Masterarbeit

Enhancing Data Efficiency for Object Detection:

Strategies for Smart Dataset Reduction

zur Erlangung des akademischen Grades „Master of Science (M.Sc.)" im Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik, Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der Leibniz Universität Hannover

vorgelegt von

Mirco Oliver Höhne

[redacted]

Hannover, den 22. Oktober 2024

Prüfer: Prof. Dr. rer. nat. Michael H. Breitner

Contents

2. Introduction

2.1. Motivation

Deep learning methods have played a pivotal role in the evolution of object detection, increasing accuracy in contrast to conventional methods by using large amounts of data and computing resources [1], [4]. However, computational efficiency becomes a significant concern as models become complex and datasets grow. Reducing data in an intelligent way that has a minimal effect on performance is one way to increase efficiency in computing. Ensuring efficient computing is essential to make training faster and more cost-effective. In addition, using fewer resources is more sustainable and leads to a lower carbon footprint. Furthermore, it enables the deployment of models in resource-constrained environments and reduces the barrier of entry for entities with limited resources. Reducing the data may also lead to faster and better generalization since irrelevant and faulty data is removed in data reduction.

2.2. Background

While research mainly focuses on developing new architectures and models, there is a limited effort regarding the improvement of data quality [44]. But what constitutes high-quality data? A high-quality dataset should be diverse, representative, comprehensive, free of bias, well-balanced, and include high-quality annotations [44]–[47]. To craft a high-quality dataset, only useful observations that impact the models' performance must be selected. Therefore, it is necessary to critically examine the options for data reduction to get a high-quality dataset that only includes samples that impact the model's performance without introducing bias.

2.3. Structure

First, the basic concepts of the technologies used are introduced. These basic concepts are crucial for grasping the methods, algorithms, and frameworks used later in the thesis. An in-depth literature review for data reduction in 2D object detection follows the basic concepts. The methodology for the literature review is explained at the beginning of the chapter. Then, methods are proposed for data reduction in object detection, and the framework and setup for the experiments are highlighted. Afterward, the results are presented, discussed critically, and summarized.

2.4. Research Questions

This thesis proposes two research questions that represent this thesis's objective.

- How can we create a subset that minimally decreases the 2D object detection performance compared to the entire training dataset?
- Is there an evaluation metric that captures the characteristics of subsets that perform well in object detection?

2.5. Objectives

This thesis aims to bridge the gap from image classification to 2D object detection by analyzing data reduction methods from image classification, adapting the concepts to object detection, and creating new approaches. A comparative analysis will give insights into performance and allow us to analyze subsets to gain knowledge about the characteristics of datasets that perform well on the object detection task. Success is quantified by subset size, adaptability to other datasets, and performance metrics such as mAP.

8. Conclusion and Future Work

This study explored different ways of reducing datasets for 2D object detection using the nuImages and BDD100K datasets. First, established techniques from image classification were introduced, and then the overarching methods and concepts were adapted and implemented for 2D object detection, which included baseline, error-/loss-based, and geometry-based strategies.

The main result is that the best-performing dataset reduction methods (*loss hard*) can reduce the dataset size needed to train a model by 40% across datasets, with performance not only matching the entire dataset but exceeding it, even across object sizes. Other methods like *bbox max* and *area max* also perform better than random selection and exceed the performance of the entire dataset in higher dataset ratios. The recommendation for action for practitioners is to try those methods for their model training, as it might substantially decrease the needed sample size for training their model and improve the model's generalizability, especially when using Faster R-CNN.

The findings have several practical implications for 2D object detection and its applications. Using effective data reduction methods, practitioners can train object detection models more efficiently, using less computational resources and time. These data reduction methods are instrumental when working with big data or limited computing resources. Furthermore, these findings can be applied in scenarios where obtaining additional labeled data is resource-intensive, and data reduction methods enable models to learn with fewer samples. Data reduction methods can be an alternative to improve model performance in this scenario. Another scenario that benefits from these findings is resource-constrained environments, like real-time systems, where models can be trained on reduced datasets while retaining performance. The results showed that improving the model generalization for 2D object detection is possible with these methods, making it worthwhile even in scenarios where resources are not as scarce.

However, this work has limitations, including methodological adaptations, dataset scope, model selection, experimental design, evaluation metric, computational considerations, and potential biases. These may limit the findings' applicability and generalizability to other models, datasets, and scenarios.

These limitations serve as the starting point for further research. Further research should cover additional backbones and models, such as SSD [65], YOLO [6] or transformer-based models [7], for the 2D object detection task using the proposed methods to see if the methods generalize across architectures. Furthermore, these methods should be extended to other datasets and scenarios to show generalization capability. Especially regarding the ID, different datasets should be tested to see if the hypothesis holds that mAP and ID are generally correlated across datasets. If that is the case, it would significantly impact the field since reduction methods could be tested without training a model, saving many resources. Other future directions include exploring advanced techniques like active learning or semi-supervised learning to obtain reduced datasets. Another interesting study could be made regarding computational efficiency, where time and resource costs could be evaluated. Furthermore, other diversity metrics could be explored to gain insights into data reduction methods.

Given the results of this work in light of the limitations, dataset reduction in 2D object detection presents a promising field of research for advancing more efficient, sustainable, and accessible models.