KI DELTA LEARNING
Scalable AI for Automated Driving

MultiTask-CenterNet (MCN): Efficient and Diverse Multitask Learning using an Anchor Free Approach

Falk Heuer, Syed Saqib Bukhari,
Georg Schneider

WP 3.4

Networks

Use-Case in Delta Learning

Multitask Networks can solve several tasks with a single network architecture. This has several beneficial properties for scalable automotive AI solutions. The most important one is filling the delta between multiple viable tasks for automotive AI such as segmentation, detection or pose estimation. A mutlitask network can solve these faster and with a generally simpler architecture than several single task networks. When using a shared backbone, a large quantity of weights is used mutually and compute resources can be saved. Exchangeable backbones can fill the gap between small, fast networks and more precise, large networks.

Technical Problem

When creating a solution for automated driving, it is often necessary to abstract the environment into several different tasks. Convolutional neural networks are usually chosen to perform this abstraction. However, in particular in the early layers, they extract similar information from the sensor space which creates unnecessary redundancies.

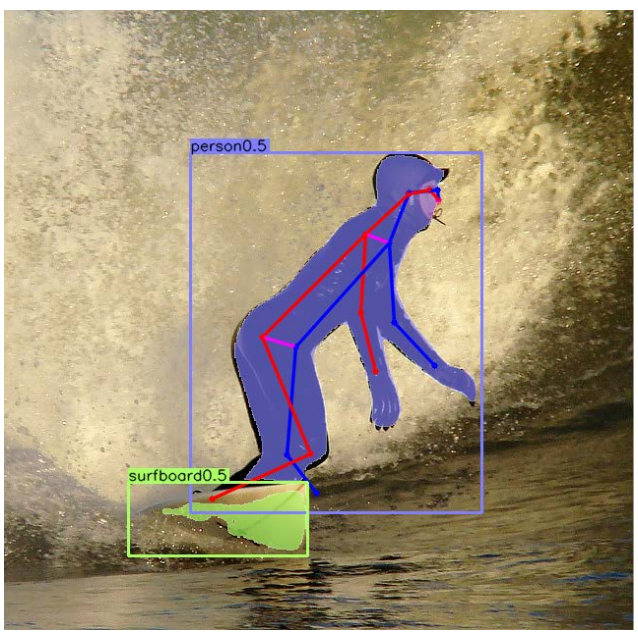

Figure 1: Visualization of a multitask network for object detection, semantic segmentation, and human pose estimation [1]

Technical Solution

In our published paper [1], we implement a multitask network for the tasks of segmentation, detection and pose estimation. A multitask network receives as the same image input as each individual specialist network would, and is built with several heads to create the relevant respective output for each task (Figure 2). The network itself processes the majority of information in a shared backbone and only splits into multiple heads towards the end of the layer pipeline. This way information between layers can be largely mutually shared and redundancies are minimized.

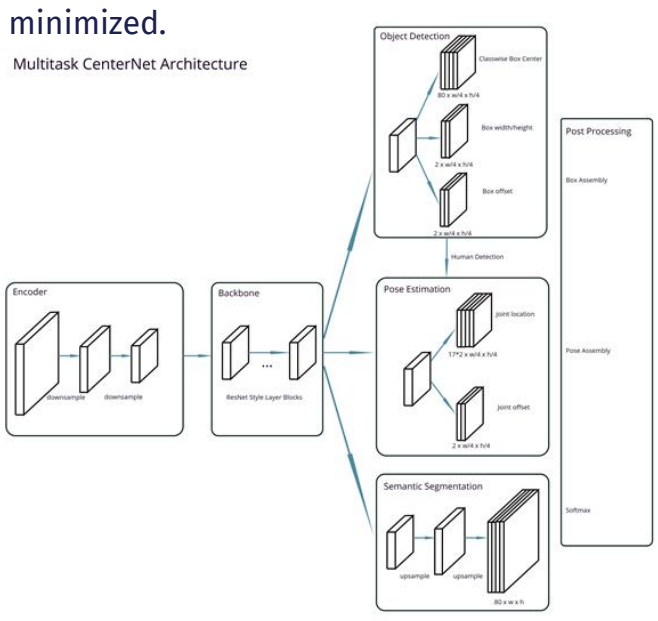

Figure 2: Depiction of the proposed Multitask-CenterNet (MCN). The multiple tasks are performed with a single backbone, saving computation time and resources [1].

Evaluation

Figure 1 shows a successful multitask network prediction for segmentation, human pose estimation and object detection. When evaluating performances for detection and segmentation, a multitask network is on par with or superior to single task networks (Table 1). In terms of latency, a multitask network for segmentation, detection and human pose estimation is more than twice as fast (34.4 fps) as a composition of single task networks for these tasks (16.3 fps).

Network	Segmentation mIoU	Detection mAP
Only Segmentation	49,0%	NA
Only Detection	NA	36.3%
Segmentation + Detection	48,9%	36.7%

Table 1: Performance of Singletask Networks (Detection and Segmentation) compared to a Multitask Network [1].

References:

[1] Falk Heuer, Sven Mantowsky, Syed Saqib Bukhari, and Georg Schneider. Multitask-centernet (mcn): Efficient and diverse multitask learning using an anchor free approach. ICCV - Embedded and Real-World Computer Vision in Autonomous Driving (ICCV-ERCVAD, 2021).

For more information contact:
falk.heuer@zf.com

Partners

External partners

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Energy.

Supported by: Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

www.ki-deltalearning.de @KI_Familie KI Familie