CSE541

Computer Vision

Ahmedabad University

Project 1: Object detection techniques (in case of small objects) on AU Drone dataset

Group: 11

AU2140122 - Kushalkumar Suthar

AU2140023 - Rohit Rathi

AU2140064 - Dhruvin Prajapati

AU2140039 - Nish Parikh

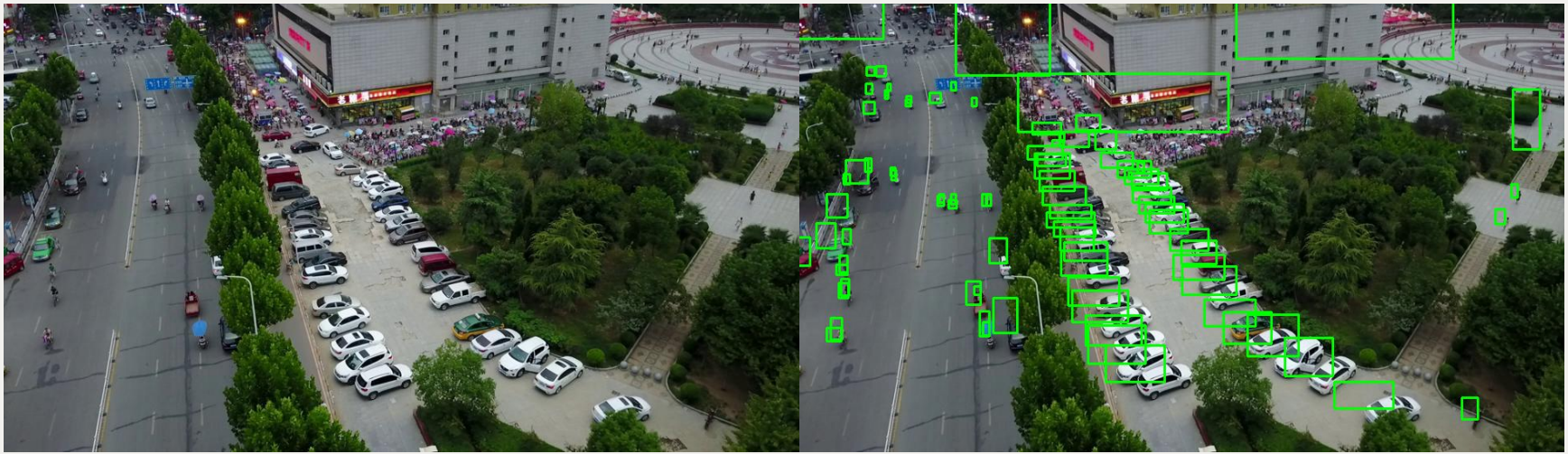
Data Set Explanation: Visdrone2019

There are eight entities:

1. Xc
2. Yc
3. W
4. H
5. Score
6. Object category ID
7. Truncation
8. Occlusion

Data Set Explanation: Visdrone2019

Xc	Yc	W	H	Score	Object ID	Truncation	Occlusion
684	8	273	116	0	0	0	0



Comparison:

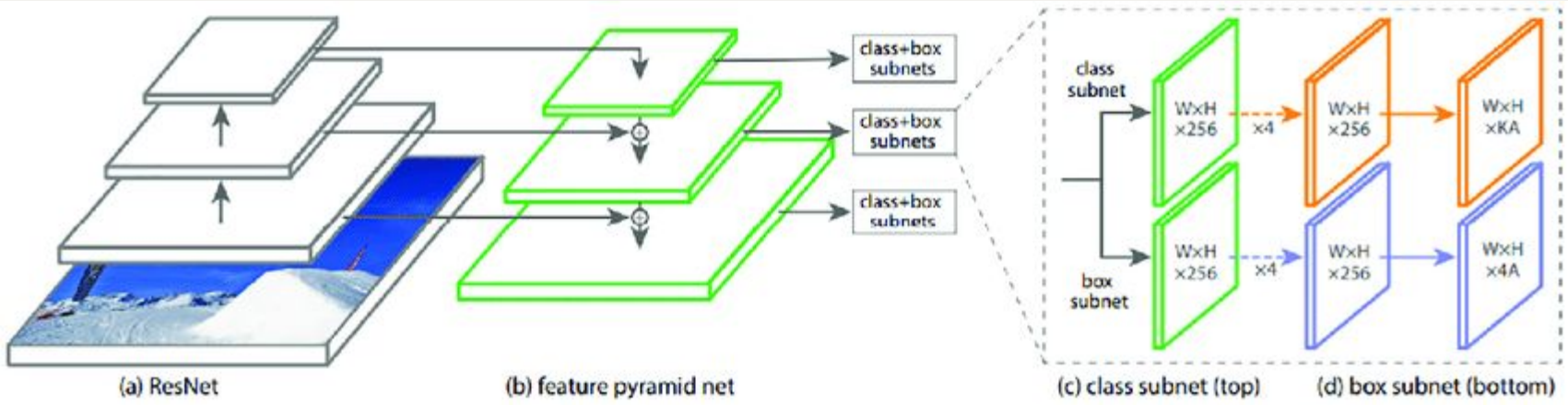
Method	AP[%]
QueryDet	33.91
RetinaNet	33.95

Author links open overlay panel Onur Can Koyun a et al., “Focus-and-detect: A small object detection framework for aerial images,” Signal Processing: Image Communication, https://www.sciencedirect.com/science/article/pii/S0923596522000273 (accessed Apr. 8, 2024).

Model Description:

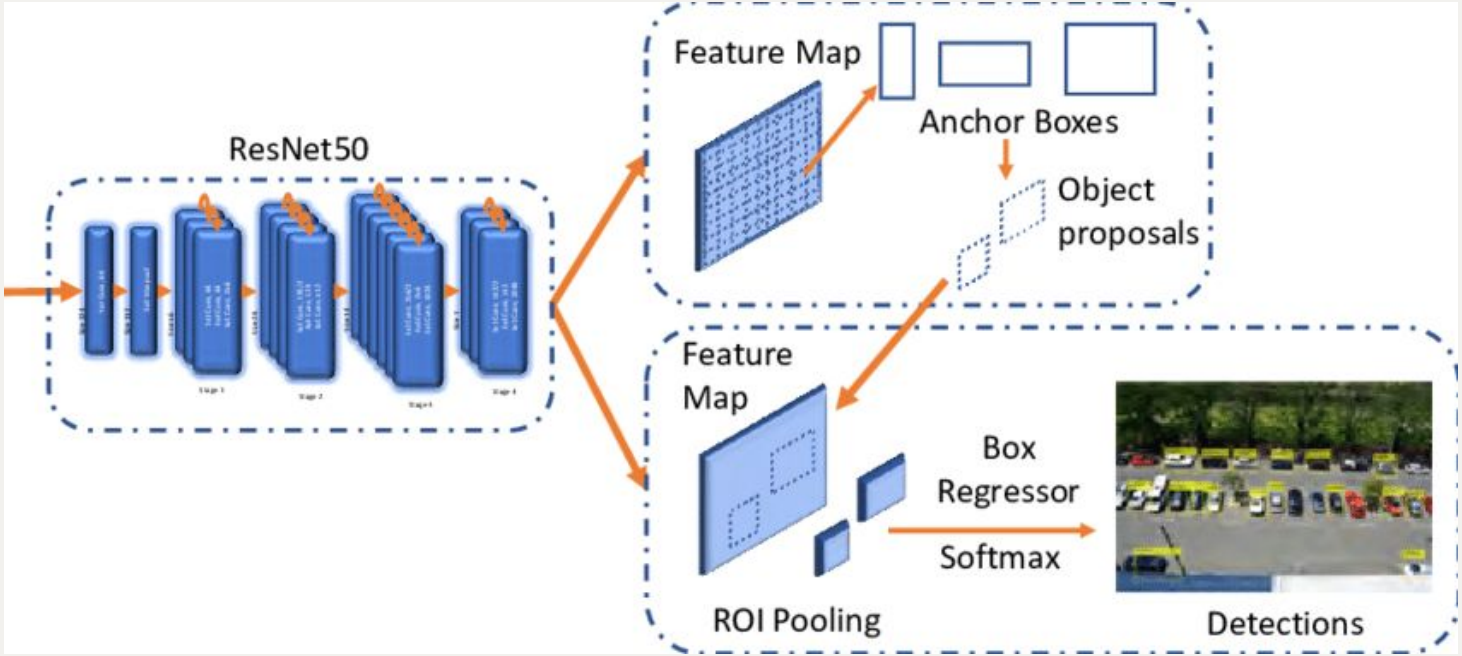
RetinaNet Architecture	Faster-RCNN Architecture
Single Stage detector	Two State detector
Backbone : ResNet50	Backbone : ResNet50
Focal Loss function	Cross Entropy loss function

RetinaNet Architecture



https://www.researchgate.net/figure/The-network-architecture-of-RetinaNet-RetinaNet-uses-the-Feature-Pyramid-Network-FPN_fig1_327737749

Faster RCNN Architecture



[1]“Architecture of Faster R-CNN with ResNeXt as the backbone,” *ResearchGate*. https://researchgate.net/figure/Architecture-of-Faster-R-CNN-with-ResNeXt-X-101-as-the-backbone_fig4_359721197

RetinaNet: on Visdrone-2019 dataset

Ahmedabad University

Faster-RCNN: on Visdrone-2019 dataset

Ahmedabad University

RetinaNet: on Visdrone-2019 dataset

Faster-RCNN: on Visdrone-2019 dataset

Results:

RetinaNet

Model	Backbone	AP[%]	AP50[%]	AP75[%]	AP95[%]
RetinaNet	ResNet-50	35.85	41.19	36.26	27.14

```
============RetinaNet============
||                              ||
|| Average Accuracy  = : 43.0165 ||
||                              ||
```

Faster-RCNN

Model	Backbone	AP[%]	AP50[%]	AP75[%]	AP95[%]
Faster -RCNN	ResNet-50	49.53	62.13	48.68	33.98

```
========Faster RCNN========
||                        ||
Average Accuracy  = : 74.6716
```

RetinaNet- on AU drone dataset

Ahmedabad University

Faster-RCNN- on AU drone dataset

Possible improvements:

- Training model over a dataset with different perspectives will help improve the accuracy
- Data Augmentation and Regularization
- Backbone Replacement or Modification
- Hyperparameter Tuning
- Feature Pyramid Networks (FPN):

  Loss functions can be tried and tested for better detection

References

Wikimedia Foundation. (2024, January 2). *Small object detection*. Wikipedia. https://en.wikipedia.org/wiki/Small_object_detection

Author links open overlay panelOnur Can Koyun a et al., “Focus-and-detect: A small object detection framework for aerial images,” Signal Processing: Image Communication, https://www.sciencedirect.com/science/article/pii/S0923596522000273 (accessed Apr. 8, 2024).

Chenhongyi Yang, Zehao Huang, and Naiyan Wang. QueryDet: Cascaded sparse query for accelerating high resolution small object detection. arXiv preprint arXiv:2103.09136, 2021.

Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision, pages 2980–2988, 2017.

Jingtao Xu, Yali Li, and Shengjin Wang. AdaZoom: Adaptive zoom network for multi-scale object detection in large scenes. arXiv preprint arXiv:2106.10409, 2021.

VisDrone, “VisDrone/Visdrone-Dataset: The dataset for drone based detection and tracking is released, including both image/video, and annotations.,” GitHub, https://github.com/VisDrone/VisDrone-Dataset (accessed Apr. 8, 2024).

Thank you