

Towards Generalization Across Depth for Monocular 3D Object Detection

Short explanatory video

Andrea Simonelli Samuel Rota Bulò Lorenzo Porzi Elisa Ricci Peter Kotschieder

FACEBOOK



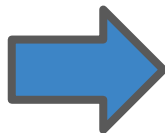
UNIVERSITY
OF TRENTO



Monocular 3D Object Detection

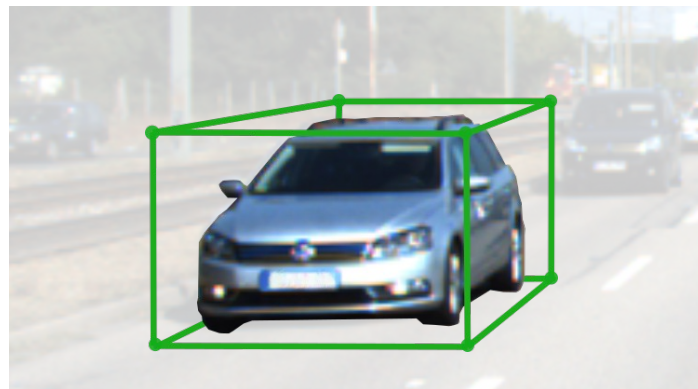
Input:

single RGB image



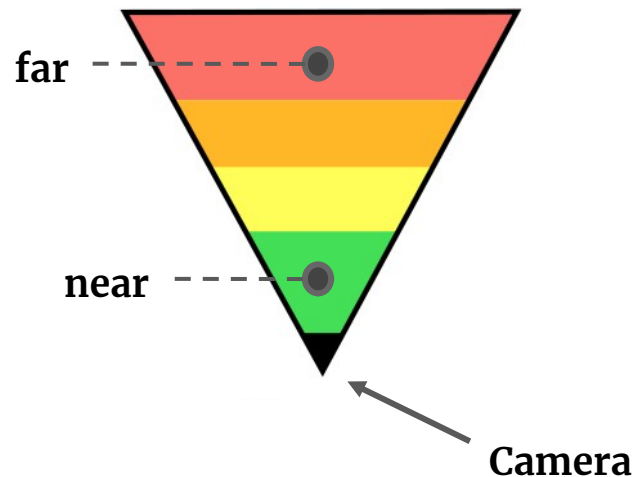
Output:

3D bounding boxes



Distance issue

Distance makes objects appear very **different**



Model **memorizes** distance-specific representations

Our 1st contribution: Virtual Views

Convert the input image into Virtual Views



Object appearance becomes **invariant** with respect to distance

Our 2nd contribution: MoVi-3D

A **single-stage**, lightweight architecture



Fully-convolutional with fast inference (~45 ms)

Results

- **State-of-the-art** on the popular KITTI3D
- Particular improvements on **far-away objects**

Car			
3D Average Precision		Bird's Eye View	
Moderate	Hard	Moderate	Hard
+12.3%	+24.8%	+24.6%	+33.5%

Towards Generalization Across Depth for Monocular 3D Object Detection

Check out our extended video!

Andrea Simonelli Samuel Rota Bulò Lorenzo Porzi Elisa Ricci Peter Kotschieder

FACEBOOK



UNIVERSITY
OF TRENTO

