KI DELTA LEARNING
Scalable AI for Automated Driving

Learning Cascaded Detection Tasks with Weakly-Supervised Domain Adaptation

Niklas Hanselmann, Nick Schneider, Benedikt Ortelt, Andreas Geiger

WP 2.4

Environment

Use-Case in Delta Learning

Modern deep learning-based perception systems depend heavily on the availability of large amounts of labelled data. Acquiring this data at the required scale and diversity can be challenging in practice due to the cost of human annotation. A promising solution is to collect data in simulation, which is fully controllable and comes with free ground-truth annotations. However, because of the simulation-to-real delta, models trained solely on synthetic data typically perform poorly in the real world.

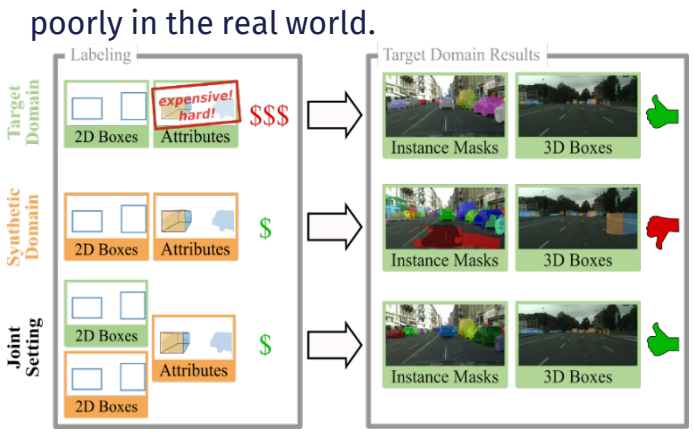

Figure 1: ***Proposed Setting.*** *We propose a weakly-supervised domain-adaptation setting that enables learning cascaded detection tasks at reduced annotation effort while still achieving competitive performance by utilizing weak 2D bounding box labels in both domains.*

Technical Problem

We propose and investigate a novel weakly-supervised domain adaptation setting (fig. 1) for cascaded detection tasks such as instance segmentation or monocular 3D detection. Here, the task is to learn a model that transfers well from a source domain to a target domain with access to cheaper weak annotations in the form of 2D bounding boxes in both domains at training time, while the full annotations (e.g. instance masks or 3D bounding boxes) are only available in the source domain.

Technical Solution

A large part of the technical solution is the proposed weakly-supervised domain adaptation setting itself, as it allows for a favorable trade-off between annotation cost and target domain performance. To fully leverage this setting we use a cascaded detection architecture and train the 2D detection stage jointly on both domains using the available weak supervision. Additionally, we employ class-wise adversarial feature distribution alignment to ensure that the learned features will be domain-invariant. The full pipeline is illustrated in figure 2.

Method	Instance Segmentation (mAP)	Monocular 3D Detection (mDS)
Source Only	17.0	13.2
DAFRCNN[1]	17.5	15.4
SWDA [2]	17.5	14.8
Ours [3]	31.3	23.4
Target Only	33.6	25.0

Table 1: ***Results.*** *„Source Only“, „DAFRCNN“ and „SWDA“ are trained without access to annotations in the target domain, „Ours“ is trained with weak annotations and „Target Only“ is the oracle with access to full annotations in both domains.*

Evaluation

We evaluate the approach on two exemplary cascaded detection tasks: instance segmentation and monocular 3D detection. We transfer from the synthetic *Synscapes* dataset to the real *Cityscapes* dataset, using its official benchmark metrics for the respective task. Evaluation on more datasets and adaptation settings can be found in the full publication [3]. As shown in table 1, models trained in our setting are competitive with the oracle on both tasks.

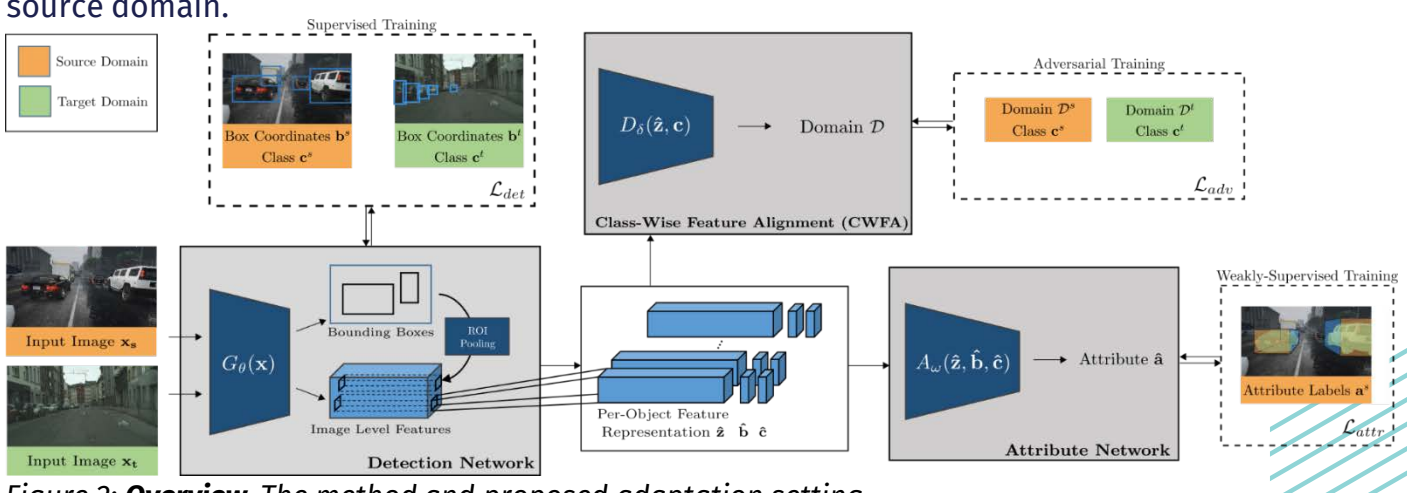

Figure 2: ***Overview.*** *The method and proposed adaptation setting.*

References:
[1] Chen et al., “Domain Adaptive Faster R-CNN for Object Detection in The Wild “, *CVPR 2018*
[2] Saito et al., “Strong-Weak Distribution Alignment for Adaptive Object Detection”, *CVPR 2019*
[3] Hanselmann et al.. “Learning Cascaded Detection Tasks with Weakly-Supervised Domain Adaptation”, *IV 2021*

BOSCH

Mercedes-Benz

For more information contact:
benedikt.ortelt@de.bosch.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.

KI FAMILIE

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