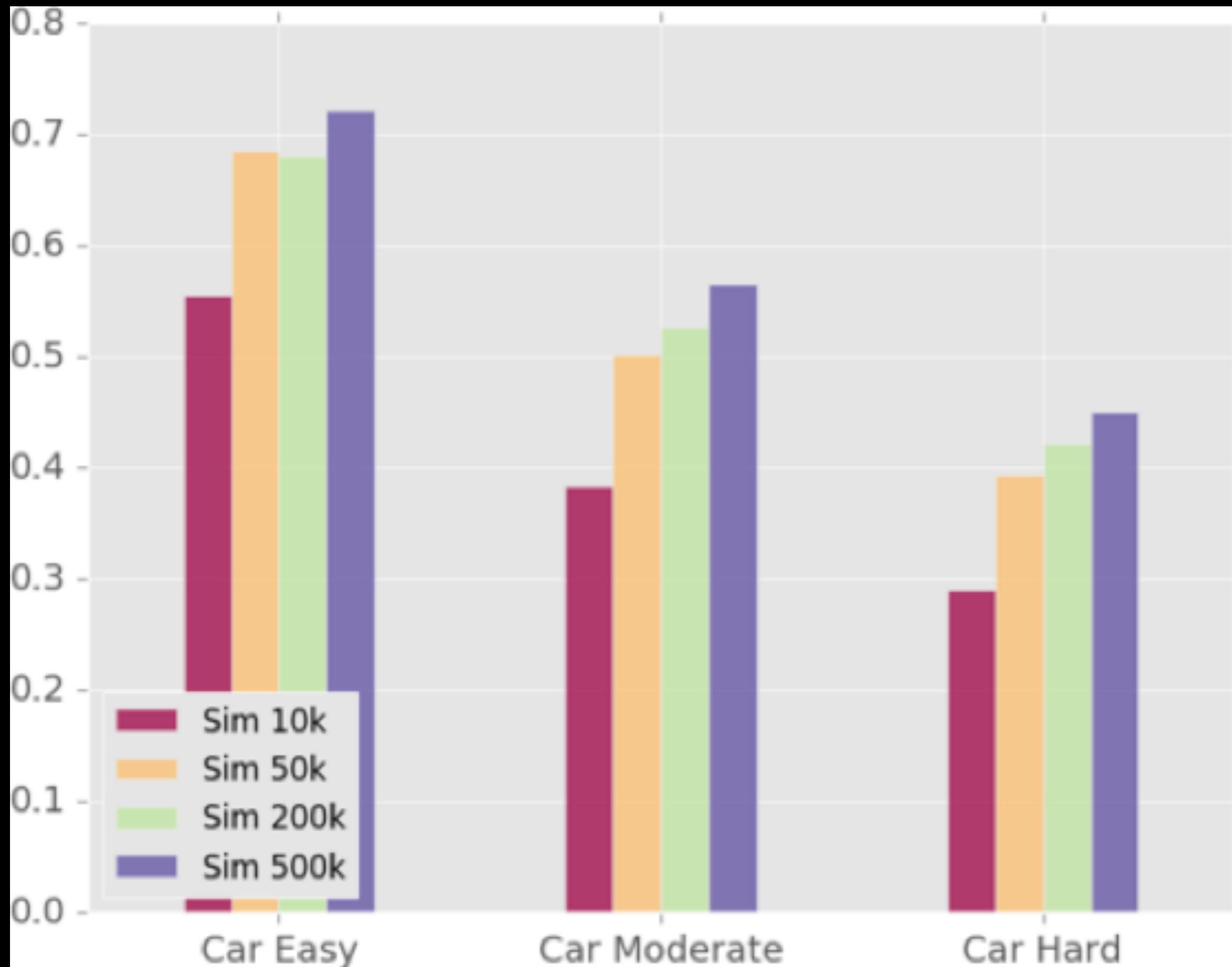
Cityscapes	2,975
KITTI	7,481
Simulation	200,000 Now 500,00



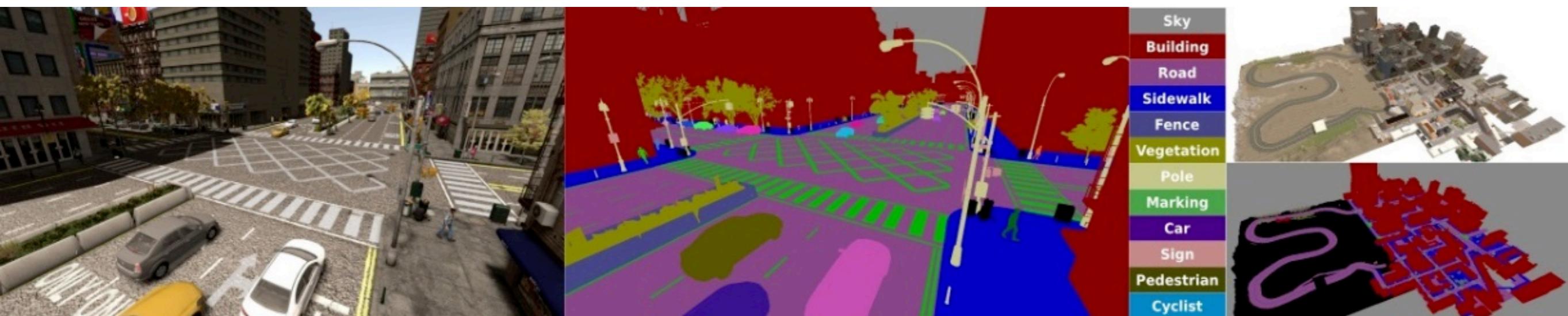
(a) Cityscapes



(b) 200k



	Easy	Moderate	Hard
Sim 10k	0.55	0.38	0.29
Sim 50k	0.68	0.50	0.39
Sim 200k	0.68	0.52	0.42
Sim 500k	0.72	0.56	0.45
Cityscapes	0.62	0.43	0.36



German Ros and Laura Sellart and Joanna Materzynska and David Vazquez and Antonio Lopez

The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes

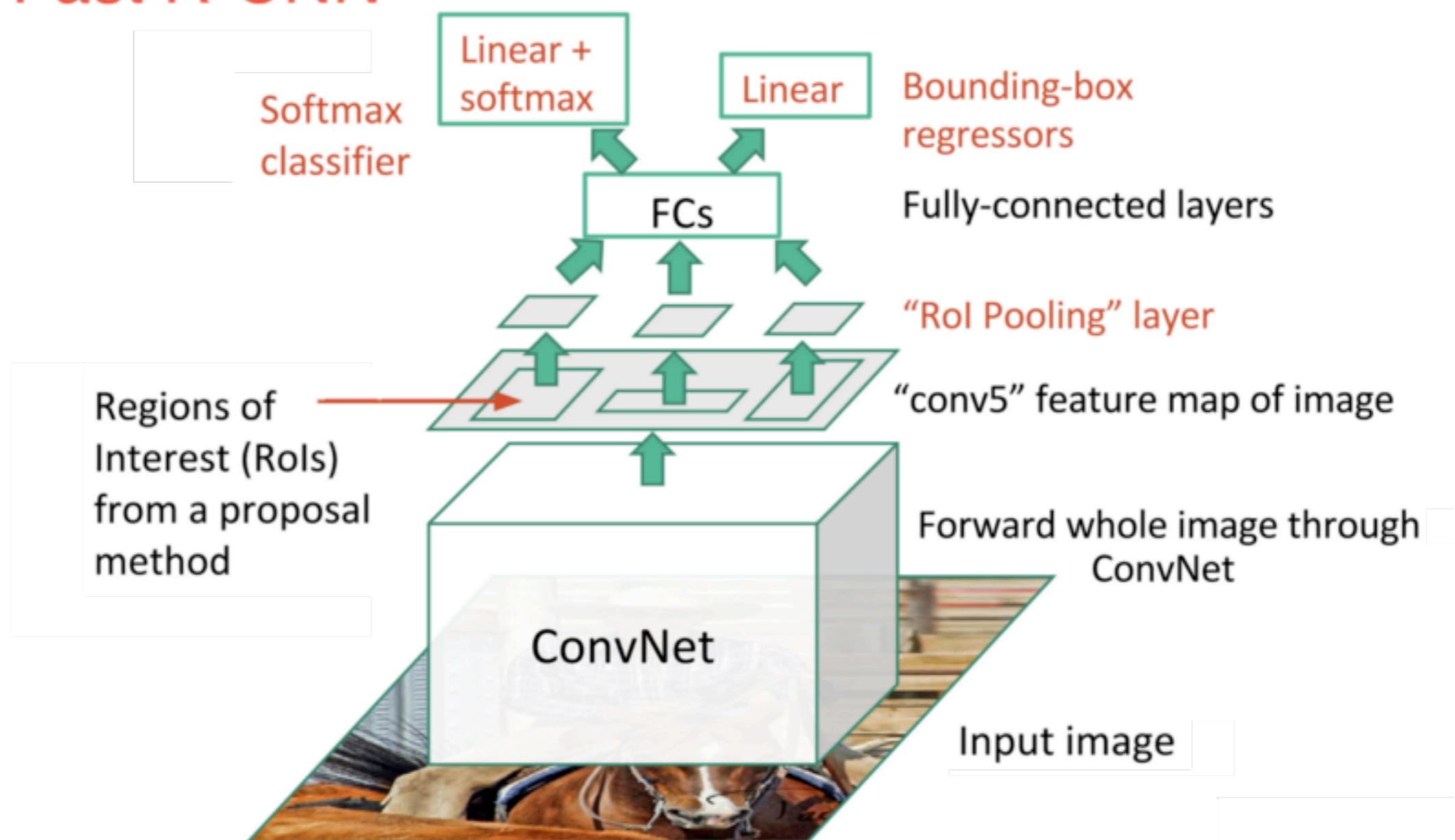
CVPR

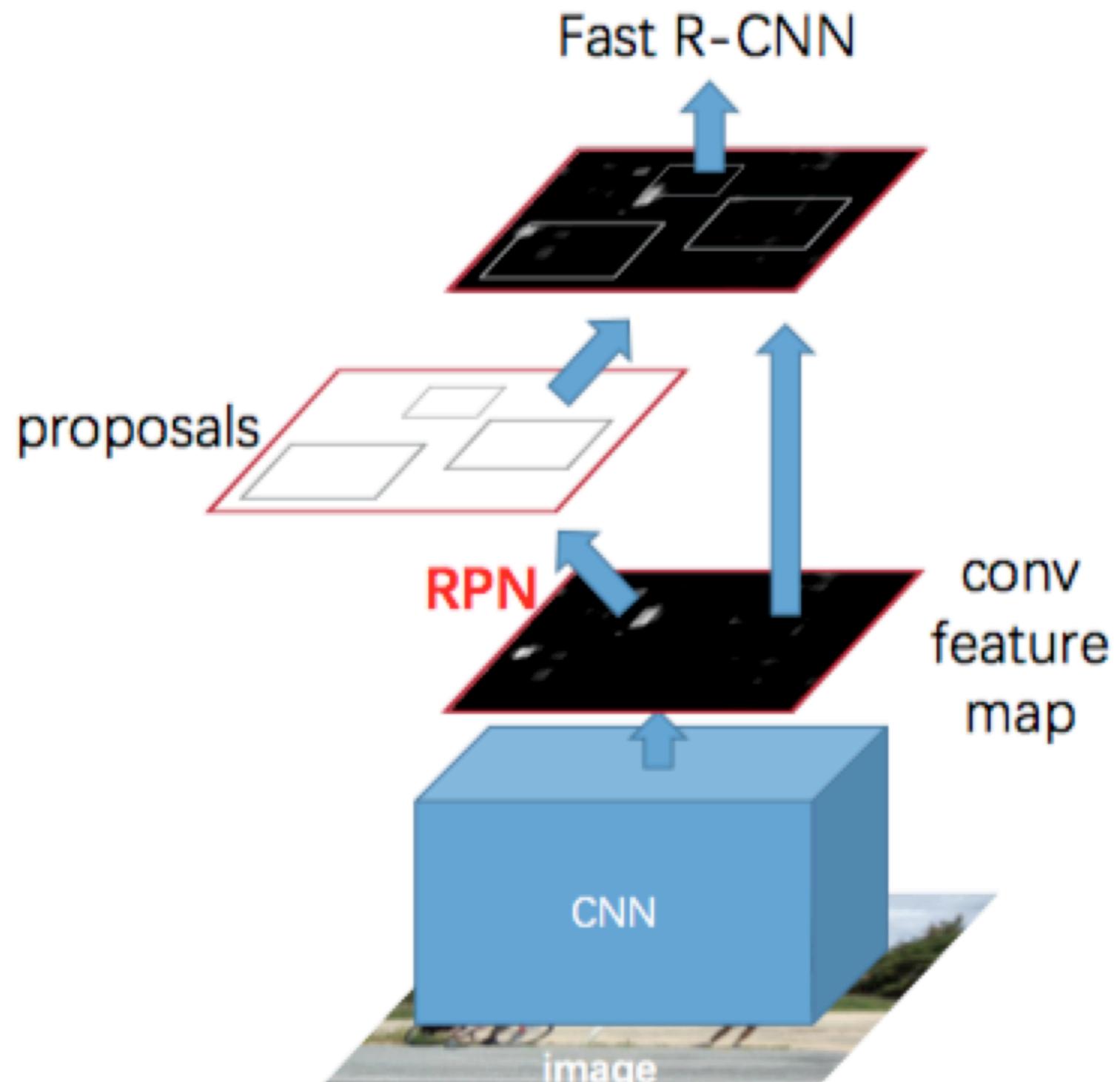


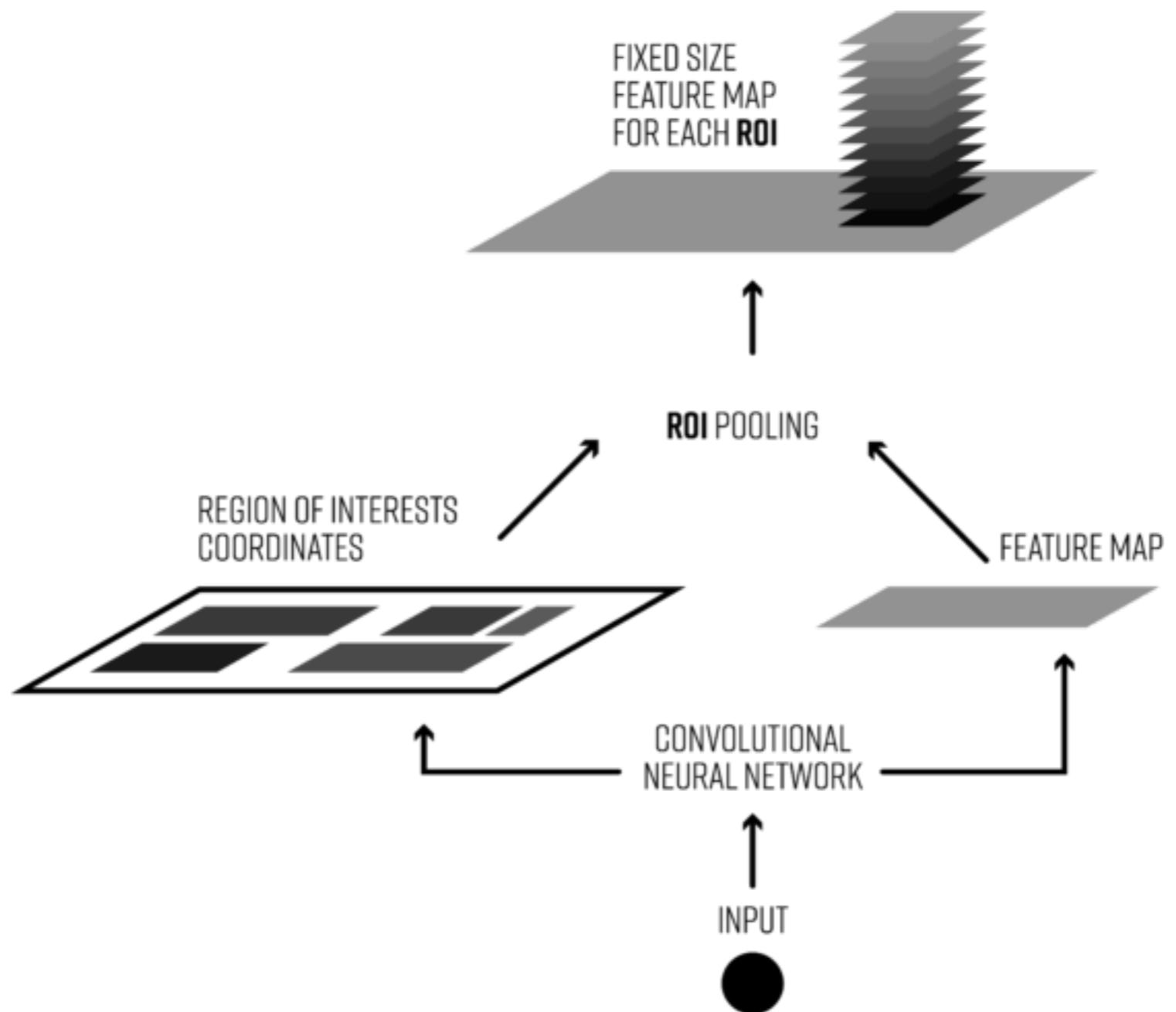
oan Driving Simulator Alexey Dosovitskiy , German Ros, Felipe Codevilla, Antonio Lopez , a

Object Detection

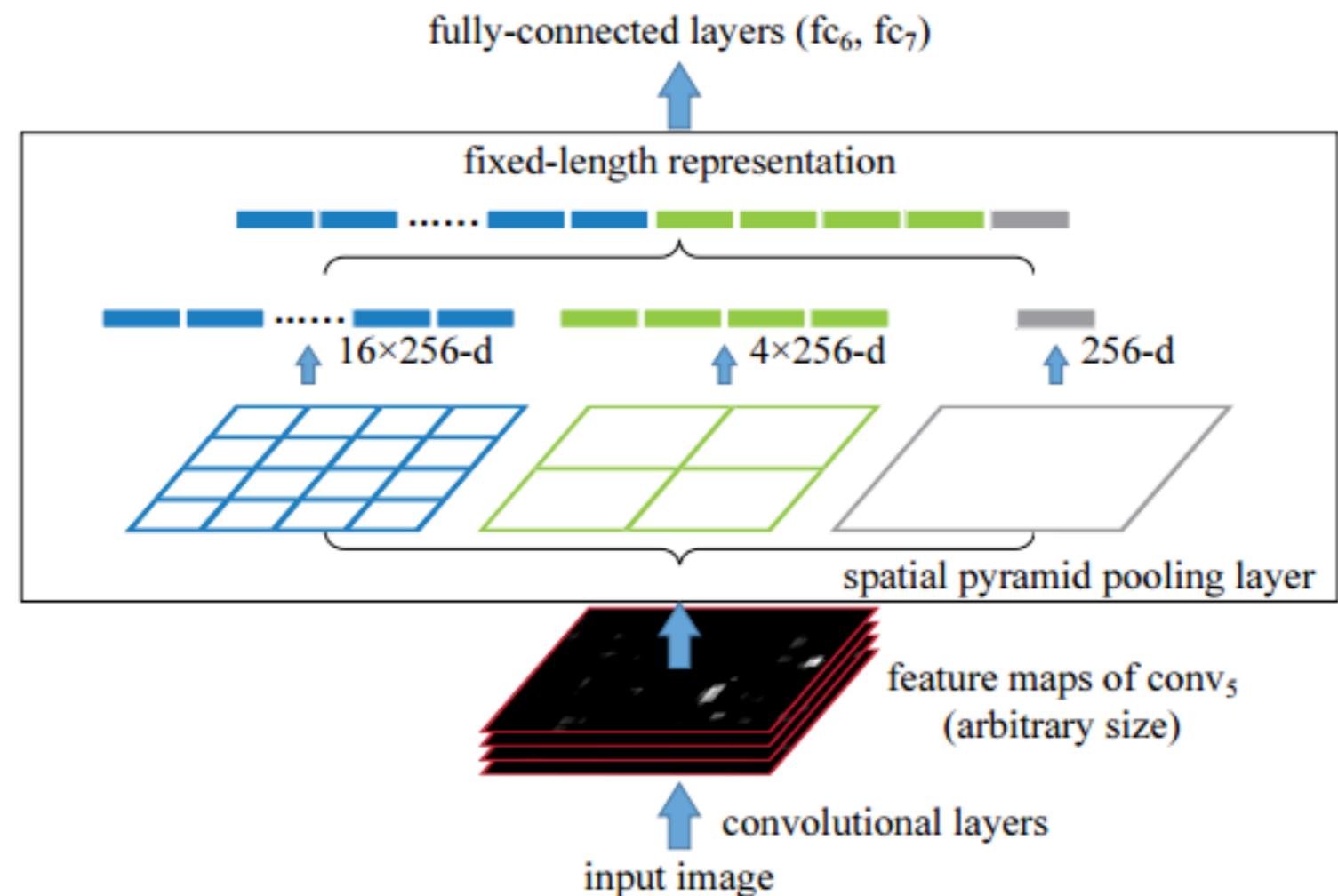
Fast R-CNN

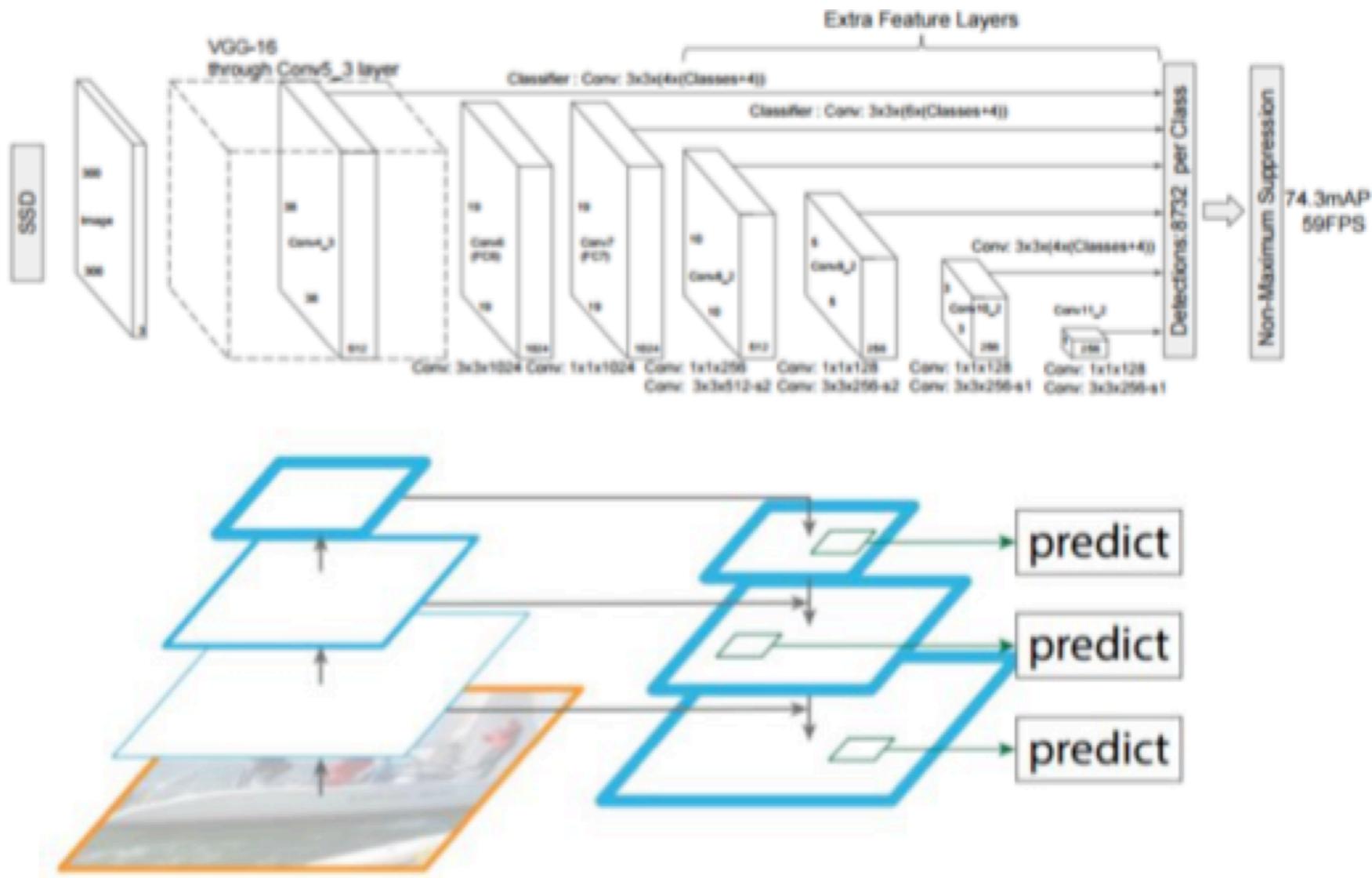


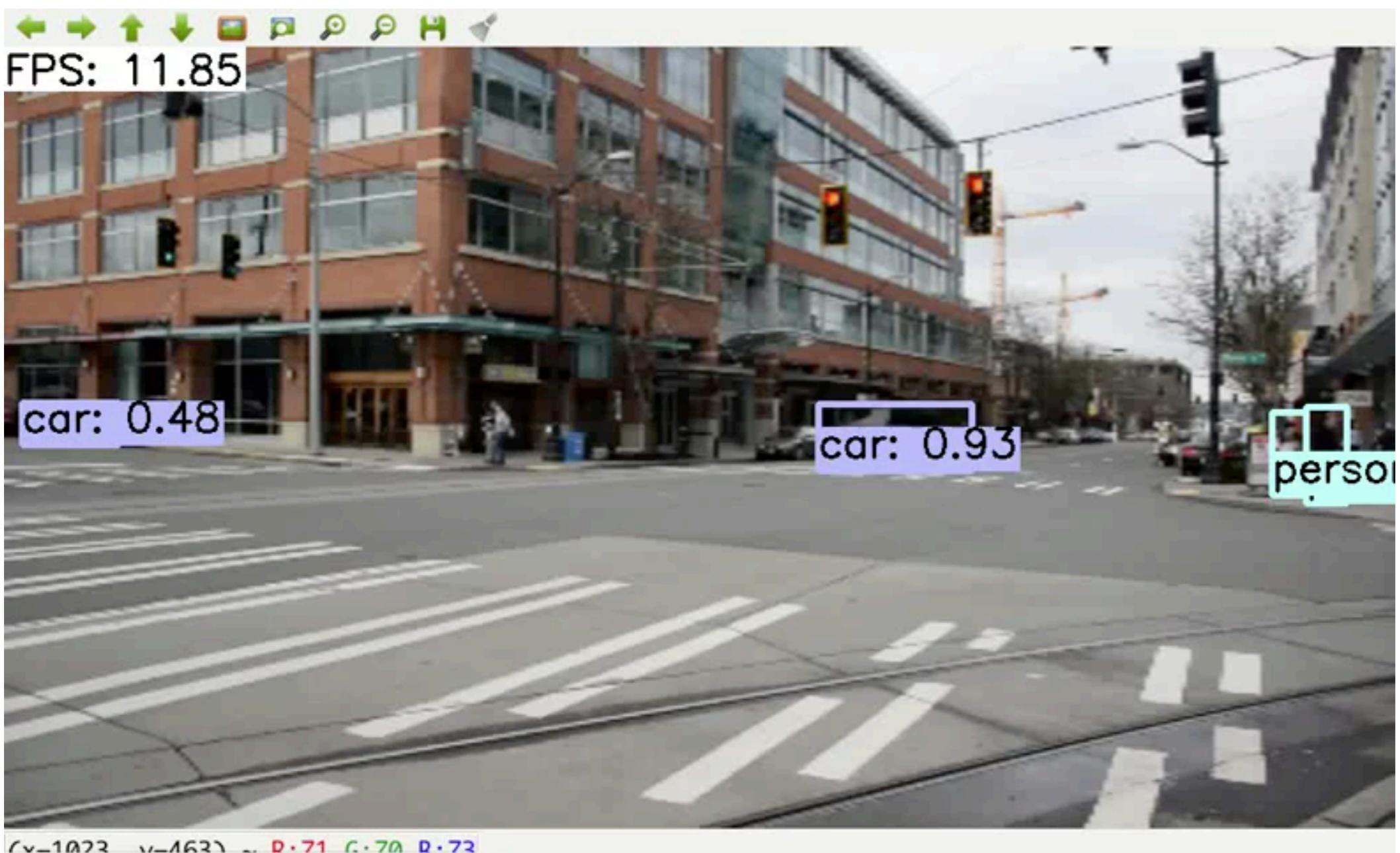








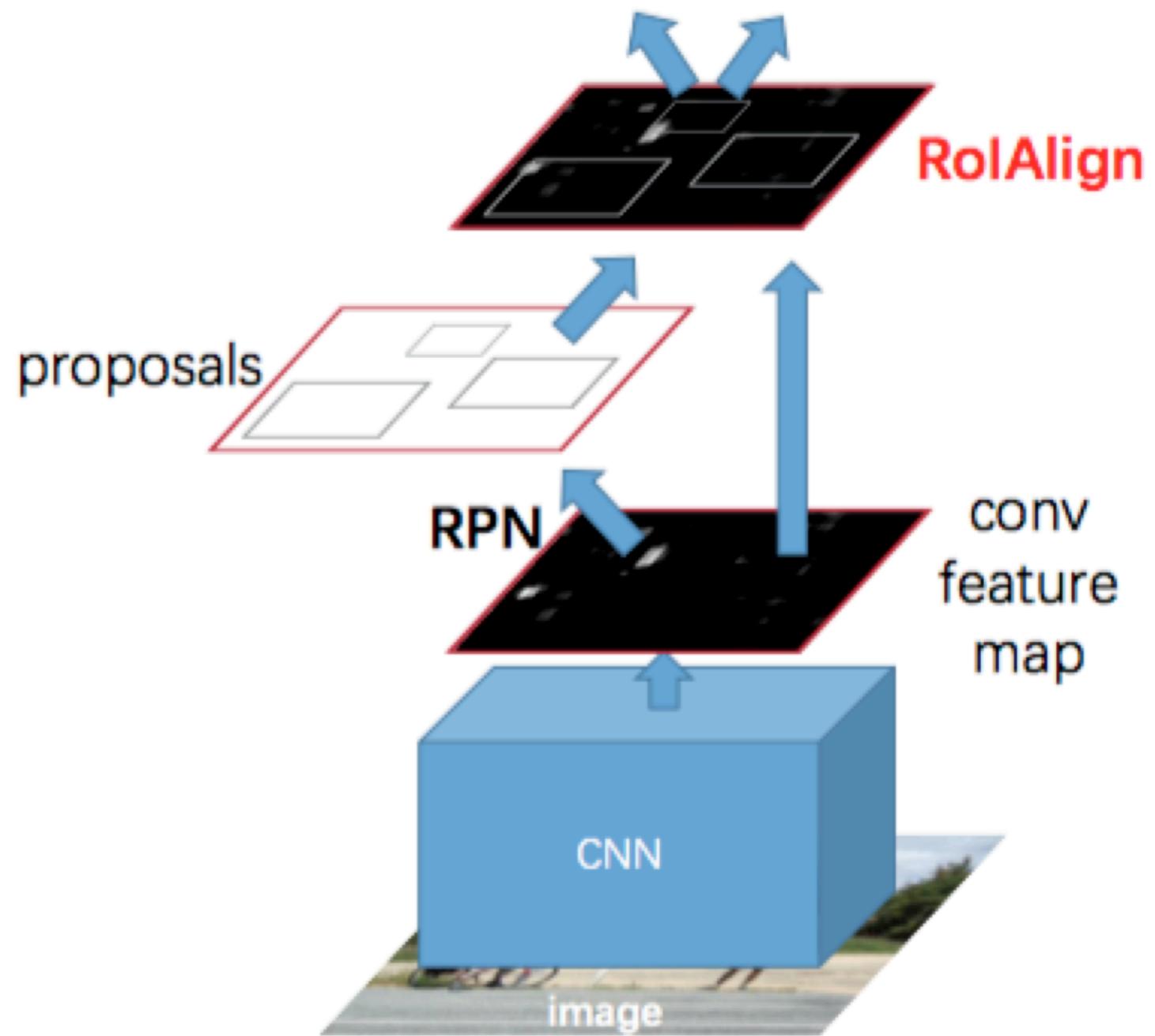


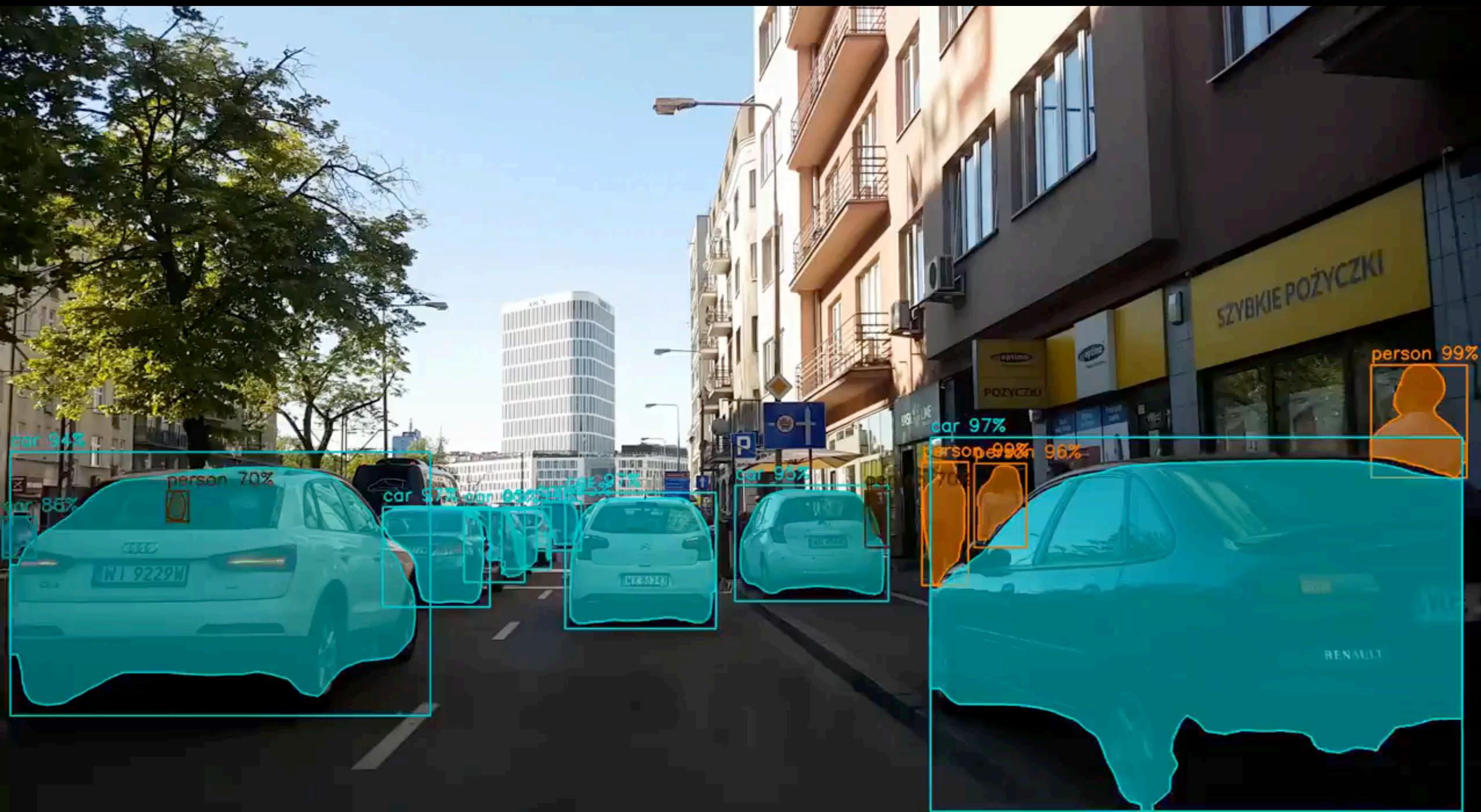


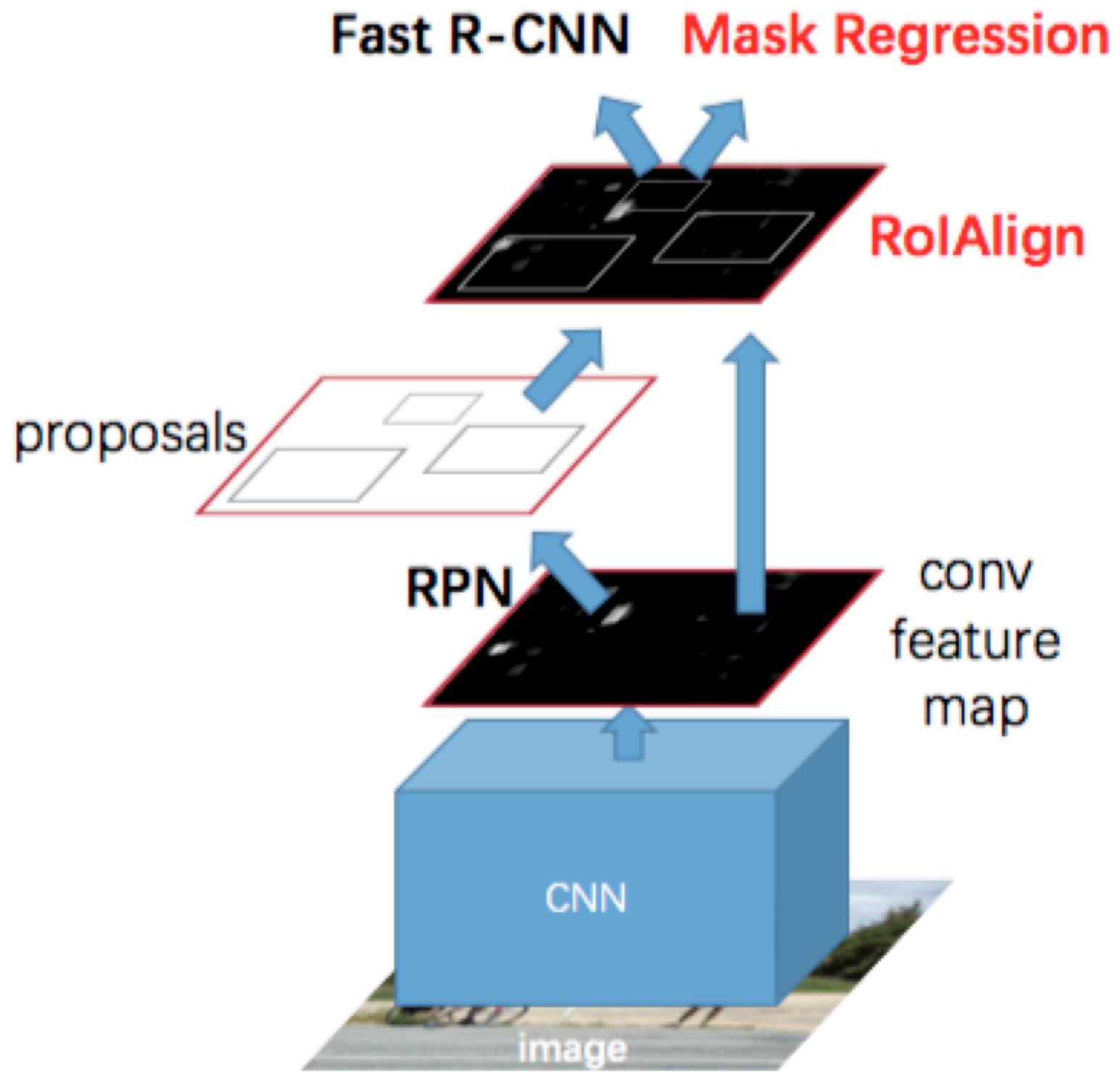
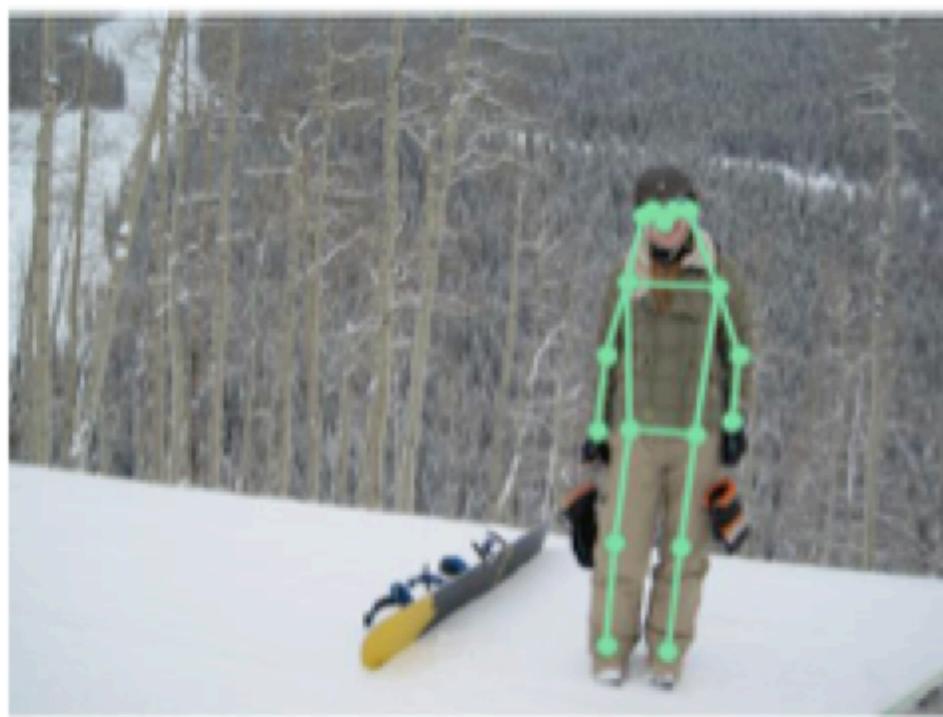
cv-1023 vu-1621 - D: 71 G: 70 R: 72



Fast R-CNN Mask Regression



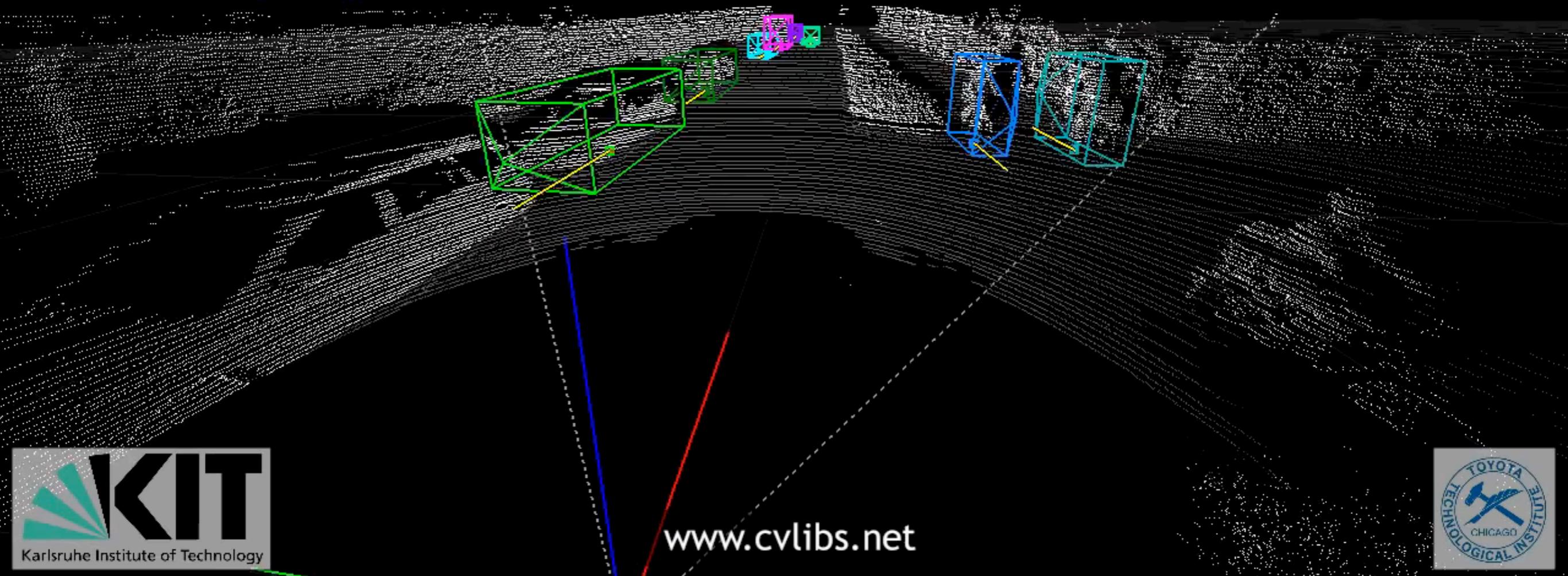




Real-time Multi-Person 2D Pose Estimation Using Part Affinity Fields

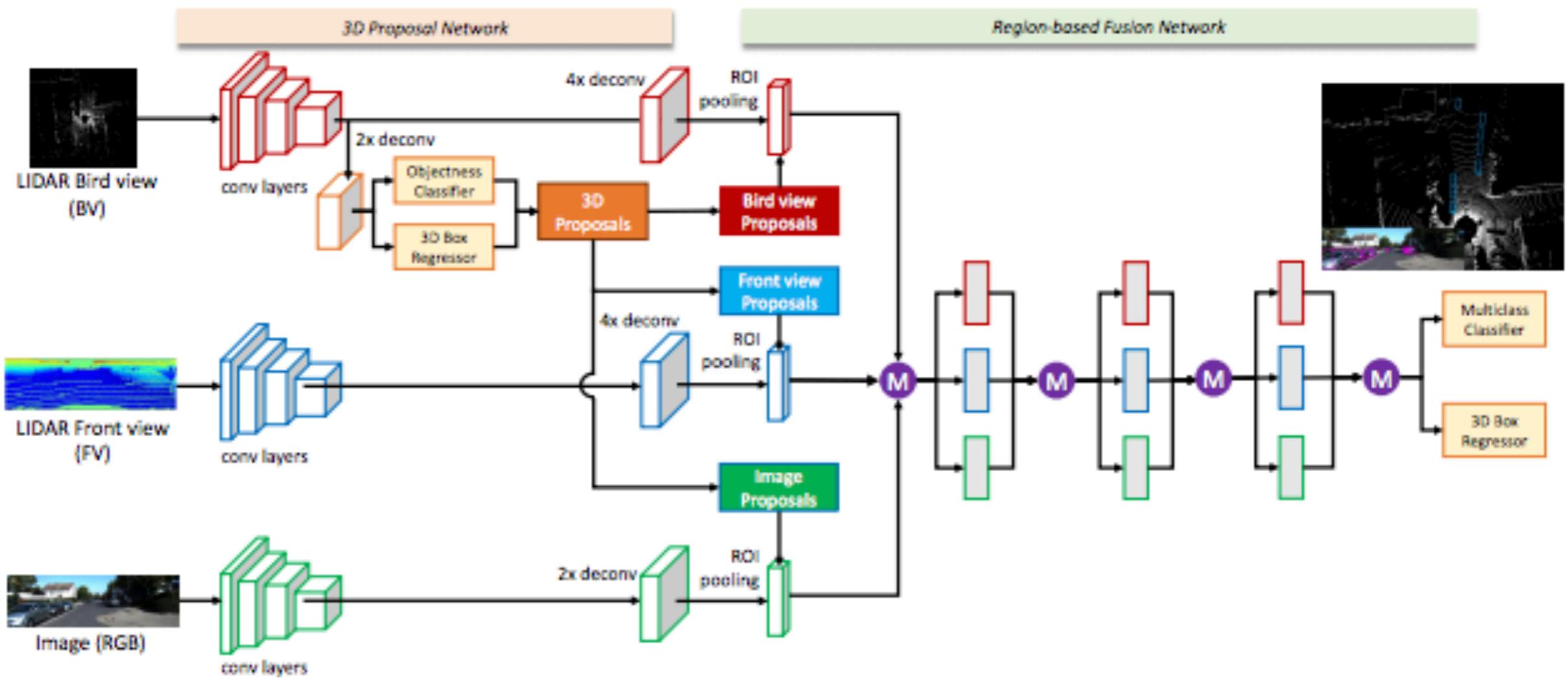
Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh
Carnegie Mellon University

What about 3D?

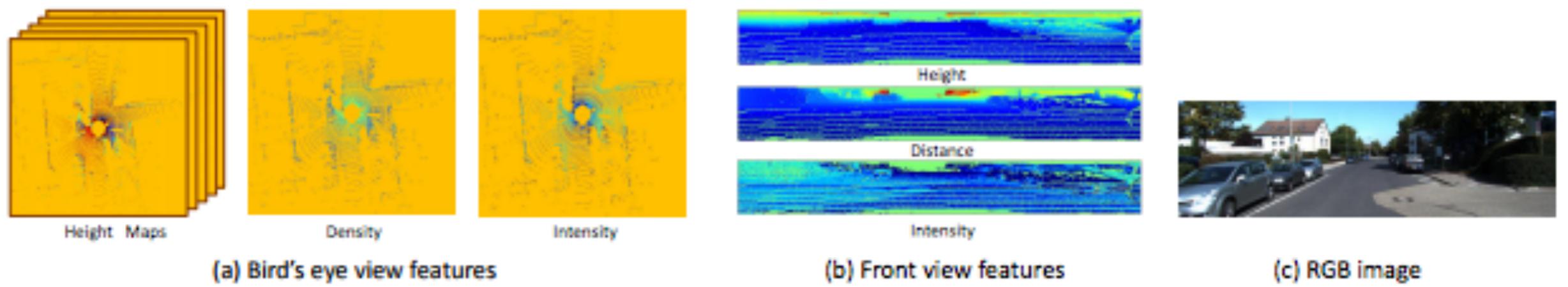


www.cvlabs.net

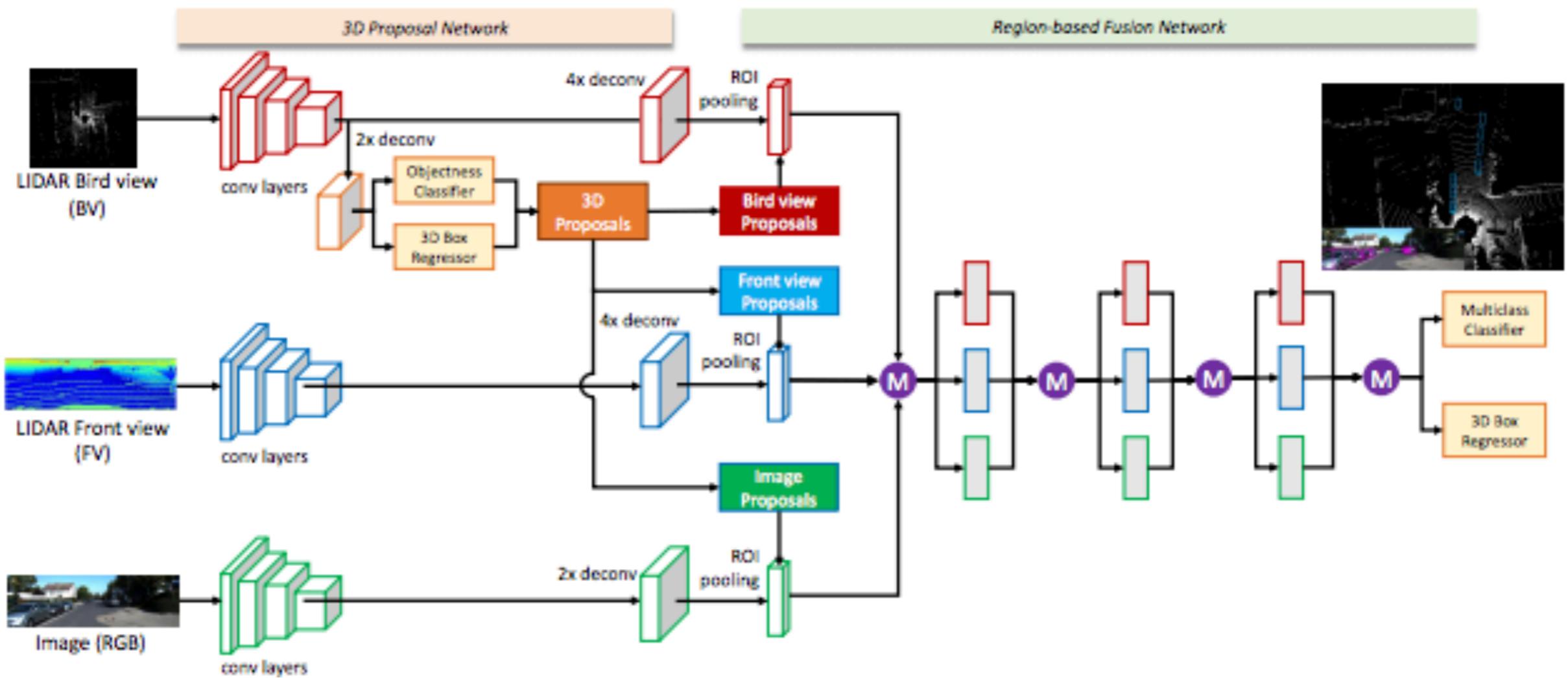




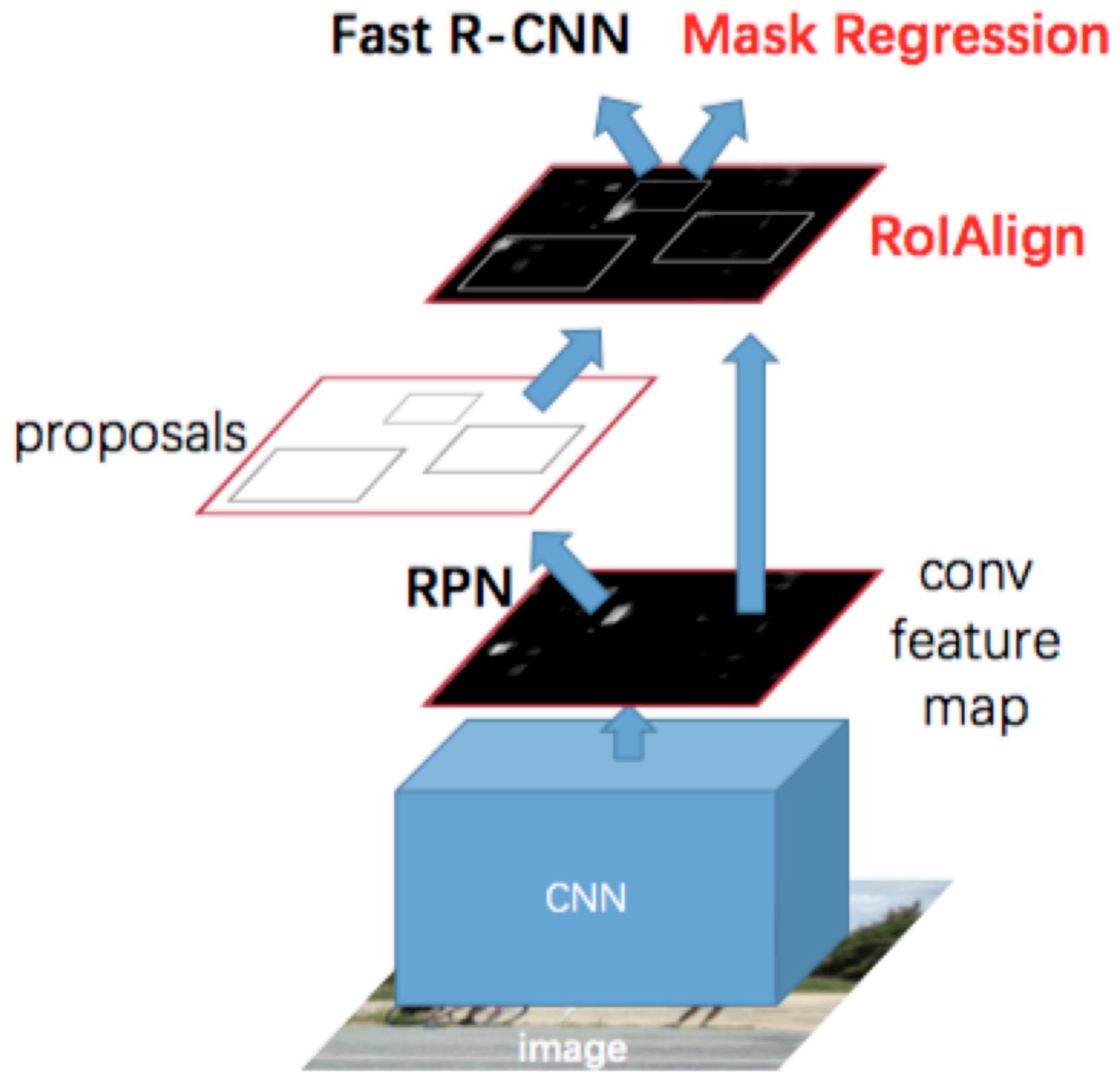
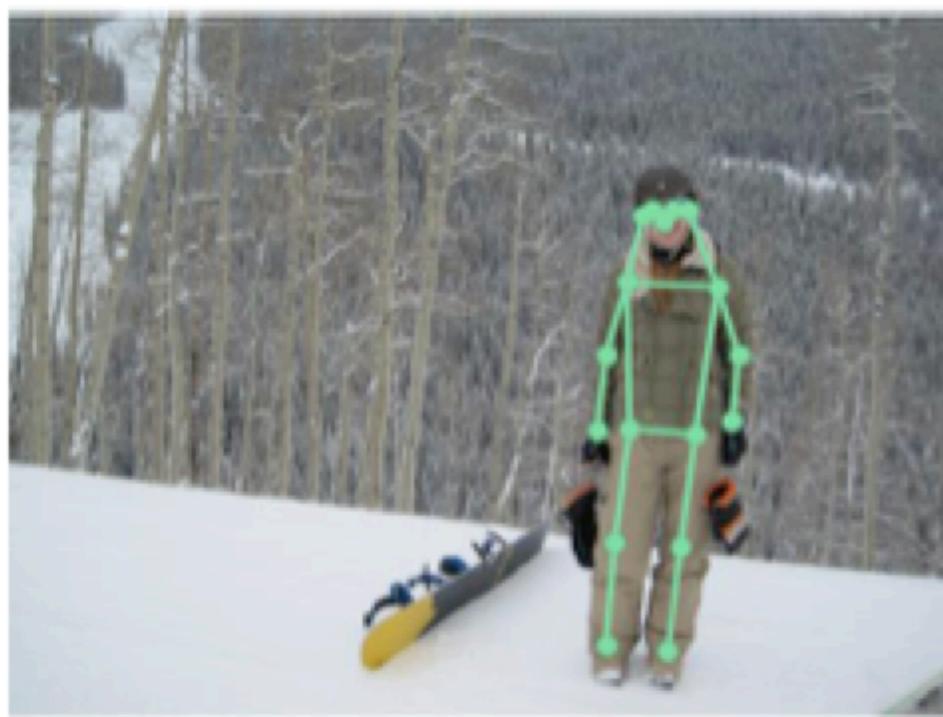
Multi-View 3D Object Detection Network for Autonomous Driving
Chen et. al CVPR 2017



Multi-View 3D Object Detection Network for Autonomous Driving
Chen et. al CVPR 2017



Multi-View 3D Object Detection Network for Autonomous Driving
Chen et. al CVPR 2017



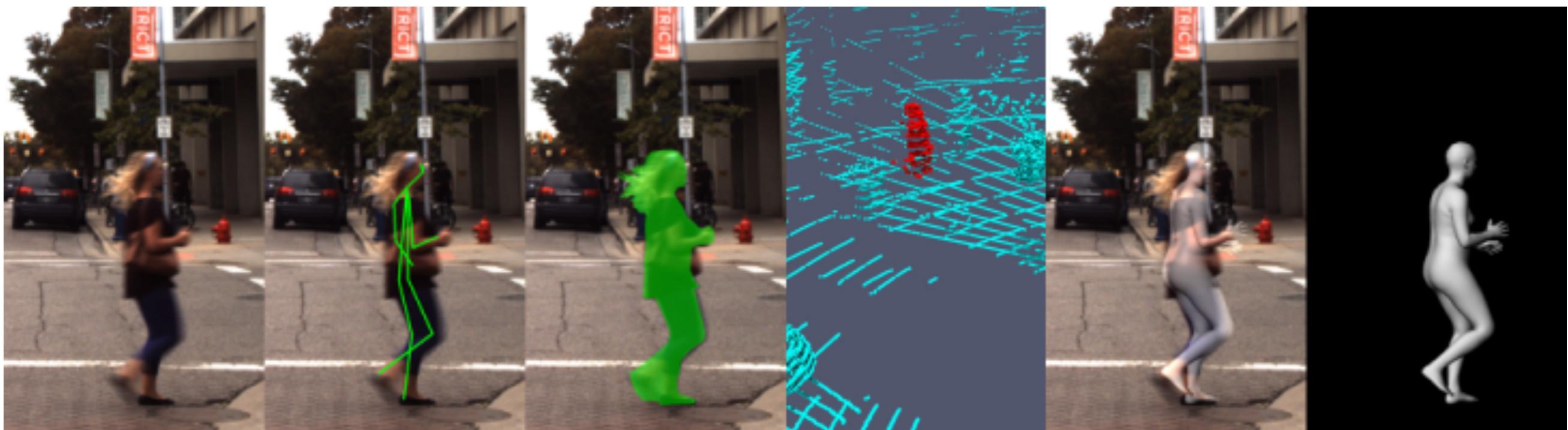




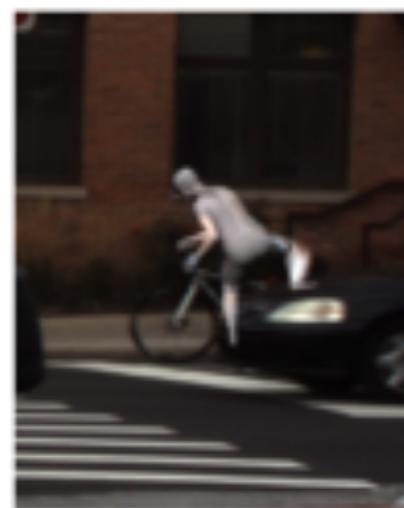
Figure 2. Examples of walking pedestrians with various appearance, surrounding context and gaze.



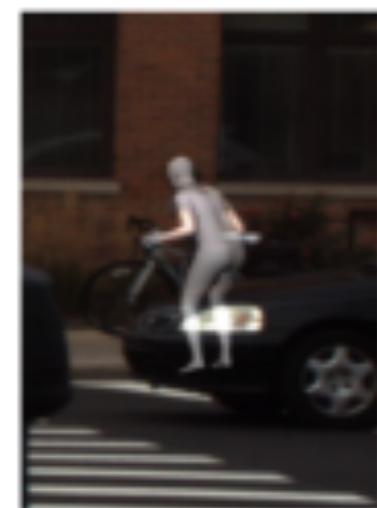
(a) Cycling



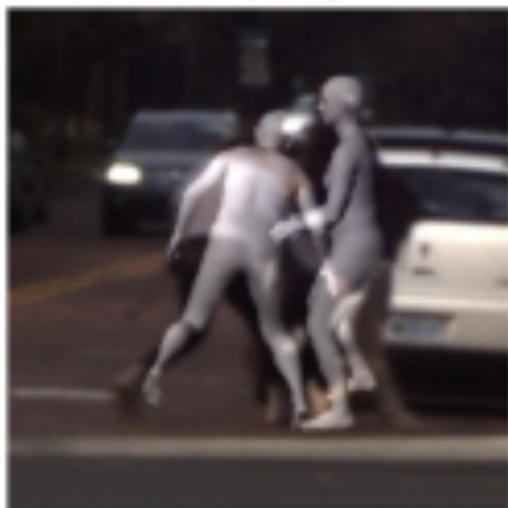
(b) Sitting on a wheelchair



(c) Getting off from the bike



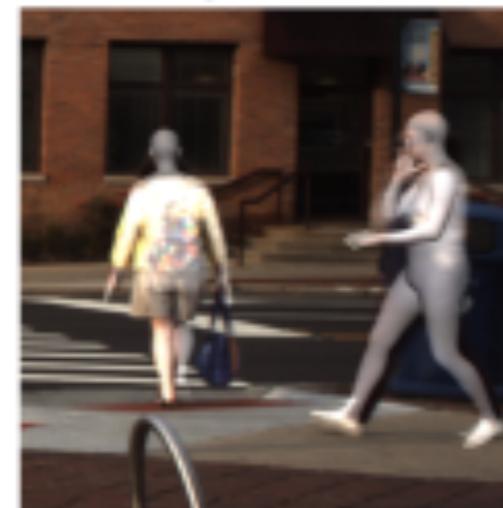
(d) Moving the bike



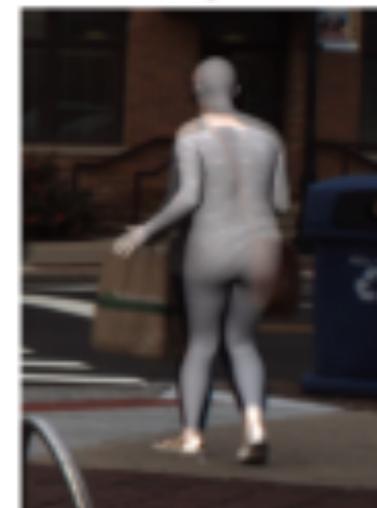
(e) Leaning down



(f) Running



(g) Talking on phone



(h) Carrying a bag

Figure 3. Interesting poses. (a) to (e) are cases in further distances, (f) to (h) are cases in closer distances.

Time: 0.000000







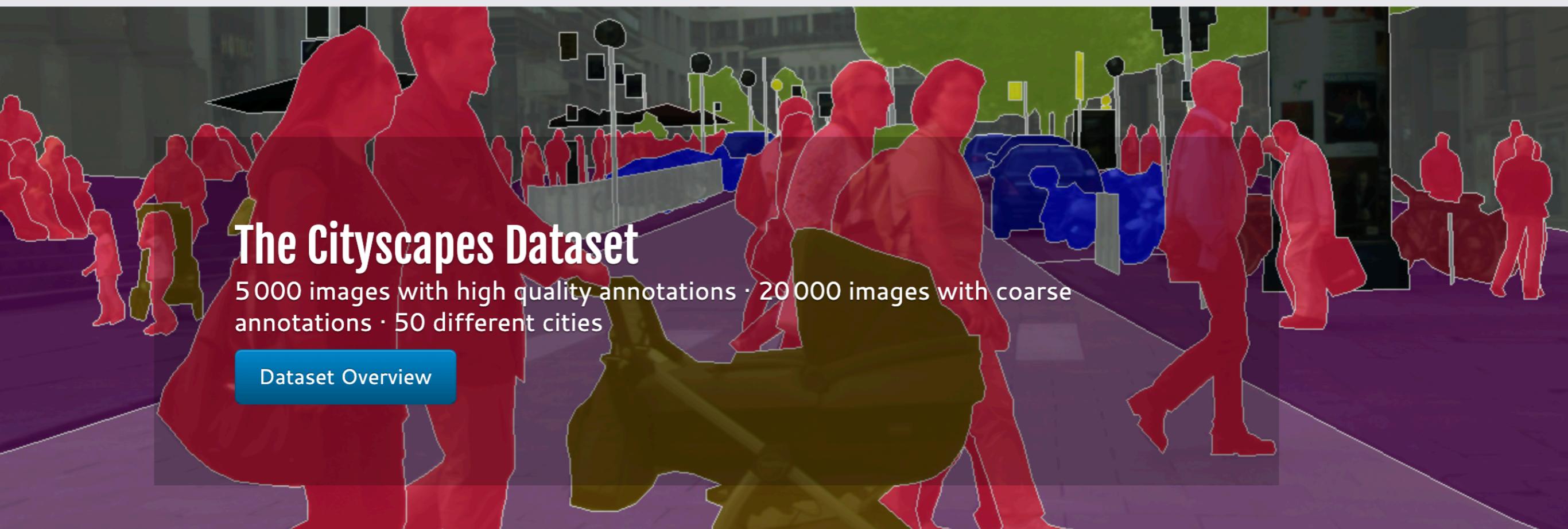
Semantic Seg





Semantic Understanding of Urban Street Scenes

News Overview ▾ Examples ▾ Benchmarks ▾ Download Submit Citation Contact ▾



The Cityscapes Dataset

5 000 images with high quality annotations · 20 000 images with coarse
annotations · 50 different cities

[Dataset Overview](#)

Fully Convolutional Networks for Semantic Segmentation

Jonathan Long*

Evan Shelhamer*

Trevor Darrell

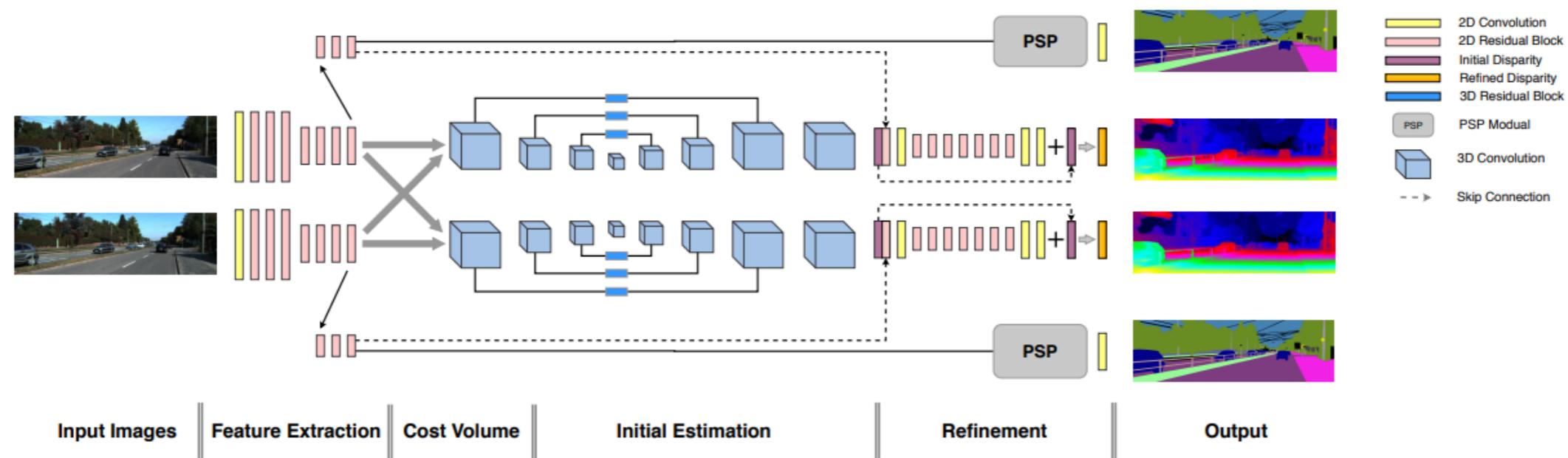
UC Berkeley

{jonlong, shelhamer, trevor}@cs.berkeley.edu

Depth Estimation

DispSegNet: Leveraging Semantics for End-to-End Learning of Disparity Estimation from Stereo Imagery

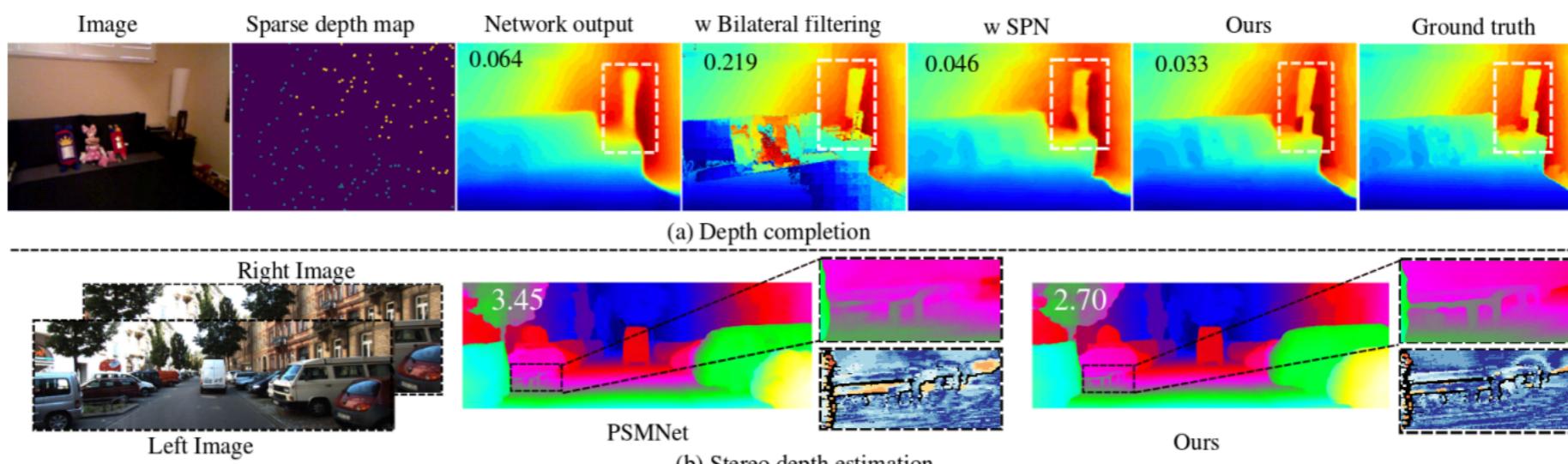
Junming Zhang¹, Katherine A. Skinner², Ram Vasudevan³ and Matthew Johnson-Roberson⁴

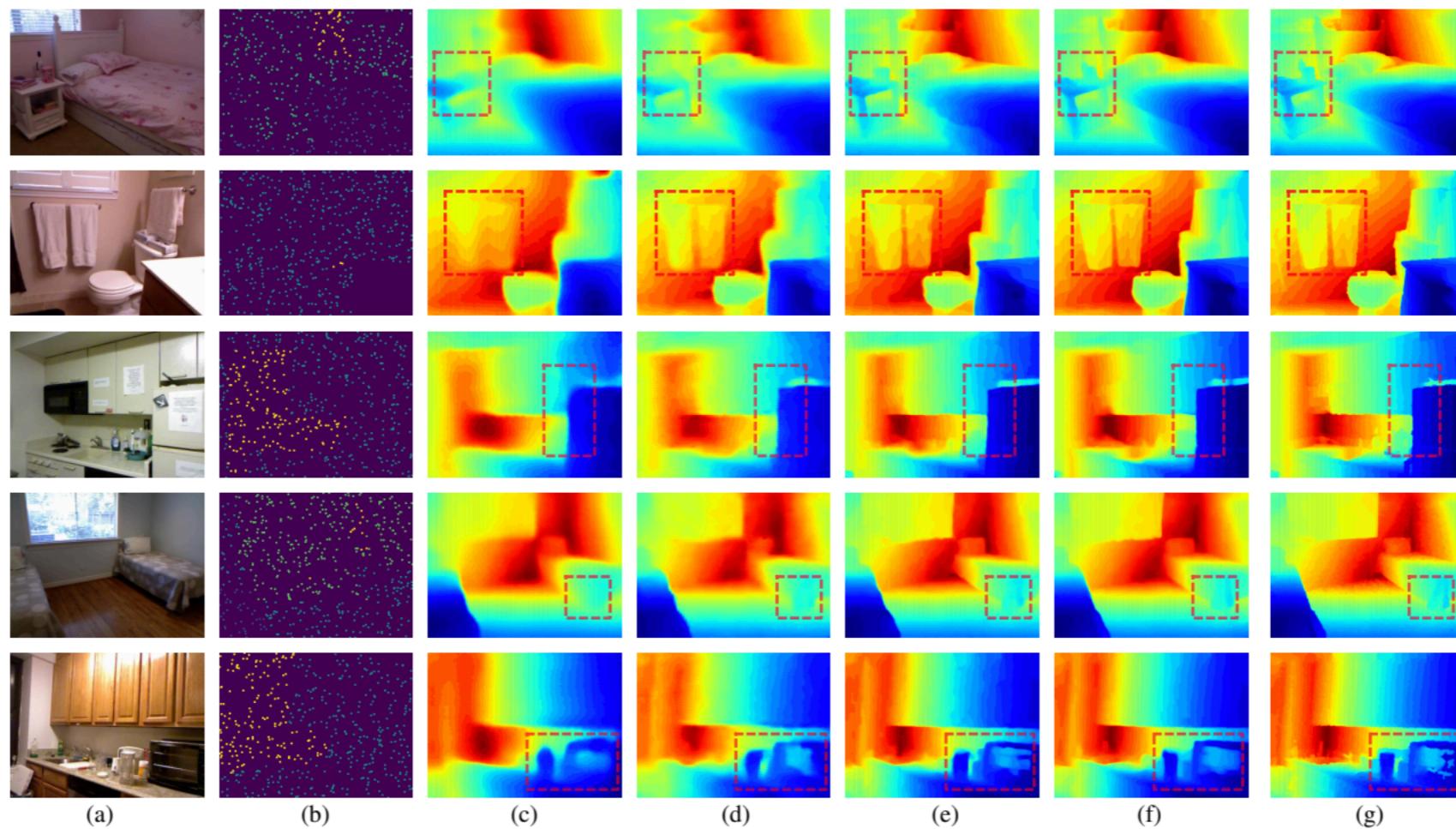


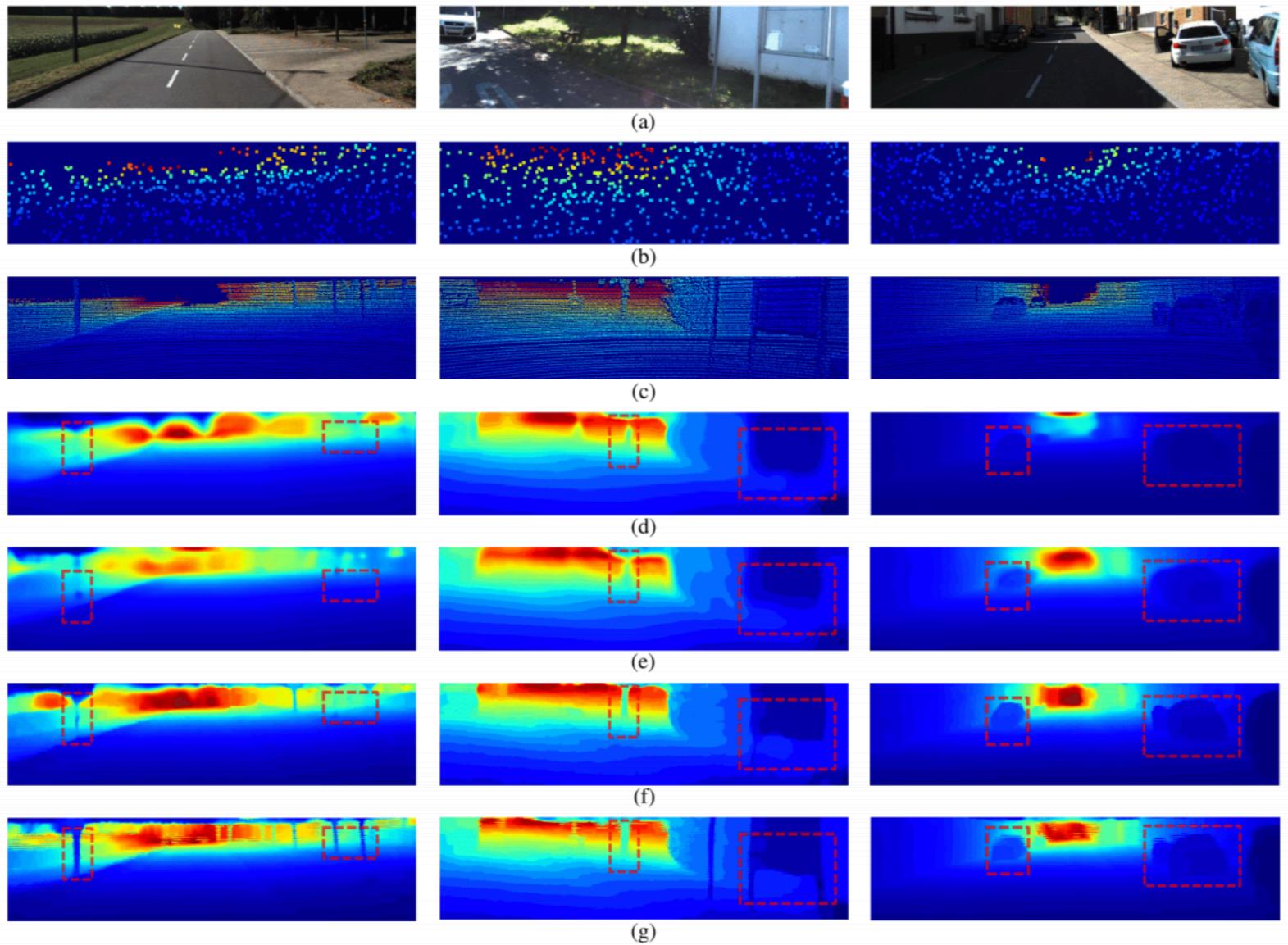
Depth Completion

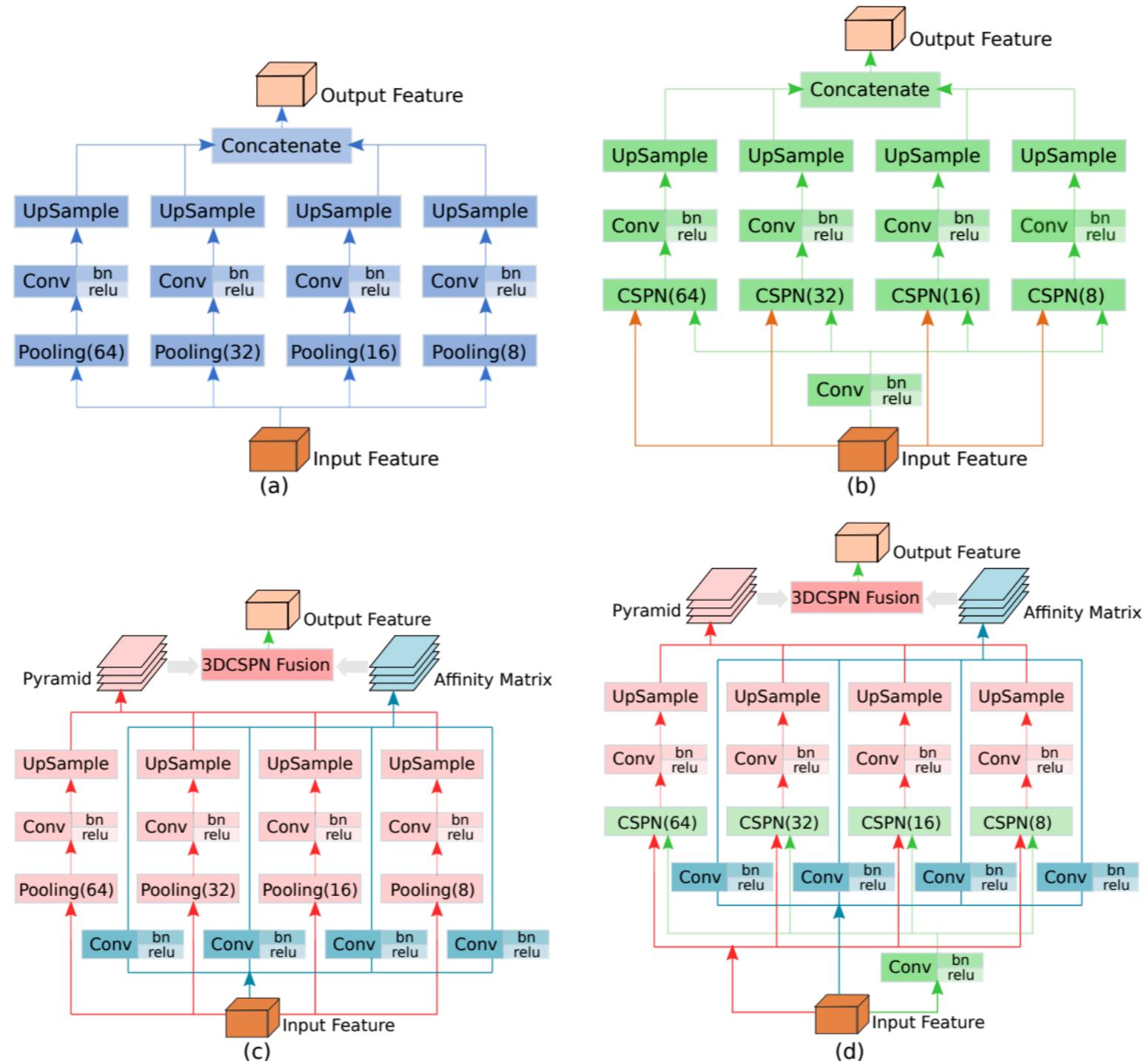
Learning Depth with Convolutional Spatial Propagation Network

Xinjing Cheng, Peng Wang and Ruigang Yang, *Senior Member, IEEE*







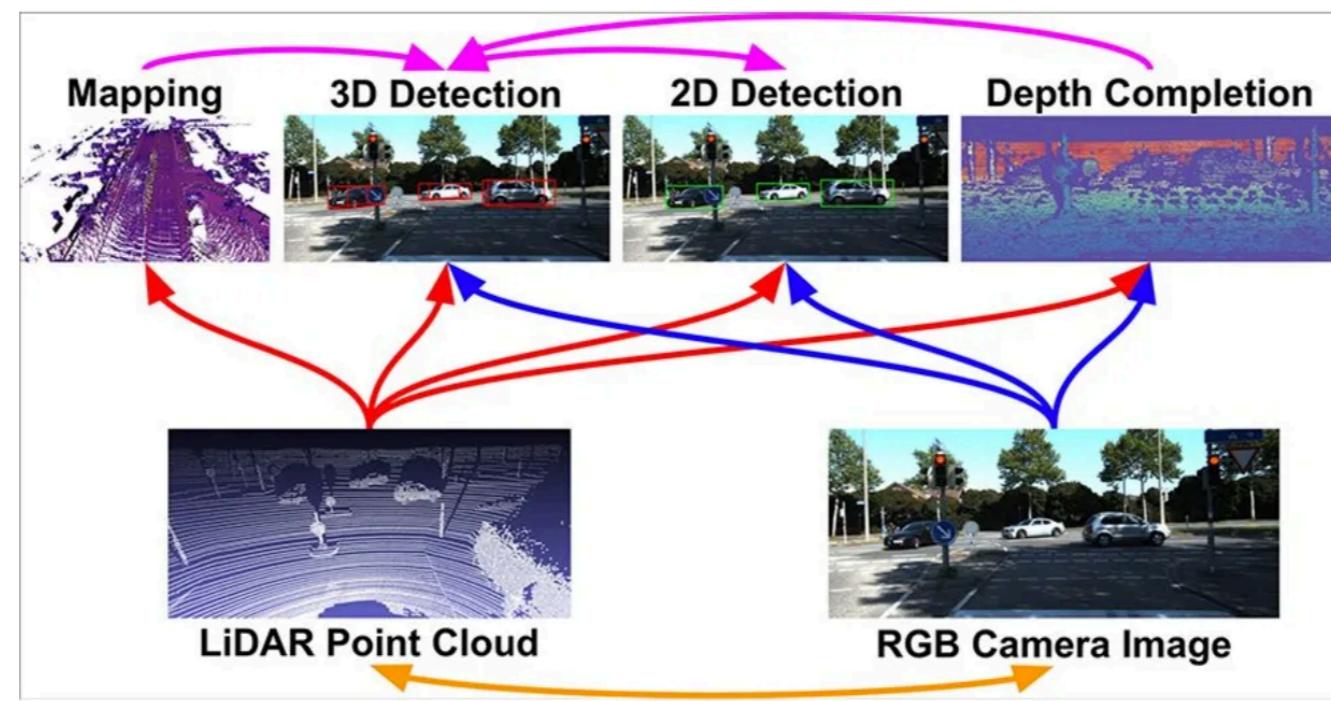


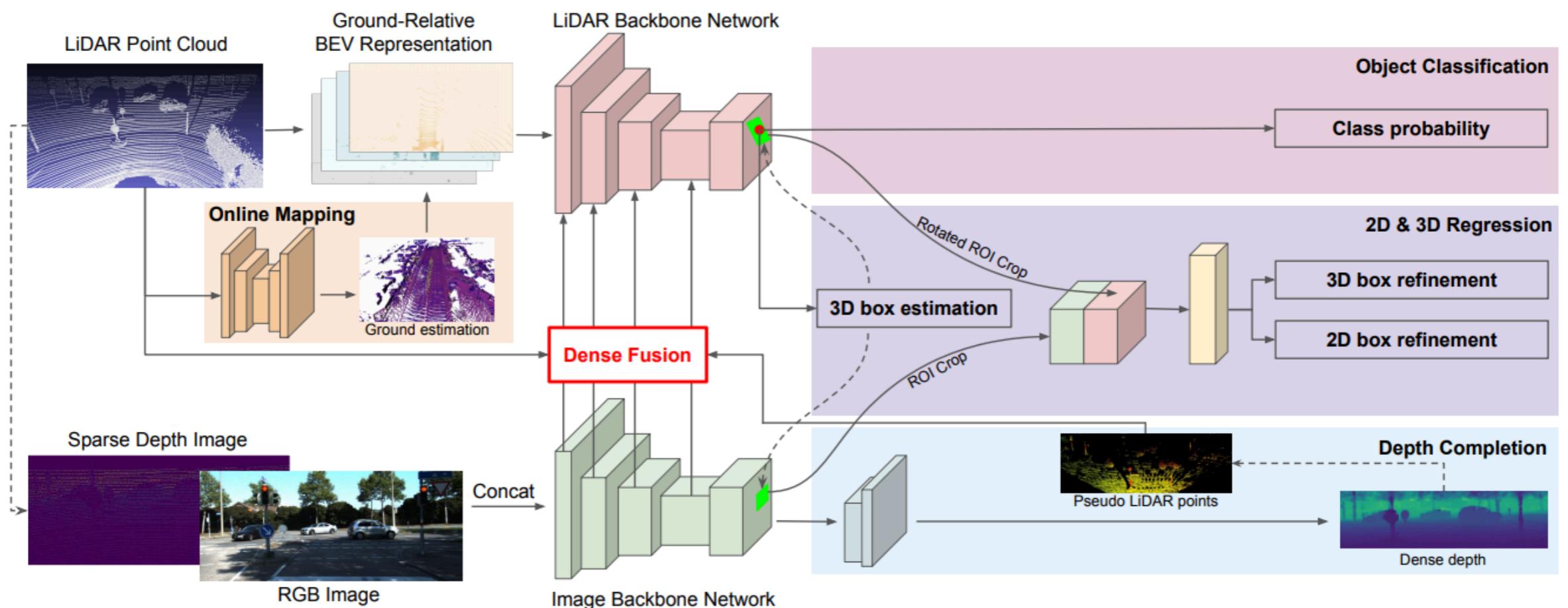
Multi-Task Multi-Sensor Fusion for 3D Object Detection

Ming Liang^{1*} Bin Yang^{1,2*} Yun Chen^{1†} Rui Hu¹ Raquel Urtasun^{1,2}

¹Uber Advanced Technologies Group ²University of Toronto

{ming.liang, byang10, yun.chen, rui.hu, urtasun}@uber.com





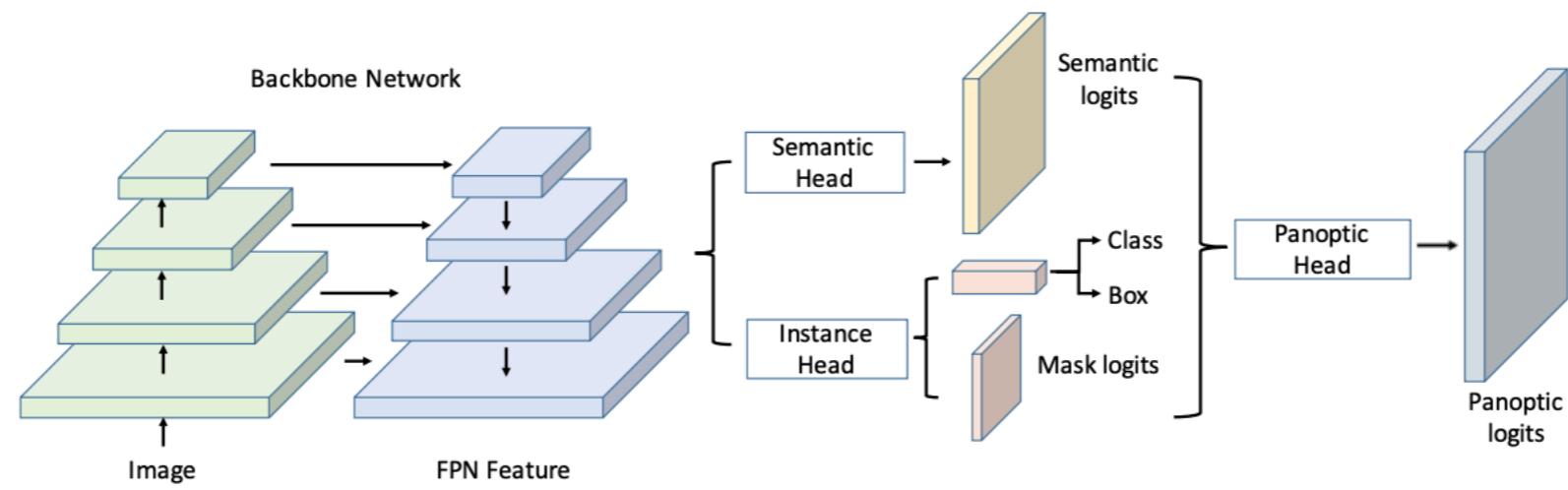
UPSNNet: A Unified Panoptic Segmentation Network

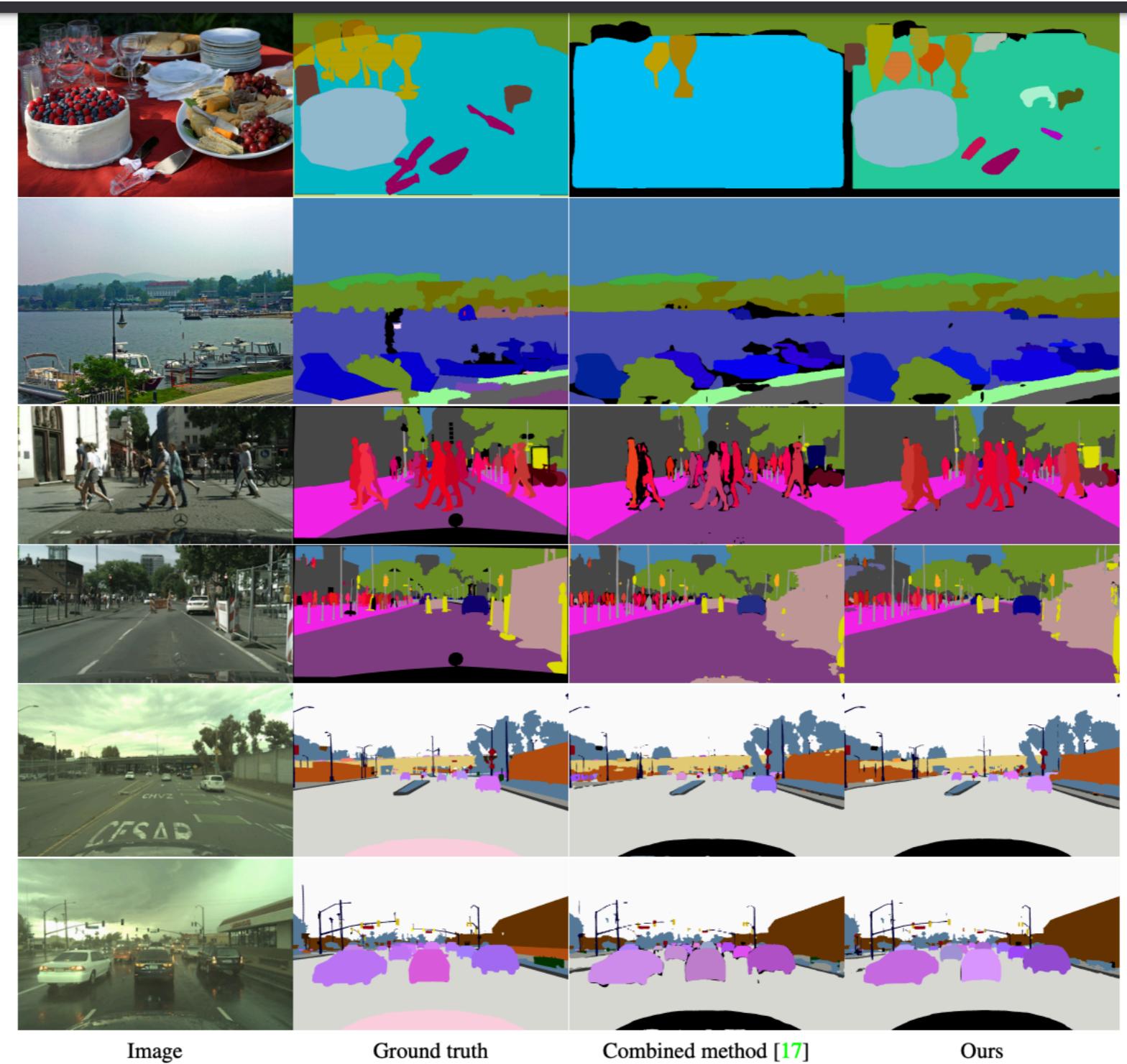
Yuwen Xiong^{1,2*} Renjie Liao^{1,2*} Hengshuang Zhao^{3*†},
Rui Hu¹ Min Bai^{1,2} Ersin Yumer¹ Raquel Urtasun^{1,2}

¹Uber ATG ²University of Toronto ³The Chinese University of Hong Kong

{yuwen, rjliao, rui.hu, mbai3, yumer, urtasun}@uber.com

hszhao@cse.cuhk.edu.hk

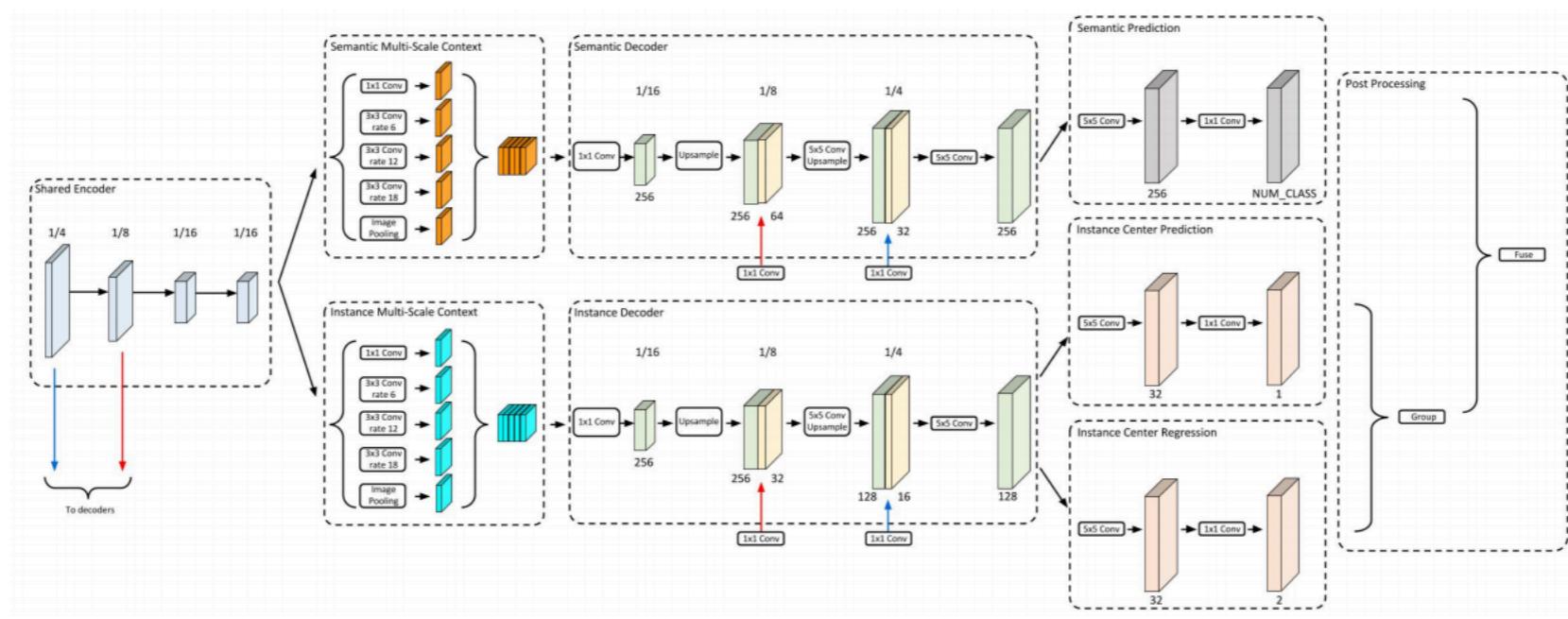




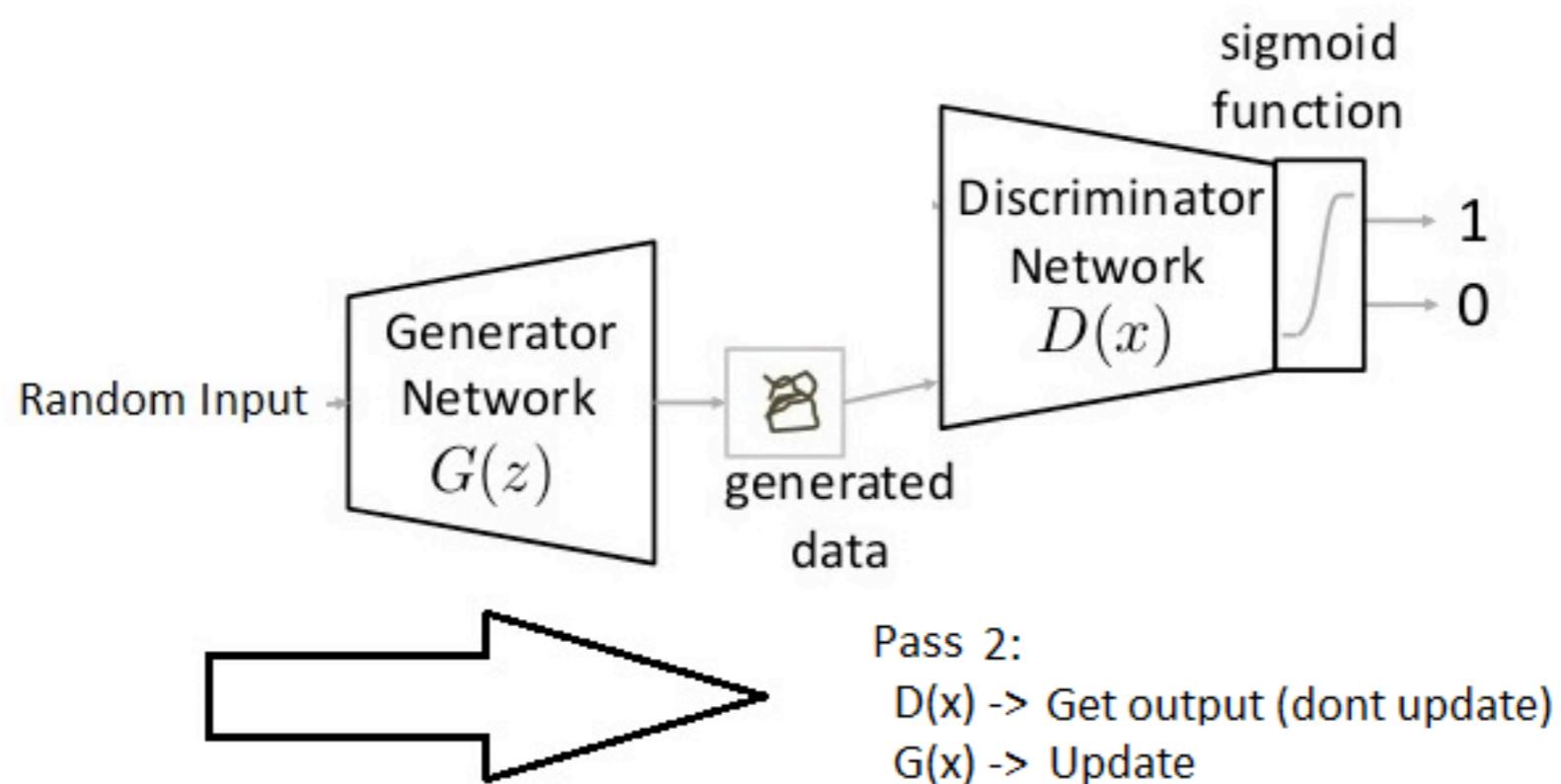
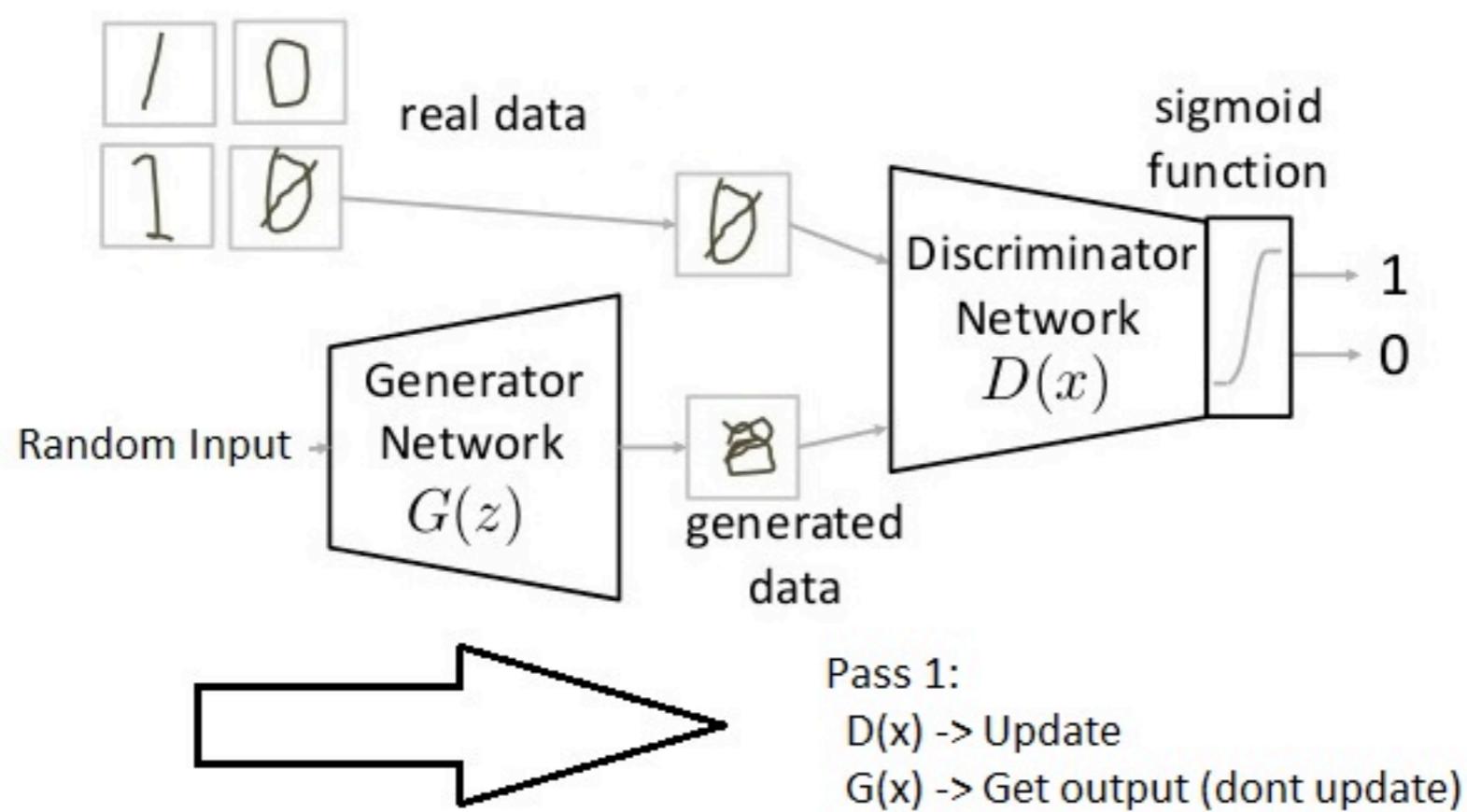
Panoptic-DeepLab

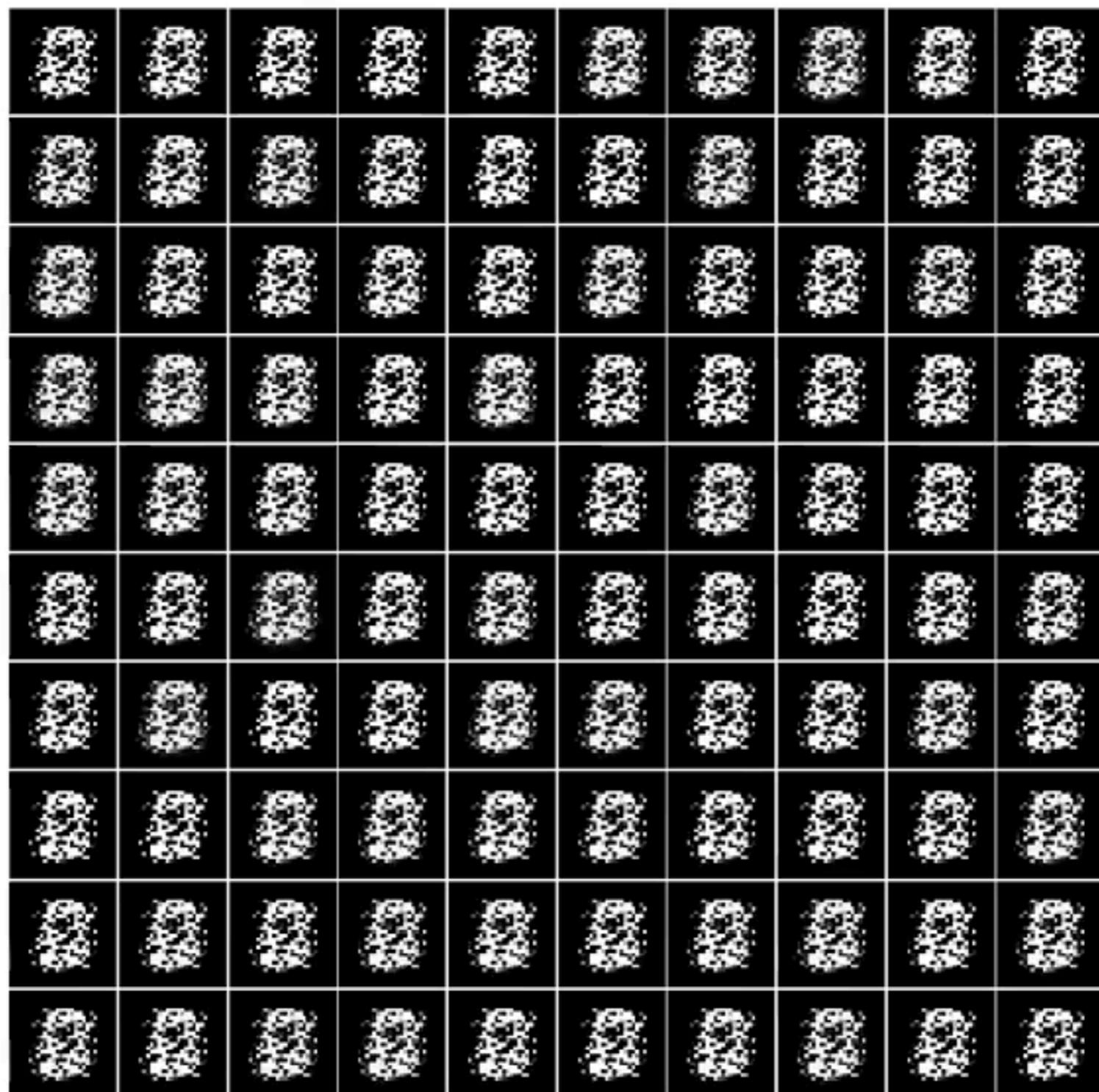
Bowen Cheng^{1,2}, Maxwell D. Collins², Yukun Zhu², Ting Liu²,
Thomas S. Huang¹, Hartwig Adam², Liang-Chieh Chen²

¹UIUC ²Google Inc.

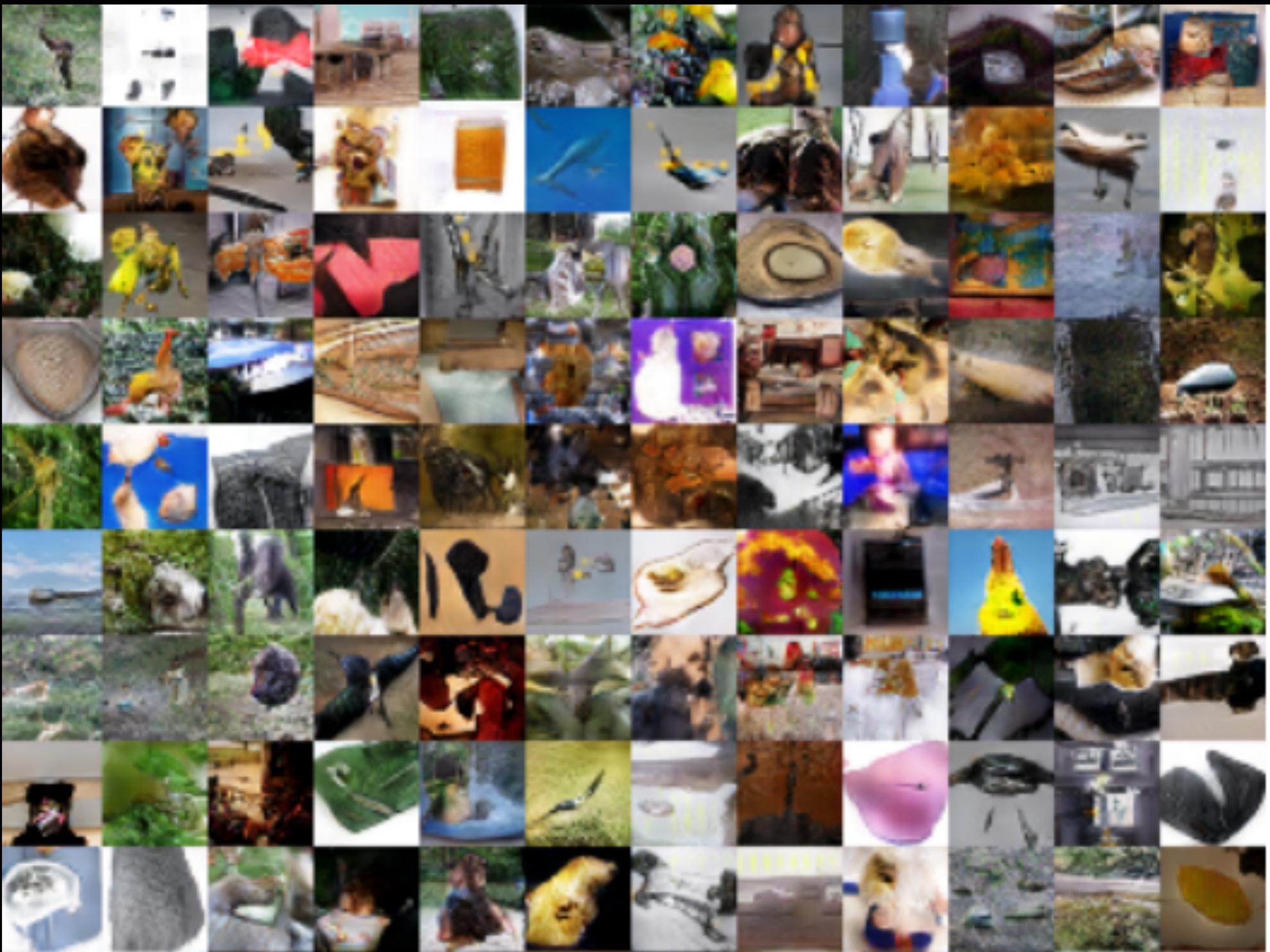


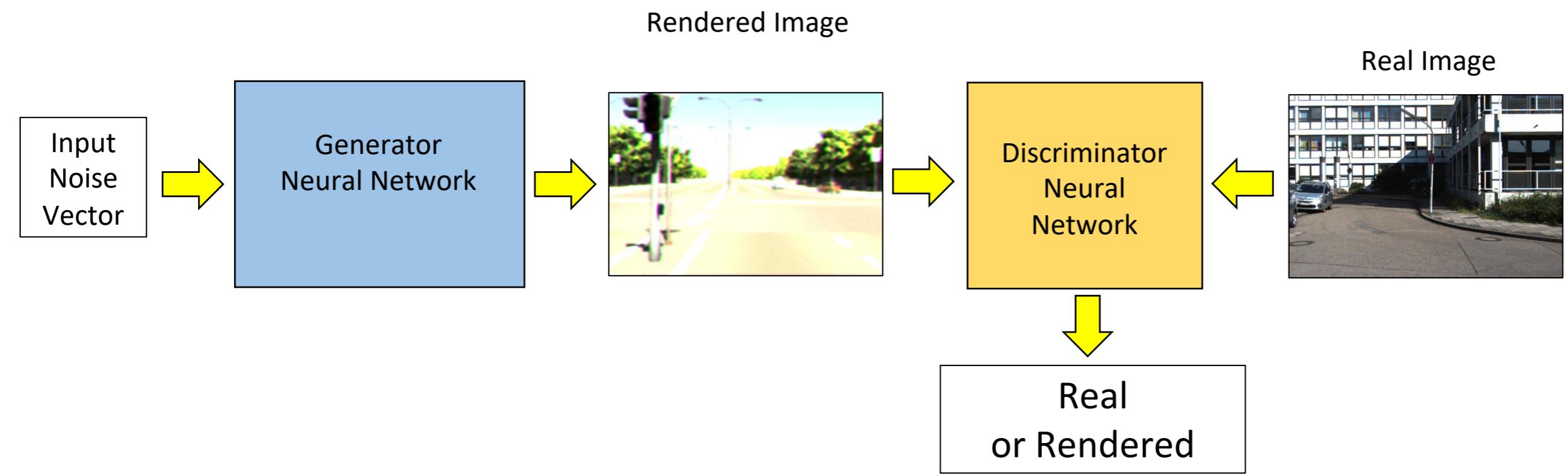
Generative Adversarial Networks











PROGRESSIVE GROWING OF GANs FOR IMPROVED QUALITY, STABILITY, AND VARIATION

Submitted to ICLR 2018

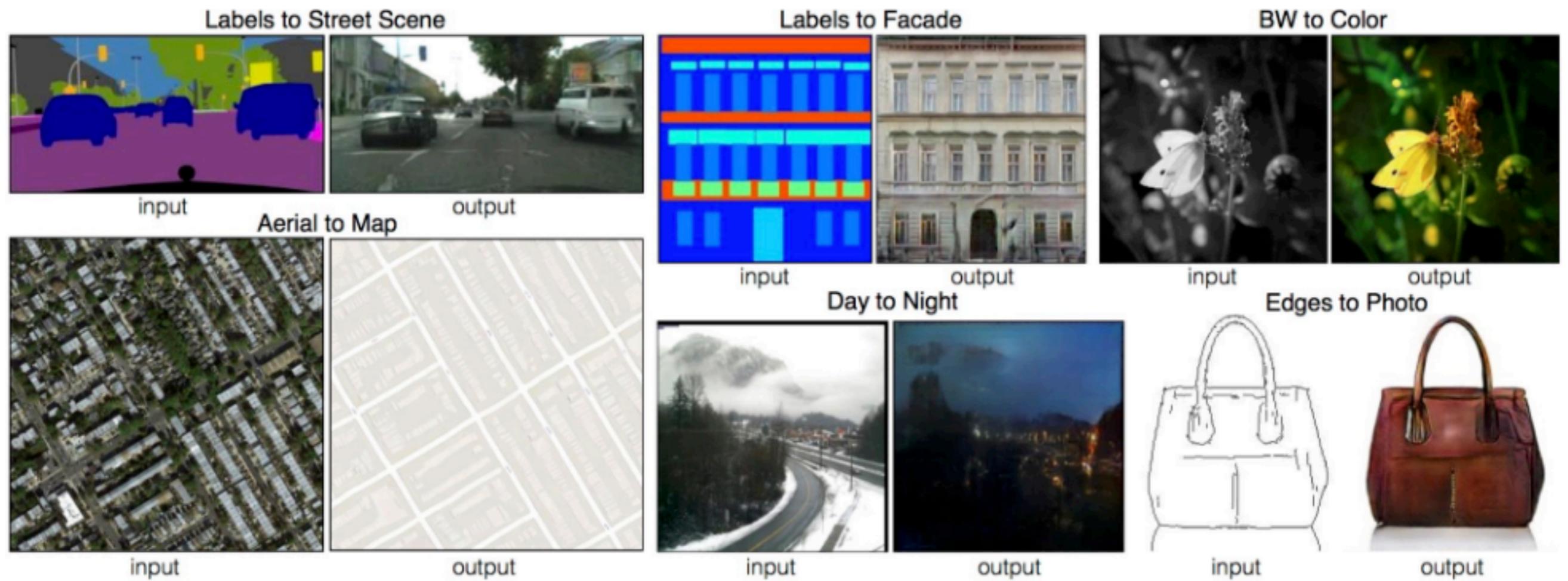


Image-to-Image Translation with Conditional Adversarial Nets

Phillip Isola
Jun-Yan Zhu
Tinghui Zhou
Alexei A. Efros

Monet Photos



Monet → photo

Zebras Horses



zebra → horse

Summer Winter



summer → winter



photo → Monet



horse → zebra



winter → summer



Photograph



Monet



Van Gogh



Cezanne



Ukiyo-e

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

[Jun-Yan Zhu*](#) [Taesung Park*](#) [Phillip Isola](#) [Alexei A. Efros](#)

[UC Berkeley](#)

Domain Transfer



Night



Winter



Summer



Image Gen

Synthesized image



High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs

Ting-Chun Wang¹ Ming-Yu Liu¹ Jun-Yan Zhu² Andrew Tao¹ Jan Kautz¹ Bryan Catanzaro¹
¹NVIDIA Corporation ²UC Berkeley



High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs

Ting-Chun Wang¹, Ming-Yu Liu¹, Jun-Yan Zhu², Andrew Tao¹,
Jan Kautz¹, Bryan Catanzaro¹

¹NVIDIA Corporation ²University of California, Berkeley

Other Robots

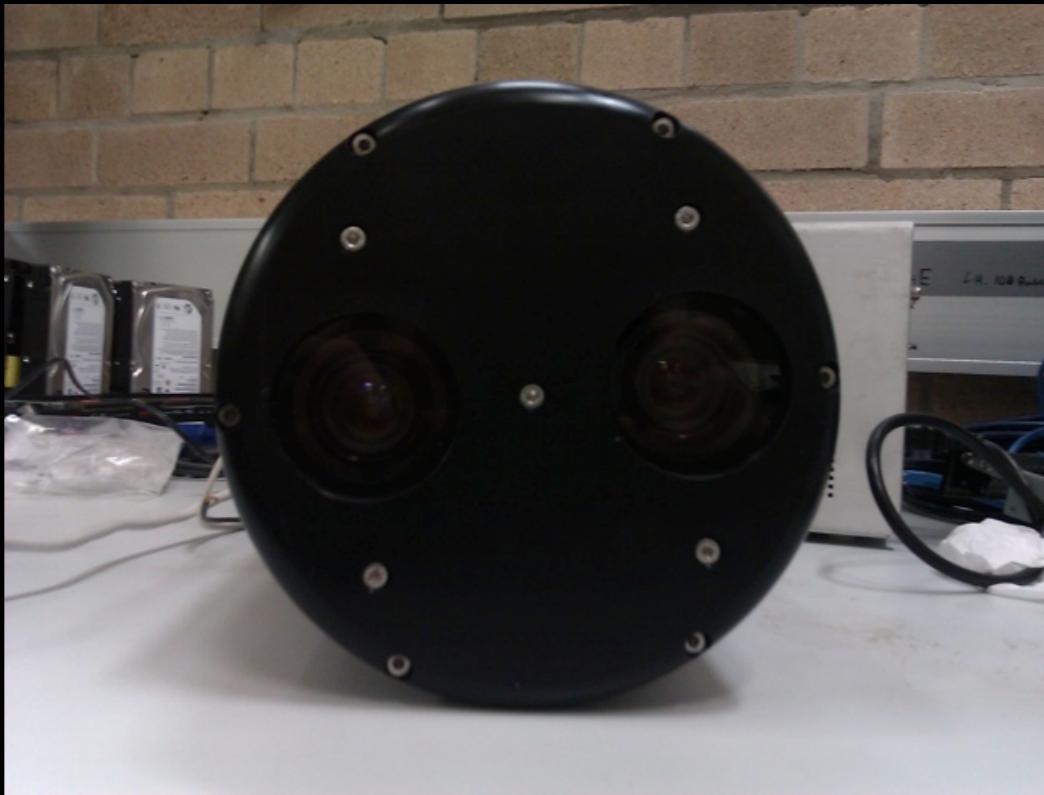


SENTRY



Woods Hole Oceanographic Institution





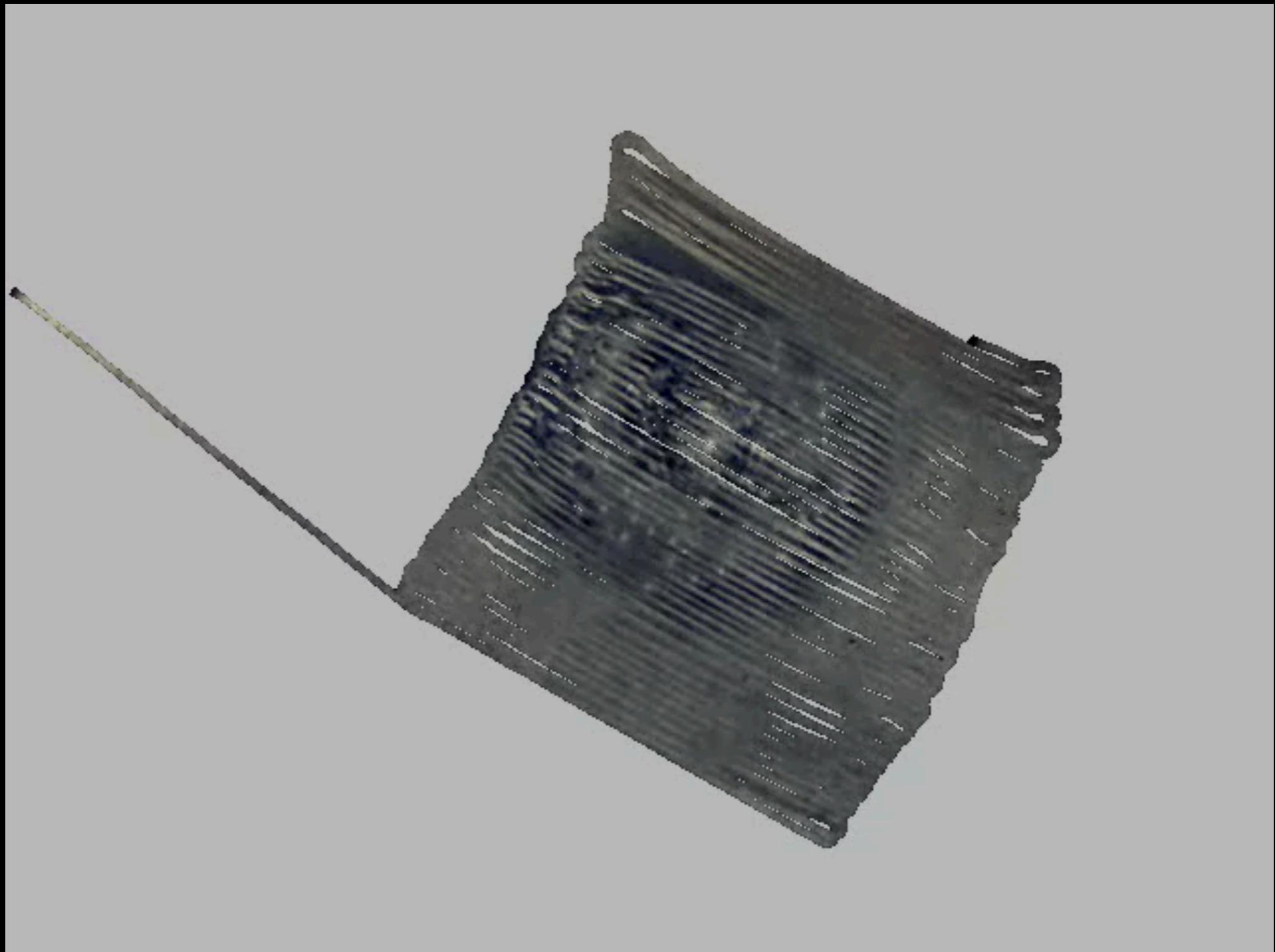
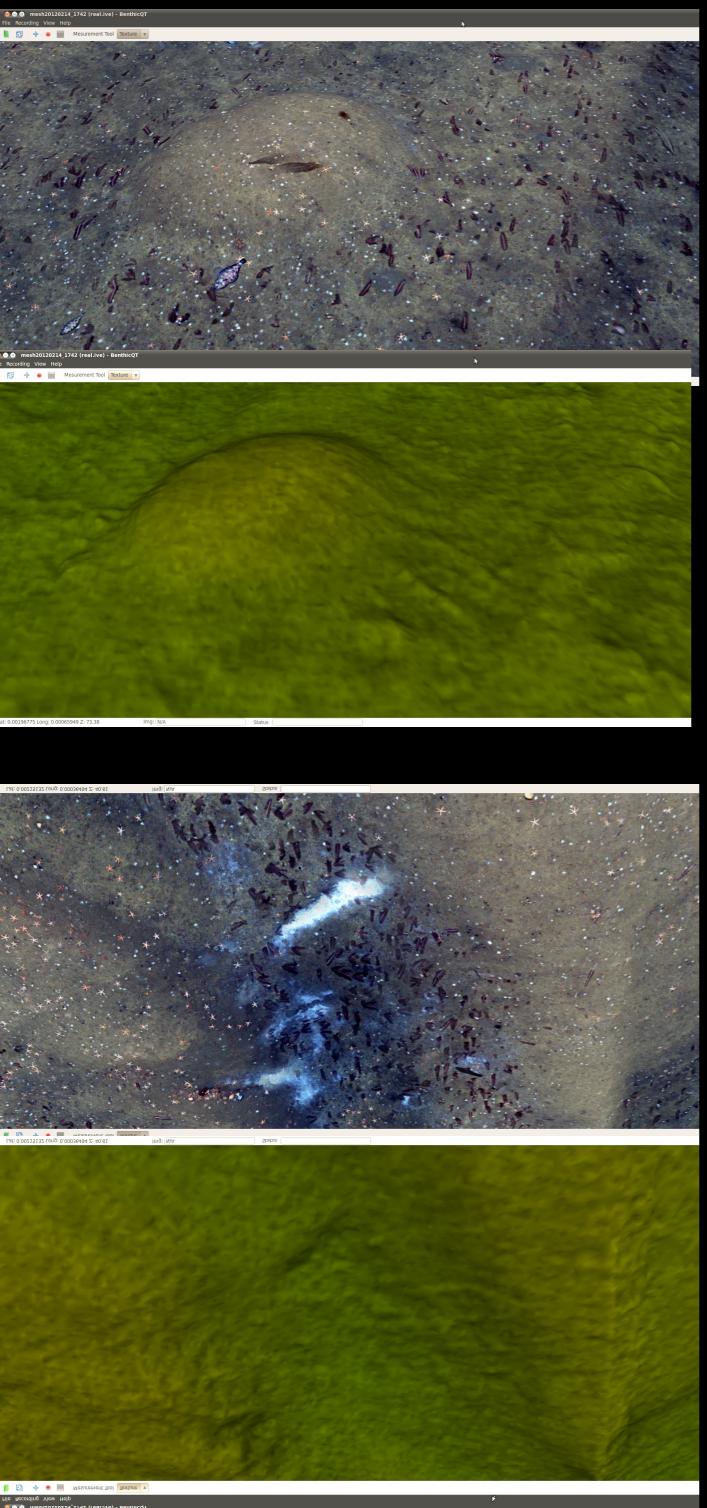




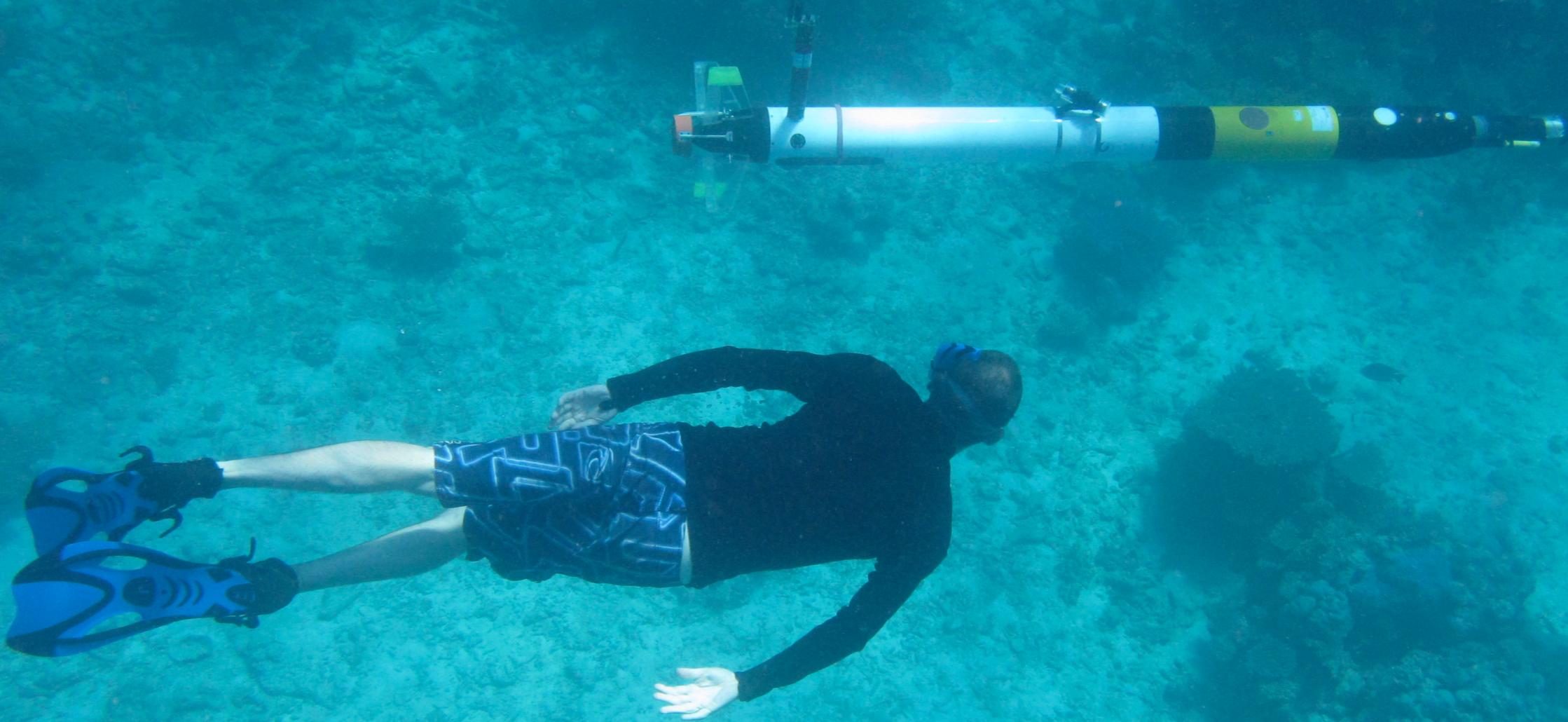


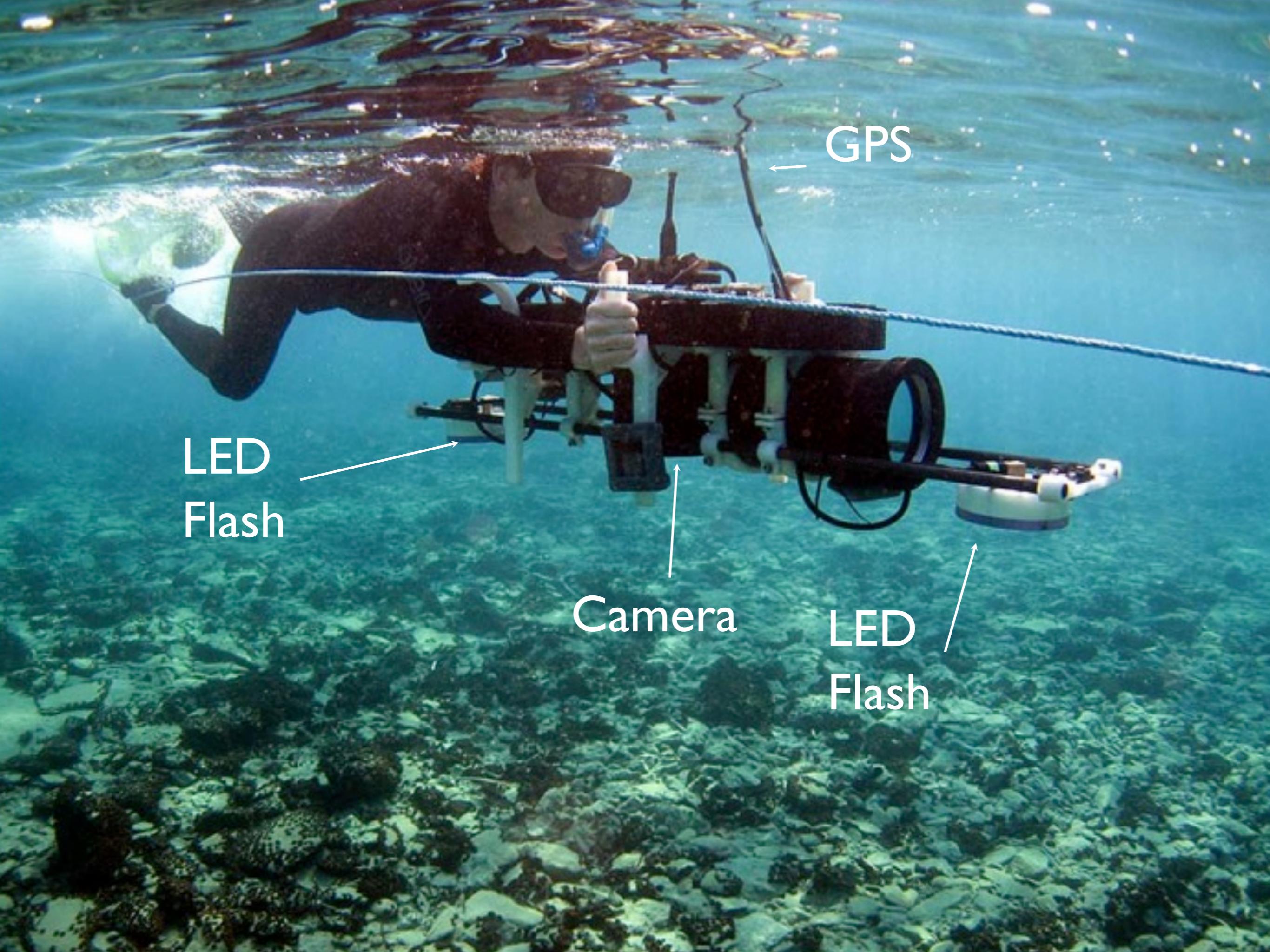




Image © 2005 EarthSat
Image © 2005 DigitalGlobe





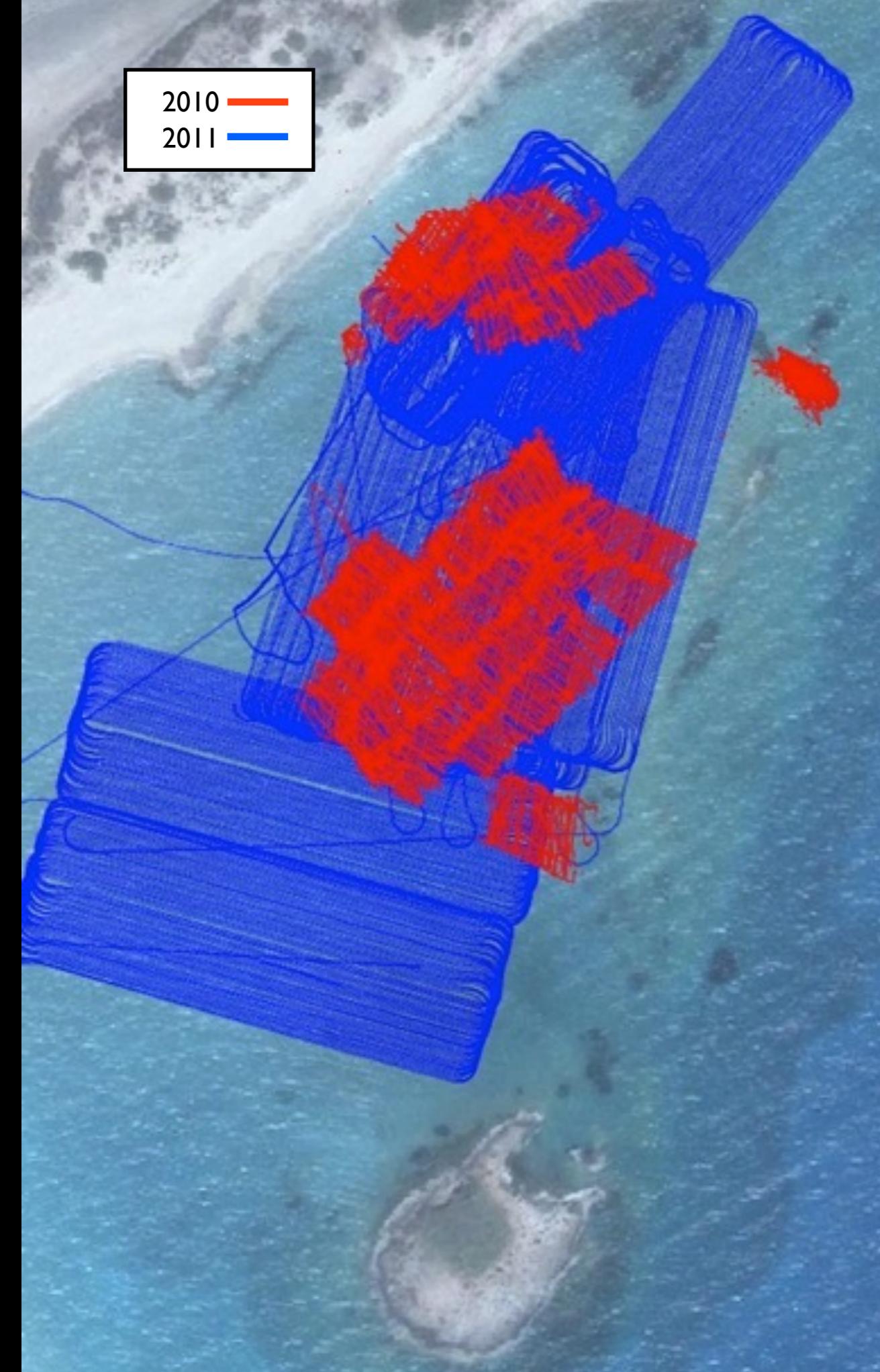


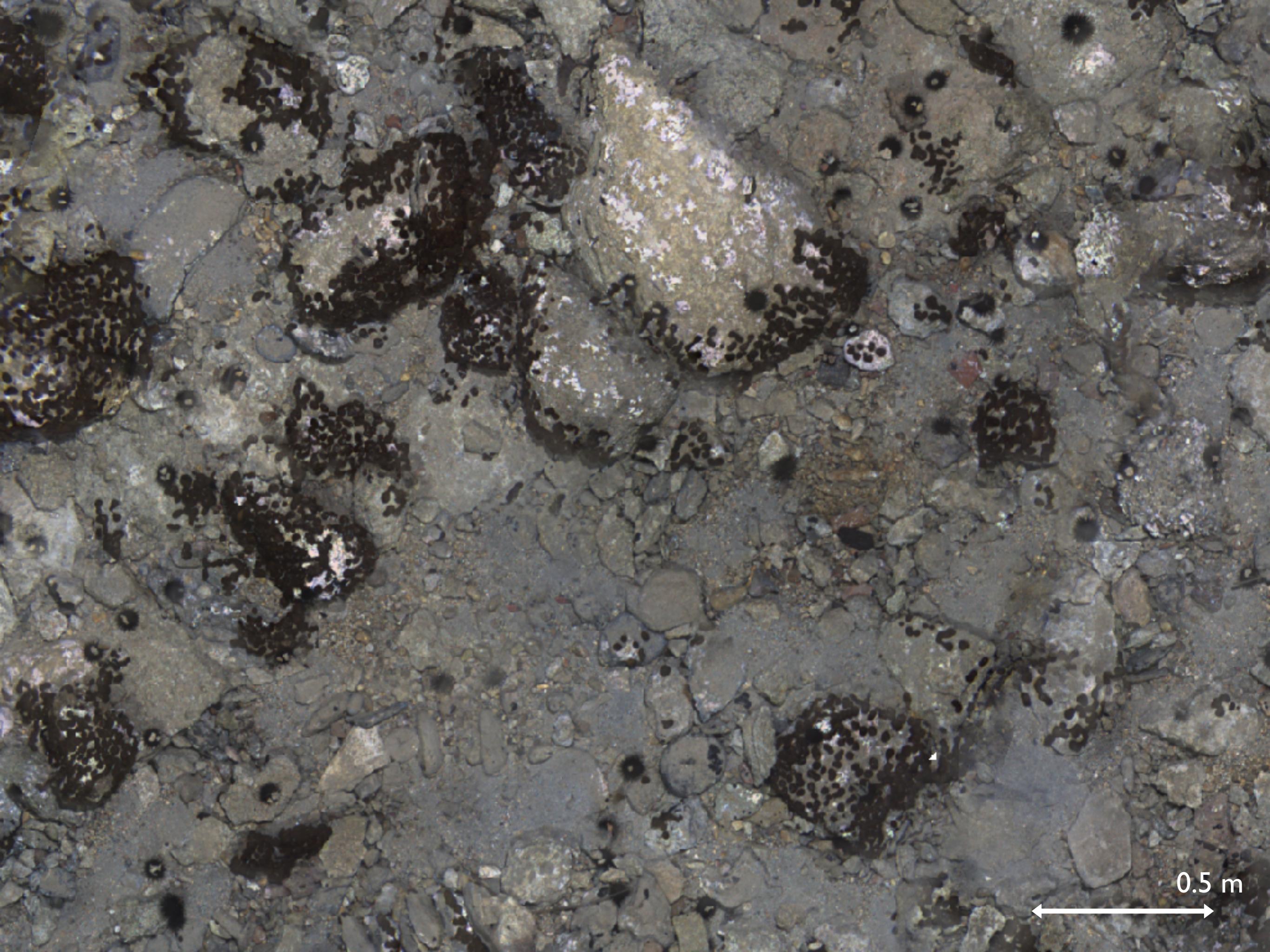
GPS

LED
Flash

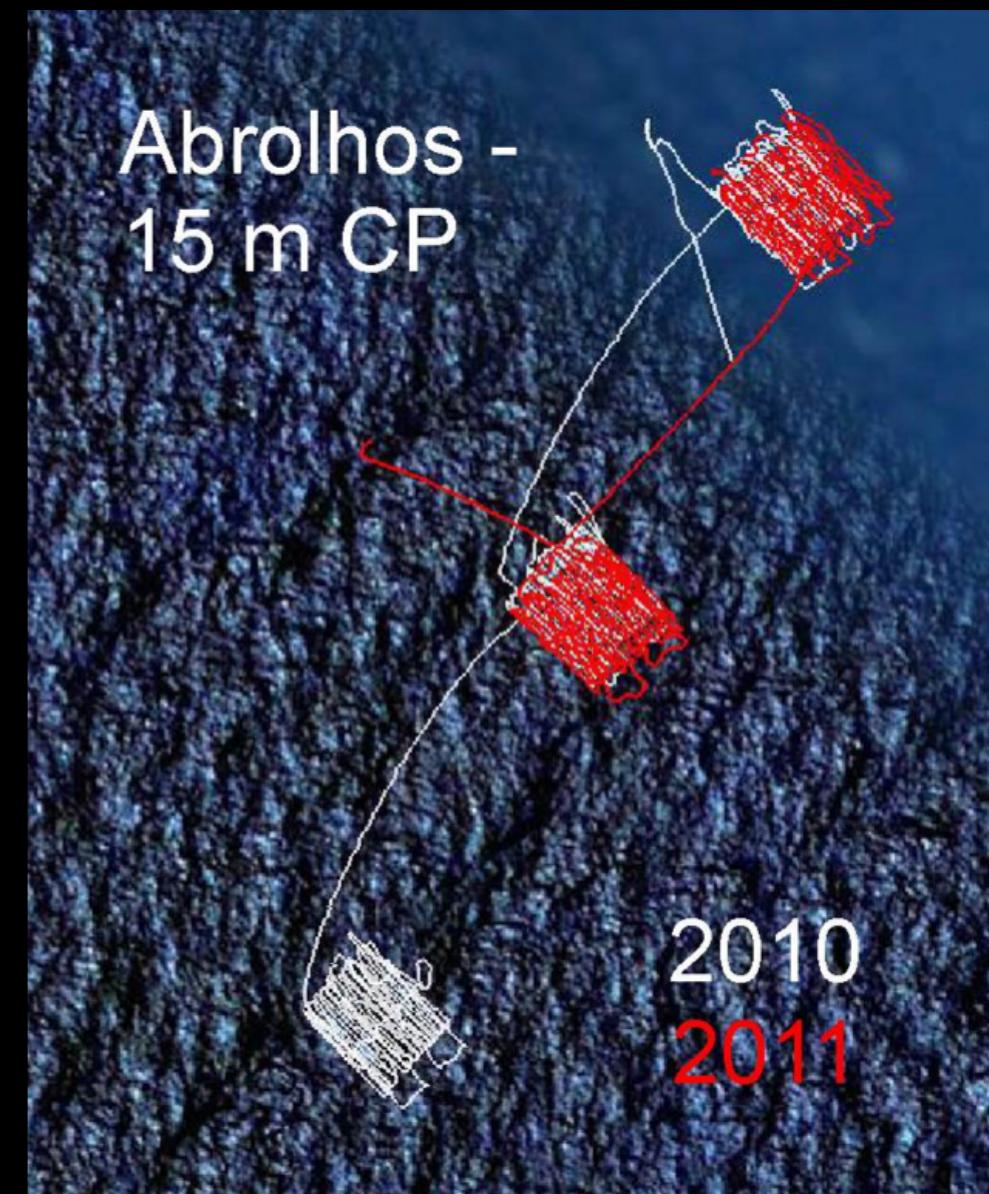
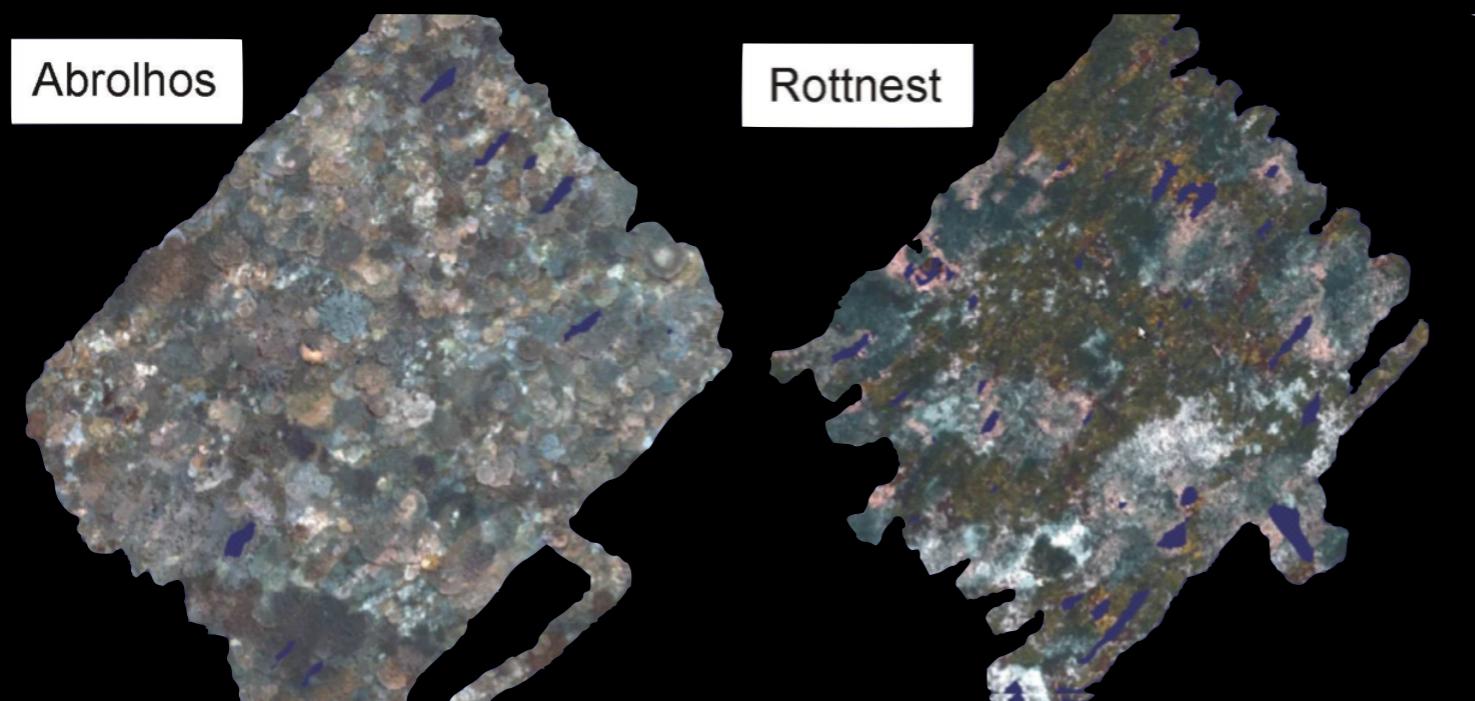
Camera

LED
Flash

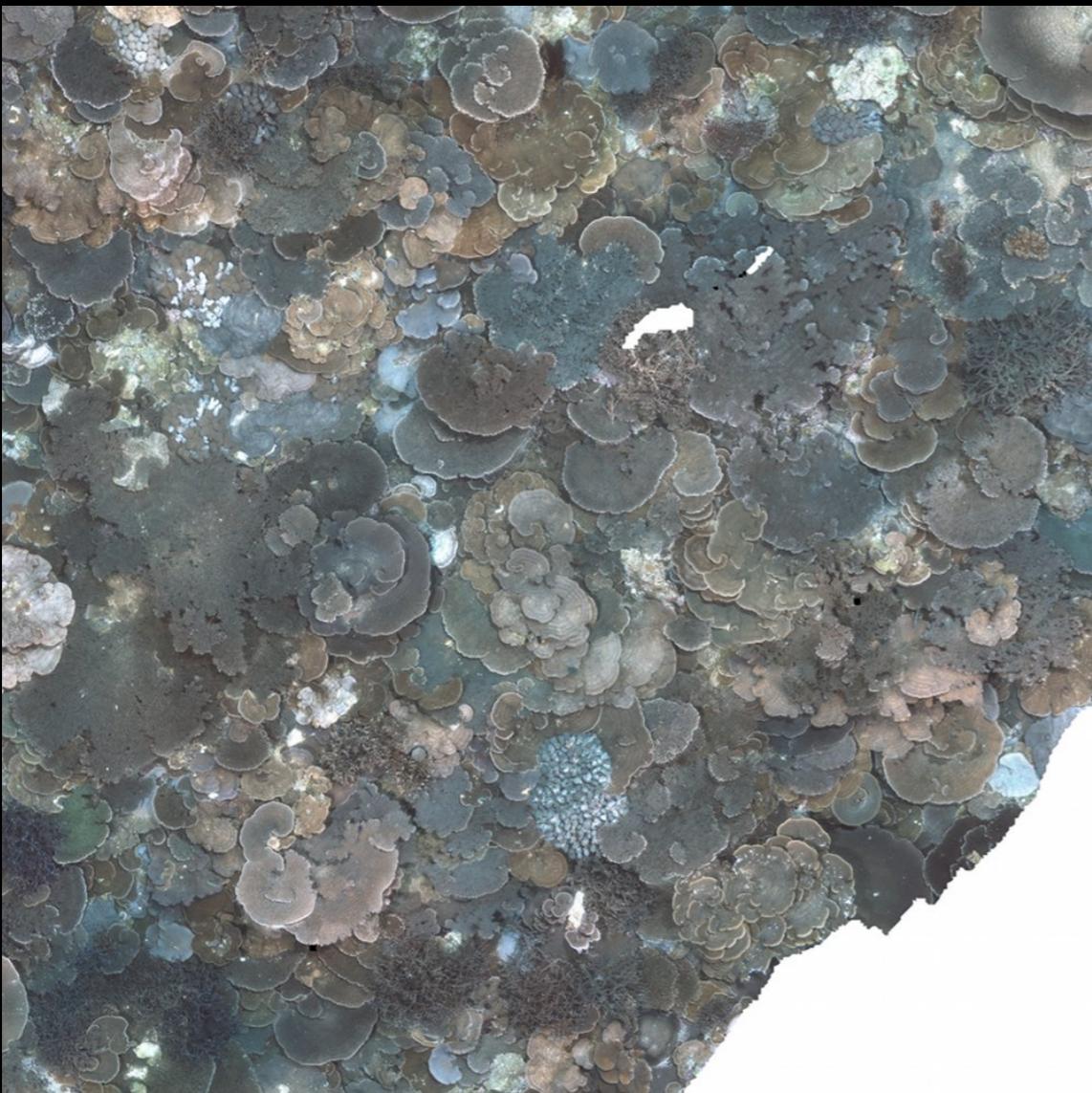




0.5 m

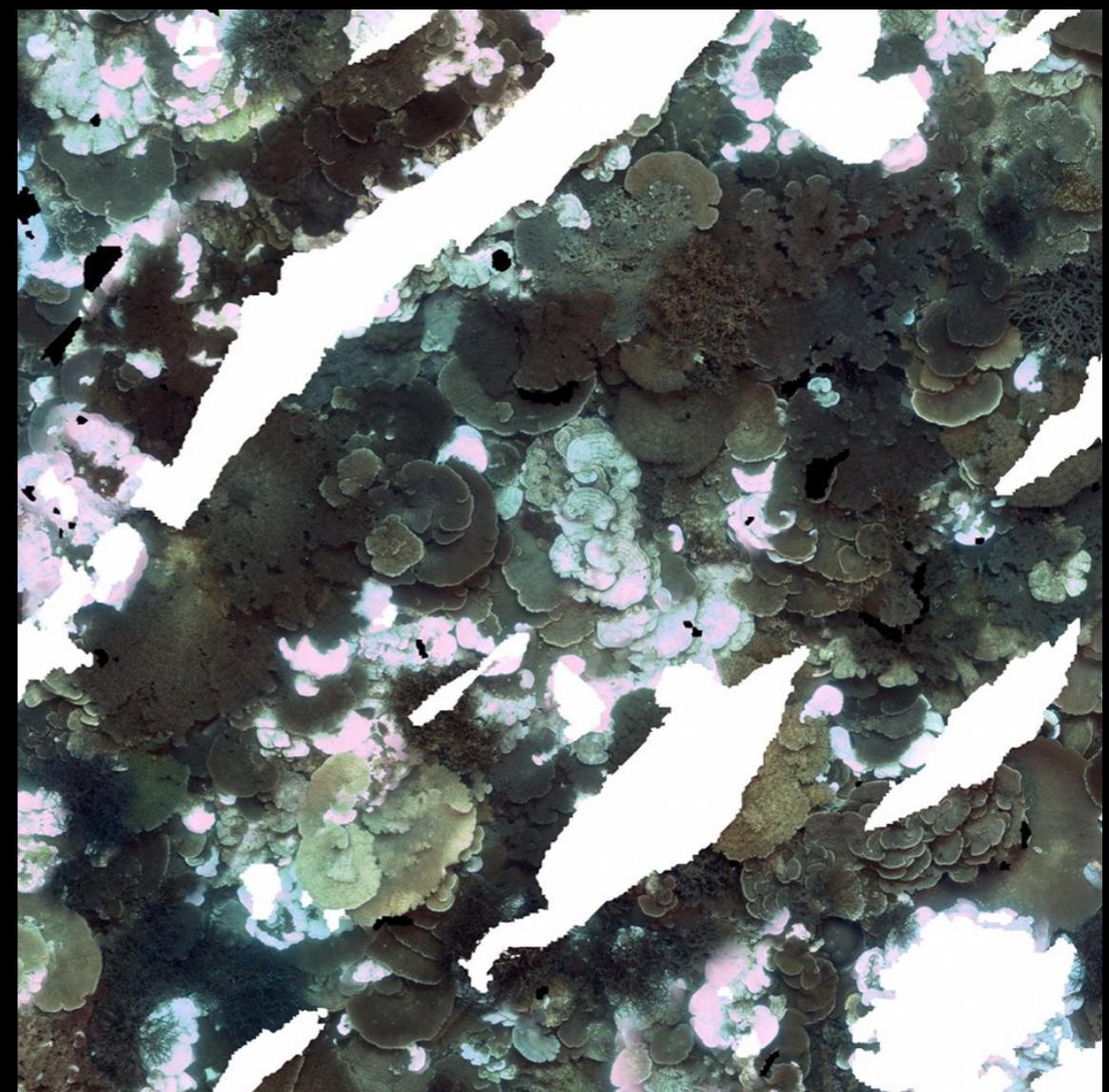


April
2010

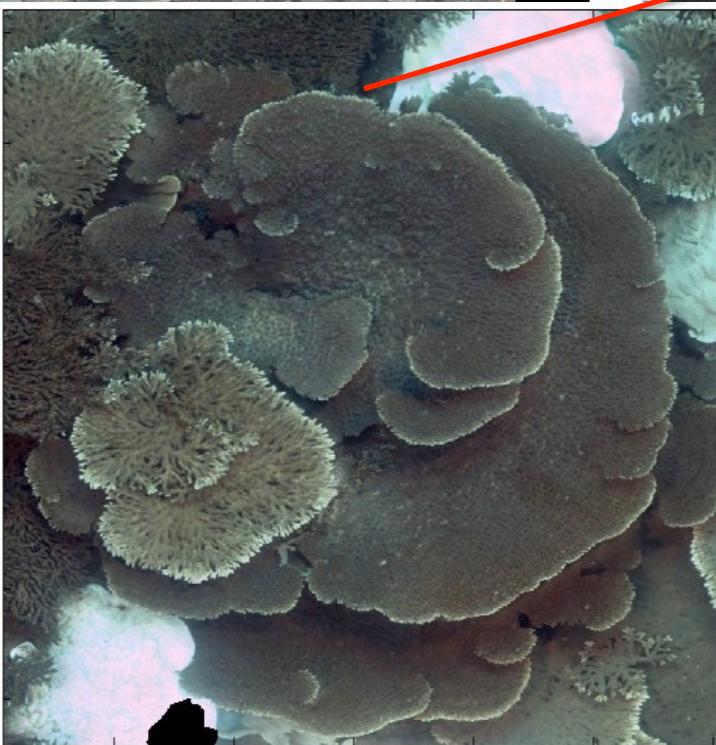


2m

April
2011

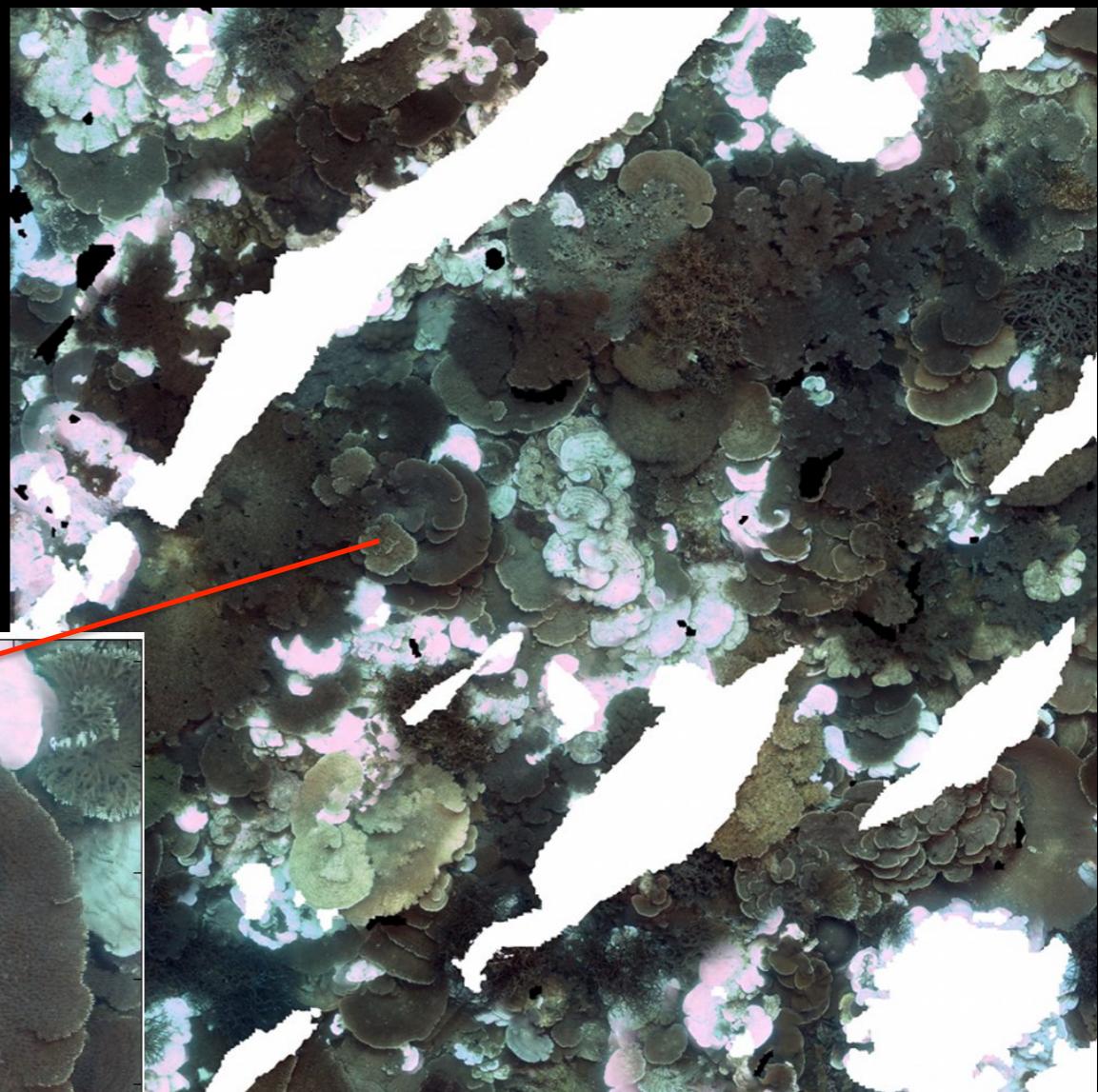


April
2010

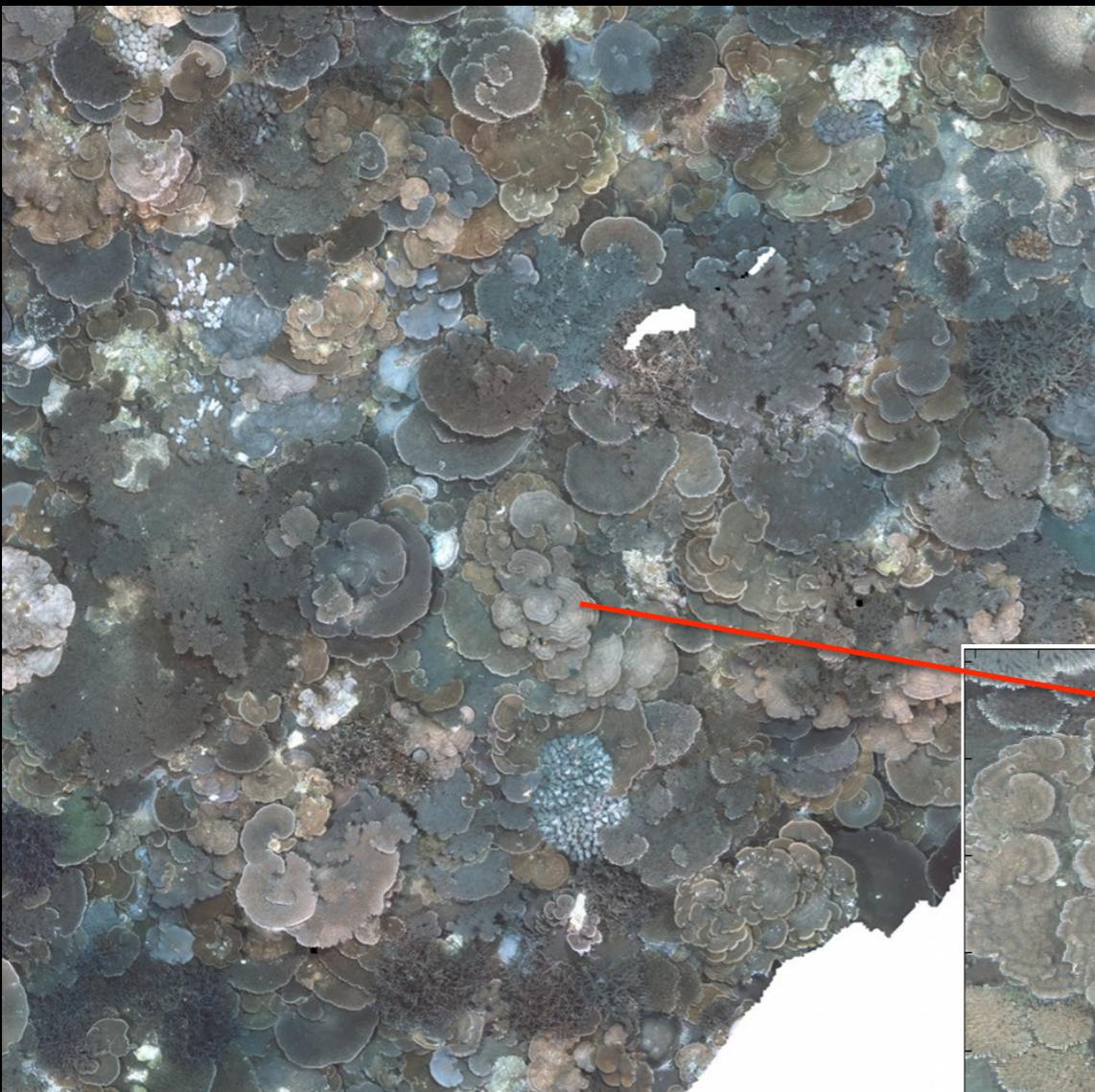


2m

April
2011

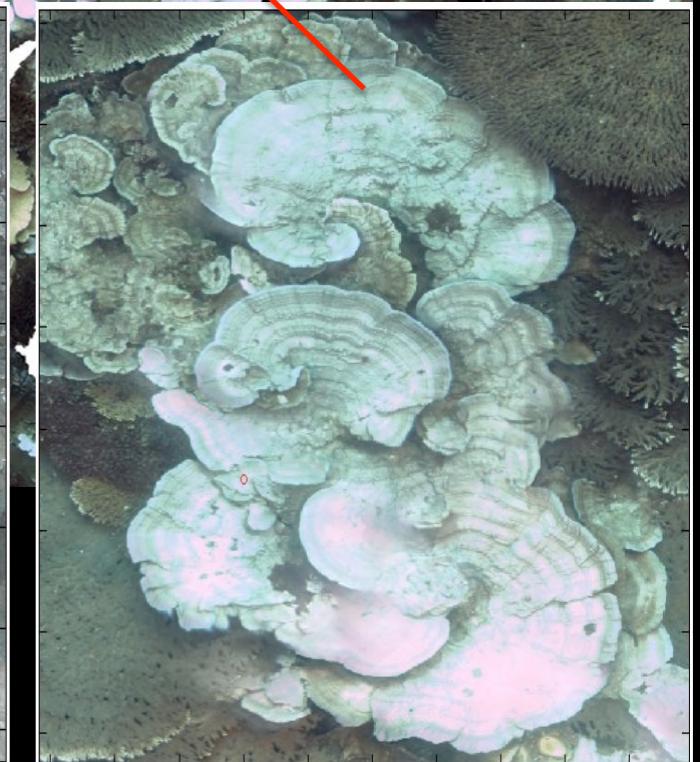
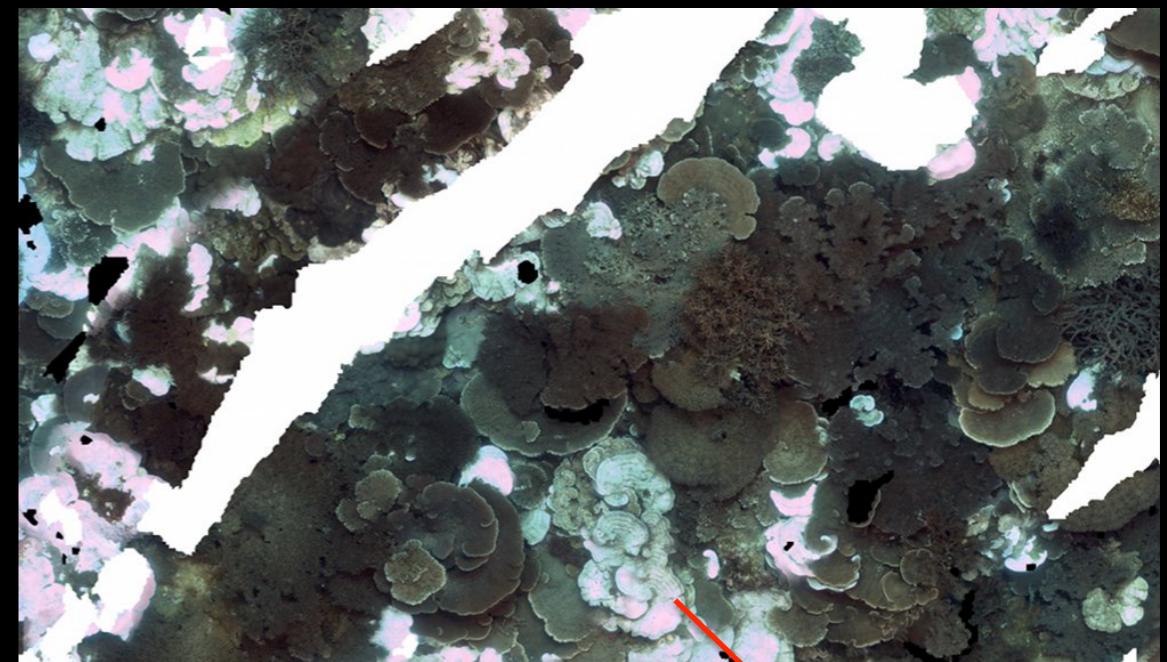


April
2010



2m

April
2011



Participation
Fill out teaching evaluation

Thanks