**Detection and segmentation of wind turbine blade faults using Mask R-CNN, YOLOV7, And YOLOV8 with different Intersection of union**

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

**Keywords**: Object detection, semantic segmentation, wind turbine blade faults, Mask R-CNN, YOLOV7, And YOLOV8

**Abstract**

Wind turbines are at the heart of sustainable electricity generation from renewable sources. These blades, however, are constantly exposed to a myriad of environmental conditions, rendering them prone to wear and tear. This vulnerability underscores the importance of consistent monitoring not just to reduce risks, maintenance costs, and downtime but also to safeguard against structural damages that could compromise the constant supply of power. In light of this, our research delves into the intricate task of detection and segmentation of prevalent faults in turbine blades, specifically surface damage and edge erosion. Furthermore, we've extended our investigation to two crucial components: the lightning receptor and the VG panel. To ensure the highest efficiency possible, we conducted a comparative analysis of three state-of-the-art models: Mask R-CNN, YOLOv7, and YOLOv8. We also researched the YOLOv8 model with three different Intersections of Unions, and they are IoU@0.6, IoU@0.7, and IoU@0.8.

**Introduction**

The primary driver of greenhouse gas emissions stems from fossil fuel-based energy sources. This research helps to minimize emissions and resolve environmental issues; it's essential to integrate renewable energy sources into our everyday routines. The Paris Agreement, signed by 165 nations in 2015, marked a significant milestone in fortifying the global commitment to curbing carbon dioxide emissions and bolstering investments in renewable energy solutions [1]. Following the agreement, there was a significant increase in expenditures for green energy infrastructure. Notably, solar and wind energy emerged as front-runners in the competition for renewable energy production. According to the United States Energy Information Administration, solar electricity accounts for 9% of the nation's energy supply, while wind generation contributes 24% [2]. Advancements in technology have revolutionized the design and production of wind turbines, resulting in a reduction in the cost of power generation compared to other renewable sources. The efficiency of a wind turbine is nearly 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 1-2 cents per kilowatt-hour [4]. Moreover, it is worth noting that harnessing 1 megawatt of wind energy can potentially offset approximately 2600 tons of carbon dioxide emissions [5].

The wind turbines are built on land (onshore) or in large bodies of water like oceans and lakes (offshore). Irrespective of the categories, the wind turbine architecture consists of many mechanical and electrical components such as a rotor, blades, generator, controller, and gearbox. Notably, the turbine blades, typically measuring an impressive 100-140 meters in length, play a crucial role in electricity generation. Its efficiency in rotating determines the amount of electric power generated. These gigantic blades are subjected to aerodynamic and gravitational loads while operating under extreme climatic conditions, which causes vibration forces that result in structural damages such as cracks on the surface, erosion of the edge, pitch angle, and twisting blades[ 4][ 5].

These forces can induce vibrations that manifest as structural damages, such as surface cracks, edge erosion, and even issues with pitch angles or twisted blades. Traditional inspection of these blades relies on time-based maintenance techniques. This approach is not only inefficient but also fraught with challenges. For instance, rope-based inspections pose significant risks to personnel, and telephotography methods can overlook microscopic structural damages due to human limitations. Furthermore, faults in wind turbines don't just compromise efficiency; they also pose environmental hazards. The operation and maintenance (O&M) can account for up to 30% of the costs of wind electricity generation, so an efficient and reliable inspection method becomes indispensable. 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. Our approach leans on cutting-edge object detection algorithms, namely Mask R-CNN, YOLOV7, and YOLOV8, to navigate the challenges of manual inspections. We'll begin by discussing our data sources.

**Data Acquisition**

In this research, we developed a dataset containing 2,127 augmented drone images. These images capture four distinct categories: edge erosion, surface damage, VG panel, and lighting receptor. These images encompass various backgrounds and fault variations. It's worth noting that while VG panels and lighting receptors aren't specific fault types, they are external components typically visible during wind turbine blade inspections [7]. We systematically divided the dataset into training, validation, and test subsets, consisting of 2,127, 181, and 41 images. Addressing concerns of underfitting and overfitting, we harnessed 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, we incorporated transfer learning techniques, resulting in a notable enhancement in model performance.

**Transfer learning**

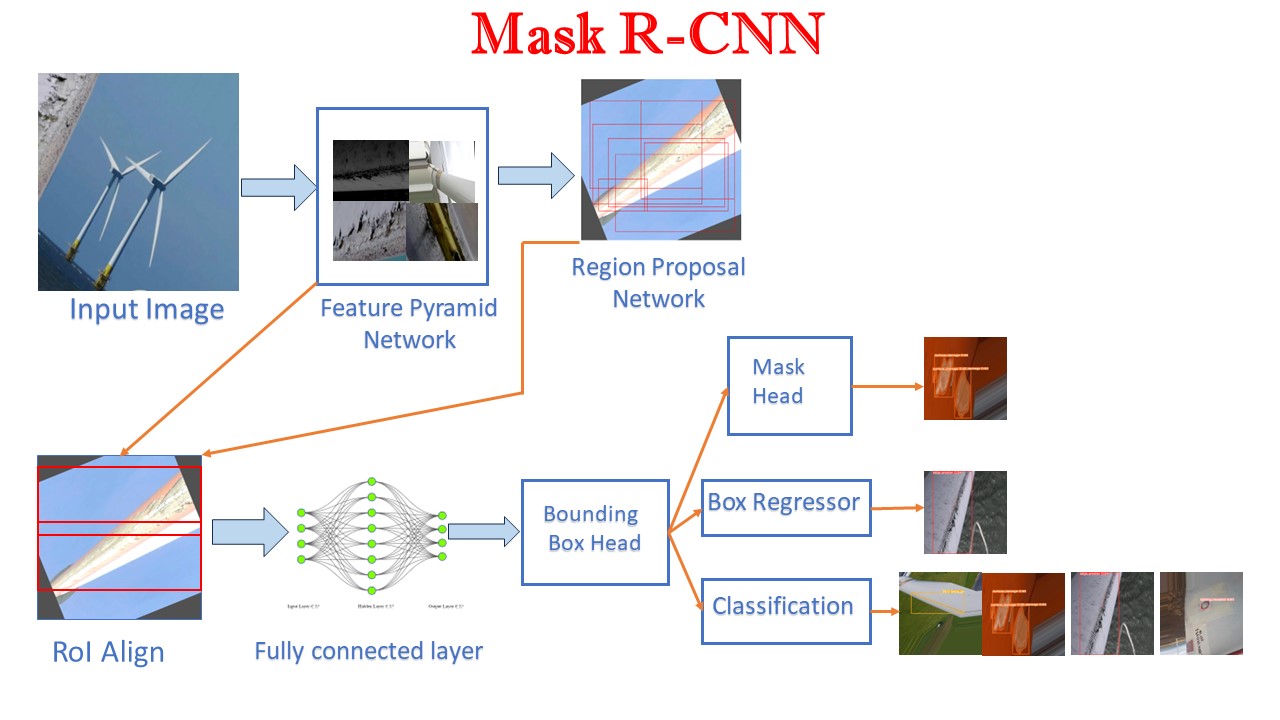
Transfer learning is used to leverage pre-trained models, reducing the need for extensive data and computational resources. Recognizing the challenge of amassing several labeled datasets, we adopted the transfer learning method. Specifically, we utilize a pre-trained model from the Microsoft Common Objects in Context (MS COCO) dataset [8]. Additionally, our study includes an analysis of the performance of transfer learning, with experiments involving the freezing of layers in the initial stages of the Mask R-CNN, YOLOv7, and YOLOv8 models. In the next section, we will delve into the algorithms utilized in this project.

**Algorithms and methodology:**

The Mask R-CNN algorithm segments the image by precise features by different classes. Masked R-CNN involves a multi-step process. In the first phase, the images are given to the model, and they extract important features from the image and generate a proposal network. In the second stage, it filters and classifies the suggestions to generate bounding boxes and masks for defects [Figure 1].

**Mask R-CNN**

Mask 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 location accuracy. The RoI is divided into 2×2 sub-windows or bins, and bi-linear interpolation is used to consider interpolated values of generated features within each RoI bin. The layer then aggregates the results using the max operation [10]. The predicted bounding box and mask corresponding to the highest class score is the final prediction for each region. The total architecture is shown below in Figure 1.

****

**Figure 1**

A diagram of a diagram

Description automatically generated

Figure 3

YOLOv7 (CSPDarknet53 with E-ELAN, CSP, PAN, BiFPN )

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 expand, shuffle, and merge 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 may be combined to increase their degree of information[20]. YOLOv7's architectural advancements don't stop there. 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. On the other hand, the lead head focuses on final detection outcomes, forecasting precise bounding boxes and class probabilities. Further, 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 notable 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 behind 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.

**YOLOv8 (**CSPDarknet53 with ShuffleNetV2 and ResBlockV2**,** C2f, SPPF**)**

YOLOv8 is the newest iteration of the YOLO object detection model. It is Anchor free. It combines both Feature Pyramid Network (FPN) and Path Aggregation Network (PAN). This combination enhances its ability to recognize object shapes and textures across varied scales, leading to higher accuracy. The introduction of a state-of-the-art model accommodating different resolutions and even an instance segmentation model reminiscent of YOLACT. The backbone of YOLOv8 consists of four sections, each prefaced by a single convolution and followed by a c2f module, which is an innovative addition to CSPDarknet53. This module involves a bottleneck segment featuring two 3x3 convolutions with residual connections.

YOLOv8 has several key strengths:

1. A lightweight network architecture.
2. Efficient feature fusion techniques.
3. Enhanced detection accuracy by combining features from several real-time object detectors.

It still leverages the CSP idea from YOLOv5 and incorporates feature fusion (PAN-FPN) and SPPF modules. The main advancements in YOLOv8 include b) While retaining the c2f module concept, its design was influenced by the ELAN structure in YOLOv7. c) A revamped detection head, which separates classification and detection. d) Improved loss computations, utilizing BCE Loss for classification and a combination of CIOU Loss + DFL for regression. The DFL approach models the box position as a general distribution, with VFL introducing an asymmetric weighting operation [17][18].

A screenshot of a computer

Description automatically generated

**Our research and results**

Applied Mask R-CNN, YOLOV7 algorithms on wind turbine blades for detection and segmentation. Achieved mAP@IoU (0.5) 86.30% for detection and mAP@mask (0.5) for segmentation is 84.56% with Mask R-CNN. With mAP@IoU (0.5) YOLOv7 accomplished 95.80% for detection, 96.30% mAP@mask (0.5) for segmentation. Researched the project by changing Intersection of Union [IoU] to 0.4, 0.5, and 0.6 using YOLOV8 algorithm.

**Sample Images:**

A close-up of a white surface

Description automatically generated A person wearing a harness

Description automatically generated

Table1. Performance comparison between Mask R-CNN. YOLOv7, and YOLOV8

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Performance Metrics** | **Mask R-CNN** | **YOLOv7** | **YOLOV8** | **Improvement** |
| mAP@IoU(0.5) | 86.30% | 95.80% | 98.4% | +2.60% |
| mAP@mask(0.5) | 84.56% | 96.30% | 96.3% | 0% |

|  |  |  |
| --- | --- | --- |
| **Performance of different IoU’s** | **Detection** | **Segmentation** |
| mAP@IoU@default | 98.4% | 95.3% |
| mAP@mask(0.6) | 94%/94.9% | 95.3% |
| mAP@IoU(0.4) | 94.9% | 96.3% |

**Conclusion**

**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

Wind turbines are pivotal in harnessing electricity sustainably from renewable resources. However, their blades face a multitude of environmental conditions, making them particularly vulnerable to wear and tear. Recognizing these vulnerabilities is crucial, as it not only underscores the importance of regular monitoring to mitigate risks but also highlights the need to minimize maintenance costs and time. Effective research in this area can provide proactive solutions to prevent structural damage, ensuring a consistent and uninterrupted power supply.

In our recent study, we zeroed in on the detection and segmentation of specific faults in turbine blades, namely surface damage and edge erosion. Additionally, we delved into two key components: lightning receptors and VG panels. Our approach was comprehensive; we compared the performance of three leading models—Mask R-CNN, YOLOv7, and YOLOv8—to determine the most proficient tool for detecting and segmenting these issues, aiming to bolster the longevity and reliability of wind turbines.

Certainly, let's integrate these ideas more seamlessly to ensure flow and engagement:

"Wind turbines are at the heart of sustainable electricity generation from renewable sources. Their blades, however, are constantly exposed to a myriad of environmental conditions, rendering them prone to wear and tear. This vulnerability underscores the importance of consistent monitoring, not just to reduce risks, maintenance costs, and downtime, but also to safeguard against structural damages that could compromise the consistent supply of power.

In light of this, our research delves into the intricate task of detection and segmentation of certain prevalent faults in turbine blades, specifically surface damage and edge erosion. Furthermore, we've extended our investigation to two crucial components: the lightning receptor and the VG panel. To ensure accuracy and efficiency in this endeavor, we've utilized and compared three state-of-the-art models: Mask R-CNN, YOLOv7, and YOLOv8. The goal? To ascertain which tool offers the most reliable results in both detection and segmentation, and ultimately, ensure the longevity and efficacy of wind turbines."