EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

FACULTY OF SCIENCE
COMPUTER SCIENCE DEPARTMENT
COGNITIVE SYSTEMS · PROF. DR. A. ZELL

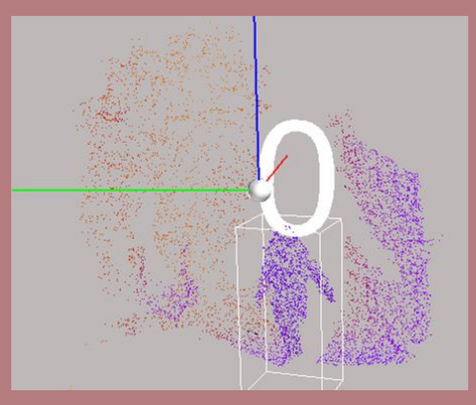
Yolo+FPN: 2D and 3D Fused Object Detection With an RGB-D Camera

Ya Wang, Andreas Zell

We propose a new deep neural network system, called Yolo+FPN, which fuses both 2D and 3D object detection algorithms.
It achieves better real-time object detection results and faster inference speed, to be used on real robots.
In order to satisfy real-time requirements, a trade-off between accuracy and efficiency is needed.
We have improved training and test accuracies and lower mean losses on the KITTI object detection benchmark.
And, we have competitive average precision on 3D detection of all classes in three levels of difficulty compared to other state-of-the-art methods.
Also, we implemented Yolo+FPN using an RGB-D camera, and compared the speed of object detection using different GPUs.
For the real implementation we focus on person detection, which is the most challenging and important among the three classes.

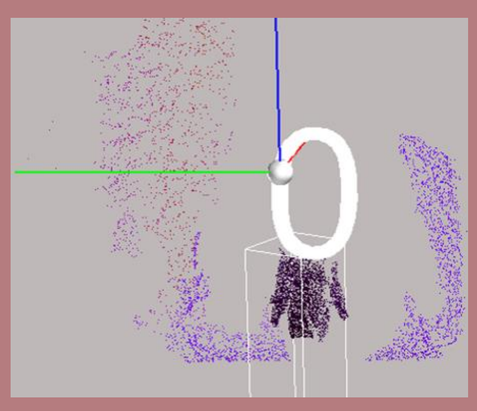
Qualitative Results

(a) 2D visualization (global)



(b) 3D visualization (global)

(c) 2D visualization (local)



(d) 3D visualization (local)

Fusion strategies

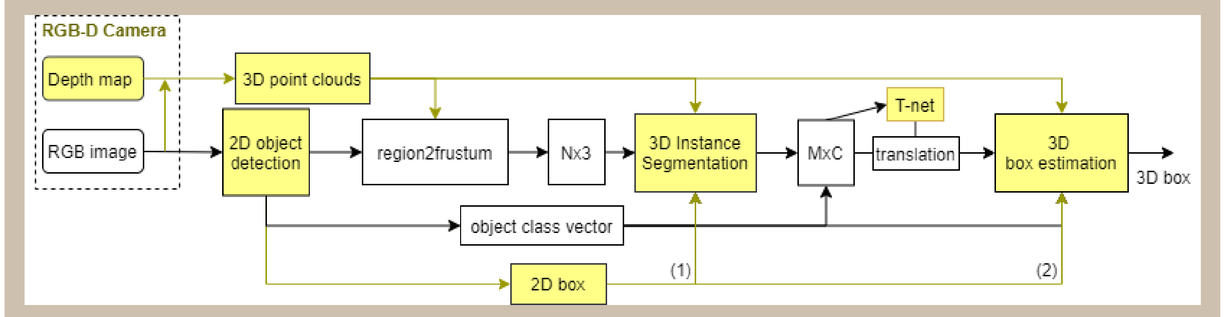
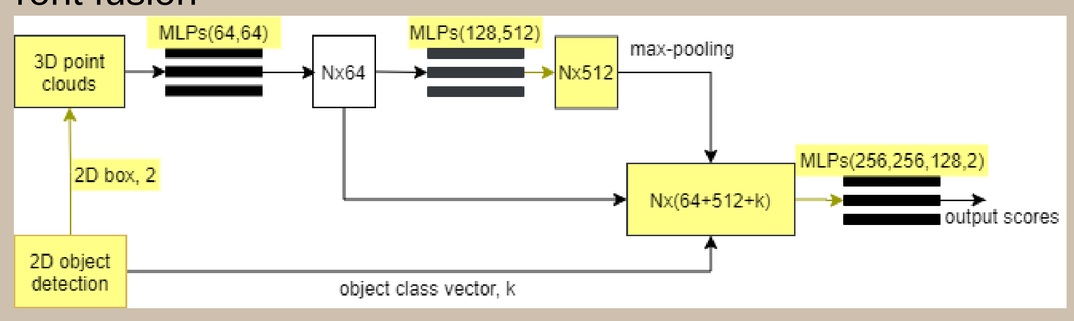


Fig. 2: Yolo+FPN structure.

(a) Front fusion



(b) End fusion

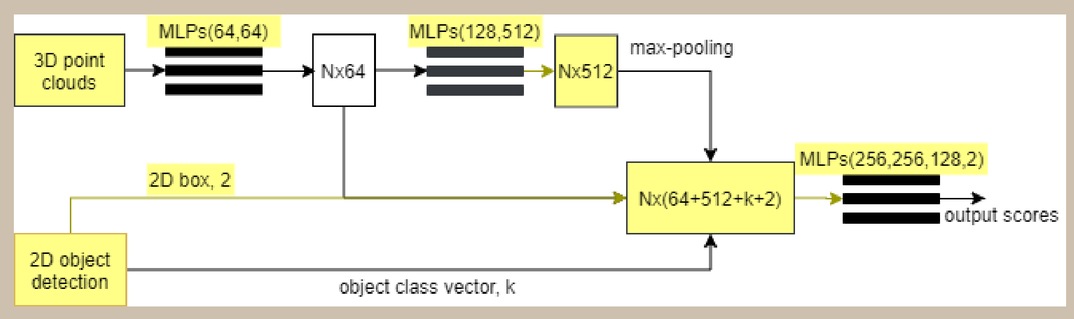
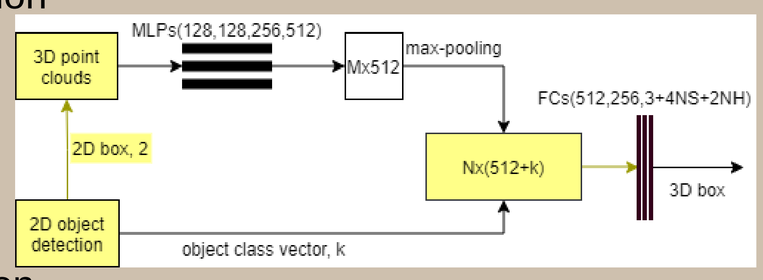


Fig. 3: Fusion strategies for 3D instance segmentation.

(a) Front fusion



(b) End fusion

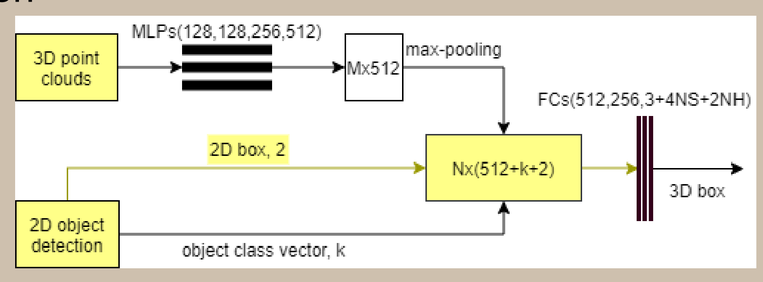
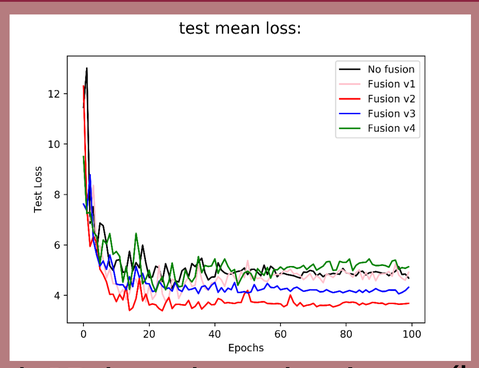
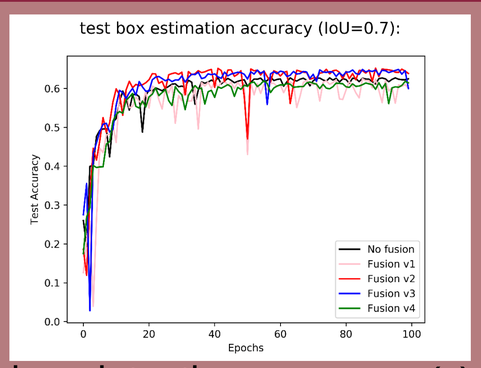


Fig. 4: Fusion strategies for 3D box estimation.

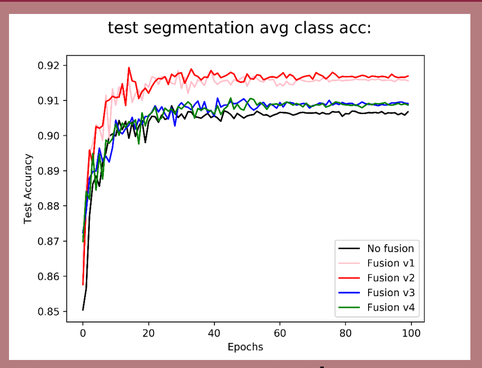
Quantitative Results



(a) 3D box detection loss



(b) 3D box detection accuracy



(c) 3D instance segmentation accuracy

References

[1] C. R. Qi, W. Liu, C. Wu, H. Su, and L. J. Guibas. Frustum PointNets for 3D object detection from RGB-D camera. CVPR, pages 918-927, 2018.

[2] J. Redmon, and A. Farhadi. Yolov3: an incremental improvement. In arXiv preprint arXiv: 1804.02767, 2018.

[3] Y. Wang, S. Xu, and A. Zell. Real-time 3D object detection from point clouds using an RGB-D camera. ICPRAM, pages 407-414, 2020.

University of Tuebingen · Sand 1 · 72076 Tübingen
Phone: +49-7071-29-78978
francis.wang@uni-tuebingen.de
www.cogsys.cs.uni-tuebingen.de