Comparative study of small objects detection using Tiny-ML and the DOTA database

Jonas Vilasboas Moreira

***Abstract*—The abstract goes here.**

***Index Terms*—IEEE, IEEEtran, journal, LATEX, paper, template.**

I. Introduction

DETECTING small objects in images is a significant challenge in computer vision due to their low resolution and high risk of overlapping with the background. Applications such as recognizing vehicles, people, and other objects in aerial, urban, or satellite images often deal with small objects, which require models capable of capturing details at small scales. Evaluating and comparing detection models specific to small objects is essential for advances in areas such as surveillance, environmental monitoring, and public safety.

II. Literature Review

Models such as BBAVectors, RVSA, and ReDet have been developed to address the challenges of small object detection, offering different approaches to improve accuracy. BBAVectors uses vector-based representations to define the orientations of objects, which is especially useful when objects are small and rotated. RVSA employs a Transformer-based visual attention mechanism to capture complex spatial relationships, while ReDet focuses on rotation-invariant convolutions, ideal for detecting objects with different orientations. Such techniques are essential for detecting objects with a small area that are easily confused with the background in complex images.

III. Comparison Proposal

The goal of this project is to compare the performance of three models focused on detecting small objects: BBAVectors, RVSA and ReDet. The proposal is to evaluate how each of these models performs in the task of identifying small objects in images, using the DOTA dataset (Dataset for Object Detection in Aerial Images), widely adopted and supported by all the chosen models.

Metrics such as average precision (mAP) for small objects, inference time and memory usage will be considered.

IV. Methodology

Three models focused on small object detection will be evaluated: BBAVectors [1], RVSA [2] [3] and ReDet [4]. The BBAVectors model uses vector representations to define the orientation of objects, which makes it especially effective in scenarios with varying rotations. RVSA adopts an architecture based on visual attention with Transformers, designed to capture complex spatial relationships. ReDet incorporates rotationally invariant convolutions, which makes it robust to detections with different angles and scales.

The database used will be version 1.0 of DOTA (Dataset for Object Detection in Aerial Images) [5], which contains high-resolution images with detailed annotations of objects in different sizes, classes and orientations. DOTA is widely recognized and adopted as a reference for evaluating small object detection models, and is fully compatible with the three models proposed for analysis.

The models will be evaluated based on three main metrics: average precision (mAP) for small objects, which measures the quality of detection; inference time, which evaluates the efficiency of the model during execution; and memory usage, which indicates the consumption of computational resources during the detection process. These metrics will allow an objective comparison between models, considering both accuracy and practical performance.

V. Conclusion

The conclusion goes here.

References

[1] J. Yi, P. Wu, B. Liu, Q. Huang, H. Qu, and D. Metaxas, "Oriented object detection in aerial images with box boundary-aware vectors," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2021, pp. 2150–2159.

[2] D. Wang, Q. Zhang, Y. Xu, J. Zhang, B. Du, D. Tao, and L. Zhang, "Advancing plain vision transformer toward remote sensing foundation model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023.

[3] D. Wang, J. Zhang, B. Du, G.-S. Xia, and D. Tao, "An empirical study of remote sensing pretraining," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–20, 2023.

[4] J. Han, J. Ding, N. Xue, and G.-S. Xia, "Redet: A rotation-equivariant detector for aerial object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2021, pp. 2786–2795.

[5] G.-S. Xia, X. Bai, L. Zhang, S. Belongie, J. Luo, M. Datcu, and M. Pelillo, "Dota: A large-scale dataset for object detection in aerial images," 06 2018.

PLACE PHOTO HERE

Michael Shell Biography text here.

John Doe Biography text here.

Jane Doe Biography text here.