Title: Comparative Study on Self-Supervision Methods for Autonomous Driving

Author: Francisco Gonçalves Cerqueira

Supervision: Ricardo Pereira de Magalhães Cruz

Date: November 27, 2023

Abstract

Deep learning models demand a colossal volume of data for effective generalization and achieving superior performance, involving not only data collection but, more critically, the process of data annotation. This labor-intensive task of labeling each element within the data, be it vehicles, pedestrians, or other entities, requires a substantial workforce. Consequently, the amount of unlabeled data typically surpasses the amount of labeled data, and it becomes advantageous to find ways to leverage this untapped resource.

Semi-supervised and self-supervised learning tackle this challenge by incorporating unlabeled data during the training process under the assumption that the prediction of a label should remain consistent despite transformations applied to the corresponding sample. For example, while not sure whether an image depicts a pedestrian or a cyclist, it is reasonably expected that a slight rotation of the image should not alter its classification. Semi-supervision methods exploit labeled and unlabeled data and incorporate this insight by using the predictions as pseudo-labels and attributing them to any image transformations when the model's confidence in those predictions is sufficiently high. Self-supervision frameworks operate in the absence of annotated data, introducing a term in the loss function that penalizes inconsistent classifications for the same modified image. Approaches like supervised learning can later utilize the annotated data to fine-tune the model, helping it further refine its understanding and representation of the data, ultimately improving its performance on downstream tasks.

The study aims to make innovative contributions through an extensive comparative analysis of semi-supervision and self-supervision methods, with the primary objective of assessing their practical application and performance within the challenging domain of Autonomous Driving – several public datasets are available, with KITTI being one of the most popular in research. A practical software package will be developed as an additional objective, enabling other projects and fields to benefit from these techniques.

Keywords: semi-supervised learning, self-supervised learning, label efficiency, autonomous driving, computer vision, deep learning, semantic segmentation

ACM Classification:

- Computing methodologies → Artificial intelligence → Computer vision
- Computing methodologies → Machine learning → Machine learning approaches → Neural networks
- Computing methodologies → Machine learning → Learning settings → Semi-supervised learning settings
- Computing methodologies → Machine learning → Learning paradigms → Unsupervised learning

References

[1] Florian Alexander Schmidt. Crowdsourced production of AI Training Data: How human workers teach self-driving cars how to see. Working Paper Forschungsförderung 155, Düsseldorf, 2019.

[2] Olivier Chapelle, Bernhard Scholkopf, and Alexander Zien. Semi-supervised learning (chapelle, o. et al., eds.; 2006)[book reviews]. *IEEE Transactions on Neural Networks*, 20(3):542–542, 2009.

[3] Xiao Liu, Fanjin Zhang, Zhenyu Hou, Li Mian, Zhaoyu Wang, Jing Zhang, and Jie Tang. Self-supervised learning: Generative or contrastive. *IEEE transactions on knowledge and data engineering*, 35(1):857–876, 2021.

[4] Samuli Laine and Timo Aila. Temporal Ensembling for Semi-Supervised Learning. In *International Conference on Learning Representations*, 2017.

[5] David Berthelot, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A Raffel. MixMatch: A holistic approach to semi-supervised learning. *Advances in neural information processing systems*, 32, 2019.

[6] Zeyu Feng, Chang Xu, and Dacheng Tao. Self-Supervised Representation Learning by Rotation Feature Decoupling. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[7] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pages 1597–1607. PMLR, 2020.