Paper Presentation

Masud Ahmed

Introduction

Computer Vision

Shihavuddin, A. S. M., Xiao Chen, Vladimir Fedorov, Anders Nymark Christensen, Nicolai Andre Brogaard Riis, Ki Branner, Anders Bjorholm Dahl, and Rasmus Reinhold Paulsen. "Wind Turbine Surface Damage Detection by Deep Learning Aided Drone Inspection Analysis." MDPI *Energies* 12, no. 4 (2019). Technical University of Denmark (DTU)

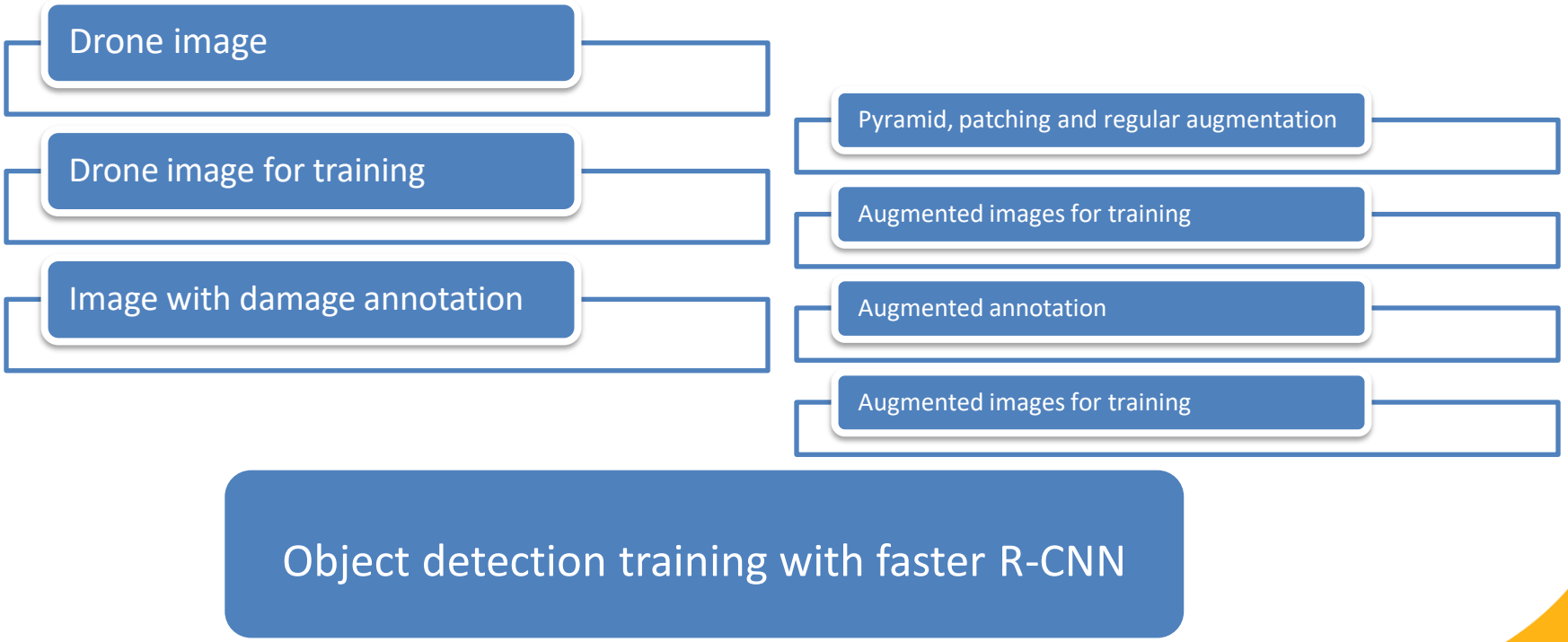
Why Drone?

- Low-cost and frequent inspections
- High-resolution optical image acquisition
- Minimal human intervention

Main Contribution

1. Automated suggestion system for damage detection in drone inspection images
2. Higher precision in the suggestion model achieved through advanced data augmentation
3. Publication of the wind turbine inspection dataset

Basic Structure

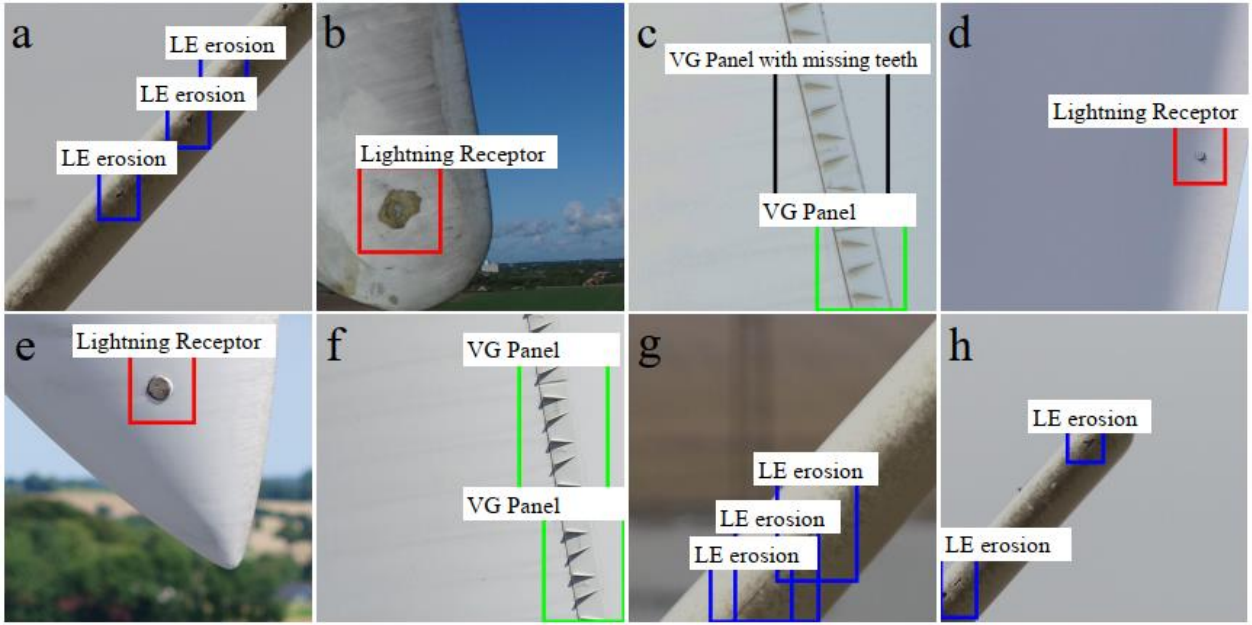


UMBC

Raw Data

Types of Damage

1. Leading Edge (LE) erosion
2. Vortex Generator panel (VG)
3. Vortex Generator with Missing Teeth (VGMT)
4. Lightning Receptor (LR)



Instance Number

	Name	Train Instant	Test Instant
1.	Leading Edge (LE) erosion	135	67
2.	Vortex Generator panel (VG)	126	65
3.	Vortex Generator with Missing Teeth (VGMT)	27	14
4.	Lightning Receptor (LR)	17	7

Result in Raw data

CNN Model	LE	VG	VGMT	LR	All
Inception-V2	20.34	13.15	4.21	41.39	19.77
ResNet-50	37.51	17.54	7.02	40.88	25.74
ResNet-101	29.61	20.29	4.21	49.33	25.86
Inception-ResNet-V2	31.54	18.35	5.96	54.14	27.50

Image Augmentation

1. Regular Augmentations
2. Pyramid Augmentations
3. Patching Augmentations

Regular Augmentations

- Perspective transformation for the camera angle variation simulation
- Left-to-right flip or top-to-bottom flip to simulate blade orientation: e.g., pointing up or down
- Contrast normalization for variations in lighting conditions
- Gaussian blur simulating out of focus images

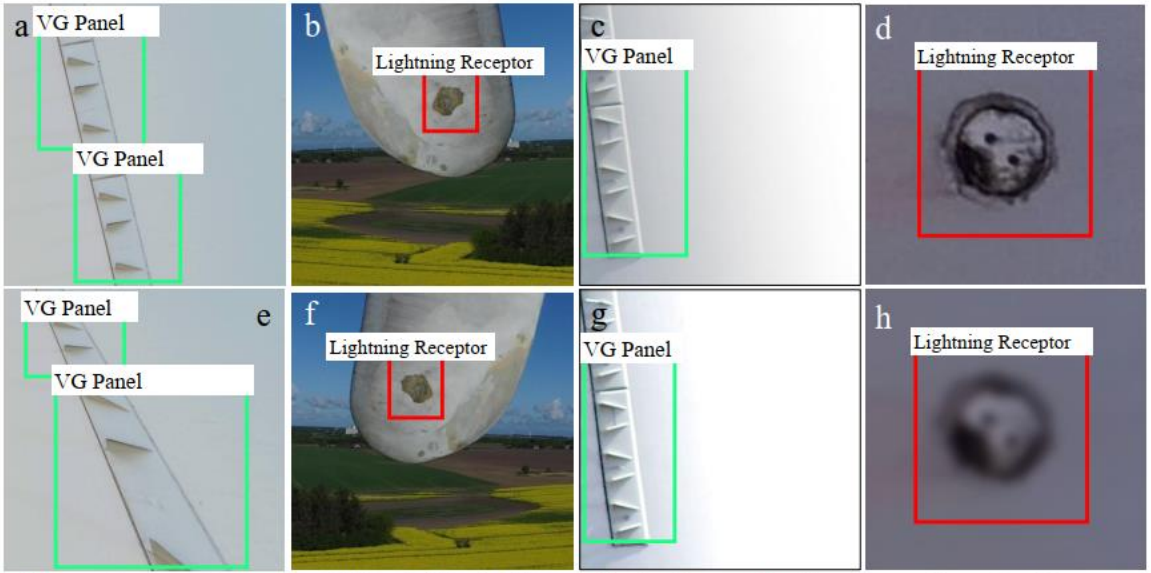
Pyramid Augmentations

- Images are scaled from 1.00× to 0.33×, simulating from the highest to the lowest resolutions
- Resolution conversions were performed through the linear interpolation method

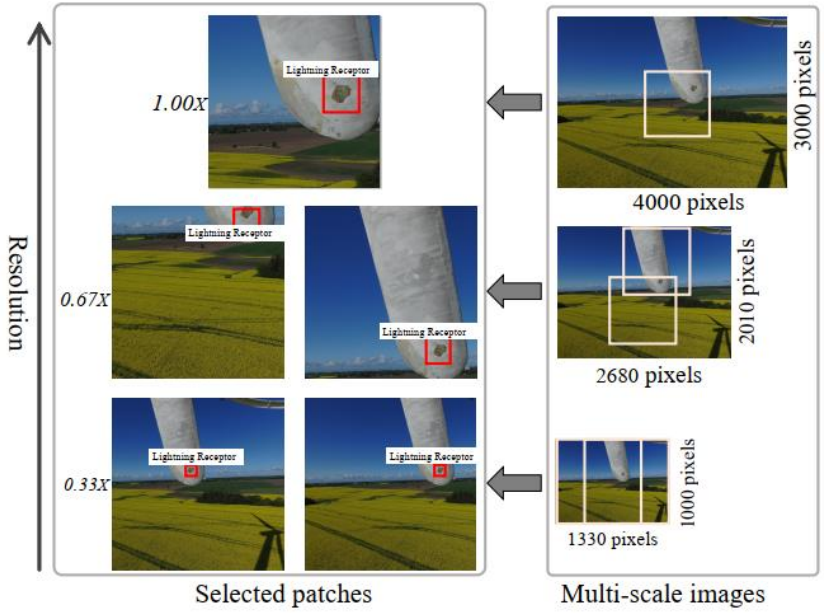
Patching Augmentations

- The acquired full resolution image is 4000 × 3000 pixels, where the lightning receptor only occupies around 100 × 100 pixels
- When fed to CNN during training in full resolution, it would be resized to the pre-defined network input size, where the lightning receptor would be occupying a tiny portion of 33 × 33 pixels

Augmentation Example



Augmentation Example



Deep Learning Network

- Crack detection performed using faster R-CNN with pretrained model:
  1. Inception-V2
  2. ResNet-50
  3. ResNet-101
  4. Inception-ResNet-V2

Evaluation Process

PC = Per class precision for each image
APC = Precision over all the images in the dataset

And Threshold is 0.3

$$\text{IoU} = \frac{P \cap GT}{P \cup GT}$$

$$\text{P}_C = \frac{N(\text{True Positives})_C}{N(\text{Total Objects})_C}$$

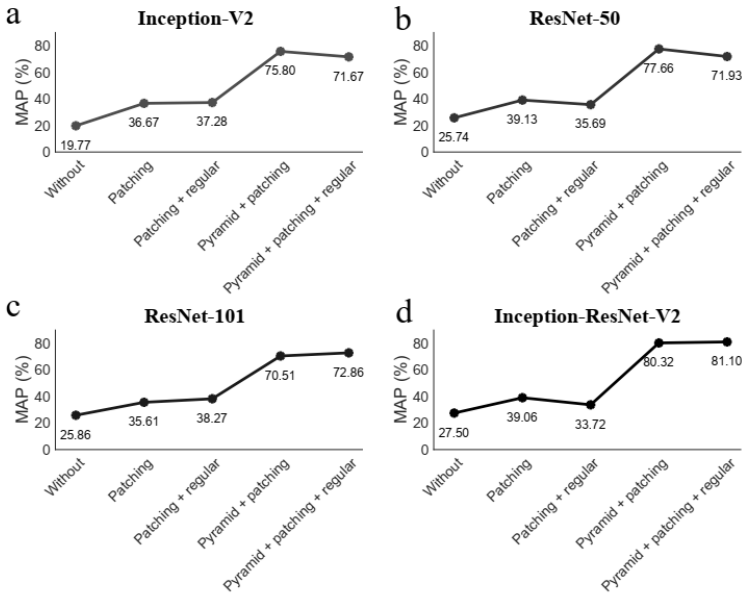
$$\text{AP}_C = \frac{\sum \text{P}_C}{N(\text{Total Images})_C}$$

$$\text{MAP} = \frac{\sum \text{AP}_C}{N(\text{Classes})}$$

Result

CNN	Augmentation	MAP (%)				
		LE Erosion	VG	VGMT	Lightning Receptor	All
Inception-V2	Without	20.34	13.15	4.21	41.39	19.77
	Patching	37.98	53.86	32.74	22.12	36.67
	Patching + regular	36.90	52.58	36.08	23.54	37.28
	Pyramid + patching	88.11	86.61	58.28	70.18	75.80
	Pyramid + patching + regular	89.94	84.15	71.85	40.76	71.67
ResNet-50	Without	37.51	17.54	7.02	40.88	25.74
	Patching	34.48	54.92	32.17	34.96	39.13
	Patching + regular	35.06	47.87	24.88	34.96	35.69
	Pyramid + patching	90.19	85.15	61.43	73.85	77.66
	Pyramid + patching + regular	90.47	88.84	35.27	73.15	71.93
ResNet-101	Without	29.61	20.29	4.21	49.33	25.86
	Patching	34.79	47.35	25.35	34.96	35.61
	Patching + regular	36.06	51.41	32.16	33.46	38.27
	Pyramid + patching	88.86	87.46	41.53	64.19	70.51
	Pyramid + patching + regular	86.77	86.15	41.53	77.00	72.86
Inception-ResNet-V2	Without	31.54	18.35	5.96	54.14	27.50
	Patching	36.30	44.43	33.37	42.12	39.06
	Patching + regular	32.73	47.59	19.58	34.96	33.72
	Pyramid + patching	90.22	91.66	69.85	69.56	80.32
	Pyramid + patching + regular	90.62	89.08	73.11	71.58	81.10

Result Visualization



Summary

Table 3. Summary of the experimental results.

	Mean Average Precision (MAP) in Percentage		
Classes	**Suggestions**	**Human**	**Suggestions Aiding Human**
LE erosion	90.62	68.00	86.60
VG panel (VG)	89.08	85.91	81.57
VG with Missing Teeth (VGMT)	73.11	80.17	91.37
Lightning receptor	71.58	98.74	81.71
Overall precision	81.10	83.20	85.31
Average processing time per image	2.11 s	200 s	131 s

Conclusion

Explore image augmentation for data scarcity

Dataset Link:
https://data.mendeley.com/datasets/hd96prn3nc/2

Github link for image augmentation:
https://github.com/aleju/imgaug

THANK YOU