Adaptive Instance Distillation for Object Detection in Autonomous Driving

Qizhen Lan Qing Tian
Bowling Green State University, Bowling Green, Ohio 43403, USA

Introduction

Knowledge Distillation (KD) has been widely used to derive efficient classification models. However,

- **Only a limited number of studies have applied knowledge distillation to object detection, especially in time-sensitive autonomous driving scenarios.**
- **Most previous Knowledge Distillation (KD) methods treat all instances equally even though the teacher does not learn the instances equally well.**

Contribution

1. We propose **Adaptive Instance Distillation (AID)** to allow students to discern the reliability of the teacher's knowledge regarding a particular instance based on the teacher's prediction.
2. Our AID re-weighing method can be applied to each layer of FPN to achieve scale-wise selection in knowledge distillation.
3. Our AID can also improve the original architecture itself by self-distillation.

Methods

Goal: To improve knowledge distillation performance through re-weighting instances based on the teacher's prediction.

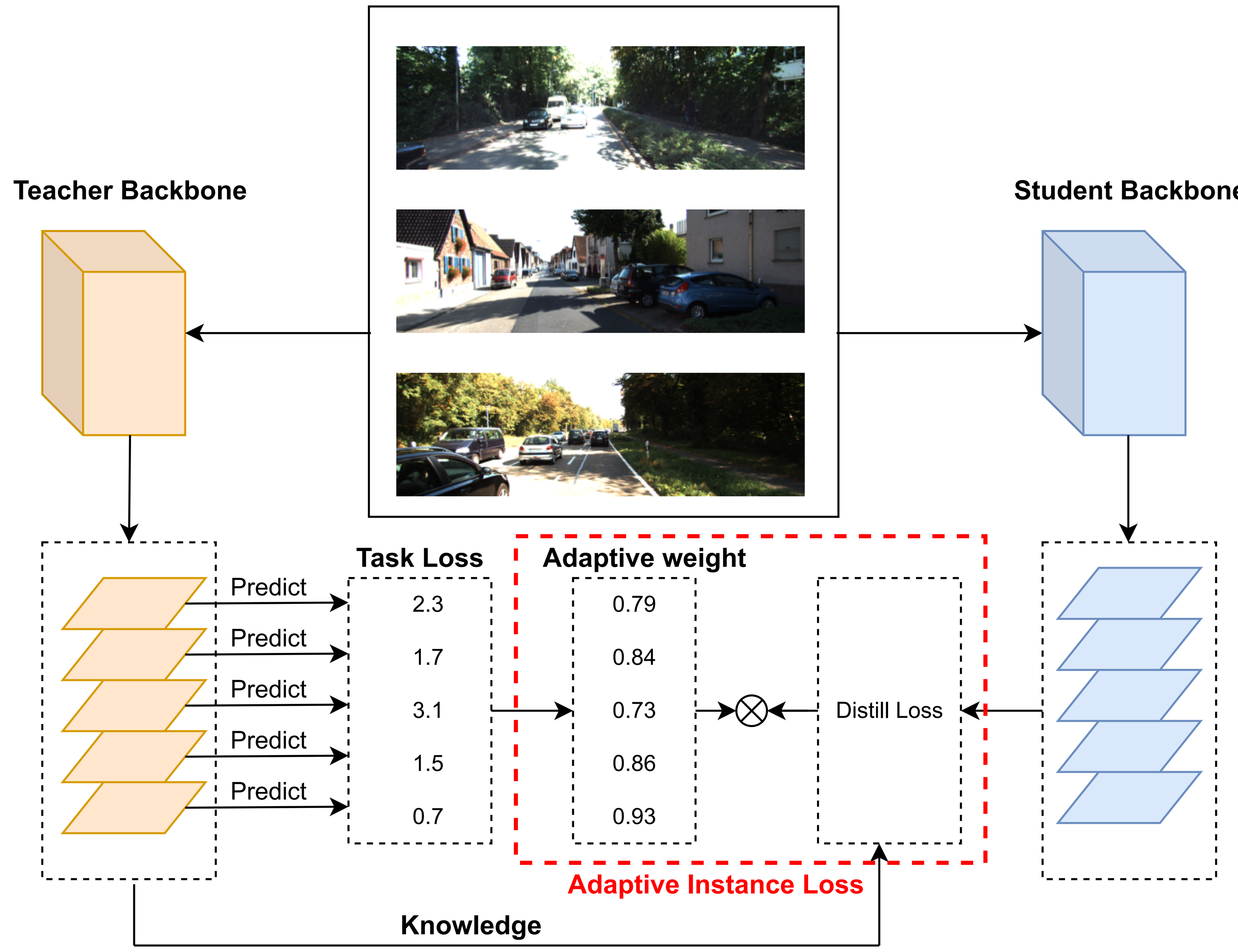

Figure 1. Overall Framework of our Adaptive Instance Distillation Method

Overall Loss:

$$\mathcal{L}_i^{\mathcal{S}} = \mathcal{L}_{task,i}^{\mathcal{S}} + \lambda \exp^{-\alpha \mathcal{L}_{task,i}^{\mathcal{T}}} \mathcal{L}_{distill,i}^{\mathcal{S},\mathcal{T}} \quad (1)$$

- **Our AID reweighs an instance based on the teacher's original loss**, which reflects the teacher's confidence in that instance. In those instances where the teacher performs bad, the student will rely more on itself to learn instead of being misled by the teacher.
- Our AID adaptively weighs **not only the instance-wise knowledge but also the scale-wise knowledge.**

Results

Table 1. Results on the KITTI dataset

Teacher-Backbone	Student-Backbone	Method	mAP (GFL)	mAP (Faster R-CNN)
ResNet-101	ResNet-50	Student-Baseline	85.1	88.9
		Zhang et al. [1]	86.4	89.0
		Ours (AID)	**88.0**	**89.6**
	ResNet-101	Teacher-Baseline	89.4	89.6
		Zhang et al. [1]	88.2	89.7
		Ours (AID)	**89.7**	**89.8**

Table 2. Results on the COCO traffic dataset

Teacher-Backbone	Student-Backbone	Method	mAP (GFL)	mAP (Faster R-CNN)
ResNet-101	ResNet-50	Student-Baseline	67.6	65.3
		Zhang et al. [1]	69.7	67.9
		Ours (AID)	**70.1**	**68.7**
	ResNet-101	Teacher-Baseline	71.0	67.1
		Zhang et al. [1]	72.5	67.8
		Ours (AID)	**72.6**	**68.1**

Table 3. Model complexity

Model	Backbone	Params(M)	GFLOPs
RetinaNet+	ResNet-50	32.06	10.07
	ResNet-101	51.05	13.79
Faster R-CNN	ResNet-50	41.17	23.40
	ResNet-101	60.17	27.13

Figure 2. Qualitative analysis of random samples.

References

[1] Linfeng Zhang and Kaisheng Ma. Improve object detection with feature-based knowledge distillation: Towards accurate and efficient detectors. In *ICLR*, 2021.