

PhD yearly assessment 2018/2019

Filippo Aleotti

Structural and Environmental Health Monitoring and
Management SEHM2

About me

Filippo Aleotti, 26 years old

Computer Science Engineer

PhD Student of the 34th-cycle

Research project Title: **Depth Reconstruction from monocular cameras**

Supervisor: Stefano Mattoccia

Contacts: filippo.aleotti2@unibo.it



PyDNet

Towards real-time unsupervised monocular depth estimation on CPU

Self-supervised monocular network. Thanks to its design, it was the first deep-learning based solution in literature able to run even on low-power device, (e.g., Raspberry Pi 3)

Published at IROS 2018

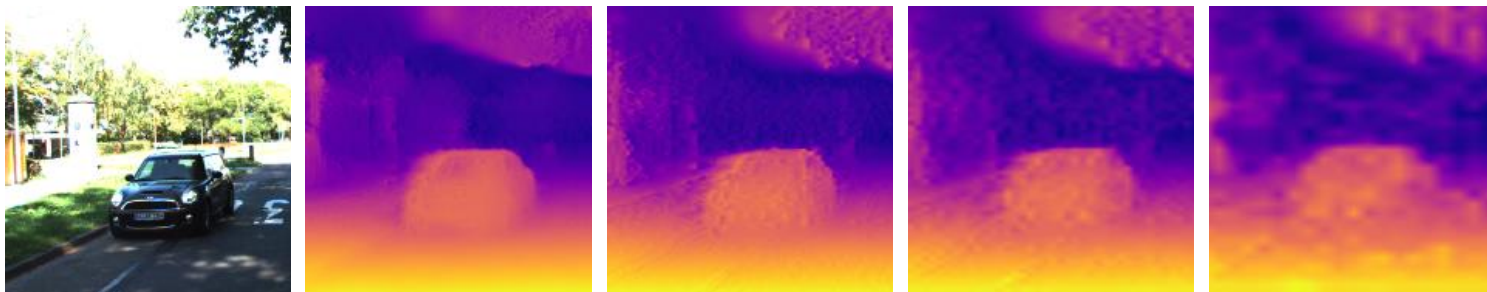
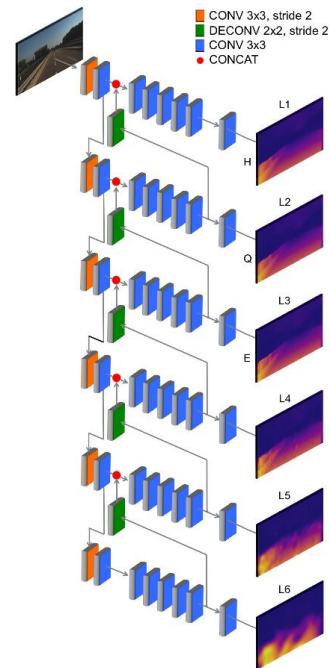


Fig: depth predictions at different resolutions. From left to right Full, Half, Quarter and Height



MonoResMatch

Learning monocular depth estimation infusing traditional stereo knowledge

We exploited traditional stereo algorithms to improve the predictions of a monocular depth network

Published at CVPR 2019

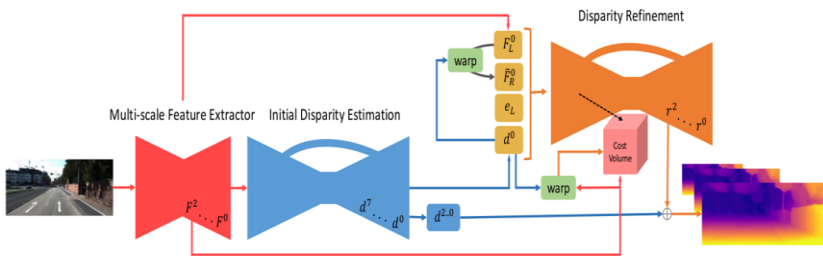


Fig: on top, the network architecture, while the figure on the right shows the predicted depth for the scene

PyDNet2

Real-time monocular depth estimation without GPU

Extension of **Towards real-time unsupervised monocular depth estimation on CPU**, exploiting the strategies used in **Learning monocular depth estimation infusing traditional stereo knowledge**.

Results have been considerably improved, with a lower number of parameters!

Under review at IEEE Transactions On Cybernetics



Fig: qualitative comparison between PyDNet2, PyDNet and a state-of-the art network Monodepth

Dwarf

Learning end-to-end scene flow by distilling single tasks knowledge

Lightweight network for Scene Flow estimation, using stereo pairs consecutive in time

Accuracy similar to state-of-the-art (requiring 2X GPU) but suited even for embedded systems

Under review at AAAI 2020

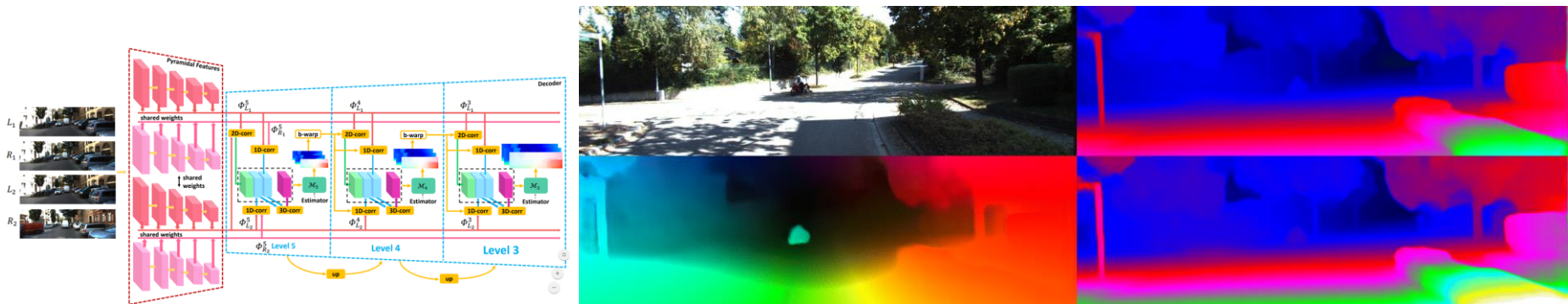


Fig: outputs of the network. On top disparity, while on bottom respectively the optical flow and the disparity change of the scene

μ PyDNet

Monocular Depth Perception on Microcontrollers (joint work with Politecnico di Torino)

We proposed an extremely compact network for monocular depth estimation. This network is able to run directly on ARM microcontrollers

This system can be used, for instance, for traffic monitoring



Image available [here](#)

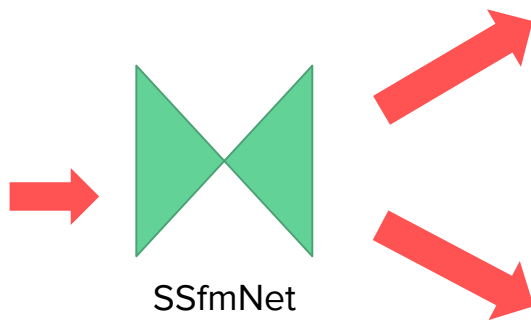


Fig: on the left, car detection in a traffic monitoring scenario using depth, on the right an ARM microcontroller

SSfmNet

Semantic self-supervised monocular depth estimation

I am currently working on a single shallow network able to infer both depth and semantic labels for each pixel in the scene. This network is trained without depth labels, exploiting directly videos taken by a monocular camera.



SSfmNet

depth



semantic

Fig: given an image, our network is able to predict both the depth and the semantic of the scene

Additional projects

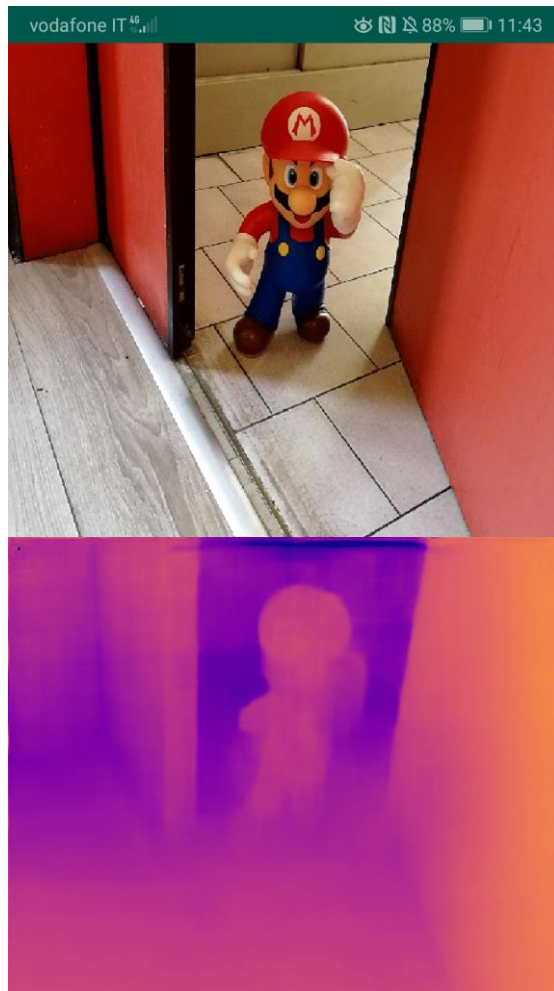
Depth estimation on iOS and Android devices

Porting of **Real-time monocular depth estimation without GPU network** to iOS and Android devices

Thanks to this application, we are able to infer depth maps on mobile phones, even on devices equipped with a single camera!

Several applications (eg, augmented reality, autonomous vehicles)

The code is [open-source](#)



Publications and Demos

- Learning monocular depth estimation infusing traditional stereo knowledge, IEEE CVPR 2019
- Real-time monocular depth estimation without GPU, under review at IEEE Transactions On Circuits and Systems for Video Technology
- Learning end-to-end scene flow by distilling single tasks knowledge, under review AAAI 2020
- Real-time monocular depth estimation without GPU, under review at IEEE Transactions On Cybernetics
- Monocular Depth Perception on Microcontrollers, under review at IEEE Transactions On Circuits and Systems for Video Technology
- Real-Time Monocular Depth Estimation Without GPU, **Demo** at CVPR 2019
- Towards real-time unsupervised monocular depth estimation on CPU, IROS 2018