Transformers and Attention-Based Deep Networks Assignment - 2
Object Detection

Middle East Technical University
Graduate School of Informatics
Due: 20.04.2025

April 7, 2025

Introduction

In this assignment we will be detecting objects in a remote sensing dataset. Object detection is a fundamental computer vision task that aims to locate objects and classify them simultaneously within an image frame. The spatial location of objects of interest are typically represented by a bounding box that encloses each object instance. Bounding boxes are typically represented with coordinates (x_min, y_min, x_max, y_max) or (x_center, y_center, width, height). After locating objects of interest, each object is assigned a class label.

Deliverables

Your submission is expected to be uploaded to ODTUCLASS in a single zip file with the naming convention "DI725_assignment_2_STUDENTNUMBER.zip".
Report: Prepare a report that clearly explains your steps in this assignment. The report shall have the following sections: abstract, introduction, dataset, method, evaluation, results, discussion, conclusion.
Code: The code that is to be shared must be working and reproducible, submit your code to odtuclass, and also upload it to your public GitHub repo.
Version Control: Your implementation shall be traceable from your github repo. Your repo must be public.
Experiment Tracking: Experiments in this assignment are to be recorded to WANDB. Your WANDB experiments page or WANDB report page must be public.

Format

The report document must be a PDF file. The report shall be in IEEE format. You can find the IEEE template here. Please follow this path to select the correct format: Conferences, Original Research, Word or Latex.

Duration

This is a two-week assignment that should not be rushed in the last day. Any submission after the deadline will have a penalty of 10 points for each day. You will be unable to submit after 2 days.

Assignment

Apply any transformer model for object detection task. Refer to the dataset and the baseline obtained with non-transformer models in the research report. You are expected to fine-tune or train small transformer models to improve over the baseline performance of the existing object detectors (YOLO and MobileNet).

Dataset

The dataset we are going to use is the AU-AIR dataset, as shown in Figure 1 and is available here. Download the images and annotations of AU-AIR dataset from the repository. You can read the details of the dataset from the official research report.
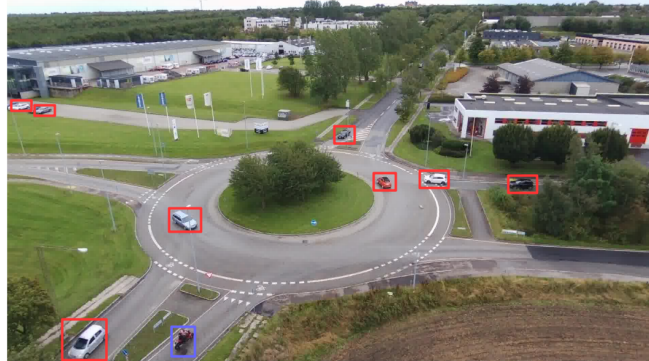

Figure 1: AU-AIR dataset

Table 1: Category-wise average precision values of the baseline networks.

Model	Human	Car	Truck	Van	M.bike	Bicycle	Bus	Trailer	mAP
YOLOV3-Tiny	34.05	36.30	47.13	41.47	4.80	12.34	51.78	13.95	30.22
MobileNetV2-SSDLite	22.86	19.65	34.74	25.73	0.01	0.01	39.63	13.38	19.50

Report

Your report shall contain the following information:
Abstract: prepare an abstract (100 words max)
Introduction: Briefly introduce the task.
Dataset: Explain your pre-processing steps briefly.
Method: Explain your method briefly.
Evaluation: Explain your evaluation metrics.
Results: Share the results of the experiments.
Discussion: Discuss your findings and results.
Conclusion: Conclusion and final remarks.

Grading

1. **Version control (Git) (10 pts)**: your GitHub repository shall be public and contain at least 3 commits in different days. If you encounter an issue or a bug while doing this assignment you shall open an issue describing it. If you leave this assignment to the last day you will not get any points from version control.

2. **Experiment Tracking (WANDB) (10 pts)**: your wandb experiments or report page shall be public. If it is not public you will not get any points from experiment tracking.

3. **Method (40 pts)**: Your task is to detect objects in the given dataset. Apply common clean coding practices and comment where necessary to get full points. If your code is messy and hard to understand you will not get full points.

4. **Report (40 pts)**: Your report shall be 2 pages maximum, if you exceed this limitation you will get -5 points for each extra page. Your report is your only way of communication for your assignment, it shall explain clearly all necessary steps you have taken. Your report shall have the following parts: abstract, introduction, dataset, method, evaluation, results, discussion, conclusion.