KI DELTA LEARNING
Scalable AI for Automated Driving

Learnable Duplicate Removal using Graph Neural Network for Single Stage Object Detector

Shuxiao Ding, Eike Rehder, Lukas Schneider, Marius Cordts, Juergen Gall

WP 3.4

Use-Case in Delta Learning

Deep learning based object detectors usually generates multiple detections for individual objects. Non-Maximum Suppression is employed to remove these duplicates. However, NMS is a heuristic algorithm which solely relies on the overlap between detections as the criterion for removal with the consequence that either too many detections are retained or removed. Following the deep learning paradigm of end-to-end optimization, we propose a learnable duplicate removal for Single Shot Detector (SSD) due to its efficiency.

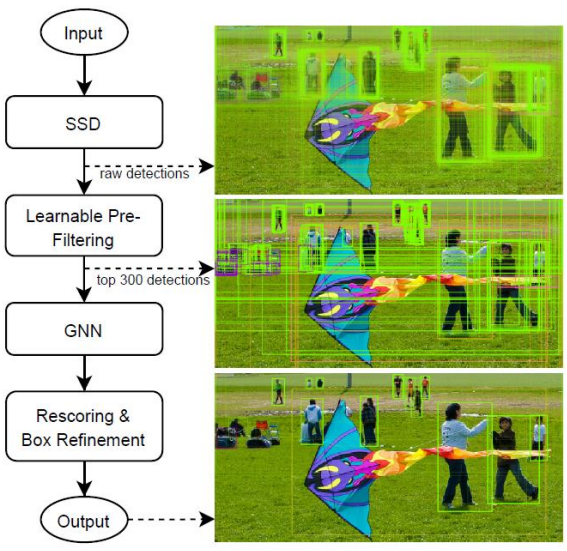

Figure1: Proposed learning NMS pipeline.

Technical Problem

Existing learned duplicate removal works are either a separate network which processes raw detections with handcrafted features [1] from an existing object detector or can be embedded into Faster- RCNN [2]. However, they are not directly applicable to SSD for two reasons. First, SSD provides way more raw detections (top image of Figure 1) compared to the pre-filtered proposals in RCNN. Second, the image features with lower dimensionality in SSD are less discriminative.

Technical Solution

To cope with the large amount of detections, we first use a learned pre-filtering step which filters the raw detections from SSD in an early stage. The set of filtered detections is regarded as an undirected graph, where an edge exists between two detections (nodes) if they overlap. Node features are propagated along the edges in the graph using a simple Graph Convolutional Network (GCN) [3] in order to obtain a single refined and rescored detection per object. In training, we use a bipartite-matching to assign ground-truth labels which considers the localization and classification quality simultaneously.

Evaluation

We compare the results of our approach with classical NMS on the MS-COCO test-dev set. To validate the generalization ability, we use two different networks as base SSD model: EfficientDet-D0 and RetinaNet-ResNet50. We train our network in two stages: first, the vanilla SSD model is trained alone. In the second stage, our proposed module is trained jointly with the underlying SSD with a lower learning rate. As shown in Table 1, our approach achieves better accuracy, especially in AP75. Our method also runs faster than using classical NMS.

		AP	AP50	AP75	Time(ms)
Efficient Det-D0	NMS	32.0	50.5	33.8	47.3
	ours	32.7	49.8	36.1	35.7
RetinaNet -ResNet50	NMS	34.2	51.7	37.2	59.7
	ours	34.3	50.4	38.2	48.0

Table 1: Experiment results. AP50, AP75 are Average Precision at IoU=0.5 and IoU=0.75.

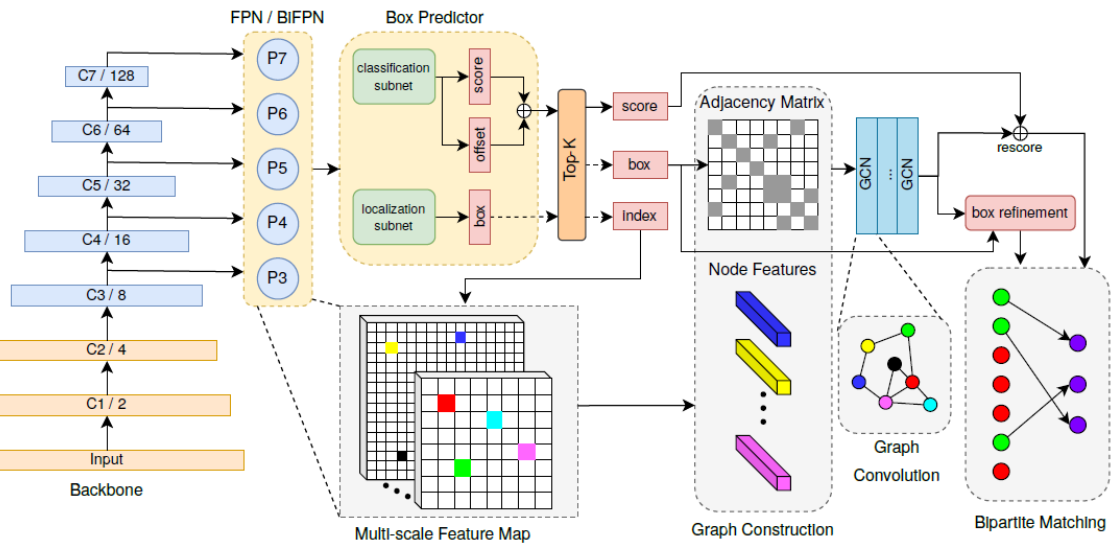

Figure 2: ***System Overview.***

References:
[1] Hosang et al., "Learning non-maximum suppression.", *CVPR 2017*
[2] Hu et al.. "Relation networks for object detection.", *CVPR 2018*
[3] Kipf et al.. "Semi-Supervised Classification with Graph Convolutional.", *ICLR 2017*

Networks

Mercedes-Benz

For more information contact:
shuxiao.ding@daimler.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