Small Object Detection on AU Drone Dataset

Jash Parikh [AU2140108], Kathan Dave [AU2140113], Paridhi Jain [AU2120226], Saahil Doshi [AU2140106]

***Abstract*—The challenge of detecting small objects in large scenes, such as those captured by drones, is a significant hurdle in computer vision, affecting applications in surveillance, environmental monitoring, and beyond. This research explores the enhancement of small object detection on the VisDrone dataset, which encompasses a comprehensive collection of aerial images. By studying the application of advanced deep learning models—YOLOv8, YOLOv5, and EfficientDet—this work highlights the effectiveness of model architecture and preprocessing techniques, such as image slicing, in improving detection rates. Notably, the integration of the SAHI slicing method with YOLOv5 demonstrates a remarkable improvement in performance, underlining the potential of image slicing in small object detection tasks. Challenges related to EfficientDet's annotation format conversion and preliminary results across models are discussed.**

***Index Terms*—Object detection, Aerial imagery, Small objects, Deep Learning, Visdrone 2019 detection, SAHI, EfficientDet**

I. Introduction

In the field of computer vision, especially for UAVs and drone detection of small objects aerial imaging remains a major problem. This challenge becomes more complicated with the natural difficulties in the contrast between small objects and the background, the large range in the sizes of the objects, and the complexities added on because of the occlusions and varying lighting conditions.[1] VisDrone dataset, a complete set of aerial images is a useful tool to support the evaluation of object detection models and to improve small objects detection models.

Among all, the YOLO (You Only Look Once) series of models, such as YOLOv5 and YOLOv8, have been gaining huge popularity for their efficiency and accuracy in object detection tasks. However, these models are quite appealing because of their simplicity, ease of training, and high-performance capabilities, which make them a preferred choice for both researchers and practitioners.[2] Nevertheless, the distinction of small objects is still a critical issue and calls for developing new approaches and approaches through which detection rates will improve.

A technique that makes use of SAHI (Slicing Aided Hyper Inference) is the most innovative approach applied to increasing the efficiency of object detection algorithms for large-scale and high-resolution image processing. By that way SAHI splits images into small parts, run the object detection on each little part and combines the outcomes back into one big picture.[3] Enabled by this approach is a more efficient use of computational resources and also an improved accuracy of detection by use of smart algorithms to merge the overlapping detection boxes at the stitching stage. The ease with which SAHI can be integrated with YOLO models, as shown in the studies, indicates its prospect to bring notable advances in the detection of objects, especially those which are small.[4]

This project centers on the usage of SAHI with YOLO models while looking into VisDrone dataset. Through incorporating SAHI with YOLO models, our goal is to explore the use of that technique to improve small object detection in the aerial image. While doing that, we aim to be a part of this process and help making it better to distinguish objects in drone images. It will assist in development of new drone applications such as surveillance, environmental monitoring and so on.

II. Methodology

A. Dataset Overview

The VisDrone dataset, comprising approximately 6000 training images, 500 validation images, and 1500 test images, serves as the foundation of this study. It presents a variety of aerial image resolutions and dense object placements, posing unique challenges for object detection algorithms.

B. Model Selection and Preprocessing

The selection of YOLOv8, YOLOv5, and EfficientDet as the focus of this study stems from their proven capabilities in object detection tasks. The integration of the SAHI slicing technique with YOLOv5 aims to assess its impact on enhancing small object detection within dense aerial images. For EfficientDet, we are in the process of converting the annotation data to the COCO format to facilitate its evaluation. However, results for EfficientDet are pending due to challenges encountered with this formatting transition.

III. Results

Fig. 1. Object detection by using yolov5 before slicing.

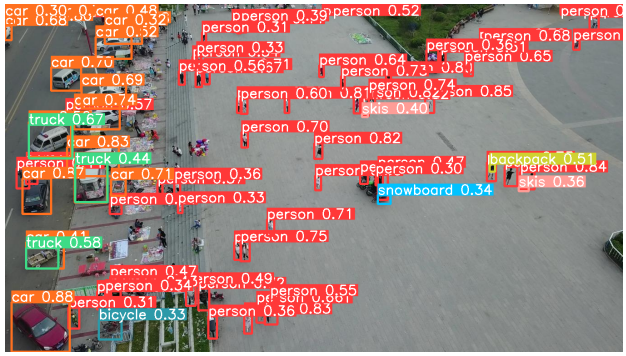

Fig. 2. Improved object detection by using yolov5 with slicing using Sahi(slicing aided hyper inference).

A. Impact of SAHI Slicing Applying SAHI slicing to YOLOv5 has led to a notable improvement in detection accuracy, showcasing the potential advantages of preprocessing techniques in this context. However, the study also highlights how lower-resolution images in the VisDrone dataset may result in blurred outcomes and possible false detections after the application of slicing.

B. Comparative Analysis While promising results have been observed with YOLOv8 and YOLOv5, the evaluation of EfficientDet has been delayed due to difficulties in converting the dataset annotations to the COCO format. A thorough comparison of these models, accounting for their performance with and without preprocessing techniques, is planned for future work.

IV. Discussions

This research emphasizes the significant role of preprocessing techniques, such as SAHI slicing, in enhancing small object detection in aerial imagery. However, the potential drawbacks, especially concerning lower-resolution images, necessitate a careful and tailored approach to their application. The ongoing challenges with EfficientDet's annotation format underscore the importance of flexible and adaptable preprocessing and evaluation strategies.

V. Conclusion

Our findings underscore the promise of combining advanced deep learning models with preprocessing methods to improve small object detection in the VisDrone dataset. Future research will explore the integration of further models with SAHI and seek solutions to the specific challenges posed by lower-resolution images. By refining these methods, we aim to contribute valuable insights and techniques to the field of small object detection in aerial and similar complex imaging scenarios.

VI. References

[1] Keleş, M. C., Salmanoğlu, B., Güzel, M. S., Gürsoy, B., Bostancı, G. E. (2023). Evaluation of YOLO Models with Sliced Inference for Small Object Detection. arXiv preprint arXiv:2203.04799. Retrieved from https://arxiv.org/pdf/2203.04799

[2] Ultralytics. (n.d.). Sliced Inference with SAHI. Retrieved from https://docs.ultralytics.com/guides/sahi-tiled-inference/

[3] OBSS. (n.d.). SAHI: Slicing Aided Hyper Inference. Retrieved from https://github.com/obss/sahi

[4] F. C. Akyon, S. Onur Altinuc and A. Temizel, "Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection," 2022 IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 2022, pp. 966-970, doi: 10.1109/ICIP46576.2022.9897990.