Object Detection Using Knowledge Distillation

1. Introduction

Object detection is one of the most crucial tasks in computer vision, with applications ranging from autonomous driving to surveillance and robotics. State-of-the-art object detection models like **Faster R-CNN, YOLO**, and **RetinaNet** are large and computationally demanding, which makes them challenging to deploy on resource-constrained devices such as mobile phones or embedded systems. Knowledge Distillation (KD) offers a solution by transferring knowledge from a large, pre-trained teacher model to a smaller student model, making it computationally efficient while maintaining high accuracy.

The goal of this project is to apply **knowledge distillation** to **object detection**, creating a lightweight student model that achieves comparable performance to the teacher model but is optimized for efficiency. This project will focus on training and evaluating the student model using metrics such as mean **Average Precision (mAP), inference speed, and model size**.

2. Objective

The objective of this project is to implement an object detection system using Knowledge Distillation. The project will:

1. Select a pre-trained object detection model as the teacher.
2. Develop a lightweight student model.
3. Use knowledge distillation techniques to train the student model.
4. Evaluate the performance of the student model in comparison to the teacher model based on mAP, inference speed, and model size.

3. Methodology

1. **Teacher Model Selection**: The teacher model will be a state-of-the-art object detection model such as **Faster R-CNN or YOLOv5**, pre-trained on the **COCO dataset**. These models have demonstrated high accuracy in detecting multiple object classes in complex images.
2. **Design of the Student Model**: A lightweight version of the teacher model will be designed to act as the student. This model will have fewer layers or reduced parameters to decrease computational complexity. For instance, we may use **MobileNet** or another efficient backbone for the student model.

3. **Knowledge Distillation Process**: Knowledge Distillation will involve training the student model using both ground-truth annotations and the softened output from the teacher model. The student will learn to mimic the teacher's outputs, including object bounding boxes and class probabilities.
4. **Training Dataset**: The **COCO dataset**, with over **200,000 labeled images** and **80 object categories**, will be used to train and evaluate both the teacher and student models. The dataset includes annotations for object detection, segmentation, and keypoint detection, providing a comprehensive test bed for object detection tasks.
5. **Evaluation Metrics**: The student model will be evaluated based on:
   - **Mean Average Precision (mAP)**: Measures the precision and recall of the object detection performance.
   - **Inference Time**: Time taken by the student model to process each image.
   - **Model Size**: The number of parameters and memory footprint.
   - **Energy Efficiency**: Power consumption during inference (optional).

4. Literature Review

Knowledge Distillation (KD) has been applied successfully across various domains, including image classification and object detection. Some key papers that inform this project are:

1. **"Distilling the Knowledge in a Neural Network" by Hinton et al. (2015)**: This paper introduced the concept of knowledge distillation and outlined how a small student model can be trained to mimic a larger teacher model's softened output. The work laid the groundwork for many modern applications of KD.
2. **"Data-Free Knowledge Distillation for Object Detection" by Liu et al. (2021)**: This paper presents a novel approach for data-free knowledge distillation in object detection, where the student is trained without access to the original dataset. Instead, images are synthesized based on the pre-trained teacher's output. This work is especially relevant when data privacy is a concern.

3. **"Knowledge Distillation for Efficient Image Super-Resolution" by Lee et al. (2020)**: Though focused on image super-resolution, this paper demonstrates the effectiveness of KD in reducing model size and inference time without sacrificing too much performance. The methodology can be adapted for object detection, where spatial accuracy is similarly critical.
4. **"Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks" by Ren et al. (2017)**: This paper proposes Faster R-CNN, one of the foundational models for object detection. It serves as an excellent candidate for the teacher model in this project, given its proven high accuracy and applicability in many object detection tasks.

5. Expected Outcome

By the end of the project, we aim to achieve a lightweight object detection model (the student) that:

- Maintains high accuracy comparable to the teacher model (within a small margin of mAP loss).
- Has a significantly smaller size (in terms of parameters and memory usage).
- Achieves faster inference speeds, making it suitable for real-time object detection on resource-constrained devices.
- Performs efficiently on the COCO dataset.

6. Conclusion

This project seeks to leverage knowledge distillation to optimize object detection models for real-time applications. By transferring knowledge from a large teacher model to a small, efficient student model, we aim to make state-of-the-art object detection technology accessible for deployment on mobile and embedded devices.

References

1. G. Hinton, O. Vinyals, and J. Dean, "Distilling the Knowledge in a Neural Network," *NIPS Deep Learning and Representation Learning Workshop*, 2015.
2. X. Liu, S. J. Pan, and Q. Yang, "Data-Free Knowledge Distillation for Object Detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 2414-2423.
3. H. Lee, K. H. Kim, and W. J. Lee, "Knowledge Distillation for Efficient Image Super-Resolution," in *IEEE Transactions on Multimedia*, vol. 22, no. 11, pp. 2794-2808, 2020.
4. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.

NORTH SOUTH UNIVERSITY

Committed to the highest standards of academic excellence

Course: CSE465(Section: 1)

Summer 2024

Efficient Object Detection Using Knowledge Distillation for

Real-time Applications

Submitted to

Dr. Nabeel Mohammed (NbM)

Department of Electrical & Computer Engineering(ECE)

Name	Id
Md Al Amin	1811904042
Tanzim Alam Fahim	1831917642