

How to satisfy the Thirst for Data?

Andreas Geiger

Autonomous Vision Group
University of Tübingen / MPI for Intelligent Systems

September 9, 2018



University of Tübingen
MPI for Intelligent Systems

Autonomous Vision Group



The Deep Learning Revolution

What is the problem?

Data

Data Data Data Data

Data Data Data Data Data Data Data Data Data Data
Data Data Data Data Data Data Data Data Data Data
Data Data Data Data Data Data Data Data Data Data
Data Data Data Data Data Data Data Data Data Data
Data Data Data Data Data Data Data Data Data Data
Data Data Data Data Data Data Data Data Data Data
Data Data Data Data Data Data Data Data Data Data
Data Data Data Data Data Data Data Data Data Data

The word "Data" repeated ten times per row across eight rows.

Labels

[illegible]

Money

[illegible]

Image Classification



[Deng et al., CVPR 2009]

Semantic Segmentation



[Cordts et al., CVPR 2016]

Amazon Mechanical Turk

[Your Account](#)[HITS](#)[Qualifications](#)

Already have an account?
Sign in as a [Worker](#) | [Requester](#)

[Introduction](#) | [Dashboard](#) | [Status](#) | [Account Settings](#)

Mechanical Turk is a marketplace for work.

We give businesses and developers access to an on-demand, scalable workforce.
Workers select from thousands of tasks and work whenever it's convenient.

192,676 HITS available. [View them now.](#)

Make Money by working on HITS

HITS - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITS now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work



or [learn more about being a Worker](#)

Get Results from Mechanical Turk Workers

Ask workers to complete HITS - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Get Started.](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITS completed in minutes
- Pay only when you're satisfied with the results



Data Annotation Industry



BOSCH

Technik fürs Leben

Baidu 百度



MOBILEYE®

 **Mapillary**

 **Mighty Ai**



DATATANG



PALLASLUDENS
VISION LAB

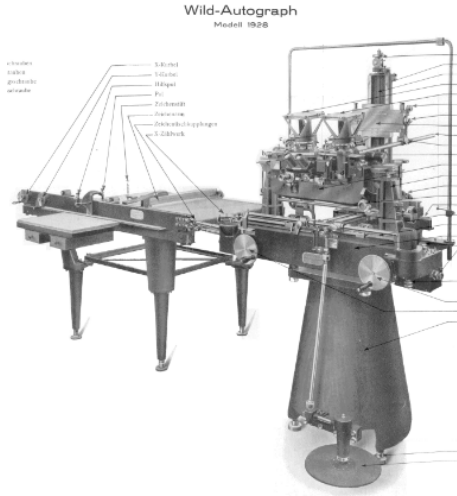


CrowdFlower

 **SURFING TECH**

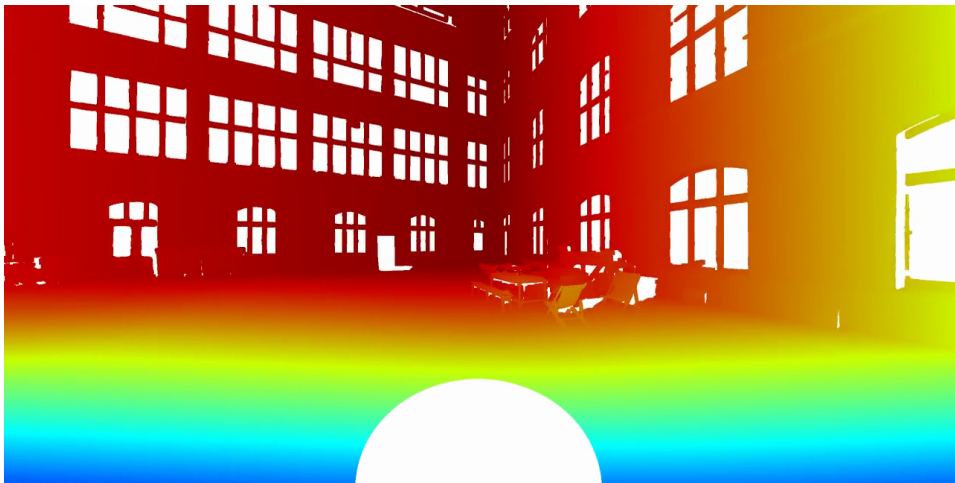
What about other tasks?

3D Reconstruction



www.wild-heerbrugg.com/photogrammetry1.htm

3D Reconstruction



www.eth3d.net

Optical Flow



J. Gibson, 1950



MPI Sintel, 2012

What can we do?

6 Strategies

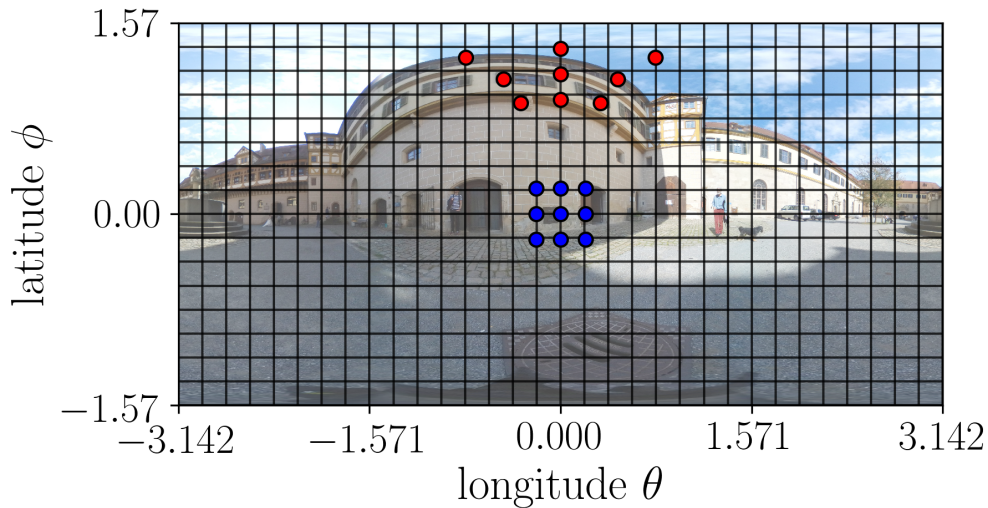
Strategy 1

Specialized Architectures

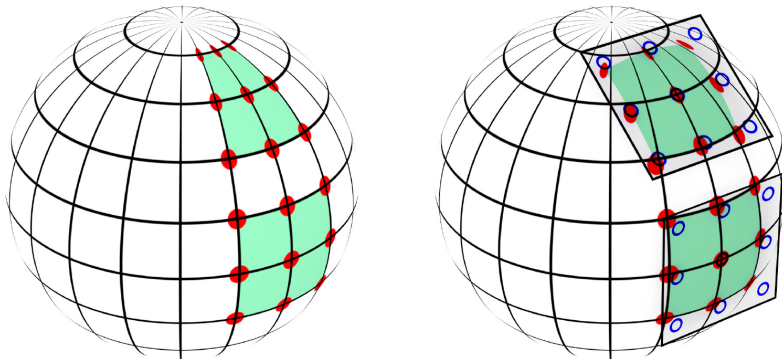
SphereNet: Learning Spherical Representations for Detection



SphereNet: Learning Spherical Representations for Detection

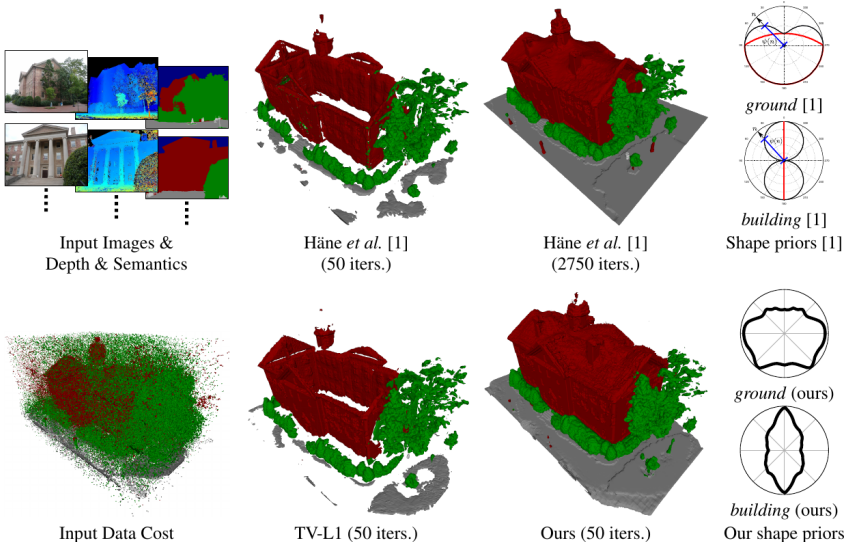


SphereNet: Learning Spherical Representations for Detection

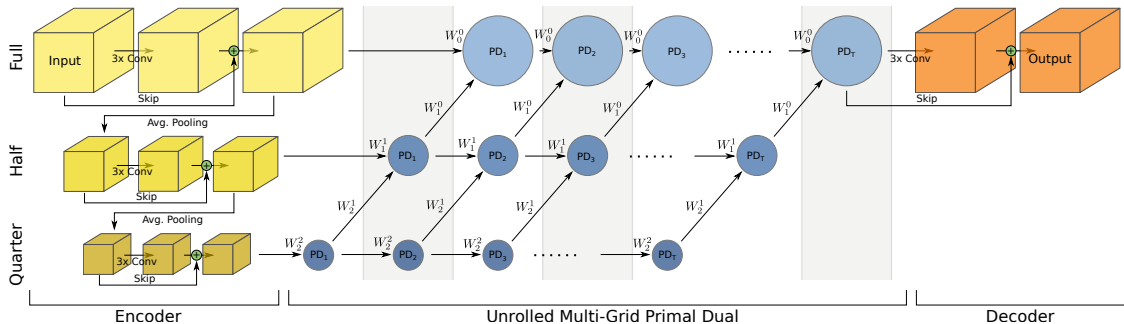


- Allows for easy transfer learning between camera models

Learning Priors for Semantic 3D Reconstruction

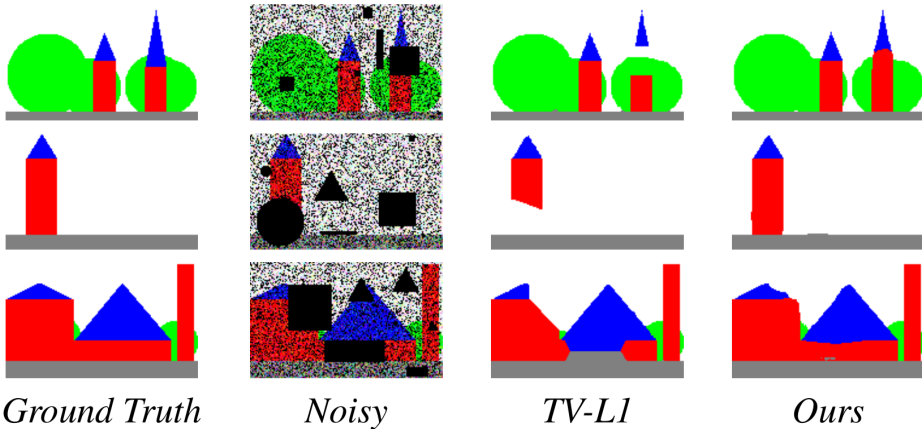


Learning Priors for Semantic 3D Reconstruction



$$\underset{u}{\text{minimize}} \quad \int_{\Omega} (\|Wu\|_2 + fu) \, dx \quad \text{subject to} \quad \forall \mathbf{x} \in \Omega : \sum_{\ell} u_{\ell}(\mathbf{x}) = 1$$

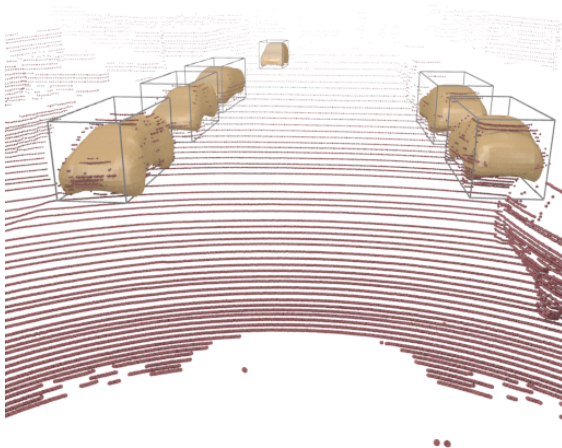
Learning Priors for Semantic 3D Reconstruction



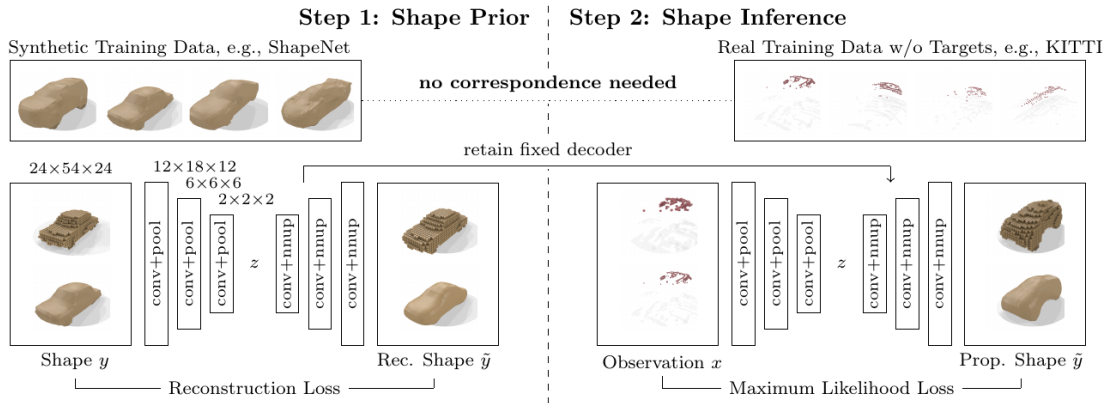
Strategy 2

Strong Priors

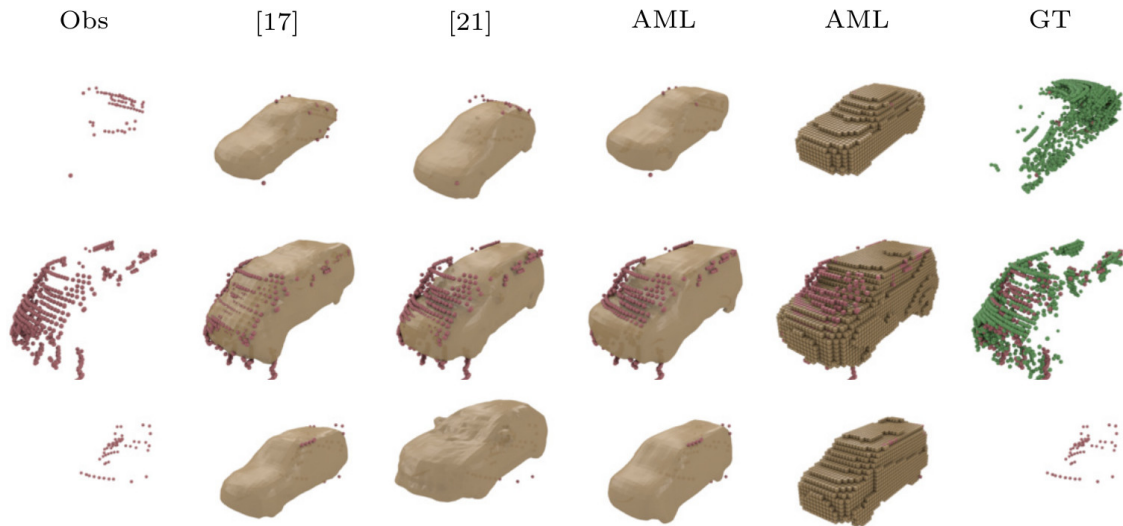
Learning 3D Shape Completion with Weak Supervision



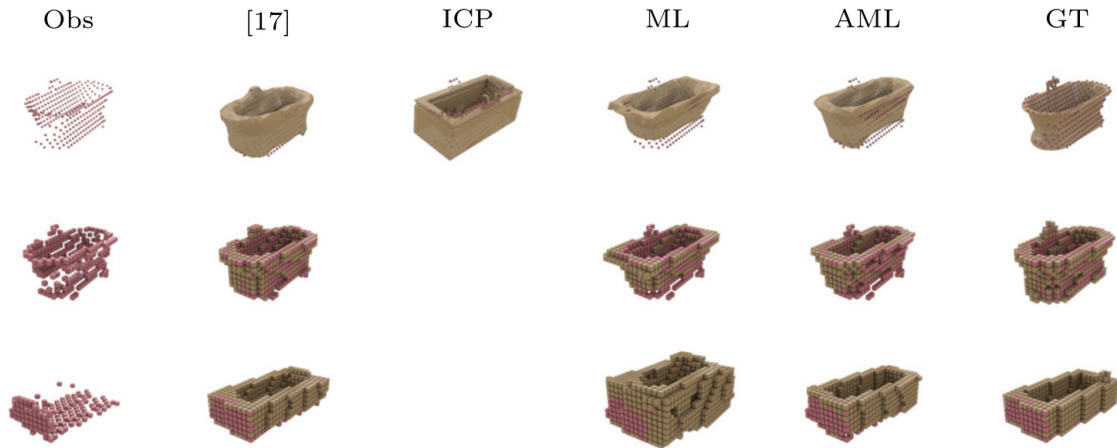
Learning 3D Shape Completion with Weak Supervision



Learning 3D Shape Completion with Weak Supervision



Learning 3D Shape Completion with Weak Supervision

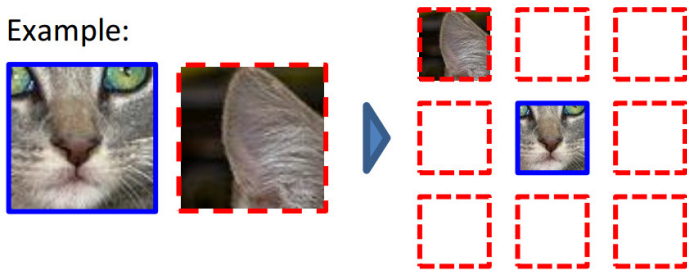


Strategy 3

Self Supervision

Unsupervised Visual Representation Learning by Context Prediction

Example:



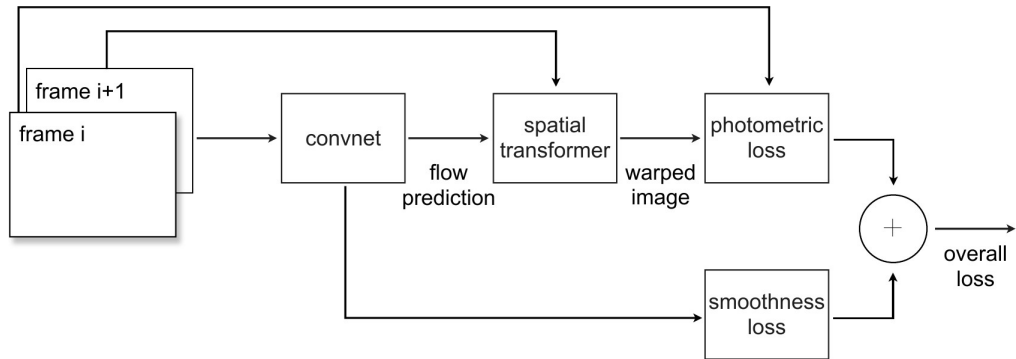
Question 1:



Question 2:

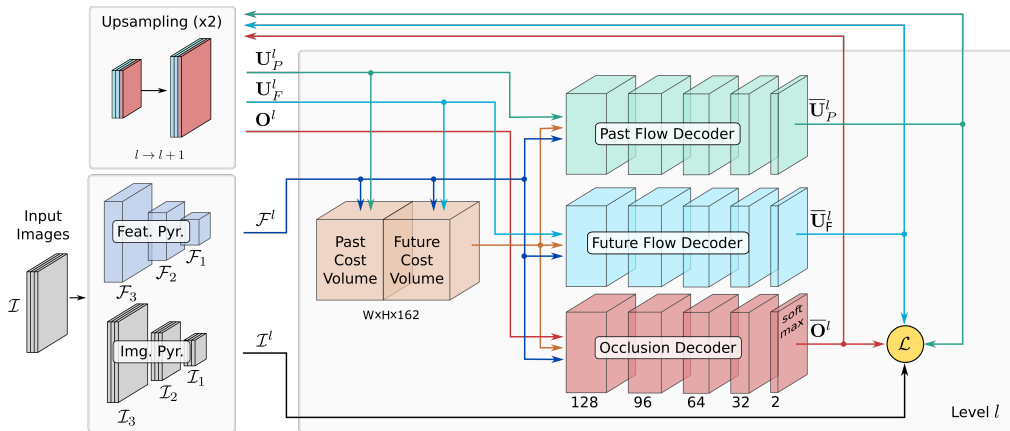


Unsupervised Learning of Multi-Frame Optical Flow with Occlusions



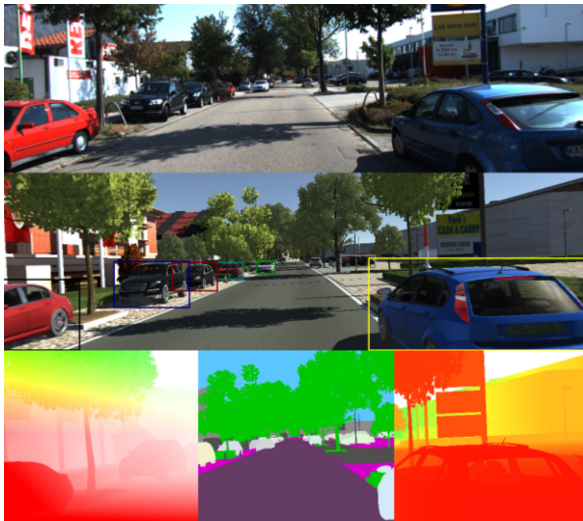
[Courtesy: Jason J. Yu]

Unsupervised Learning of Multi-Frame Optical Flow with Occlusions



Strategy 4 Simulation

Virtual Worlds as Proxy for Multi-Object Tracking Analysis



Augmented Reality Meets Deep Learning

Real scene (KITTI)

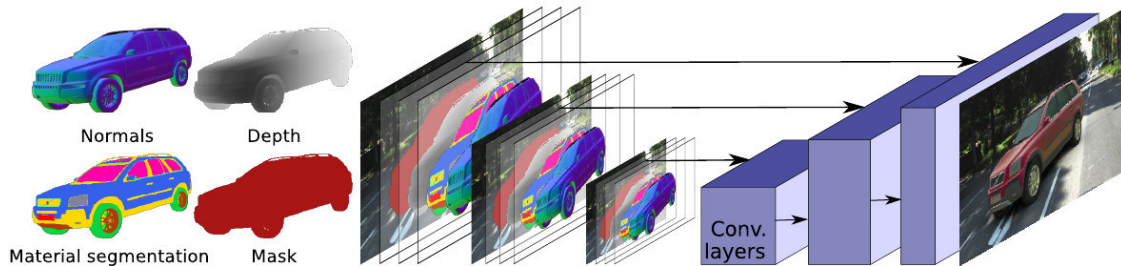


Synthetic scene (Virtual KITTI)



Real scene augmented with synthetic cars (Ours)

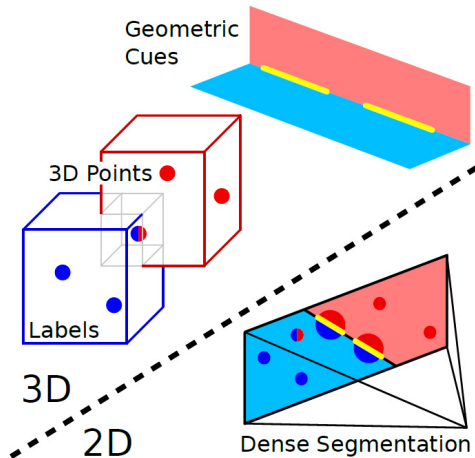
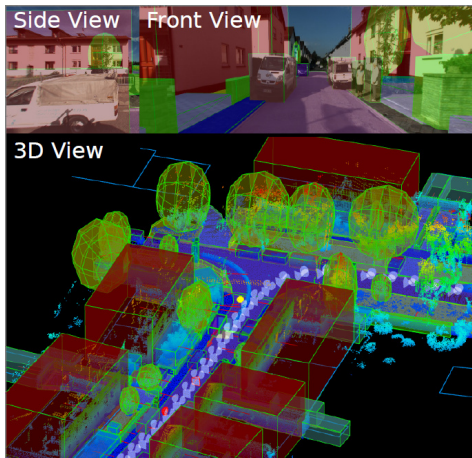
Geometric Image Synthesis



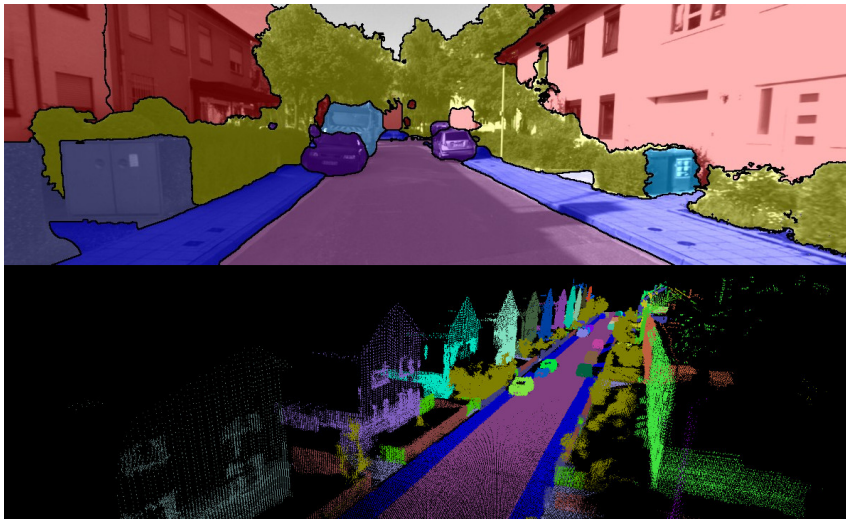
Strategy 5

Label Transfer

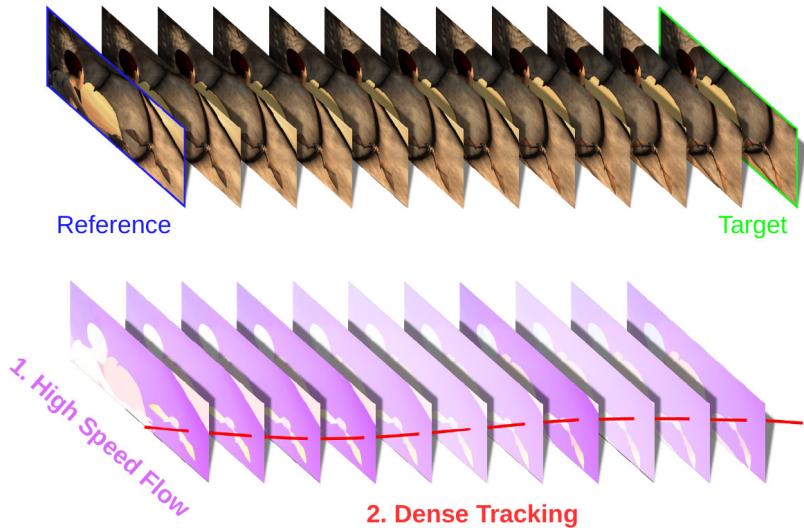
3D to 2D Semantic Label Transfer



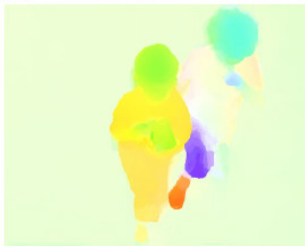
3D to 2D Semantic Label Transfer



Slow Flow



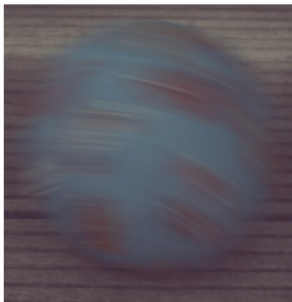
Slow Flow



Slow Flow



Input Image

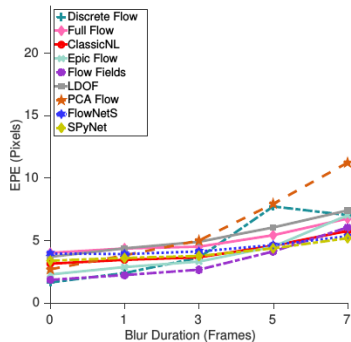


High frame rate

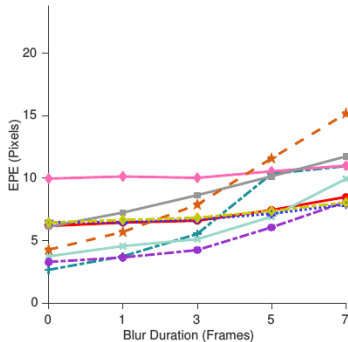


Low frame rate

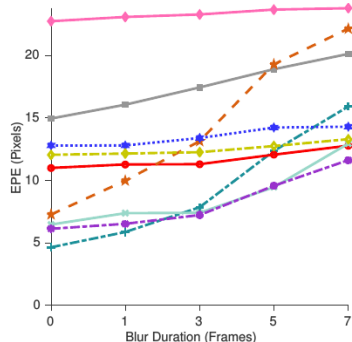
Slow Flow



(a) 100px Flow Magnitude



(b) 200px Flow Magnitude



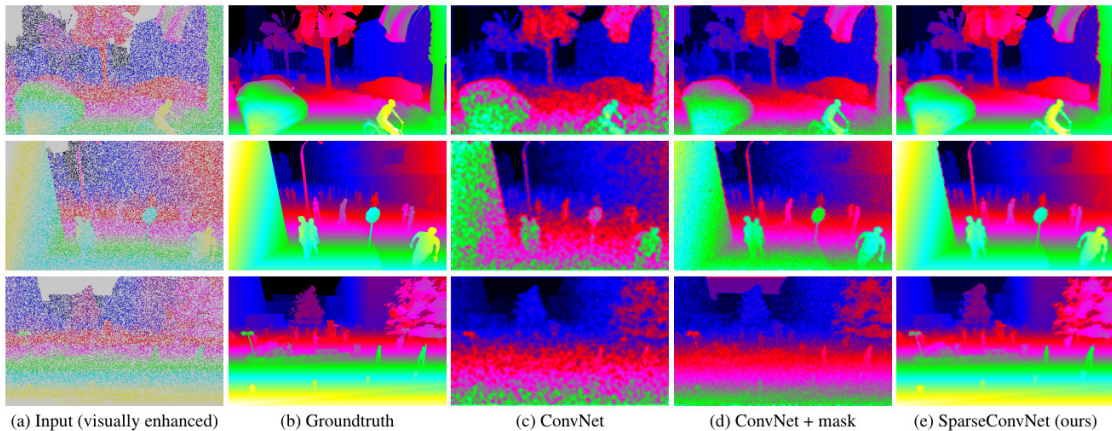
(c) 300px Flow Magnitude

Figure 4: State-of-the-art comparison on the generated reference data wrt. motion magnitude and blur.

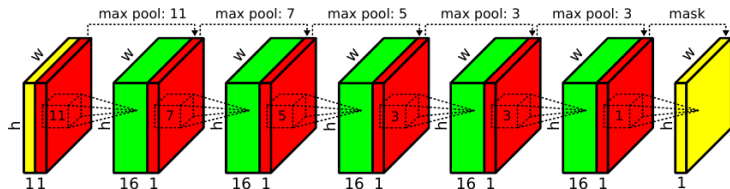
Strategy 6

Model Generalization

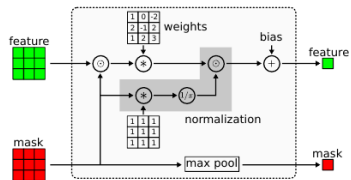
Sparsity Invariant CNNs



Sparsity Invariant CNNs



(a) Network Architecture



(b) Sparse Convolution

Thank you!