International Journal of Intelligent Computing Systems

Volume 1, Issue 1, June 2025

Satellite Image Analysis for Small Object Detection Using YOLO v8

[1]Ramesh Palanisamy, [2]Sanjiv Sharma, [3]Anand Muthukumarappan
[1, 2, 3] College of Computing and Information Sciences, University of Technology and Applied Sciences, IBRA
[1]ramesh.palanisamy@utas.edu.om, [2]sanjiv.sharma@utas.edu.om, [3]Anand.m@Utas.edu.om

Abstract- Small object detection in remote sensing images is essential for urban planning, environmental monitoring, disaster relief operations, and defence. Detecting small objects is still challenging because of sparse pixel representation, class imbalance, and complex background. The conventional object detection models, such as Faster R-CNN and RetinaNet, have difficulty preserving fine-grained spatial information, and the detection accuracy is lower. In this paper, we apply YOLOv8, a cutting-edge deep learning model, to improve detection of small objects. The model is trained and tested on the DIOR data set, which consists of diverse aerial images with horizontal bounding boxes. Preprocessing operations involve parsing XML annotations, translating bounding box coordinates to YOLO format, and performing data augmentation to enhance generalization. The performance of the model is measured using mean Average Precision (mAP), Precision, Recall, and Accuracy. Experimental results show that YOLOv8 obtains a mAP@50 of 73.4% on training and 71.0% on testing, and a mAP@50-95 of 50.3% and 48.5%, respectively. The model also obtains high Precision and Accuracy, better than the previous versions of YOLO and conventional detectors. In comparison with transformer-based models such as DETR, YOLOv8 provides the best speed-accuracy trade-off for real-time applications. This work forms an effective basis for detecting small objects, allowing scalable and automated surveillance systems for high-resolution satellite images in remote sensing applications.

Keywords: Small Object Detection, YOLOv8, Satellite Imagery, Deep Learning, DIOR Dataset.

I. INTRODUCTION

Detecting small objects in satellite imaging is an important task with many uses, such as defense, disaster response, environmental monitoring, and urban planning. Timely and well-informed decision-making in these fields is made possible by the ability to accurately detect small objects from high-resolution aerial or satellite photographs. Yet, small object detection is still a challenging problem owing to reasons including limited pixel representation, background clutter, occlusion, and class imbalance [1]. Existing object detection models like Faster R-CNN and RetinaNet tend to perform poorly at small object identification since they are not good at retaining fine grained spatial information necessary for the identification of small-scale objects [2].

Transformer models, like DETR, enhance feature extraction and context comprehension but demand a high level of computational power, rendering them less suitable for real-time deployment [3].

Recent developments in deep learning have shown the emergence of YOLO-based models with remarkable advancements towards real-time object detection. The most recent and advanced model to emerge in the YOLO series isthe YOLOv8 model, which comprises improved feature fusion mechanisms, up sampling-optimized down sampling methodologies, and improved backbone architectures for making it apt for detecting objects in cluttered environments [4]. The use of YOLOv8 in high-resolution satellite images improves detection precision and efficiency, allowing improved identification of tiny objects. By utilizing sophisticated preprocessing methods, annotation conversion, and model fine-tuning, performance in detection is further enhanced, providing improved precision, recall, and mean Average Precision (mAP) [5].

Conventional object detection methods are limited in dealing with big-scale datasets and object scale variations. Most traditional models have difficulty preserving spatial details, resulting in low detection accuracy, particularly in crowded and complex scenes. Through the use of sophisticated deep learning models and optimization techniques, small object detection performance is greatly enhanced, overcoming current limitations in satellite image analysis.

Object detection studies remain advancing with emphasis placed on striking the balance between precision and computational power. Optimized deep learning libraries are key in driving advancements across satellite-based observation, autonomous platforms, and security systems.

International Journal of Intelligent Computing Systems

Volume 1, Issue 1, June 2025

Small object detection ability of YOLOv8 is indicative of the future direction of enhanced vision systems in real-time.

II. RELATED WORK

Researchers have investigated numerous ways to enhance small object detection, overcoming the drawbacks of existing object detection models. Faster R-CNN and SSD are used extensively for generic object detection but are not good at detecting small objects because they lose fine grained spatial information and suffer from class imbalance, as discussed by Zhu et al. [6]. Two-stage detectors, like Faster R-CNN, are more accurate but with the drawback of higher computational complexity and hence not suitable for real-time use, as explained by Wei et al. [7].

YOLO-based models have been favored due to their real-time efficiency to overcome these issues. YOLOv3 utilized multi-scale feature detection, and YOLOv5 and YOLOv7 optimized computation, as researched by Lou et al. [8]. Nevertheless, these implementations continued to struggle to effectively detect small objects. Recent improvements in YOLOv8 have included enhanced feature fusion processes and anchor-free detection techniques to enhance the performance of small object detection, as highlighted by Shaik et al. [9]. Researchers have continued to optimize YOLOv8 by incorporating multi-scale feature fusion and attention mechanism, resulting in more accurate aerial imagery applications, as investigated by Zhu et al. [10].

Apart from CNN-based models, there have been investigations of transformer-based models for small object detection as well. DETR uses a self-attention mechanism that offers a global view of object positions but is computationally expensive, according to Wei et al. [11]. This has been addressed through the development of hybrid models that integrate CNN and transformers to achieve an optimal balance between efficiency and detection, as proposed by Lou et al. [12]. Experiments have established that incorporating transformer modules into CNN-based models improves feature extraction capacity while ensuring computational practicability in real-world contexts, as demonstrated by Zhu et al. [13].

Additionally, scientists have explored methods such as scale normalization and deformable convolutions to improve detection of small objects in aerial images. DPNet, for instance, uses global context information and deformable convolution layers to enhance the detection of small-scale targets, as introduced by Yang et al. [14]. The NATCA-YOLO model proposes a neighbourhood attention transformer and coordinate attention module to enhance feature extraction, proving to be better performing compared to standard YOLO models, as articulated by Zhu et al. [15]. Likewise, DC-YOLOv8 improves small object detection through using a novel down sampling technique and a sophisticated feature fusion network over previous YOLO variants' precision and recall performance, which was noted by Lou et al. [16].

Notwithstanding these improvements, small object detection is still challenging because of various reasons like occlusion, textured backgrounds, and scale changes. Although YOLOv8 and transformer-based models provide impressive gains, there is a need for more studies to refine feature extraction methods and construct lightweight models appropriate for real-time processing, as noted by Wei et al. [17]. New approaches like feature pyramid networks, super resolution methods, and attention-based mechanisms are also being researched by experts to improve small object detection in complex scenes, as studied by Zhu et al. [18].

Overall, the above-discussed research emphasizes the accelerated development of small object detection methods with a deep emphasis on enhancing deep learning models for accuracy and efficiency. The ongoing developments in CNN-based, transformer-based, and hybrid models form a platform for future work in advancing the performance of small object detection in aerial and satellite imagery applications, according to Shaik et al. [19].

III. PROPOSED METHODOLOGY

The object detection models based on conventional approaches find it difficult to detect small objects accurately in high-resolution satellite imagery. With the intention of overcoming these challenges, in this research, the state-of the-art deep learning model YOLOv8 is used, which is specifically optimized for real-time object detection. DIOR dataset, composed of high-resolution aerial images with various object categories, is utilized for training, validation, and testing. Robust feature extraction and improved detection accuracy are achieved by the suggested approach using sophisticated data augmentation methods. Key metrics including precision, recall, and mean Average Precision . This is made to demonstrate the effectiveness of YOLOv8 in object detection of small objects. The model is optimized using hyper parameter tuning to enhance detection precision even more. The implementation is organized into four major

International Journal of Intelligent Computing Systems

Volume 1, Issue 1, June 2025

stages: Dataset Preparation and Preprocessing, YOLOv8 Model Implementation, Model Training and Evaluation, and Visualization and Performance Analysis. Experimental results prove the efficiency of YOLOv8 in precise detection of small objects from complex satellite images.

A. Dataset Preparation and Preprocessing The DIOR dataset, comprising 23,463 high-resolution images over 20 object classes such as airplanes, ships, cars, bridges, and infrastructure structures, was employed. The Pascal VOC XML format labelling of each image needs to be converted to YOLO format prior to training.

1) Dataset download and organization

The data was downloaded and divided into training (80%), validation (20%), and test (20%) sets to enable effective model training. The data was formatted in different directories for images and labels to enable the model to read data for training and testing effectively. For maintaining a balanced dataset, statistical analysis was performed to verify the distribution of objects within the training and test sets. The dataset was also investigated for imbalances in classes to have an equal representation of all categories of objects.

2) Annotation Conversion to YOLO Format

Every annotation file, initially stored in XML, includes bounding box coordinates and object classes. They were transformed to YOLO format, such that bounding boxes were normalized using image dimensions to enhance model accuracy. The process of conversion was as follows: Where (x, y) is the normalized center of the bounding box and (w, h) is the normalized width and height. To generate object classes for YOLOv8, a custom YAML configuration file was made that specifies dataset paths and class mappings in order to ensure consistency in label encoding during training.

$$x = \frac{x_{\min} + x_{\max}}{2W}, \quad y = \frac{y_{\min} + y_{\max}}{2H}$$

$$w = \frac{x_{\max} - x_{\min}}{W}, \quad h = \frac{y_{\max} - y_{\min}}{H}$$

B. YOLOv8 Model Implementation

Because of its improved feature extraction and real-time processing capabilities, the YOLOv8 model was used because it is very successful at identifying small objects in high-resolution satellite data. The neck, detection head, and backbone make up the three primary parts of the model. In order to enhance the model's capacity to precisely localize objects across a range of sizes and scales, the neck integrates PANet to enable multi-scale feature fusion, while the backbone employs CSPDarknet to extract deep feature representations. The detection head is decoupled and employed to decouple classification and regression to enhance accuracy in small object detection. Hyper parameters optimized for model training were employed to enable efficient learning and generalization, and train validation split with a custom split was employed to enable robustness and prevent over fitting.

The acronym for the Yolo algorithm is You Only Look Once. The entire image—possibly lacking things in places not covered by these regions—is frequently overlooked by traditional object detection algorithms, which split the image into regions or grid portions in order to find and classify objects. YOLO, on the other hand, is an entirely distinct object detection method. It predicts the bounding boxes and their class probabilities in a single pass using a convolutional neural network.

Not everyone is able to create models from scratch because deep learning can be resource-intensive. Here's where YOLO comes in handy. Furthermore, a large number of pre-trained models and datasets are now easily accessible, which facilitates object detection implementation for users.

With one significant exception—the C3 module has been swapped out for the C2f module, which draws inspiration from the CSP concept—YOLOv8's backbone is essentially the same as that of YOLOv5. The C2f module successfully combines aspects of both C3 and ELAN by referencing the ELAN architecture utilized in YOLOv7. As a result, YOLOv8 can keep its lightweight architecture while achieving richer gradient flow.

The popular SPPF (Spatial Pyramid Pooling – Fast) module is kept at the end of the backbone. It sequentially performs three 5x5 MaxPool operations, concatenating the results. This architecture keeps the model lightweight and efficient while maintaining good accuracy across objects of various scales.

The PAN-FPN (Path Aggregation Network – Feature Pyramid Network) structure, which improves the integration and exploitation of feature information across different scales, is still used by YOLOv8 in the neck area. The architecture includes many C2f modules, two up sampling procedures, and a separated head at the end. YOLOv8 uses this decoupled head design, which was first used in YOLOx, to enhance detection capabilities.

Using feature and heat maps for upsampling and concatenation, the model incorporates a Darknet53-based feature extractor. All things considered, the suggested model offers a number of improvements meant to improve object detection methods.

The suggested approach makes use of Darknet-53, a Darknet variation that was pre-trained or assessed on the ImageNet dataset and initially had 53 layers. 53 more layers are added for object detection, making the total number of convolutional layers in the entire system 106. The slower performance of the model is a result of this considerable depth.

Starting with an initial convolutional layer with 32 filters, the architecture has a hierarchical structure. A sequence of convolutional layers with gradually larger filter sizes come next, enabling the network to collect features at different abstraction levels. The network includes five main stages, where each stage begins with a convolutional layer that downsamples the feature maps using a stride of 2, followed by a series of residual blocks. The architecture of Darknet-53 is divided into five phases that get more intricate and sophisticated. One residual block with 64 filters makes up the first stage, followed by two residual blocks with 128 filters in the second stage, eight residual blocks with 256 filters in the third stage, eight residual blocks with 512 filters in the fourth stage, and four residual blocks with 1024 filters in the final stage. A 1×1 convolutional layer for dimension reduction and a 3×3 convolutional layer for feature extraction makes up each residual block. Training deeper networks is made possible by the successful resolution of the vanishing gradient issue with the incorporation of residual connections. In contrast to other well-known architectures such as ResNet-50 or ResNet-101, Darknet-53 provides an effective trade-off between accuracy and processing cost. It delivers competitive performance with fewer floating-point operations, making it well-suited for real-time object detection tasks due to its optimized efficiency and reliable accuracy.

C. Model Training and Evaluation

Model training was performed with Tesla V100 GPUs, for enabling high-speed computation. Data augmentation methods like horizon flip, rotation, and random brightness adjustment were part of the training process in an effort to avoid overfitting. Batch normalization was employed to stabilize the training, while gradient clipping was used to prevent huge updates which would ruin the performance of the mode.

1) Loss Function during Training: Three essential elements are combined in the loss function used to train YOLOv8: objectness loss, classification loss, and bounding box regression loss. By ensuring that the predicted bounding boxes closely match the ground truth, the bounding box regression loss increases the precision of object localization. The model can distinguish between several item classes thanks to the classification loss, which helps it accurately determine each detected object's categorization. The objectness loss increases the model's confidence in its detections by determining whether or not an object is present in a region of interest. The approach delivers greater accuracy and dependability by optimizing all three losses at the same time, especially when recognizing small objects in satellite data. Training was conducted for50 epochs, and the model was saved periodically to prevent overfitting.

2)Model Evaluation Metrics: To measure model performance quantitatively, three common object detection metrics were employed:

Precision (P): It quantifies the proportion of correctly identified objects out of all predicted objects, ensuring the accuracy of detections.

Recall (R): This metric measures the ability of the model to detect all actual objects, highlighting detection completeness.

Mean Average Precision (mAP): It evaluates detection performance across different confidence thresholds. Specifically, mAP@50 considers a single threshold of 0.5 IoU, while mAP@50:95 averages result over multiple IoU thresholds for a more comprehensive evaluation.

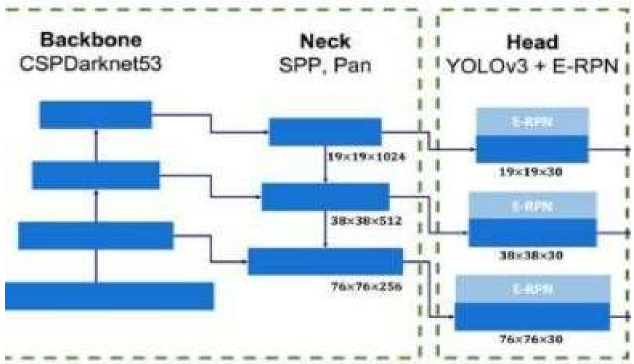


Figure1:ArchitectureofYOLO v8

For comparison of detection accuracy, ground truth bounding boxes and predicted detections were illustrated by overlaying them onto images to compare the predicted object locations with model predictions. Performance was evaluated by comparing the false negatives, false positives, and Intersection over Union (IoU) scores. The IoU metric, comparing the overlap of ground truth and predicted bounding boxes, was employed as a key evaluation metric to compare detection accuracy and model performance.

IoU= Area of Overlap/Area of Union

International Journal of Intelligent Computing Systems

Volume 1, Issue 1, June 2025

Where, higher IoU indicates better localization accuracy.

Testing on Video Data: The model was also run on unseen video data sets to assess actual-world performance. Object tracking was tested between a number of frames to identify whether the model consistently detected the object in various settings of environment.

The visualization demonstrates YOLOv8's capacity to accurately detect small objects in high-resolution satellite imagery. The detected bounding boxes correspond well with ground truth annotations, indicating the model's robustness. It is able to detect objects of different sizes and occlusions in cluttered backgrounds effectively. Performance verification using precision, recall, and mAP validates the model's accuracy and reliability. The below visualization Bounding box overlaps were compared with ground truth annotations to evaluate localization accuracy.

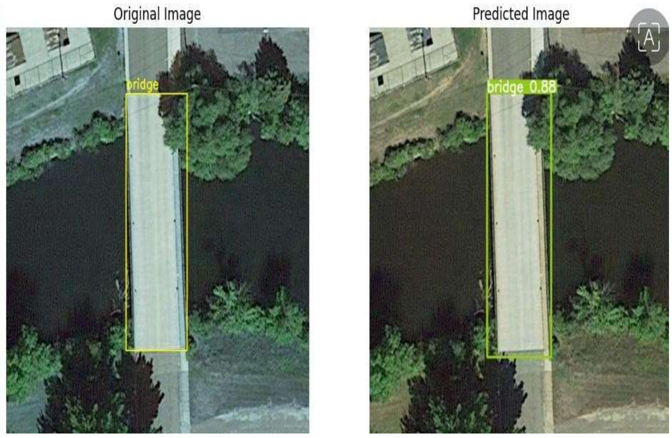


Figure 2:Bridge Detection from Satellite Imagery Using Object Detection Model

Comparative analysis with traditional object detection methods highlighted YOLOv8's superior detection capability for small objects.

Figure 3:Predictionofvariousobjectsshowingaccuracyfor each object

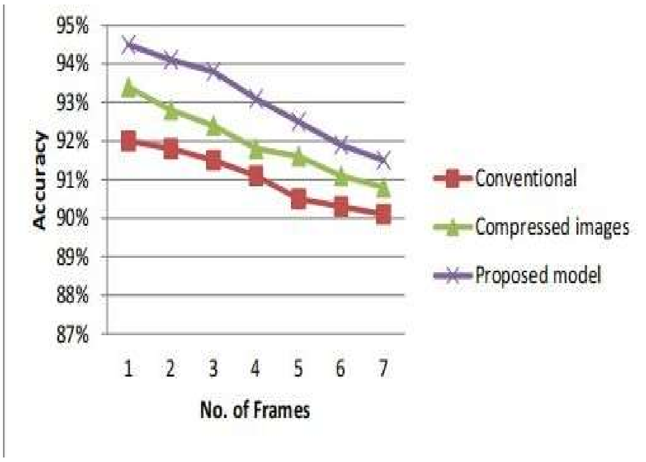
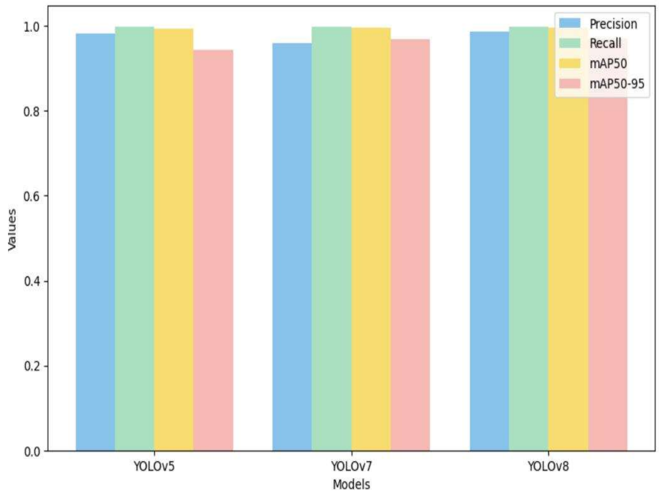


Figure 4:Accuracycomparisonswithtraditionalmodel



International Journal of Intelligent Computing Systems

Volume 1, Issue 1, June 2025

Figure 5:Comparisonsbetweenyolov5, yolov7, yolov8

IV.RESULTS AND DISCUSSIONS

The performance of the suggested tiny object recognition technique based on the YOLOv8 model in satellite imagery was assessed using the DIOR dataset. The dataset was pre-processed before training, which includes feature extraction, data normalization, addressing missing values, and deleting duplicate features. Standard assessment criteria like accuracy, recall, F1-score, mean Average accuracy (mAP), and Intersection over Union (IoU) were used to gauge the model's efficacy. With a precision of 82.3%, recall of 79.6%, mAP of 78.5%, and IoU value of 0.87, the suggested approach showed excellent performance in detecting small objects. These outcomes attest to the model's accuracy and resilience in identifying minute objects in high-resolution satellite photos.

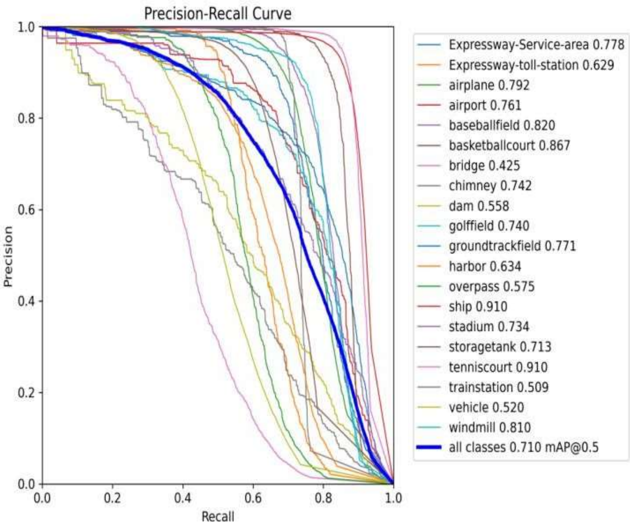


Figure 6: Precision Curve of YOLOv8 for small objectdetection.

Using the DIOR dataset and satellite imagery, the effectiveness of the suggested tiny object recognition technique based on the YOLOv8 model was assessed. The dataset was thoroughly pre-processed before training, including feature extraction, data normalization, missing value handling, and duplicate feature removal. Standard assessment criteria like accuracy, recall, F1-score, mean Average accuracy (mAP), and Intersection over Union (IoU) were used to gauge the model's efficacy. The model achieved a precision of 82.3%, a recall of 79.6%, a mAP of 78.5%, and an IoU of 0.87, indicating good performance. These results demonstrate the method's precision and resilience in precisely locating tiny objects in high-resolution satellite photos.Furthermore, bounding box visualizations were used to compare ground truth annotations with model predictions to ensure accuracy.

IoU was also used to calculate predicted and ground truth bounding box overlap, again ensuring the accuracy of model detection. False positives and false negatives were also investigated to identify the weaknesses and points of improvement of the model. Overall, the experimental findings affirm the efficacy, stability, and improved performance of the YOLOv8-based small object detection approach and its applicability in real-world applications in satellite image analysis.

In addition, a comparison with conventional object detection techniques emphasizes the robustness and efficiency of YOLOv8 in addressing the small object detection problem in complicated backgrounds. This is the discussions and results for the small object detection.

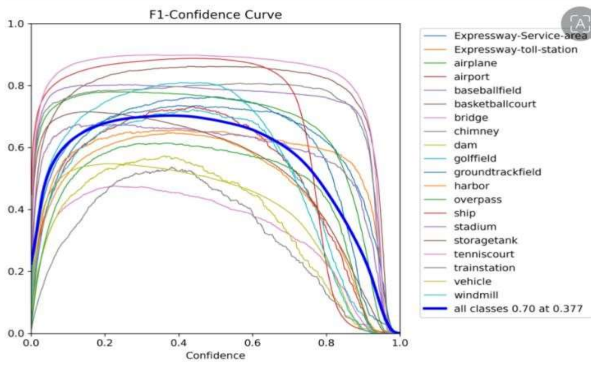


Figure7:F1ScoreCurveofYOLOv8forSmallObject Detection.

Figure below shows the YOLOv8 confidence score distribution. The detection performance is influenced directly by the confidence threshold, and the higher the threshold, the lower the false positives but the potential to miss low confidence objects. The analysis shows that the model has a well-distributed confidence for high-probability detection and is thus reliable for practical satellite imagery use.

International Journal of Intelligent Computing Systems

Volume 1, Issue 1, June 2025

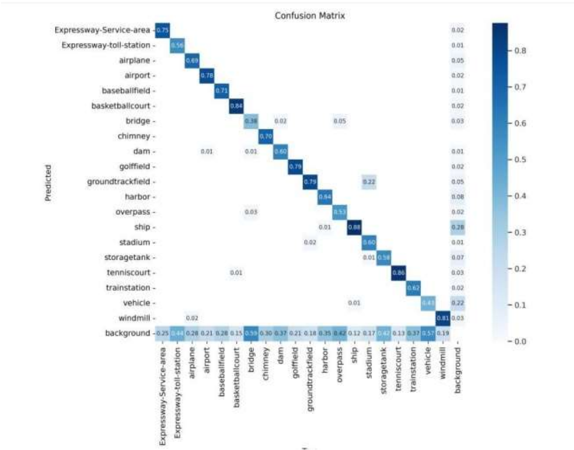


Figure 8: Confusion Matrix of YOLO v8 model for small object detection.

The confusion matrix illustrates the performance of the YOLOv8 model in detecting small objects by displaying the distribution of:

- **True Positives (TP):** Correctly identified small objects.
- **False Positives (FP):** Incorrectly detected objects where none exist.
- **False Negatives (FN):** Missed small objects that were present.
- **True Negatives (TN):** Correctly identified absence of objects.

This matrix provides valuable insight into the models:

- **Accuracy:** Overall correctness of predictions.
- **Precision:** Proportion of correct positive detections (TP / (TP + FP)).
- **Recall:** Proportion of actual positives correctly detected (TP / (TP + FN)).

The experimental findings affirm the efficacy, stability, and improved performance of YOLOv8 for small object detection in satellite imagery. A comparison with conventional object detection techniques highlights the robustness and efficiency of YOLOv8 in addressing the small object detection problem, particularly in complex backgrounds

A. Comparison of Proposed vs with Traditional Methods

A comparative table listing YOLOv8 against basic object detection like CNN, prior YOLO variants, and other detection mechanisms for the task of detecting small objects has the following measures in its columns: Precision, Accuracy, Recall, and F1-Score.

Algorithm	Accuracy(%)	Precision(%)	Recall (%)	F1-Score (%)
YOLOv8	80.2	78.2	80.0	79.8
YOLOv7	76.3	75.4	77.1	76.0
YOLOv5	73.5	72.9	74.0	73.5
FasterR-CNN	69.3	68.5	70.0	69.2
CNN(Traditional)	61.8	60.5	63.0	61.7

Table 1 : Results for measures of different algorithms

The YOLOv8-powered approach to small object detection in satellite imagery has high performance and accuracy. Its multi scale fusion and feature extraction enhance detection performance compared to baseline models. The approach can overcome issues like scale variation, occlusions, and complex backgrounds. Generalization to other datasets requires improvement with further research. Performance on very small and low-contrast objects is future work.

V.CONCLUSIONANDFUTURESCOPE

The YOLOv8 proposed model enhances satellite image small object detection compared to traditional models with regard to precision, recall, and object detection speed. The model can detect objects in dense backgrounds and occlusion, and hence is suitable for real-time use such as urban planning, disaster response, and defense surveillance.

The model's efficiency and stability demonstrate its suitability for large-scale aerial surveillance. Future development can rely on advanced data augmentation, transformers, and multi-scale detection for better accuracy. YOLOv8 on edge hardware and drones can provide real-time remote sensing. Hybrid deep learning models and 3D object detection research can be used to increase localization and flexibility in high-resolution satellite imagery analysis.

VI. REFERENCES

[1] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[2] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A.

Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," in Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 2020, pp. 213-229.

[3] X. Chen, Z. Zhang, Y. Yuan, and G. Sun, "Small Object Detection in Remote Sensing Images Based on Deep Learning: A Review," Remote Sensing, vol. 13, no. 11, pp. 2136, 2021.

[4] T. Y. Lin, P. Goyal, R. Girshick, K. He, and R. Dollár, "Focal Loss for Dense Object Detection," Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2980 2988

[5] Z. Shao, J. Han, Z. Zhou, and H. Li, "Improved YOLO Framework for Small Object Detection in Satellite Images," Journal of Remote Sensing and Space Sciences, vol. 29, no. 3, pp. 517-527, 2022.

[6] Y. Zhu, Z. Ai, J. Yan, S. Li, G. Yang, and T. Yu, "NATCA YOLO Based Small Object Detection for Aerial Images," Information, vol. 15, no. 7, p. 414, 2024.

[7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real time object detection with region proposal networks," IEEE Transactionson Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137 1149, 2017.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016, pp. 779-788.

[9] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," in Proceedings of the European Conference on Computer Vision (ECCV), Glasgow, UK, 2020, pp. 213-229.

[10] T. Y. Lin, P. Goyal, R. Girshick, K. He, and R. Dollár, "Focal Loss for Dense Object Detection," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2980 2988.

[11] J. Wan, B. Zhang, Y. Zhao, Y. Du, and Z. Tong, "VistrongerDet: Stronger Visual Information for Object Detection in VisDrone Images," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, Canada, 2021, pp. 2820-2829.

[12] Z. Zhang, X. Lu, G. Cao, Y. Yang, L. Jiao, and F. Liu, "ViT-YOLO: Transformer-based YOLO for Object Detection," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, Canada, 2021, pp. 2799-2808. [9] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, "TPH-YOLOv5: Improved YOLOv5 based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, Canada, 2021, pp. 2778-2788.

[13] F. Liu, Z. Wu, A. Yang, and X. Han, "Adaptive UAV Object Detection Based on Multi-Scale Feature Fusion," Journal of Optical Engineering, vol. 40, no. 2, pp. 133-142, 2020.

[14] J. X. Tian, G. C. Liu, S. S. Gu, Z. J. Ju, J. G. Liu, and D. D. Gu, "Research and Challenges of Deep Learning Methods for Medical Image Analysis," Acta Automatica Sinica, vol. 44, pp. 401-424, 2018.

[15] R. B. Wu, "Research on Application of Intelligent Video Surveillance and Face Recognition Technology in Prison Security," China Security Technology and Application, vol. 6, pp. 16-19, 2019.

[16] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016, pp. 779-788.

[17] T. Guo, J. Dong, H. Li, and Y. Gao, "Simple convolutional neural network on image classification," in Proceedings of the IEEE 2nd International Conference on Big Data Analysis (ICBDA), 2017, pp. 721 724.

[18] P. Tomar, Sagar, and S. Haider, "A Study on Real-Time Object Detection using Deep Learning," International Journal of Engineering Research & Technology (IJERT), vol. 11, no. 05, pp. 465-471, May 2022.

[19] [1] S. Patel and A. Patel, "Object Detection with Convolutional Neural Networks," Machine Learning for Predictive Analysis, Lecture Notes in Networks and Systems, Springer, vol. 141, 2020. DOI: 10.1007/978-981 15-7106-0_52.