Introduction to Machine Learning Assignment 3 Proposal

Matthijs Prinsen (s4003365)
Marinus van den Ende (s5460484)
Group 54

December 13, 2024

1 Proposal

We propose fine-tuning an object detection model, specifically YOLO v7(Labs, 2023), to classify items in images of a user's fridge or grocery cabinet. While YOLO v7 is currently trained on a diverse range of objects, such as people, bicycles, and furniture, we aim to adapt it for recognizing common supermarket products typically found in an average household.

At present, we do not have a finalized dataset, but we have identified several promising options on Hugging Face (Face, 2024). Regardless of the source, we will ensure high-quality annotations by defining target variables with bounding boxes, adhering to best practices (Labs, 2021). Consequently, our dataset may comprise two components: one sourced externally and another created in-house for training.

For evaluation, we plan to utilize the following metrics:

1. **Intersection over Union (IoU):** This metric measures the ratio of the overlap area to the total area between predicted and ground truth bounding boxes, providing a quantitative measure of the model's prediction accuracy (Labs, 2023).

2. **Average Precision (AP):** This metric is calculated as the area under the precision-recall curve, offering a class-wise average precision score (Labs, 2023).

Our implementation will involve fine-tuning the YOLO v7 model using **PyTorch**. We will draw inspiration from WongKinYiu's example project on YOLO v7 (WongKinYiu & Contributors, 2024), which builds on the research paper that introduced the model (Wang, 2023).

References

Face, H. (2024). Hugging face – the ai community building the future [Accessed: 2024-12-13]. https://huggingface.co

Labs, V. (2021). Annotating with bounding boxes: Quality best practices [Accessed: 2024-12-13]. https://www.v7labs.com/blog/bounding-box-annotation

Labs, V. (2023). Yolo object detection: A guide to the latest advances [Accessed: 2024-12-13]. https://www.v7labs.com/blog/yolo-object-detection

Wang, C.-Y. (2023). Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 7464–7475. https://doi.org/10.1109/cvpr52729.2023.00721

WongKinYiu & Contributors. (2024). Yolov7: The github repository [Accessed: 2024-12-13]. https://github.com/WongKinYiu/yolov7