International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 15s,2025
https://theaspd.com/index.php

Fuzzy Logic-Assisted Damage Detection On Wind Turbine Blades Using Mask R-CNN

Clark Allen[1], Reed Pratt[2], Mohammad A.S. Masoum[3]
[1,2,3]Machine Learning and Drone Lab, Utah Valley University, Orem, United States,
ClarkA@uvu.edu[1], ReedP@uvu.edu[2], m.masoum@ieee.org[3]

***Abstract*—** This study explores the use of deep learning and drones to facilitate automated inspections of wind turbine blades. Three Mask Region-Based Convolutional Neural Network (R- CNN) models with VGG19, Xception, and ResNet-50 backbones were constructed. The models were trained on an annotated dataset of 3,000 RGB images (size 300×300 pixels) which contain 10,391 defect annotations consisting of cracks, holes, and edge erosion. The dataset of images was captured and annotated at Utah Valley University (UVU). To improve defect detection performance, a simple and practical Single-Variable Fuzzy (SVF) voting system was proposed and implemented. This method demonstrated superior accuracy compared to the individual models. The best-performing standalone model, Mask R-CNN with Xception, achieved an mAP of 77.48%, while the SVF system increased the mAP to a score of 81.89%.
***Keywords*—** Defect Detection, Fuzzy Voting, Mask R-CNN, Wind Turbine Blades

I. INTRODUCTION

The year 2023 marked a pivotal milestone for the global wind energy sector, achieving unprecedented growth with a 50% increase in installations compared to the previous year. This expansion reflects the accelerating transition to renewable energy, driven by international commitments such as the COP28 agreement to triple renewable capacity by 2030. [1].Wind turbines play a key role in the transition to clean energy, but their operation exposes them to harsh weather conditions, including storm winds, rain, hail, and lightning. Among turbine components, the blades are particularly vulnerable, enduring high mechanical stress and environmental wear [2]. Studies show that blade failures contribute to 19.4% of all turbine failures, highlighting the urgent need for advanced inspection and maintenance solutions [3].Traditional inspections rely on manual rope teams, a labor-intensive, time-consuming, and hazardous approach. Recent advancements in unmanned aerial systems have enabled drone-based inspections, a safer and more efficient alternative [4]. Studies suggest that drone-based inspections can reduce costs by up to 70% and cut downtime-related revenue losses by 90% compared to traditional inspection methods [5].These new inspection methods are complemented by advancements in deep learning and computer vision, which have led to model architectures that can be applied to defect detection. Convolutional Neural Networks (CNNs) have demonstrated strong performance in anomaly detection [6]. Advanced architectures for object detection including YOLO, Faster R-CNN, and Mask R-CNN have proven effective in localizing and segmenting damage on wind turbine blades [7], [8]. This study builds on these advancements and previous work [9] by focusing on enhancing the performance of Mask R- CNN for defect detection on wind turbine blades through a novel post-processing system. We propose a Single-Variable Fuzzy (SVF) voting system designed to enhance detection ac- curacy without unnecessary complexity. The proposed voting system processes outputs from multiple Mask R-CNN models (with VGG19, Xception, and ResNet-50 backbones, in this research) and utilizes techniques from fuzzy logic to make practical voting decisions. Increasing the accuracy of defect detection also improves the system's robustness and may help facilitate adoption into wind turbine inspection practices across the globe. Various voting systems have been applied to CNNs in the past [10] and Mask R-CNN [11], including a Multi-Variable Fuzzy (MVF) voting system from our previous work [9]. The key contributions of this paper are:

- The design and implementation of a Single-Variable Fuzzy (SVF) voting system to further enhance detection accuracy by leveraging outputs from multiple Mask R- CNN models.
- The comparison and analysis between the SVF and MVF voting systems.

The remainder of this paper is organized as follows: A brief overview of Mask R-CNN using CNN backbone architectures, VGG19, Xception, and ResNet is presented in Section II. The annotated RGB dataset with cracks, holes, and edge erosion is described in Section III. The main innovation of this paper, the SVF voting system, is proposed in Section IV. Our previous Multi-Variable Fuzzy Voting system [9] is briefly described in Section V. The experimental results and comparisons are presented in Section VI followed by the conclusions and suggestions for future work in Section VII.

II. Mask Region-Based Convolutional Neural Network

Mask R-CNN is an extension of the Faster R-CNN introduced by Kaiming He et al. in 2017 [12]. This architecture adds a branch for predicting segmentation masks in parallel to the bounding box predictions. The architecture consists of two main stages:

- Stage one consists of the backbone network, often built on established architectures like ResNet, and a Region Proposal Network (RPN). The backbone extracts high- level feature maps from input images, while RPN slides over these feature maps to generate region proposals. It predicts objectness scores and bounding box coordinates for anchor boxes of different scales and aspect ratios, enabling identification of potential object locations [13]. In this research three backbone networks were used: VGG19, Xception, and ResNet-50 [14]-[16]. These networks were used in this study to build on our previous work [9] and were originally chosen due to their success in [6] and [17].
- Stage 2 processes the proposed regions. Each region undergoes classification, bounding box regression, and mask prediction. The mask branch, implemented as a fully convolutional network, generates binary masks for each detected instance, providing pixel-level segmentation for the identified objects [12].

III. Annotated RGB Dataset with Cracks, Holes, and Edge Erosion

The Mask R-CNN deep learning models used in this re- search were trained on a custom-made dataset consisting of 3,000 images, each with a size of 300×300 pixels [9]. The small-scale wind turbine in the dataset is a 12V Primus Air Max, featuring artificial damage of three different defect types: cracks, holes, and edge erosion. Originally, 6,000 images were captured at Utah Valley University (UVU), with 3,000 being photographs of the undamaged wind turbine, and the other half having the artificial damage applied. Of the 6,000 images, 4,000 were taken indoors with a smartphone, and the remaining 2,000 images were captured using a DJI Mini SE and DJI Mini 3 Pro drone outdoors [6]. However, to focus on

defect detection, the dataset was filtered to include only the 3,000 images with visible damage.

Fig. 1 Samples from the dataset captured at UVU with the raw images (top) and annotated images (bottom).

To allow for use with the Mask R-CNN segmentation architecture, the images in the dataset were annotated with pixel-level binary masks for each visible defect. The complete dataset contains 5,335 cracks, 1,632 holes, and 3,424 edge erosion masks, totaling 10,391 defect annotations. The 3,000 images were divided into 2,100 training images, 600 test images, and 300 validation images [9]. The validation images are what are used for evaluation and comparison between the different Mask R-CNN models. Example images from the dataset, with and without annotations, are shown in Figure 1.

IV. Proposed Single-Variable Fuzzy (SVF) Voting for Defect Detection Using Mask R-CNN with Xception, ResNet50, and VGG19 Backbones

Building on our previous work [9], we propose a Single-Variable Fuzzy (SVF) voting system designed to improve defect detection on wind turbine blades by integrating predictions from multiple Mask R-CNN models. This is a three stage system consisting of the following: an initial grouping mask grouping stage, the fuzzification stage, then the processing and defuzzification stage.

A. Grouping of Mask Predictions

In the grouping stage, the Intersection over Union (IoU) is calculated between all predicted instances passed into the SVF system. These IoU values are then used to group overlapping masks together. This groups together predictions which lie in the same area of the image. Each group is then used as an individual unit for fuzzy evaluation.

B. Fuzzification of Input Masks

Once the masks are grouped, the SVF system maps the crisp inputs to a fuzzy variable defined as follows:

- Overlap Score (μ): The average IoU of the group is calculated as seen in (1)

The μ from each group is then processed through two sequen- tial step functions to determine its membership and how it will be defuzzified.

C. Defuzzification

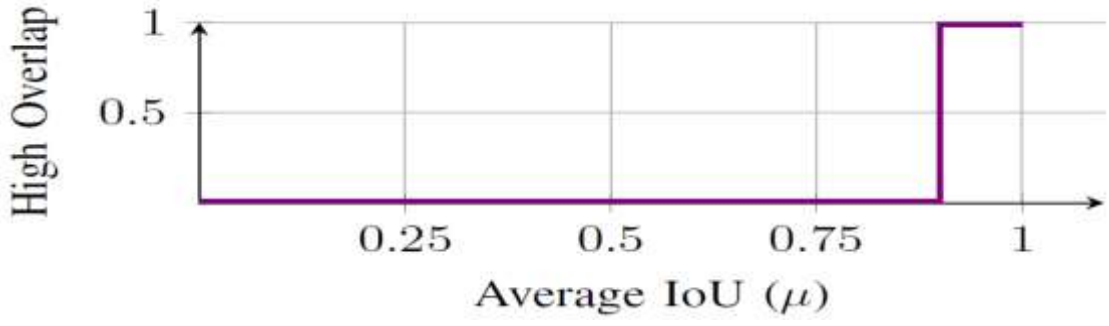
The first membership function checks for High Overlap, as described in (2) and illustrated in Figure 2. If a group's μ is greater than or equal to the threshold of 0.9, it is classified as a High Overlap group. Otherwise, the group's μ is evaluated against the membership function for Low Overlap, see (3) and Figure 3. Groups that do not meet the 0.5 threshold for Low Overlap are discarded.

$$\mu = \frac{\mathrm{IoU}_1 + \mathrm{IoU}_2 + \cdots + \mathrm{IoU}_n}{n} \tag{1}$$

$$H(\mu) = \begin{cases} 1 & \text{if } \mu \geq 0.9 \\ 0 & \text{if } \mu < 0.9 \end{cases} \tag{2}$$

$$L(\mu) = \begin{cases} 1 & \text{if } \mu \geq 0.5 \\ 0 & \text{if } \mu < 0.5 \end{cases} \tag{3}$$

The defuzzification and processing stage differs depending on whether a group is classified as High or Low Overlap. In High Overlap groups, the predicted instances are typically similar, which allows for the prediction with the highest confidence to be selected as the crisp output. For Low Overlap groups, where there is greater disagreement between models and instances are often only partially masked. The two predictions with the highest confidence values are merged using a logical OR operation, increasing the probability that the resulting mask correctly identifies a defect.



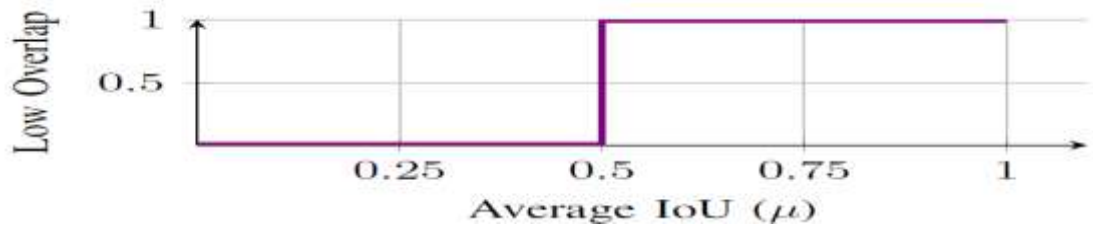


Fig. 2: Membership function for High Overlap H(x) (2). Fig. 3: Membership function for Low Overlap L(x) (3).

Algorithm 1 Single-Variable Fuzzy (SVF) Defect Mask R- CNN Voting

Input: Mask R-CNN predictions from Xception, VGG19, and ResNet-50 for one image.
Output: A new set of masks.
Grouping: Group overlapping masks together.
Fuzzification: For each group, calculate Overlap Score (μ)
    $\mu = (IoU1 + IoU2 + \ldots + IoUn)/n$
Defuzzification: Determine the output fuzzy set from the grouped predictions based on the
If H(μ):
    Return the highest confidence mask.
Else If L(μ):
    Merge the two highest confidence predictions
    Return the merged mask.

V. Multi-Variable Fuzzy (MVF) Voting for Defect Detection Using Mask R-CNN with Xception, ResNet50, and VGG19 Backbones

In our previous work [9], we introduced a Multi-Variable Fuzzy (MVF) voting system, which leverages fuzzy inference to integrate predictions from three Mask R-CNN models (Xception, ResNet50, and VGG19) and produces a refined set of masks to detect defects on wind turbine blades. The MVF voting system consists of the following stages: mask cleaning, mask grouping, fuzzification, and defuzzification [9]. In this paper, a simplified and more practical version of that system, a Single-Variable Fuzzy (SVF) voting is designed and implemented to further improve detection accuracy by leveraging outputs from multiple Mask R-CNN models. An overview of the proposed approach is illustrated in Figure 10 (see Algorithm 1).The mask cleaning stage consists of mask isolation, noise filtering, and mask merging. This will maintain a single mask per defect, discard masks smaller than a defined size threshold, and merge highly overlapping masks to prevent duplicates, respectively.The grouping stage collects sets of masks that overlap from the input predictions. Maintaining separate groups if the classifications differ between predictions.The fuzzification stage maps the crisp input masks into fuzzy variables c (confidence score) and a (agreement score) using clamped membership functions. The confidence score represents the highest confidence mask in the group, and the agreement score corresponds to the maximum IoU in the group.The defuzzification stage computes a crisp value v (validity score) from the fuzzy values calculated in the previous stage. If this value meets a specified threshold, the mask with the highest individual confidence is selected as the output for the corresponding group. Otherwise, the group is discarded from the system's outputs [9].

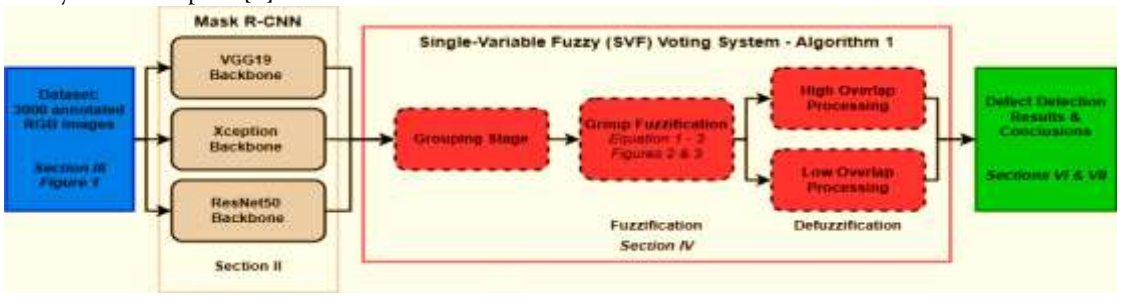


Fig. 4: Overview of the proposed SVF voting system (Algorithm 1) for detecting defects in wind turbine blades using Mask R-CNN.

VI. Experimental Results and Comparisons

The Mask R-CNN models with Xception, ResNet-50, and VGG19 backbones were trained and evaluated using the created dataset introduced in Section III. Each backbone underwent hyperparameter tuning and were used to generate the results provided in Tables I-II. Validations were performed on a separate dataset of 300 im- ages, and to maintain consistency the same base models from [9] were used with the SVF voting system. The accuracy of the three base models and the proposed voting system are presented in Table I.

- Among individual backbones, Xception achieved the highest performance with an mAP of 77.48% (Table I, row 1, column 2).
- The most notable improvement was observed in mAP, as it was the primary focus during the development of these voting systems. However, there was no significant changes in the secondary metric, the F1 score, indicating the models were already well-balanced in terms of precision and recall.
- The SVF voting system raised the mAP by over 4%, from Xception's 77.48% to 81.89% after processing the predictions, as shown in the second column of Table I.
- The best F1 score was achieved by VGG19, which increased slightly from 78.33% to 78.38% after SVF voting, as shown in the third column of Table I.

TABLE I Defect detection accuracy: mAP and F1 score Comparison.

Model / Voting system	**mAP**	**F1 Score**
Xception	77.48%	77.31%
VGG19	77.30%	78.33%
ResNet-50	72.98%	74.54%
Proposed SVF Voting	81.89%	78.38%
MVF Voting	80.10%	74.52%

TABLE III Detailed Class-Based Accuracy Metrics for the Proposed SVF Voting System.

	mAP	**Precision**	**Recall**	**F1 Score**
Cracks	77.48 %	75.95%	88.40 %	77.31%
Holes	77.30 %	90.09%	91.83 %	78.33%
Erosion	72.98 %	64.17%	73.51 %	74.54%
Overall	81.89 %	73.86%	83.50 %	78.38%

TABLE IIIII Detailed Class-Based Accuracy Metrics for the Proposed MVF Voting System [9].

	mAP	**Precision**	**Recall**	**F1 Score**
Cracks	79.93 %	68.47%	86.00 %	76.24%
Holes	89.80 %	88.26%	90.38 %	89.31%
Erosion	70.58 %	57.73%	77.57 %	66.19%
Overall	80.10 %	67.17%	83.67 %	74.52%

A speed benchmark was conducted to compare the pro- cessing speeds of the SVF and MVF voting systems. This was done by recording the processing time for each image in the validation set, then the average computation time was calculated. In the case of MVF, the mask cleanup stage was timed separately, as it required significantly more time than the voting itself. The average times for each system are compared in Table IV.

TABLE IVV AVERAGE RUN TIME (SECONDS) TO PROCESS ONE IMAGE WITH THE MVF AND SVF SYSTEMS.

	MVF Voting	**SVF Voting**
Cleaning	0.239	N/A
Voting	0.104	0.163
Total	0.343	0.163

On average, the SVF system processes images at roughly twice the speed of the MVF voting system. This speed increase was achieved by eliminating the preprocessing stage that the MVF system uses in favor of a mask merging condition in the SVF defuzzification stage. The mask cleaning stage of the MVF system was a computationally intensive process that scales poorly with large image sizes because of the increased matrix sizes. The input matrices are of size NxNxD where N is the pixel dimensions of the image and D is the number of defects. Using 3 models, each outputting 5-12 defects per image, the number of necessary computations increases drastically along with the image size.

VII. CONCLUSIONS

Three Mask R-CNN models were developed and trained using different CNN backbones: VGG19, Xception, and ResNet-50. The models were trained on a novel dataset comprising 3000 RGB images (300×300 pixels) annotated with defects such as cracks, holes, and edge erosion. To further improve defect detection accuracy, a fuzzy voting method was proposed and implemented: Single-Variable Fuzzy (SVF) voting. The results showed that the voting system significantly outperformed the individual models, providing a promising step toward safer and more efficient automated inspections of wind turbine blades. The best individual backbone, Xception, achieved an mAP of 77.48%, while the proposed SVF voting system achieved an mAP score of 81.89%. Our future research will aim to improve the fuzzy system and explore new voting mechanisms capable of integrating more than three input models. Additionally, the defect analysis could be extended to include fault size computation using fuzzy logic, enhancing the depth of inspection insights.

ACKNOWLEDGEMENT

We extend our gratitude to Associate Professor Mohammad Shekaramiz for his oversight and management of the project (USHE Grant 20210016UT) prior to his untimely passing in December 2024. We would also like to thank Kaden Clements and Nathan Archer for their assistance in creating the annotated dataset described in Section III.

REFERENCES

[1] Lee, J.; Zhao, F. Global Wind Energy Council. Global Wind Report 2024. Available online: https://www.gwec.net/reports/globalwindreport (accessed on 29 March 2025).
[2] Katsaprakakis, D.A.; Papadakis, N.; Ntintakis, I. A Comprehensive Analysis of Wind Turbine Blade Damage. Energies **2021**, 14, 5974. https://doi.org/10.3390/en14185974
[3] Wang, W.; Xue, Y.; He, C.; Zhao, Y. Review of the Typical Damage and Damage-Detection Methods of Large Wind Turbine Blades. Energies **2022**, 15, 5672. https://doi.org/10.3390/en15155672
[4] GEV Wind Power. 2024 Wind Turbine Blade Inspection. Available online: https://www.gevwindpower.com/blade-inspection/ (accessed on 2024-12-10).
[5] Kabbabe Poleo, K.; Crowther, W. J.; Barnes, M. Estimating the Impact of Drone-Based Inspection on the Levelised Cost of Electricity for Offshore Wind Farms. Results in Engineering 2021, 9, 100201. https://doi.org/10.1016/j.rineng.2021.100201.
[6] Altice, B.; Nazario, E.; Davis, M.; Shekaramiz, M.; Moon, T.K.; Masoum, M.A.S. Anomaly Detection on Small Wind Turbine Blades Using Deep Learning Algorithms. Energies 2024, 17, 982. https://doi.org/10.3390/en17050982

[7] Zhang, J.; Cosma, G.; Watkins, J. Image Enhanced Mask R-CNN: A Deep Learning Pipeline with New Evaluation Measures for Wind Turbine Blade Defect Detection and Classification. J. Imaging **2021**, 7, 46. https://doi.org/10.3390/jimaging7030046

[8] Zhang, J.; Cosma, G.; Watkins, J. Image Enhanced Mask R-CNN: A Deep Learning Pipeline with New Evaluation Measures for Wind Turbine Blade Defect Detection and Classification. J. Imaging **2021**, 7, 46. https://doi.org/10.3390/jimaging7030046

[9] Pratt, R.; Allen, C.; Masoum, M.A.S.; Seibi, A. Defect Detection and Classification on Wind Turbine Blades Using Deep Learning with Fuzzy Voting. Machines, **2025**, 13, 283. https://doi.org/10.3390/machines13040283,

[10] Ko, H.; Ha, H.; Cho, H.; Seo, K.; Lee, J. Pneumonia Detection with Weighted Voting Ensemble of CNN Models. Proceedings of the 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), 2019, pp. 306–310. doi:10.1109/ICAIBD.2019.8837042.

[11] Zhang, R.; Cheng, C.; Zhao, X.; Li, X. Multiscale Mask R-CNN–Based Lung Tumor Detection Using PET Imaging. Molecular Imaging **2019**, 18, 1536012119863531. https://doi.org/10.1177/1536012119863531

[12] He, K.; Gkioxari, G.; Dolla´r, P.; Girshick, R. B. Mask R-CNN. arXiv 2017, arXiv:1703.06870 https://doi.org/10.48550/arXiv.1703.06870

[13] Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. arXiv **2016**, arXiv:1506.01497 https://doi.org/10.48550/arXiv.1506.01497.

[14] Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-scale Image Recognition. arXiv **2014**, arXiv:1409.1556.

[15] Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807. https://doi.org/10.1109/CVPR.2017.195.

[16] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. arXiv **2015**, arXiv:1512.03385, https://doi.org/10.48550/ arXiv.1512.03385

[17] Ward, Z.; Miller, J.; Engel, J.; Masoum, M.A.S.; Shekaramiz, M.; Seibi, A. Fuzzy-Based Image Contrast Enhancement for Wind Turbine Detection: A Case Study Using Visual Geometry Group Model 19, Xception, and Support Vector Machines. Machines 2024, 12, 55. https://doi.org/10.3390/machines12010055

[18] Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. Nature **2020**, 585, 357–362. https://doi.org/10.1038/s41586-020-2649-2.