**Autonomous detection and segmentation of wind turbine blades using YOLOv8**

**Pavan Sai Prasanth Sabnaveesu and Dr. Md Monirul Islam**

**Keywords**: Autonomous detection and segmentation, wind turbine blades, Mask R-CNN, YOLOv7, And YOLOv8

**Abstract**

This study delves into sustainable electricity generation, focusing on wind turbines' pivotal role in meeting escalating power demands via renewable energy sources. Vigilant monitoring of turbine blades exposed to environmental conditions ensures uninterrupted power supply. Employing cutting-edge object detection and segmentation algorithms such as Mask R-CNN, YOLOv7, and YOLOv8 enhances operational efficiency. Extensive drone-based data acquisition yields a diverse dataset of 2349 augmented images. Through thorough analysis, YOLOv8 emerges as the top performer in wind turbine fault detection and segmentation, consistently demonstrating excellence across varying thresholds. Notably, YOLOv8 achieves a detection mAP@IoU (0.7) of 93.7% and a segmentation mAP@mask (0.8) of 97.60%, highlighting the significant contribution of this research to the evolution of renewable energy infrastructure and paving the way for more reliable and sustainable wind power solutions.

**Introduction**

Fossil fuel-based energy sources constitute the primary contributors to greenhouse gas emissions. To effectively mitigate these emissions and address pressing environmental concerns, the integration of renewable energy sources into our daily lives is imperative. The Paris Agreement, signed by 165 nations in 2015, marked a substantial milestone, fortifying the global commitment to curbing carbon dioxide emissions and boosting investments in renewable energy solutions [1]. Following the agreement, there was a substantial increase in expenditure for green energy infrastructure. Notably, solar and wind energy emerged as leaders in the competition for renewable energy production. According to the US Energy Information Administration [2], solar power contributes 9%, and wind production contributes 24% to the nation's energy supply. Consequently, advancements in technology have revolutionized the manufacturing and design of wind turbines, resulting in a reduction in the cost of power generation when compared to other renewable energy sources. The efficiency of a wind turbine is double that of solar panels [3]. As of now, wind energy stands out as one of the most cost-effective renewable resources, with production costs averaging between one and two cents per kilowatt-hour [4]. Additionally, it is worth noting that utilizing one megawatt of wind energy can potentially offset approximately 2600 tons of carbon dioxide emissions [5].

Wind turbines, whether built onshore or offshore, harness wind energy through a combination of mechanical and electrical components, including blades, a rotor, a generator, a controller, and a gearbox. The blades, reaching impressive lengths of 100-140 meters, are crucial for electricity generation as their efficient rotation directly affects the amount of power produced. Adding to the complexity of inspection, wind turbine blades endure severe operational conditions. They are exposed to extreme weather phenomena and subjected to substantial aerodynamic and gravitational forces, resulting in pronounced vibrations. These vibrations, in turn, can precipitate a spectrum of structural anomalies, such as surface cracks and edge erosion [4][5]. The existing approaches for inspecting these blades, which rely heavily on scheduled maintenance, are notably inefficient and expose personnel to considerable safety hazards. For instance, rope-based inspections exemplify this inefficiency and the associated safety risks posed to personnel. Additionally, human limitations inherent to visual and manual techniques, like telephotography, can cause microscopic structural damage. Moreover, neglecting these issues isn't just a matter of lost efficiency; it can also pose environmental hazards if faults go undetected. This traditional approach, heavily dependent on human involvement, is becoming increasingly impractical as wind farms grow and complexity and are also prone to human error [27].

As wind turbine technology advances, so does the need for robust fault detection and segmentation methods to ensure optimal performance and mitigate potential failures. It is necessary for the adoption of deep-learning approaches for efficient and precise fault detection and segmentation. Mask R-CNN, YOLOv7, and YOLOv8 are advanced object detection models capable of both bounding box detection and instance segmentation. So, this research addresses these challenges by proposing an autonomous detection system using Mask R CNN, YOLOv7, and YOLOv8 algorithms. The goal is to precisely identify faults, thereby reducing O&M costs while elevating safety standards. This approach leans on innovative object detection algorithms, namely Mask R-CNN, YOLOV7, and YOLOV8, to navigate the challenges of manual inspections. First, let's start with contributions and then data acquisition.

**Contributions**

* This project made a significant contribution by acquiring an extensive dataset of turbine blade images using drones. These images were captured across diverse backgrounds to ensure the robustness of the model
* YOLOv8, a cutting-edge multi-stage detector, was chosen for its exceptional accuracy (mAP), speed, and efficient training. Its strong masking capabilities enabled precise blade segmentation, while hyper-tuning Intersection of Union thresholds further optimized performance
* Implemented and investigated the performance of three advanced deep learning algorithms for wind turbine blade fault detection and segmentation: Mask R-CNN, YOLOv7, and YOLOv8

**Data Acquisition**

In this research, created a dataset of **2349** images captured by drones and subsequently augmented to increase the data size and improve model performance. These images capture four distinct categories: edge erosion, surface damage, VG panel, and lighting receptor. These images encompass various backgrounds and fault variations. While VG panels and lighting receptors are not specific fault types, they are external components typically visible during wind turbine blade inspections [7]. Then divided the dataset into training, validation, and test subsets, consisting of 2,127, 181, and 41 images. Addressing concerns of underfitting and overfitting, applied data augmentation strategies like flipping, rotating, shearing, blurring, and cropping. Each image underwent precise annotation, employing bounding boxes and labels to spotlight specific faults in wind turbine blades using the Roboflow Annotator tool, a publicly available resource. Though initially set to a pixel dimension of 856 x 856, images are resized to a 640 x 640 resolution for both the training and testing phases. Elevating the models' effectiveness, incorporated transfer learning techniques, resulting in a significant enhancement in model performance.

**Transfer learning**

Leveraging transfer learning has revolutionized object detection, akin to how seasoned scholars guide students [28]. Transfer learning allows us to exploit this knowledge by initializing our object detection model with these weights. Fine-tuning only the final layers of the model, like a student specializing in a specific area, allows for classification and localization of objects within the target dataset [29]. This ingenious approach, exemplified by utilizing pre-trained models from the MS COCO dataset, significantly reduces training time and resource requirements by circumventing the need for extensive data collection and labeling [30]. The following section explores the specific algorithms employed and how they leverage transfer learning to address the challenges of wind turbine blade inspection.

**Algorithms and methodology:**

In this study, leverage three innovative object detection algorithms—Mask R-CNN, YOLOv7, and YOLOv8—to effectively detect and segment wind turbine faults. The approach begins with the mAP-focused segmentation tasks handled by Mask R-CNN. Subsequently, YOLOv7 and YOLOv8 for robust and real-time fault detection. The prime goal is to achieve a harmonious balance between accuracy and efficiency in identifying defects in turbine blades. Additionally, investigation investigates into the YOLOv8 model, exploring its algorithms with different Intersection of Union parameters, further enhancing understanding of its capabilities. First, chose the Mask R-CNN algorithm because of its good detection and segmentation capabilities.

**Mask R-CNN**

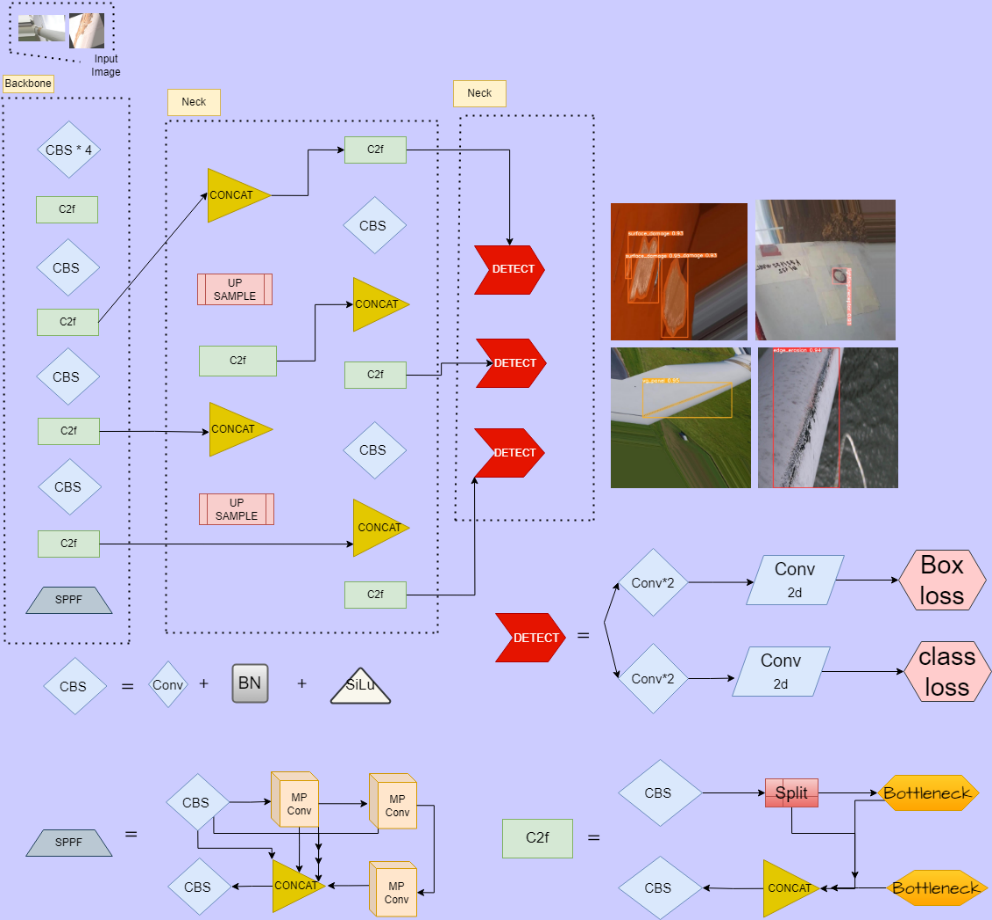
A diagram of a mask

Description automatically generatedMask R-CNN is a state-of-the-art instance segmentation model that builds on top of Faster R-CNN. It is a two-stage framework, where the first stage proposes regions of interest (RoI), and the second stage performs classification, bounding box regression, and instance segmentation on each RoI. The backbone of Mask R-CNN is Res Net, which handles the vanishing gradient problem in deep networks by introducing skip or residual connections, and it is integrated with a Feature Pyramid Network (FPN) to improve model detection accuracy and training time. Feature Pyramid Network is a top-down architecture with lateral connections developed to extract and build high-level semantic feature maps at different spatial resolutions by a bottom-up pathway, a top-down pathway, and lateral connections [19]. The feature maps produced by the Feature Pyramid Network (FPN) serve as input to the Region Proposal Network (RPN). Utilizing the concept of anchors, the RPN generates region proposals at various scales and aspect ratios for objects in the image. RPN processes all feature maps', and extracts RoI (Region of Interest) features from different sizes of the feature pyramids based on the size of the specific fault type. It operates like a sliding window and efficiently identifies areas containing objects in parallel due to its convolution operation. RoI Align is a critical layer in implementing the Mask R-CNN algorithm. It is responsible for extracting "M×M" feature maps from each RoI and unifying the output size of each RPN. Unlike RoI Pool, RoI Align eliminates aggressive quantization, significantly improving lo cation accuracy. The RoI is divided into 2×2 sub-windows or bins, and bi-linear interpolation is used to interpolate values of generated features within each RoI bin. The layer then aggregates the results using the max operation [10]. The loss function in Mask R-CNN is a combination of RoI loss, classifier loss, and mask head loss. Where represents the predicted probability that anchor point, is the target; p i represents the predicted value of the corresponding real area label; lb represents the log loss function [23]. signifies the four parameterized coordinate vectors of the predicted frame; signifies the coordinate vector corresponding to the border of the real area [24]. The total loss function as = + + , where , , + , represent the classification loss, bounding box loss and the average binary cross-entropy loss respectively [22]. The total architecture is above in the above figure. Next, chose YOLOv7 because of its speed and efficiency.

**YOLOv7**

YOLOv7 introduces several innovative features to enhance detection accuracy. One standout is its multi-head architecture, which incorporates a unique approach. Additionally, it implements the E-ELAN network, surpassing its predecessors in efficiency by focusing on layer aggregation. This network consists of residual blocks, each housing expands, shuffles, and merges cardinality operations. In the design of YOLOv7, a series of actions, including expansion, channel mixing, and merging, work in harmony to significantly enhance the network's learning capacity. The integration of the Bidirectional Feature Pyramid Network is particularly noteworthy. This design aims to optimize the flow of information throughout the network, significantly improving object detection accuracy. Building on this, the Path Aggregate Network enhances the structure by strengthening connections between feature pyramid levels. As a result, the features at each level of the feature pyramid combined to increase their degree of information [20]. It introduces a dual-head system for detection. The auxiliary head, being the first, is pivotal during training phases in intermediate layers, as it predicts coarse bounding boxes and class probabilities. Then, the lead head focuses on final detection outcomes, forecasting precise bounding boxes and class probabilities. Later in the process, refinement comes through a deliberate model re-parameterization strategy. Especially the identity connection is omitted in RepConv. This strategic design enhances gradient diversity for different feature maps, optimizing the network structure. Another main aspect of YOLOv7 is its dynamic feature, which reviews model outcomes and ground truths, assigning soft labels to detected objects. These soft labels, derived from ground truth, contribute to fine-tuning predictions. Lastly, YOLOv7 collaborates with YOLACT, known for its rapid instance segmentation capabilities. While YOLACT autonomously handles instance segmentation tasks on feature maps without requiring bounding box regression, it's important to note that its accuracy slightly lags multi-stage algorithms like Mask R-CNN. The combination of YOLOv7 and YOLAC makes YOLOv7 the most precise and fastest instance segmentation algorithm for Wind turbine blade fault detection and instance segmentation. The Loss function of Yolov7 is = + + , BCE cross-entropy loss formula is represented as =−  [\*logσ()+(1−)\*log(1−σ())] and σ()=where σ() represents the sigmoid function,  signifies the average of the results, and  signifies the real sample label. Intersection of Union (𝐼𝑜𝑈)=, where A represents the predicted frame, B represents the real frame. YOLOV7 uses CIoU loss. Α = , V= (, - + αv, = 1- . Here, v is to measure the aspect ratio's consistency, α is the weight function. The projected bounding box's width and height are represented by 𝑤 and ℎ, respectively, the ground truth bounding box's width and height are represented by 𝑤gt and ℎgt. The diagonal distance of a bounding box that may at least contain the two bounding boxes is denoted by c, and the center point distance between two bounding boxes is represented by 𝜌2 (b, bgt) [25]. Finally, YOLOv8 is the latest version of the YOLO family whose accuracy and speed are higher than previous versions. Above figure is YOLOv7 architecture.

**YOLOv8**

YOLOv8 is the latest version of the YOLO family. The architecture of YOLOv8 is state-of-the-art, capable of accommodating different resolutions, and even includes an instance segmentation model reminiscent of YOLACT. It is also anchor-free and a constituent of the Feature Pyramid Network (FPN) Path Aggregation Network (PAN), and SPPF modules. These enhance the ability to discern object shapes and textures across various scales and lead to improved accuracy. The backbone of the model is composed of four sections, each introduced by a single convolution and succeeded by an innovative C2f module [12]. This new addition to CSPDarknet53 involves a bottleneck segment featuring two 3x3 convolutions with residual connections. The model features a revamped detection head that separates classification and detection tasks. The loss computations have been refined, utilizing BCE Loss for classification and a combination of Complete Intersection of Union (CIOU) Loss + Distributed Focal Loss (DFL) for regression. The DFL approach models the box position as a general distribution, while VFL introduces an asymmetric weighting operation. These enhancements contribute to the robust performance of YOLOv8 in object detection tasks. [17][18]. Above Figure is YOLOv8 architecture. Below are equations of DFL [26]. Regression Target using Dirac Delta y= . Proposal to learn General Distribution =, Discretization of the Distribution = , Distribution of Focal loss : , = - () log() + (y-, where , are probabilities. Global minima solution for ,  *;* = , =

**Our research and results**

We applied advanced detection and segmentation algorithms—Mask R-CNN, YOLOv7, and YOLOv8—to wind turbine blades, achieving noteworthy results. Specifically, Mask R-CNN exhibited a detection mAP@IoU (0.5) of 86.30% and a segmentation mAP@mask (0.5) of 84.56%. YOLOv7 performed with a detection mAP@IoU (0.5) of 95.80% and a segmentation mAP@mask (0.5) of 96.30%. Subsequently, this research extended to the implementation of the YOLOv8 algorithm, introducing the Intersection of Union (IoU) tuning at thresholds of 0.6, 0.7, and 0.8. YOLOv8 outperformed with a detection mAP@IoU (0.7) of 97.30% and a segmentation mAP@mask (0.8) of 97.60%. This comprehensive exploration showcases the nuanced performance of these algorithms in the realm of wind turbine fault detection and segmentation.

**Sample Image outputs:**

A close-up of a wall

Description automatically generated A close-up of a crack in a wall

Description automatically generated A close-up of a paint stain

Description automatically generated A close-up of a plane wing

Description automatically generated

**Table:1** Performance comparison between Mask R-CNN, YOLOv7, and YOLOV8

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Performance Metrics** | **Mask R-CNN** | **YOLOv7** | **YOLOV8** | **Improvement** |
| mAP@IoU0; detection | 86.30% | 95.80% | 97.3% | +1.50% |
| mAP@IoU; mask | 84.56% | 96.30% | 97.6% | +1.30% |

**Table2:** Performance of YOLOv8 algorithm with different IoU’s

|  |  |  |
| --- | --- | --- |
| **Performance of different IoU’s** | **Detection** | **Segmentation** |
| mAP@IoU (0.7) | 97.3% | 97.2 % |
| mAP@IoU (0.6) | 97.1% | 97% |
| mAP@IoU (0.8) | 97.1% | 97.6 % |

**Conclusion**

This research successfully employed deep learning models for automated detection and segmentation of wind turbine blade faults, paving the way for improved wind farm inspections. YOLOv8 surpassed Mask R-CNN and YOLOv7 in accuracy and efficiency, achieving an impressive mAP@IoU of 97.3% for detection and 97.6% for segmentation. YOLOv8 demonstrated robustness to different IoU thresholds, suggesting its adaptability to diverse detection requirements. Integrating these models into real-world systems can significantly reduce maintenance costs, downtime, and safety risks associated with manual inspections. Overall, this study highlights the effectiveness of deep learning algorithms for automated wind turbine blade fault detection and segmentation, contributing to a more efficient and sustainable wind energy future. The proposed research potentially revolutionized wind turbine maintenance by enabling automated, accurate, and cost-effective fault identification.

**References**

1. Seo, S. N. (2017). Beyond the Paris Agreement: Climate change policy negotiations and future directions. Regional Science Policy & Practice, 9(2), 121-140.
2. (June 10, 2022). U.S. Energy Information Administration - EIA - Independent Statistics and Analysis. U.S. energy facts explained - consumption and production - U.S. Energy Information Administration (EIA). https://www.eia.gov/energyexplained/us-energy-facts/
3. Wiser, R., Bolinger, M., Hoen, B., Millstein, D., Rand, J., Barbose, G., and Paulos, B. (2022). Land-based wind market report: 2022 edition. Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).
4. Stetco, A., Dinmohammadi, F., Zhao, X., Robu, V., Flynn, D., Barnes, M., and Nenadic, G. (2019). Machine learning methods for wind turbine condition monitoring: A review. Renewable Energy, 133, 620-635.
5. Kusiak, A., and Li, W. (2011). The prediction and diagnosis of wind turbine faults. Renewable Energy, 36(1), 16-23.
6. Nielsen, J. J., and Sørensen, J. D. (2011). On risk-based operation and maintenance of offshore wind turbine components. Reliability Engineering & System Safety, 96(1), 218-229.
7. Shihavuddin, A. S. M., Xiao Chen, X., Fedorov, V., Nymark Christensen, A., Andre Brogaard Riis, N., Branner, K., and Reinhold Paulsen, R. (2019). Wind turbine surface damage detection by deep learning aided drone inspection analysis. Energies, 12(4), 676.
8. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., and Ramanan, D., and Zitnick, CL (2014, September). Microsoft COCO: Common objects in context. In European Conference on Computer Vision (pp. 740-755).
9. Roboflow Annotator. Roboflow. <https://roboflow.com/>.
10. Kaiming He Georgia Gkioxari Piotr Dollar Ross Girshick Facebook AI Research (FAIR) Mask R-CNN
11. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors Chien-Yao Wang1, Alexey Bochkovskiy, and Hong-Yuan Mark Liao1 1 Institute of Information Science, Academia Sinica, Taiwan
12. Real-Time Flying Object Detection with YOLOv8 Dillon Reis\*, Jordan Kupec, Jacqueline Hong, Ahmad Daoudi Georgia Institute of Technology
13. Fu, Y., Wu, J., Hu, Y., Xing, M., and Xie, L. (2021, January). Desnet: A multi-channel network for simultaneous speech dereverberation, enhancement and separation. In *2021 IEEE Spoken Language Technology Workshop (SLT)* (pp. 857-864). IEEE.
14. Soudy, M., Afify, Y., and Badr, N. (2022). RepConv: A novel architecture for image scene classification on Intel scenes dataset. *International Journal of Intelligent Computing and Information Sciences*, *22*(2), 63-73.
15. Bolya, D., Zhou, C., Xiao, F., and Lee, Y. J. (2019). Yolact: Real-time instance segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9157-9166).
16. Jacob Solawetz and Francesco. What is yolov8? the ultimate guide., 2023. 04-30-2023. 1, 5, 8
17. DC-YOLOv8: Small size Object detection algorithm based on camera sensor Haitong Lou1, Xuehu Duan1, Junmei Guo1, Haiying Liu1 \*, Jason Gu2, Lingyun Bi1, Haonan Chen 1 7 April 2023 doi:10.20944/preprints202304. 0124.v1
18. DCF-Yolov8: An Improved Algorithm for Aggregating Low-Level Features to Detect Agricultural Pests and Diseases by Lijuan Zhang 1,2, Gongcheng Ding 1,2, Chaoran Li 3 and Dongming Li 1,\*ORCIDAgronomy 2023, 13(8), 2012; https://doi.org/10.3390/agronomy13082012.
19. Feature Pyramid Networks for Object Detection Tsung-Yi Lin, Piotr Dollar , Ross Girshick , Kaiming He, Bharath Hariharan, and Serge Belongie
20. Path Aggregation Network for Instance Segmentation Shu Liu† Lu Qi† Haifang Qin§ Jianping Shi‡ Jiaya Jia†,[ †The Chinese University of Hong Kong §Peking University ‡SenseTime Research [YouTu Lab, Tencent arXiv:1803.01534v4
21. Representation Learning: A Statistical Perspective Jianwen Xie1 , Ruiqi Gao2 , Erik Nijkamp2 , Song-Chun Zhu2 , and Ying Nian Wu2 1Hikvision Research Institute, 2Department of Statistics, University California, Los Angeles arXiv:1911.11374v1 [stat.ML] 26 Nov 2019
22. Baruah P. Types of tea, value addition and product diversification of Indian tea. 2015.
23. Chen J, Chen Y, Jin X, et al. Research on a parallel robot for green tea flushes plucking/Proceedings of the 5th International Conference on Education, Management, Information and Medicine. 2015: 22–26.
24. Hu G, Wu H, Zhang Y, Wan M. A low shot learning method for tea leaf’s disease identification. Comput Electron Agric
25. Channel Pruning-Based YOLOv7 Deep Learning Algorithm for Identifying Trolley Codes by Jun Zhang, Rongxi Zhang \*, Xinming Shu, Lulu Yu and Xuanning Xu Appl. Sci. 2023, 13(18), 10202
26. Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection Xiang Li1,2, Wenhai Wang3,2, Lijun Wu4, Shuo Chen1, Xiaolin Hu5, Jun Li1, Jinhui Tang1, and Jian Yang1∗ arXiv:2006.04388v1 [cs.CV] 8 Jun 2020
27. Xu, B., Li, W., Yang, M., & Liu, Y. (2023). Fast detection of wind turbine blade damage using Cascade Mask R-DSCNN-aided drone inspection analysis. Signal, Image and Video Processing, 17(10), 2333-2341
28. Krishnamoorthy, N., Rath, T. K., & Rao, K. M. P. (2017). Faster r-cnn object detection: A comprehensive review and future directions. arXiv preprint arXiv:1705.07678.
29. Najibi, M., Aminmansour, M., Vandali, B., & Rahmati, M. (2020). Deep learning-based object detection for remote sensing vehicles: A comprehensive review of the state-of-the-art and future trends. IEEE Geoscience and Remote Sensing Letters, 17(5), 1801-1805.
30. Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In European conference on computer vision (ECCV) (pp. 740-755). Springer, Cham