IDENTIFICATION/DETECTION OF NON-OBJECT AREAS FOR DATA MINING IN IMAGES USING YOLOV8

A Project Based Learning Report Submitted in partial fulfilment of the requirements for the award of the degree

of

Bachelor of Technology

MULTIMODAL INFORMATION PROCESSING

Course name with Course code

Submitted by

2310080022: JAGGAVARAPU AKHIL SAI REDDY

2310080007: MOHAMMED SHOAIB HANNAN

Under the guidance of

Dr. Gangamohan Paidi

Department of Electronics and Communication Engineering

Koneru Lakshmaiah Education Foundation, Aziz Nagar

Aziz Nagar – 500075

FEB - 2025.

Introduction

The rapid growth of artificial intelligence and computer vision has transformed the way digital images are analyzed and interpreted across industries. Traditionally, the majority of research in image processing and computer vision has been centered around object detection, where algorithms are trained to identify and classify visible entities such as humans, vehicles, animals, or specific items of interest. However, in many real-world scenarios, the non-object areas or background regions of an image hold equally significant value. These regions, though often ignored, can provide critical insights for applications such as environmental monitoring, medical imaging, satellite image analysis, and anomaly detection. For example, in satellite imagery, empty land, barren soil, or water bodies that do not contain man-made structures are essential for land use classification and environmental studies. Similarly, in manufacturing, the detection of irregular patterns in non-object surfaces may indicate defects.

With the advancement of deep learning, particularly convolutional neural networks (CNNs), new opportunities have emerged to utilize background information. Among the state-of-the-art models, YOLOv8 (You Only Look Once, version 8) stands out as a powerful and versatile detection framework. Unlike its predecessors, YOLOv8 offers enhanced speed, accuracy, and adaptability across tasks such as object detection, image segmentation, and classification. The strength of YOLOv8 lies in its ability to process images in real time, making it highly suitable for large-scale data mining tasks where efficiency is crucial.

The core idea of this project is to employ YOLOv8 not only for the detection of objects but also to explicitly identify and analyze non-object regions within images. These non-object areas can be segmented out once the objects are detected, thereby isolating the background for further data mining. By doing this, the project addresses a research gap: most existing detection systems overlook the hidden information in non-object areas, leading to incomplete understanding of the scene. Our approach, therefore, expands the traditional scope of object detection toward holistic image analysis, where both object and non-object features are studied.

This work has multiple practical applications. In remote sensing, background details such as vegetation patterns, deforestation, or water spread can be mined for environmental monitoring. In medical imaging, non-object regions may represent tissues, textures, or anomalies that require closer examination. In industrial inspection, the background surfaces of products may reveal flaws, cracks, or irregularities invisible to the naked eye. Furthermore, in autonomous systems, understanding non-object areas such as road boundaries, empty paths, or unsafe terrain is critical for navigation.

Thus, this project aims to develop a framework that integrates YOLOv8 with data mining techniques to systematically analyze non-object areas in images. The broader goal is to contribute toward knowledge discovery from images beyond object-centric approaches, unlocking valuable patterns and contextual information that have so far remained underutilized. By bridging object detection with non-object analysis, this project not only enhances the scope of computer vision but also provides a new dimension to data-driven decision-making in diverse application domains.

LITERATURE REVIEW

Fast R-CNN is used for object detection in images and videos, with applications in autonomous driving, surveillance, and medical imaging. It achieves 66.9% mAP on VOC 2007, 66.1% on VOC 2010, and 65.7% on VOC 2012, improving to 68.4% with extra data. While faster than older models, it is not real-time due to slow region proposals. Training is expensive, requiring powerful GPUs and long hours. The RoI pooling layer loses details, affecting small object detection and accuracy. Too many object proposals can confuse the model instead of improving detection. Multi-scale training is costly, requiring high memory and processing power. [8,11].It improves upon YOLOv3 by combining various tricks to enhance accuracy while maintaining speed. This model is useful for autonomous driving, surveillance, medical imaging, and retail automation. PP-YOLO achieves a mean Average Precision (mAP) of 45.2% on the MS COCO dataset, which is higher than YOLOv4's 43.5%. It also runs faster at 72.9 FPS, making it an efficient object detector. It still relies on anchor-based detection, which can be inefficient for objects of varying scales. Training is computationally expensive, requiring high-end GPUs.[5].

The YOLO framework is widely used for real-time object detection in fields like autonomous vehicles, healthcare, surveillance, agriculture, and industrial automation. The latest version, YOLOv11, improves both speed and accuracy, outperforming previous versions in terms of mean Average Precision (mAP) while maintaining real-time performance. However, YOLO still struggles with detecting small objects, handling complex environmental variations, and requires high computational resources for training and deployment. Despite these challenges, it remains one of the best choices for object detection due to its balance between speed and accuracy [10]. the use of Convolutional Neural Networks (CNNs) and YOLO for object detection in images and videos. Applications include autonomous vehicles, surveillance, medical imaging, and industrial automation, where fast and accurate detection is essential. The accuracy of Faster R-CNN reaches a mean Average Precision (mAP) of 76.4, but it is slower, while YOLO achieves an mAP of 78.6 with a speed of 155 FPS, making it much faster. However, YOLO struggles with detecting small objects and unusual aspect ratios and requires high computational resources [8,11]. YOLO-based object detection, which is widely used in autonomous driving, surveillance, medical imaging, agriculture, and industrial automation due to its fast inference speed. The model achieves a mean Average Precision (mAP) of 63.4% for YOLO and 70% for Fast R-CNN, but YOLO is about 300 times faster, making it preferable for real-time applications. However, YOLO struggles with detecting small objects, handling overlapping objects, and is sensitive to environmental variations [9].

DAMO-YOLO, a real-time object detection model with applications in autonomous driving, surveillance, healthcare, and industrial automation. It achieves mAP scores of 43.6% to 51.9% on COCO for different model scales while maintaining low latency on T4 GPUs. The lightweight versions for edge devices achieve 32.3% to 40.5% mAP with efficient processing on x86-CPU. However, DAMO-YOLO faces challenges such as high computational cost, difficulty detecting small objects, and reliance on complex neural architecture search (NAS) methods [4]. A novel object detection algorithm used in autonomous driving, surveillance, medical imaging, robotics, and industrial automation. It improves upon previous YOLO versions by incorporating EfficientNet-B4 as a backbone and NAS-FPN for better feature fusion. The model achieves an APAS (Average Precision Across Scales) score of 52.7 on the COCO dataset, outperforming YOLOv7's 50.3, while running at 150 FPS for real-time applications. However, YOLOv8 struggles with small object detection, occlusions, and computational demands, making it less suitable for low-power devices [1]. The ViT-YOLO model is designed for object detection in drone-captured images, making it useful for aerial surveillance, agriculture, delivery systems, and autonomous navigation. It improves upon previous YOLO versions by integrating multi-head self-attention (MHSA) and BiFPN, allowing for better small object detection and feature fusion. The model achieves a mean Average Precision (mAP) of 39.41 on the VisDrone-DET 2021 challenge, outperforming traditional CNN-based YOLO models [6].

Its various versions, which are widely used for real-time object detection in autonomous driving, surveillance, medical imaging, agriculture, and industrial automation. The accuracy of YOLO has improved across versions, with YOLOv4 achieving an mAP of 43.5% and YOLOv5 improving detection speed and efficiency. However, YOLO struggles with detecting small and overlapping objects, is sensitive to lighting variations, and requires high computational power, making deployment on low-end devices challenging [7]. YOLO with Adaptive Frame Control (AFC), which is used for real-time object detection in applications like autonomous driving, surveillance, embedded systems, and industrial automation. The proposed AFC-enhanced YOLO maintains high accuracy while optimizing frame control, ensuring real-time processing for network camera inputs. The model achieves consistent object detection without frame delays, improving real-time performance compared to standard YOLO. However, it still faces hardware dependency issues, high computational demands, and struggles with complex environments [3]. Image-Adaptive YOLO (IA-YOLO) for object detection in adverse weather conditions, with applications in autonomous driving, surveillance, and low-light environments. IA-YOLO enhances detection by using a differentiable image processing (DIP) module, improving accuracy in foggy and low-light conditions. The model achieves an mAP of 72.03% on VOC_Foggy and 37.08% on RTTS, outperforming traditional YOLOv3.It requires high computational power and struggles with extreme weather variations. IA-YOLO significantly enhances detection in challenging environments [2].

References

[1] J. Redmon and A. Farhadi, ”YOLO9000: Better, Faster, Stronger,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2017, pp. 7263–7271.https://doi.org/10.1109/ACCESS.2022.3146884

[2] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, ”YOLOv4: Op- timal Speed and Accuracy of Object Detection,” arXiv preprint arXiv:2004.10934, 2020.

[3] D. Barry et al., ”xYOLO: A Model for Real-Time Object Detection in Humanoid Soccer on Low-End Hardware,” in Proc. 2019 Int. Conf. Image Vis. Comput. New Zealand (IVCNZ), IEEE, 2019.

[4] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, ”YOLOv4: Op- timal Speed and Accuracy of Object Detection,” arXiv preprint arXiv:2004.10934, 2020. (duplicate entry, merged with [2]).

[5] F. Sultana, A. Sufian, and P. Dutta, ”A Review of Object Detection Mod- els Based on Convolutional Neural Network,” in Intelligent Computing: Image Processing Based Applications, Springer, 2020.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ”ImageNet Classification with Deep Convolutional Neural Networks,” in Advances in Neural Information Processing Systems (NeurIPS), 2012, pp. 1097–1105.

[7] S. Albelwi and A. Mahmood, ”A Framework for Designing the Archi- tectures of Deep Convolutional Neural Networks,” Entropy, vol. 19, no. 6, p. 242, 2017.

[8] T. Diwan, G. Anirudh, and J. V. Tembhurne, ”Object Detection Using YOLO: Challenges, Architectural Successors, Datasets and Applica- tions,” Multimed. Tools Appl., vol. 82, pp. 9243–9275, 2023.

[9] J. Carreira, R. Caseiro, J. Batista, and C. Sminchisescu, ”Semantic Segmentation with Second-Order Pooling,” in Proc. Eur. Conf. Comput. Vis. (ECCV), 2012.

[10] T. Diwan, G. Anirudh, and J. V. Tembhurne, ”Object Detection Using YOLO: Challenges, Architectural Successors, Datasets and Ap- plications,” Multimedia Tools and Applications, vol. 82, no. 7, pp. 9243–9275, 2023.

[11] . Carreira, R. Caseiro, J. Batista, and C. Sminchisescu, ”Semantic Segmentation with Second-Order Pooling,” in Proc. Eur. Conf. Comput. Vis. (ECCV), 2012.