Final Project Proposal: Exploring Object Detection Capabilities from UAV Aerial Perspectives with Deep Learning Tools

Background and Motivation

Unmanned Aerial Vehicles (UAVs) are increasingly used in areas such as traffic monitoring, surveillance, and disaster relief. A key challenge in these applications is accurate object detection from aerial imagery, which is often complicated by factors such as varying object scales, cluttered environments, and altitude changes. Despite these challenges, advances in deep learning, in particular the You Only Look Once (YOLO) object detection model, offer promising solutions for real-time detection.

The motivation for this project stems from the need for efficient real-time object detection in UAV-captured imagery. By implementing and adapting a YOLO-based detection system, the project aims to demonstrate the feasibility of real-time object detection and tracking in aerial perspective. Successful implementation will provide a practical tool for UAV applications while laying the groundwork for future enhancements in detection accuracy and system robustness. What's more, the project will provide insight into adapting state-of-the-art object detection models for UAV applications, with practical implications for areas such as autonomous navigation and situational awareness in disaster areas.

Aims

Develop a Reliable Object Detection System: Implement an object detection system using the YOLO model that is capable of accurately identifying objects in UAV imagery from the VisDrone dataset. Once the basic detection system is functional, attempt to optimize the model to improve accuracy and performance.

Develop a Real-time Object Detection System: After establishing a reliable detection system, optimize the model for real-time performance. The goal is to integrate the system with live video feeds from UAVs to enable real-time identification of objects such as vehicles and pedestrians while maintaining detection accuracy.

Proposed Methodology

Data and Model Selection:

The VisDrone dataset will be used as the primary source of UAV captured imagery. This dataset contains a variety of objects such as vehicles, pedestrians, and other commonly observed objects in aerial imagery.

A pre-trained YOLO model will be selected for its ability to perform object detection in real-time. Transfer learning techniques will be used to adapt the model to the VisDrone dataset to ensure it performs well on UAV-specific imagery.

Real-Time Detection Implementation:
The real-time detection pipeline will be developed using the YOLO model. This system will take live video input from UAVs and process it in real-time, identifying objects and displaying detection results directly on the video stream.
The focus will be on achieving stable performance with the live feed to ensure that objects are efficiently detected and tracked without significant delay.

Evaluation:
The performance of the system will be evaluated based on its ability to accurately detect objects in real time. Metrics such as detection speed, object tracking consistency, and frame processing rate will be used to assess the system's practicality in real-world UAV applications.

References

VisDrone Dataset. github.com/VisDrone/VisDrone-Dataset.
YOLO Documentation. docs.ultralytics.com/.