Object Detection in Drone Imagery Using YOLO v3

GitHub Link: https://github.com/saurabh-028/DL-Tutorials

Introduction

Object detection in drone photography is essential for applications including surveillance, search and rescue, and traffic monitoring. Images obtained by drones provide distinct obstacles, such as diminutive item dimensions, fluctuating illumination conditions, and congested environments. This experiment employs YOLO v3, a cutting-edge object detection algorithm, on a subset of the VisDrone dataset to assess its efficacy in identifying items from aerial viewpoints. The implementation was executed in Python, adhering to the instructor's directive to reproduce the results from the peer-reviewed article "Drone Object Detection Using Deep Learning Algorithms." Due to the absence of a specified dataset in the original study, we opted for a sample from the VisDrone dataset available on Roboflow, which offers a variety of drone imagery for comprehensive examination. The aim is to evaluate the efficacy of YOLO v3 in drone-based object recognition and investigate the possibilities of a tailored Convolutional Neural Network (CNN) model. The research emphasizes YOLO v3's enhanced efficacy compared to SSD (Single Shot Detector) at identifying distant or multiple objects, rendering it an appropriate selection for this purpose.

Methodology

Overview of the Base Paper

The document "Drone Object Detection Using Deep Learning Algorithms" evaluates SSD and YOLO for object detection in drone applications. SSD uses a convolutional neural network to forecast bounding boxes and class scores. YOLO v3 segments images into a grid, forecasting bounding boxes and classifications for each cell, thus facilitating expedited and precise identification of remote or many objects. The report designates YOLO v3 as the optimal choice for drone imagery, informing its selection for this study.

Implementation Details

The implementation was carried out in Python with the following steps:

1. **Model Configuration:** YOLO v3 configurations, comprising weights and architecture files, were sourced from the official YOLO website (https://pjreddie.com/darknet/yolo/). The model was employed in its conventional configuration for object detection.
2. **Dataset Selection:** A sample of 8000 photos was obtained from the VisDrone dataset via Roboflow (https://universe.roboflow.com/visdrone/visdrone-

lzsy1/dataset/3). The VisDrone dataset, developed by the AISKYEYE team at Tianjin University, comprises a variety of drone imagery annotated for things such as pedestrians, vehicles, and bicycles. Images were preprocessed to conform to YOLO v3's input specifications, including scaling to 416x416 pixels.

3. **Testing of YOLO v3:** The pre-trained YOLO v3 model was evaluated on the VisDrone dataset. Results were illustrated using subplots, depicting detections of entities such as pedestrians, vehicles, and bicycles.
4. **Custom CNN Development**: A custom CNN model was trained on the same VisDrone sample to investigate enhanced performance. Hyperparameters, such as learning rate and batch size, were optimized, resulting in approximately 97% accuracy.

Modifications

The architecture of YOLO v3 remained unaltered, maintaining consistency with the technique outlined in the study. The VisDrone sample from Roboflow was used as the dataset, as none was indicated in the study. The custom CNN served as an ancillary experiment to evaluate possible enhancements but did not modify the YOLO v3 implementation.

Results

The performance of YOLO v3 and the custom CNN model was evaluated on the VisDrone dataset sample, with results summarized below:

Model	**Metric**	**Outcome**
YOLO v3	Detection Performance	Strong in daylight; limited in low light and crowded scenes
YOLO v3	Objects Detected	Pedestrians, vehicles, bicycles (displayed in subplots)
Custom CNN	Accuracy	Approximately 97%

YOLO v3 Performance

- Successful detection was accomplished in brightly lighted environments, with subplots displaying precise bounding boxes for pedestrians, automobiles, and bicycles.
- Efficacy diminished in dim environments, resulting in frequent omissions or misclassifications of objects due to inadequate visibility.
- In heavily populated settings, the identification of persons was erratic, resulting in both overlooked detections and erroneous bounding boxes.

Custom CNN Performance

The bespoke CNN model, trained on the VisDrone dataset, attained an accuracy of almost 97%. The elevated performance suggests that a model specifically designed for drone footage can substantially enhance detection accuracy.

Observations

- **YOLO v3 Detection Results:** Subplots demonstrated YOLO v3's proficiency in identifying pedestrians, automobiles, and bicycles in daylight, featuring distinct and precise bounding boxes.
- **Performance Notes:** Robust performance was noted in daytime; however, low-light conditions resulted in diminished detection accuracy. In congested environments, the efficacy of people detection was significantly diminished, underscoring constraints in managing high object density.

Conclusion

The execution of YOLO v3 on a VisDrone dataset sample effectively mirrored the methods outlined in the relevant research. The investigation, conducted in Python, validated the efficacy of YOLO v3 in daylight circumstances, as demonstrated by subplots illustrating precise detections of pedestrians, automobiles, and bicycles. Nonetheless, performance was constrained in low-light and congested environments. The bespoke CNN model, attaining 97% accuracy, illustrated the capability of tailored models to improve detection in drone footage. Future endeavors may investigate advanced YOLO iterations, such as YOLOv5 or YOLOv7, or integrate methodologies like data augmentation to tackle difficulties in low-light and congested environments. This study highlights the intricacies of drone-based object detection and the significance of model tuning.

References

A. N and U. S V, "Drone Object Detection Using Deep Learning Algorithms," *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2021, pp. 1187-1192, doi: 10.1109/ICIRCA51532.2021.9544983.

VisDrone dataset: VisDrone Official Website

YOLO v3: Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv:1804.02767

VisDrone performance: Ultralytics YOLO Docs - VisDrone Dataset