FACULTY OF SCIENCE AND TECHNOLOGY

MASTER THESIS

Study programme / specialisation:

Industrial Asset Management

The spring semester, 2023

Open

Author:

Can Özkan

....
(signature author)

Supervisor(s):

Idriss El-Thalji, UiS

Thesis title:

Computer Vision in Wind Turbine Blade Inspections: An Analysis of Resolution Impact on Detection and Classification of Leading-Edge Erosion

Credits (ECTS): 30

Keywords: Leading Edge Erosion, Predictive Maintenance, Wind Turbine Maintenance, Object Detection, Damage Classification, Computer Vision, Image Resolution, Machine Learning, Blade Inspections, Drone Imagery, Industrial Asset Management

Pages: 58

+ appendix: 8

Stavanger, June 2023

Computer Vision in Wind Turbine Blade Inspections: An Analysis of Resolution Impact on Detection and Classification of Leading-Edge Erosion

By

Can Özkan

Thesis is submitted to the Faculty of Science and Technology
University of Stavanger
In Fulfillment of the Requirements for the degree of
Master of Science (MSc) in Industrial Asset Management

University of Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

University of Stavanger

2023

Abstract

Wind turbines, as critical components of the renewable energy industry, present unique maintenance challenges, particularly in remote or challenging locations such as offshore wind farms. These are amplified in the inspection of leading-edge erosion on wind turbine blades, a task still largely reliant on traditional methods. Emerging technologies like computer vision and object detection offer promising avenues for enhancing inspections, potentially reducing operational costs and human-associated risks. However, variability in image resolution, a critical factor for these technologies, remains a largely underexplored aspect in the wind energy context.

This study explores the application of machine learning in detecting and categorizing leading edge erosion damage on wind turbine blades. YOLOv7, a state-of-the-art object detection model, is trained with a custom dataset consisting of images displaying various forms of leading edge erosion, representing multiple categories of damage severity. Trained model is tested on images acquired with three different tools, each providing images with a different resolution. The effect of image resolution on the performance of the custom object detection model is examined. The research affirms that the YOLOv7 model performs exceptionally well in identifying the most severe types of LEE damage, usually classified as Category 3, characterized by distinct visual features. However, the model's ability to detect less severe damage, namely Category 1 and 2, which are crucial for early detection and preventive measures, exhibits room for improvement.

The findings point to a potential correlation between input image resolution and detection confidence in the context of wind turbine maintenance. These results stress the need for high-resolution images, leading to a discussion on the selection of appropriate imaging hardware and the creation of machine learning-ready datasets. The study thereby emphasizes the importance of industry-wide efforts to compile standardized image datasets and the potential impact of machine learning techniques on the efficiency of visual inspections and maintenance strategies. Future directions are proposed with the ultimate aim of enhancing the application of artificial intelligence in wind energy maintenance and management, enabling more efficient and effective operational procedures, and driving the industry towards a more sustainable future.

Acknowledgements

I would like to extend my sincerest gratitude to my thesis supervisor, Professor Idriss El-Thalji, for his guidance, patience, and expert advice throughout this research process. His support and presence have made this journey a rewarding experience.

Further, my heartfelt appreciation goes to Professor Andreas Delimitis, who has provided me with invaluable feedback and insight at every stage of this project. His helpful attitude and consistent encouragement were very valuable to accomplishing this research.

Likewise, I would like to express my deepest appreciation to Andreas Skaare, who has kindly provided unique value and perspective to the whole project. His cooperation has led to the addition of a new angle to the study and has made the execution of this project much smoother.

I would like to offer special thanks and gratitude to Frank Emil Moen, CEO of Energy Innovation AS. His willingness to provide the equipment and unique insight into the wind industry has essentially made this project possible. I also wish to acknowledge the team at Energy Innovation AS that were involved with the project at multiple stages, and the company as a whole for their valuable collaboration with our program and faculty. Their collective support and assistance have been significant to my work.

I would like to also congratulate my fellow researchers, students and friends in the Industrial Asset Management program at University of Stavanger, thank you for your feedback and company during this process. Your support and friendship made this journey an enjoyable one.

Lastly, I wish to express my profound gratitude to my friends, family and my partner. Your unwavering support, understanding, and encouragement have been my source of strength throughout my academic journey. Your faith in my abilities kept me motivated, and for that, I am forever thankful.

This journey has been enriched and enlightened by all your contributions. I am deeply thankful to you all.

Can Özkan
Stavanger, June 2023

Table of Contents

List of figures

List of appendices

List of abbreviations

LEE	Leading Edge Erosion
UAV	Unmanned Aerial Vehicle
O&M	Operations and Maintenance
CoE	Cost of Energy
WT	Wind Turbine
WTB	Wind Turbine Blade
CMS	Condition Monitoring System
SHM	Structural Health Monitoring
NDT	Non-destructive Testing
VI	Visual Inspection
AEP	Annual Energy Production
SCADA	Supervisory Control and Data Acquisition
CNN	Convolutional Neural Network
R-CNN	Region-based Convolutional Neural Network
YOLO	You Only Look Once
MS	Microsoft
COCO	Common Objects in Context
MW	Megawatt
GPU	Graphics Processing Unit
mAP	Mean Average Precision
IoU	Intersection over Union
FN	False Negative
FP	False Positive
LiDAR	Light Detection and Ranging

1. Introduction

Management and maintenance of industrial assets plays a significant part in achieving optimal levels of efficiency in critical energy infrastructure, such as wind turbine systems. Effective maintenance practices are essential in sustaining reliability across operational time periods, considering that these systems are exposed to a variety of environmental conditions that eventually reduce or compromise their effectiveness or lifespan to alarming levels. One such issue is leading edge erosion (LEE), which, if unchecked, can affect structural integrity as well as appropriate aerodynamic properties of the blades. It is crucial to identify and evaluate LEE proactively with emphasis on accuracy in order to enable optimal maintenance planning and reduce operational costs.

The way that inspections and maintenance activities are carried out has the potential to change and adapt substantially as a result of the integration of computer vision technologies into industrial asset management. In particular, its application to assessing leading edge erosion (LEE) damage to wind turbine blades presents a promising possibility of conducting reviews more precisely and rapidly. The quality and resolution of the input visuals, however, play an important role on how well these strategies work. This thesis explores the impact of image resolution variances from different sources on the precision of LEE damage identification and categorization using a computer vision algorithm. Ultimately, the aim is to strengthen asset management procedures for maintenance planning.

Three primary objectives guide this study and determine its methodology. They include:

- Evaluating how effectively computer vision technologies detect and classify LEE damage on wind turbine blades utilizing multiple datasets.

- Examining the effects of several picture resolution/quality levels when analyzing LEE damage.

- Determining the most useful sources of images in determining different cases of LEE.

1.1. Background and Problem Presentation

It is increasingly becoming essential to include wind energy as a key component in order to effectively meet the goals set for the production of renewable energy on a worldwide scale. Managing the effectiveness and efficiency of wind turbine systems through appropriate asset management procedures is an important part of this change. Leading-edge erosion (LEE) on the turbines' blades, however, is one of the primary obstacles encountered throughout these maintenance procedures. Failures resulting from erosion damages could lead to decreased aerodynamic capabilities and limited lifespan for these assets. Therefore, early detection of LEE is crucial for efficient maintenance planning with the objective at minimizing operational expenses.

Despite the importance of detecting and evaluating LEE, traditional methods of inspection can be time-consuming, labor-intensive, and often expose maintenance personnel to safety risks. For instance, manual visual inspections require workers to physically access turbine blades, sometimes in challenging weather conditions or at great heights.

The emergence of computer vision technology offers a promising alternative to these conventional methods. Computer vision can enable the automation of LEE damage identification and categorization with the use of machine learning techniques on high resolution imagery. The accuracy, effectiveness, and safety of wind turbine inspections are prone to improvement with such applications.

Within the potential that computer vision brings, one major component of successful application on maintenance cases is the resolution of the input images. The accuracy of LEE detection is open to be influenced to a certain degree by the quality of the images to be analyzed. The focus, then, lies on how the performance of computer vision algorithms might vary with different resolutions.

From cameras embedded on drones to everyday smartphones, given the variety of imaging devices available, there is a need to understand which sources of images are most useful in detecting and categorizing levels of LEE. In this study, a Ryze Tello drone with a built-in lens, an iPhone, and a consumer-grade digital camera will be utilized to capture images of a 2.5-meter blade section with artificially recreated LEE damages.

Overall, the problem this thesis aims to address is the gap in understanding about the impact of image resolution variances on the effectiveness of computer vision algorithms in identifying and categorizing LEE damages. Successful understanding could therefore support the selection of imaging devices for wind turbine inspections and provide a novel approach to maintenance applications in the field.

1.2. Research Objectives and Relevance

The study looks at optimal input equipment for this application as well as how image resolution affects a computer vision algorithm's ability to detect and categorize LEE damage. The goals defined to be achieved are:

- Determine the impact of using different resolutions collected from different sources for identifying and classifying LEE damage on wind turbine blades.
- Establish the ideal level of resolution that produces optimal results with the given inputs, while evaluating how variation affects performance ratings.
- Evaluate how well various imaging tools perform when it comes to capturing images that meet the criteria for carrying out computerized analyses of LEE damage.

The increasing focus on wind energy on a global scale, and the use of computer vision for asset management provides the basis for the study's significance. This research will contribute to the knowledge-base regarding how image resolution affects computer vision algorithms, which will ultimately assist stakeholders choose the best imaging systems to enhance the effectiveness and efficiency of wind turbine blade inspections.

The results of the study can also benefit fields beyond wind turbine maintenance. It can inform researchers on how to use computer vision technologies in comparable circumstances where inspections are important, such as in construction, infrastructure, or manufacturing quality control. From this perspective, the overall understanding of industrial computer vision applications can be expanded.

1.3. Research Question

The primary research question that drives the work completed in this thesis is:

"Does image resolution have an impact on the performance of customized object detection models within wind turbine blade inspections?"

This research aims to explore the link between the resolution of input images fed to computer vision algorithms and the level of accuracy achieved by specific applications in generating outputs. Due to its presence at the intersection of two areas that are quickly evolving, renewable energy and artificial intelligence, this particular focus carries significant weight. Its prospective effects on visual maintenance initiatives and asset management techniques, particularly in the wind energy sector, are particularly important.

By answering the question defined above, the study seeks to contribute novel perspectives to the body of knowledge that exists on computer vision applications and to reveal insightful information that can improve the general effectiveness of maintenance procedures across a variety of fields.

1.4. Methodology

The methodology of this study is built on an interdisciplinary approach, combining elements of physical operations, experimental procedures, and computational analysis to explore the impact of image resolution on the identification and categorization of leading edge erosion (LEE) in wind turbine blades using a computer vision algorithm.

The physical component of this research involves the utilization of a wind turbine blade section that is subjected to controlled erosion to simulate different stages of LEE damage. The blade serves as the physical representation of real-world conditions and provides the base for the collection of image data. Furthermore, various imaging devices, including a drone, a smartphone, and a digital camera, are used in capturing these representation of real-world conditions for further computational analysis.

Experimental procedures in this study encompass the controlled capture of image data from the eroded blade model using the selected imaging tools. These procedures are performed under uniform conditions to ensure the consistency of the data collected.

The core of the computational analysis involves the application of a computer vision algorithm to the collected image data. This algorithm, selected based on its relevance and efficiency in handling object detection tasks, serves to process the images, detect and categorize the LEE damage. The impact of image resolution on the performance of the algorithm is evaluated by comparing the algorithm's output with the ground truth data. Additionally, the computational part of the methodology extends to the evaluation of the imaging tools, comparing their effectiveness in capturing images that lead to the successful detection and classification of LEE. It also involves the analysis of the results, which will help understand the relation between image resolution and the performance of the computer vision algorithm.

This methodology's structure combines physical, experimental, and computational elements to offer a comprehensive approach for addressing the research question, providing an innovative insight into the interaction between image resolution and computer vision technology in the context of wind turbine maintenance.

1.5. Thesis Scope

This thesis' main objective is to investigate how variations in picture resolution might impact how well a computer vision system works to identify and classify leading edge erosion (LEE) damage on wind turbine blades. The goal is to investigate how variations in picture resolution affect the accuracy levels reached by computer vision algorithms and to determine the imaging equipment that is most suited for precisely recognizing and characterizing LEE.

Through the controlled erosion of a blade's tip portion by 2.5 meters in length, it will be possible to examine LEE damage in the context of wind turbine blades. Three different photo-taking tools are used in order to get clear photographs for analysis: a Ryze Tello drone, a smartphone camera, and a consumer-grade digital camera. The computer vision algorithm will be equipped through the use of the most recent computer vision model appropriate especially for this application and relevant machine learning techniques.

The fundamental restriction on the scope of this thesis is to build upon current computer vision models. The goal of this project does not entail creating a novel computer vision model or a commercial application. Instead, it focuses on using an existing computational technique to handle particular challenges associated with wind turbine maintenance tasks. The direct application of the findings to industrial settings is outside the scope of this study, which instead intends to establish relevant foundational information for directing future research and potential implementation strategies for stakeholders.

Finally, it is important to point out that, while there is a possibility of this study encouraging and having consequences on other fields utilizing computer vision technology, its primary focus is on using the techniques to address erosion at the leading edge of wind turbine blades. The study does not intend to thoroughly investigate potential uses in other industries.

1.6. Thesis Structure

The thesis is structured into seven primary sections that are allocated to a significant area of the research undertaken:

1) **Introduction**: This section opens with an outline of the subjects covered in the thesis. The background of the issue, the goals of the research, its importance, and the methods used are all briefly addressed. Then, a summary of this thesis's structure is given while highlighting its range of application.

2) **Theoretical Background**: The second chapter delves into existing literature and past research in the areas of wind turbine maintenance, leading edge erosion (LEE), and computer vision technology. It establishes the current understanding in these fields and identifies the gaps this research aims to fill.

3) **Methodology:** This chapter of the work aims to provide a further detailed overview in how the research and experiment was designed and executed. The creation of artificial erosion, the process of image capture through different devices, the configuration of the computer vision application, and finally the analysis of the findings will be discussed.

4) **Results:** The fifth chapter provides the output directly from the computer vision algorithm. It details the performance of the algorithm in identifying and categorizing LEE damage using images of different resolutions.

5) **Discussion:** This section provides a further in-depth analysis of the results obtained from the computer vision algorithm, while touching upon what the output indicates. It discusses the impact of image resolution variation on the effectiveness of the algorithm, and the performance of the different imaging devices used. A methodical comparison between the devices is provided, before drawing final conclusions.

6) **Conclusion and Future Work:** With the final chapter, the research is concluded while providing recommendations for future research. The main findings of the experiment are summarized and the study’s implications are discussed. Looking back on the overall achievement of the work; potential areas for improvement, suggestions on future research and the ways in which the study can be extended upon finalizes the paper.

2. Theoretical Background

This section provides a comprehensive exploration of the theoretical principles associated with wind energy generation, particularly focusing on the maintenance of wind turbine blades. It delves into the significance of maintenance strategies, the unique challenges encountered with wind turbine deployment in diverse locations, and the potential role of advanced technology in revolutionizing maintenance procedures. The discussion also considers the competitive dynamics of the industry and their implications for cooperative efforts aimed at enhancing reliability. The emergence of innovative technologies such as unmanned aerial vehicles (UAVs) and computer vision, and their potential impacts will be assessed, including the limitations of implementing these technologies, such as the influence of input resolution on the effectiveness of object detection models, within the context of wind turbine inspections. Finally, this fundamental theoretical groundwork will guide the study in the following sections, through the identification of knowledge gaps in the existing literature.

2.1. Wind Turbine Operations and Maintenance

2.1.1. Reliability, Costs and Failures

The emergence of wind power generation as a competitive alternative and the increasing pace of technological developments on wind turbines over different scales have emphasized the importance of effective operation and maintenance (O&M) practices. The reliability and efficiency of these large-scale turbines is a multifaceted issue that has a significant impact on the total cost of energy (COE) from wind power projects. Substantial research on wind turbine O&M have provided valuable insights, and present key challenges within the industry.

In their research, Walford (2006) emphasizes the importance of reliability for a project's revenue stream, points out that the reliability of wind turbines varies with the operating environment and is influenced by factors such as design assumptions, knowledge of the operating environment, and manufacturing quality control. Its highlighted that O&M costs can represent up to 20% of a project's total cost of energy (COE), with unscheduled maintenance costs accounting for 30%-60% of total O&M costs, thereby emphasizing the necessity for improved certainty in the O&M cost estimations (Walford, 2006). The study suggests that effective maintenance programs, including comprehensive condition monitoring and early identification of critical components and failure modes, can help optimize costs and prevent catastrophic failures.

Pinar Pérez et al. (2013) underscores the fact that larger turbines tend to fail more frequently than smaller ones, necessitating the implementation of robust condition monitoring systems to enhance their reliability. Their paper also details the structural and functional aspects of WTs and considers the need for sophisticated maintenance systems due to the high costs of WT machinery and infrastructure. Finally, the authors define the increasing failure rates of certain components, such as rotor blades and gearboxes, stressing the necessity of predictive and preventive maintenance strategies, especially for larger wind turbines.

Artigao et al. (2018) provide a detailed overview of the reliability analysis of wind turbines, drawing from thirteen different studies in the domain. Their work underscores the shared crucial parts between onshore and offshore uses, and also points out the variability in failure rates among different geographical areas. The authors emphasize the critical role of Operation and Maintenance (O&M) activities in enhancing turbine availability, with O&M costs constituting a substantial part of wind farm project expenditures. They point out that larger wind turbines, despite their benefits, show higher failure rates than smaller ones, underscoring the need for efficient condition monitoring. The research further unveils that condition monitoring systems (CMS) help optimize preventive and corrective maintenance, preventing unnecessary repairs and unplanned downtime. The study also underscores the significant disparities in O&M strategies between onshore and offshore wind turbines, primarily due to accessibility issues and cost variations (Artigao et al., 2018). Ultimately, this review determines that the electric and control systems, the gearbox, the generator, and the hub & blades are the most critical assemblies needing attention in CMS design, suggesting that condition-based maintenance could improve wind turbine availability and cost-effectiveness (Artigao et al., 2018).

Carroll et al. (2016) present an exhaustive analysis of failure rates, repair times, and unscheduled operations and maintenance (O&M) costs based on offshore wind turbines. Drawing from a sample of approximately 350 offshore wind turbines spanning 5 to 10 wind farms across Europe, the authors provide unique insight into failure rates of the overall turbine and its various sub-assemblies, repair times, costs, and resource requirements, and reveal that reliability and maintenance resources can constitute around 30% of the overall energy cost, with blades accounting for 6.2% of the overall failure rates (Carroll et al., 2016). Notably, their analysis identifies that offshore turbines in high wind speed areas exhibit higher failure rates. The researchers suggest that this correlation is stronger offshore than onshore, and it could be attributed to factors such as the harsher marine environment, the larger size of offshore turbines, and potential differences in maintenance standards due to accessibility (Carroll et al., 2016). This study is instrumental in understanding and optimizing O&M cost modeling, ultimately aiding in the decision-making process for O&M planners and managers.

Turbine blades are a particularly crucial component, as they contribute to for 15–20% of overall turbine costs (Ciang et al., 2008). The blades' structural health is, therefore, of primary concern due to the expensive and time-consuming nature of repairs. Ciang et al. (2008) also highlight that minor blade damage can lead to significant secondary damage to the entire wind turbine system, which could result in the collapse of the whole tower if prompt repair action isn't taken. Consequently, the study suggests regular monitoring of the blades is essential to ensure early detection and repair of potential damage, effectively reducing the total cost of repair and preventing more serious damage.

Taken together, these studies emphasize the need for improving wind turbine reliability, optimizing O&M strategies, and utilizing data and predictive analytics. They also underline the need to focus on the maintenance of wind turbine blades due to their significant contribution to turbine costs and overall cost of energy. Enhancements in public databases, structural health monitoring (SHM) systems, and O&M cost reductions are all necessary steps towards improving wind turbine reliability and efficiency, which is essential for the long-term viability and sustainability of wind power generation.

2.1.2. Damage Detection and Repairs

The growing recognition of the importance of wind turbine operation and maintenance, particularly in regards to reliability, costs, and failures, has inspired the development of innovative damage detection and assessment technologies. Du et al. (2020) and Zhang et al. (2020) emphasize the urgency of early damage detection in wind turbine blades, driven by escalating maintenance costs associated with the increasing size and complexity of these structures. McGugan and Mishnaevsky (2020) introduce a novel approach to the structural health monitoring of wind turbine blades, focusing on various physical degradation mechanisms such as surface erosion, adhesive fatigue, and laminate cracking, among others.

Mishnaevsky (2019) highlights the critical task of wind turbine blade repair in the advancement of renewable energy technologies. In a comprehensive review, the research underscores the significance of blade repair for the progression of renewable energy technologies, given that an out-of-service turbine can be exceedingly costly, with repair fees ranging from $800 to $1600 per day (Mishnaevsky, 2019). The study stresses the importance of improving and optimizing repair methodologies to reduce costs, shorten repair times, and ensure that repaired structures maintain initial performance levels.

Xu et al. (2019) provide an innovative perspective on blade surface inspection by proposing a method that utilizes deep learning and unmanned aerial vehicles (UAVs), treating blade inspection as an image recognition task. This approach is intended to overcome the limitations of traditional methods, such as stability issues, sensor installation challenges, and difficulties in data storage and processing (Xu et al., 2019). Utilizing UAVs for image acquisition, they propose improved efficiency and ensured personnel safety.

Stokkeland et al. (2019) expands on the use of UAVs in wind turbine inspection by exploring the autonomous visual navigation of UAVs. They illustrate how UAVs, especially when autonomous or remotely controlled, can approach inspection targets closely and accurately, given the large dimensions of modern wind turbines. The cost-effectiveness of UAVs over manual inspection by climbing, particularly for offshore wind turbines, was also highlighted.

Tchakoua et al. (2012) contribute to the discourse on wind turbine condition monitoring by underscoring the necessity for remote, intelligent, and integrated systems, particularly as the wind energy industry leans towards larger and more remotely located wind turbines. A key focus of the authors' work is discussing emerging trends such as non-contact and remote non-destructive testing (NDT) methods, and the automation of condition monitoring and diagnostic systems. They also discuss the role of visual inspection (VI) as a complementary method in condition monitoring, although traditional methods rely on human intervention and can benefit from further research.

Together, the insights from literature create a compelling narrative for the future of wind turbine operations. The integration of machine vision, UAVs, artificial intelligence, localized nanoscale sensors for damage detection and assessment, and advanced computational modeling presents significant potential for cost savings, reduced human risks, and enhanced operational efficiency. Despite challenges such as capturing high-definition images, extracting damage information from complex backgrounds, managing environmental factors, and ensuring that sensor deployment does not weaken the blade structure, these technological advancements signal a promising future for wind turbine maintenance and operation. Continual research and technological innovation remain vital to fully exploit these opportunities and tackle the remaining challenges.

2.2. Leading Edge Erosion

2.2.1. Impact on Energy Production

Surface material degradation at the forward facing edge of wind turbine blades, also known as leading edge erosion (LEE), is a significant issue in wind energy studies. This deterioration, brought about by elements such as wind and rain, primarily concerns researchers due to its pronounced effect on both the energy output and the structural soundness of the wind turbine blades. A common attention within the literature is given to accurately estimating the impact of LEE on annual energy production (AEP) over a turbine's lifetime, given the associated maintenance costs and lost AEP (Herring et al., 2019; Law and Koutsos, 2020; Sareen et al., 2014). The transition of the wind industry towards larger blade lengths, higher tip speeds, and new markets characterized by harsh climatic conditions, has further brought forward the issue of leading edge erosion (Herring et al., 2019). Moreover, the current aggressive expansion of the offshore wind industry, coupled with the persistent lack of a permanent solution that offers protection on the leading edge for blades' dedicated lifetime of 25-years on average, further underscores the importance of addressing LEE (Herring et al., 2019).

Understanding the complexity of leading edge erosion is not a recent effort literature. Keegan et al. (2013) conducted an in-depth examination of the potential wear and tear caused by various environmental factors, with a special emphasis on the effects of raindrop and hailstone collisions on the leading edge of the blade. The researchers studied various methods and resources like weather and climate information, lab-controlled rain and hailstone exposure tests, and computational modeling techniques to assess and alleviate the threats linked with leading edge erosion. Their review clearly showed that overcoming issues related to leading edge erosion and crafting superior materials that can withstand these conditions is a considerable challenge (Keegan et al., 2013). This difficulty is compounded by the various environmental elements involved and the increasing size of contemporary wind turbine blades.

Mishnaevsky et al. (2021) focused on the issue of leading edge erosion (LEE) in wind turbine blades, underlining the serious negative effects of LEE on the blades' aerodynamic efficacy. The study stressed that severe erosion could cut annual energy output by more than 5%

(Mishnaevsky et al., 2021), and also observed that geographical differences substantially influence the way erosion processes occur.

Contreras López et al. (2023) suggested a computational model capable of predicting the progression of leading edge erosion and its impact on energy output over time, which can be applied in operation and maintenance decision making. They tested this model using a 5MW wind turbine in the North Sea as an example, and the results showed that the greatest annual energy production losses ranged between 1.6 and 1.75% (Contreras López et al., 2023).

The impact of leading edge erosion on the aerodynamic properties of a wind turbine was explored by Sareen et al (2014). Their results suggest that even a relatively trivial stage of leading edge erosion leading to an 80% increase in drag, has to potential to decrease annual energy production (AEP) by 5%, and in more severe cases of LEE, this loss could be as high as 25%. (Sareen et al., 2014) Their research underpins the urgency of addressing LEE and devising mitigation strategies.

LEE negatively impacts the performance of turbines and necessitates expensive repairs. Over time, it has evolved from a challenge limited to a few turbines in harsh conditions to a widespread problem that impacts whole wind farms, even those located in relatively mild climates (Herring et al., 2019). Consequently, leading edge erosion is now one of the most significant concerns in the wind industry, a point also underscored by Duthé et al. (2021) in their discussion of the detrimental effects of LEE on power performance and the functionality of blades.

Moreover, the nature and evolution of leading edge erosion, influenced by variables such as the speed of the tip, the composition and form of the blade, and the surrounding environment, complicates the task of both comparing erosion rates across various blades and establishing an accurate timeline for its progression (Herring et al., 2019).. The need for early detection and intervention is crucial given this context, as noted by Mishnaevsky et al. (2021) in their discussion of the importance of predicting erosion and setting the frequency of control and maintenance events.

Existing diagnostic techniques predominantly rely on direct visual inspection, and statistical analysis of supervisory control and data acquisition (SCADA) output (Duthé et al., 2021). However, research suggests these techniques have not been effective in terms of accurately determining and standardizing the severity and spatial extent of erosion on the blades. Meanwhile, the increasing use of drone inspections reduces turbine downtime and speeds up the inspection process, offering a promising alternative to traditional methods (Law and Koutsos, 2020).

Herring et al. (2019) highlight that performing repairs on eroded blades can lead to a significant period of downtime and substantial costs due to the need for suitable wind and weather conditions, equipment, and technicians. The expenses tied to comprehensive turbine inspections often result in operators conducting these only every two to three years, which can leave repair issues undetected until the next review period (Herring et al., 2019).

In conclusion, the far-reaching impacts of leading-edge erosion on wind turbine blades necessitate continuous research and innovation. The complex mechanisms underlying LEE, the promising advancements in monitoring techniques, the need for accurate prediction models, and the vital role of timely maintenance collectively underscore the significance of this issue. With these insights, the wind energy sector continues to strive towards mitigating the economic and energy production losses associated with leading-edge erosion.

2.2.2. Inspections Against Leading Edge Erosion

As wind turbines continue to play a central role in renewable energy production, ensuring their efficient operation remains critical. Key to this is the early detection and proper management of the erosion of their leading edges. Anisimov et al. (2021) highlights the limitations of the current visual or drone-based camera system inspections, and the need for a more reliable and precise means of detecting and monitoring leading-edge erosion. The authors employ a customized long-range laser line scanner, detecting eroded and damaged areas with sub-millimeter resolution, thus moving towards a condition-based and predictive maintenance approach (Anisimov et al, 2021).

Dimitrov (2018) suggested a method for evaluating the risks associated with blade damage identified during visual checks, aiming to pinpoint the most cost-effective repairs. The research showed that the best repair strategy varies depending on the severity of the damage, and that a risk evaluation can help to find the most economically viable solution (Dimitrov, 2018). This idea corresponds closely with the call for better evaluation methods found in related research. The procedure developed for assessing the severity of an issue relied on the use of a model to estimate how fast the damage was developing.

The existing literature on wind turbine blade inspection methods against leading edge erosion damage is limited, and collectively underscores the significance of developing reliable, sensitive, high-resolution methods for detecting, monitoring, and assessing leading-edge erosion of wind turbine blades. The laser line scanner approach proposed by Anisimov et al. (2021) represents a promising avenue for achieving these goals, while the risk-based assessment procedure by Dimitrov (2018) offers a rational way of categorizing damage severity and determining optimal interventions. These integrated efforts suggest a path forward towards condition-based and predictive maintenance, though further advancements are still needed to address the challenges inherent in offshore wind turbine inspections.

2.3. Machine Learning and Computer Vision

2.3.1. Object Detection and Improvement of Learning

Object detection is a key task within the broader field of computer vision that involves identifying specific objects within a digital image or a video sequence. It differs from related tasks such as image classification and segmentation by not only categorizing what is present in an image, a task that usually is attributed to image classification, but also precisely locating each object via a bounding box or a similar method. Unlike image segmentation, which aims to assign a class to each pixel in the image for a detailed breakdown, object detection provides a higher-level overview of object locations and classes. This balance makes object detection a crucial tool in many applications such as self-driving vehicles, video surveillance, and augmented reality, where both recognizing and locating objects are important.

The comprehensive reviews by Zou et al. (2023) and Liu et al. (2020), alongside the work by Zoph et al. (2020), have offered an insightful understanding of the evolution, achievements, and challenges in the field of object detection, including the significant role played by data augmentation. In this context, the research presented by Perez and Wang (2017) provides valuable insights into the effectiveness of data augmentation in the closely related field of image classification.

Perez and Wang (2017) explore and compare various solutions to the problem of data augmentation in image classification, a challenge also underscored by Zoph et al. (2020) in the domain of object detection. Their work highlights the effectiveness of simple techniques such as cropping, rotating, and flipping images. Their conclusions align with Zoph et al.'s (2020) findings, emphasizing that such techniques, when combined with the creation of specialized data augmentation policies, can lead to improvements in the generalization performance of models trained on limited data.

Echoing Zoph et al.'s (2020) assertions about the benefits of data augmentation, Perez and Wang (2017) highlight that it offers a promising way to enhance the accuracy of classification tasks, even when the quality of data is relatively low. They posit that the more data a machine learning algorithm has access to, the more effective it can be, as long as useful information can be extracted from the original dataset.

Perez and Wang (2017) further explore the idea of taking a small, structured dataset and augmenting it to improve model performance, an approach they found to be effective in multiple problems. This notion is particularly crucial for specialized image and video classification tasks, such as object detection, which often suffer from insufficient data (Perez and Wang, 2017). The problem is potentially even more pronounced in industries where data access is heavily protected due to privacy concerns.

While both Zou et al. (2023) and Liu et al. (2020) highlight the role of large and unbiased datasets in pushing object detection research forward, Perez and Wang's (2017) work underscores the importance of techniques that can effectively augment and leverage smaller datasets. They discuss the problems of overfitting in models trained on small datasets, which do not generalize well to validation and test sets, a challenge that resonates with Liu et al.'s (2020) concerns about the limitations of fully supervised learning.

Finally, Perez and Wang (2017) introduce the concept of transfer learning, a technique closely related to data augmentation. They describe it as a method where pre-trained weights of a neural network trained on similar or more comprehensive data are fine-tuned to best solve a more specific problem. This technique, coupled with data augmentation, provides additional ways to reduce overfitting on models (Perez and Wang, 2017). Overfitting in computer vision refers to a model learning the training data too well, to the point where it performs poorly on unseen data because it has picked up on noise or irrelevant patterns in the training set, rather than generalizable features.

In summary, the reviews by Zou et al. (2023) and Liu et al. (2020), along with the contributions from Zoph et al. (2020) and Perez and Wang (2017), provide a comprehensive understanding of the field of object detection. They emphasize the need for data augmentation and transfer learning strategies, particularly for models trained on limited datasets, and highlight the necessity for ongoing research to further refine these techniques and methodologies. These efforts are all directed towards the ultimate goal of creating detection systems with abilities rivalling those of the human visual system.

2.3.2. Computer Vision in Industrial Maintenance

In the realm of industrial maintenance, the intelligent detection of defects and faults has become an increasing necessity. The applications of computer vision has found use in many different industries. This is highlighted by recent increased focus on research in the field. Xu et al. (2022) identify the limitations of traditional methods in road crack monitoring, especially with the growing road network requiring more advanced and intelligent technologies. Their study demonstrates the potential of state-of-the-art algorithms in intelligently detecting road cracks, building upon previous research in computer vision and digital image processing. However, they emphasize the need for large datasets for training and the continuous evolution of deep learning methods.

Parallelly, Wang et al. (2022) address fundamental issues in unmanned aerial vehicle (UAV)-based inspection, a key area of research in the wind energy industry. They propose an improved model for segmenting wind turbines from UAV-taken images. The results present superior performance in terms of performance metrics such as intersection over union (IoU), and recall values after 20 epochs of training, providing evidence of the effectiveness and practical utility of deep learning in industrial maintenance. (Wang et al., 2022)

Furthermore, Shihavuddin et al. (2019) explores a deep learning-based automated damage suggestion system for wind turbine blade surface inspection using drone imagery. Their approach achieved near-human-level precision in suggesting damage location and types on wind turbine blades. By utilizing a specialized model architecture within Faster R-CNN, they present a model achieving a mean average precision (mAP) of 81.10% for four different types of damages (Shihavuddin et al., 2019). The study also demonstrates the effectiveness of data augmentation in improving the generalization of the trained model and highlights the potential cost advantages of automating the inspection and analysis process.

In the field of wind turbine blade crack inspections, Wang et al. (2019) suggested a method to accurately identify cracks on the surface of wind turbine blades by examining images taken by drones. The initial stage of their approach utilizes a fundamental detection framework to accurately locate cracks. The method's reliability and efficiency is further validated through images captured by drones from a commercial wind farm.

Furthermore, Xu et al. (2019) propose the use of convolutional neural networks (CNNs) for image recognition of common wind turbine blade defect conditions, while constructing a preliminary dataset consisting of 25,773 images depicting five common wind turbine blade defect conditions and trained three different deep learning models. The authors' proposed method demonstrates advantages such as intuitive understanding of wind turbine blade conditions, reduced downtime, improved productivity, and increased economic benefits (Xu et al., 2019). The study also acknowledges the importance of considering regional variations in wind turbine blade damage.

Moreover, Ye et al. (2016) provides a comprehensive review to highlight the advantages of noncontact, nondestructive, and high-precision monitoring techniques. The review offers valuable insights into the initial efforts on the integration of machine vision technology, and highlights the challenges and limitations in the field.

These studies collectively contribute to the growing body of knowledge in the field of intelligent defect detection and maintenance, showcasing the potential of machine learning and deep learning algorithms in various industrial applications. Overall, the existing literature highlights the ongoing efforts to enhance inspection processes, reduce downtime, and improve the overall efficiency of maintenance operations through automated analysis of collected data. The continual evolution and refinement of these approaches are crucial for ensuring the reliability and safety of industrial assets in sectors such as transportation and renewable energy.

2.3.3. Real-time Object Detection and State-of-the-Art

Over time, the realm of object detection has seen substantial progress, characterized by remarkable improvements in both speed and accuracy, as demonstrated in various studies and publications. Ren et al. (2017) introduced Faster R-CNN, an innovative solution aimed at efficiently generating high-quality region proposals. This work was crucial in tackling the processing limitations of the top-tier detection systems of that period, allowing for a deep-learning-oriented object detection system to operate in near-real time. However, it was still not considered suitable for complete real time applicaitons. The resulting system was able to achieve remarkable object detection accuracy that led the field for a substantial amount of time (Ren et al., 2017).

On the other hand, Redmon et al. (2016) proposed a novel architecture, known as YOLO (You Only Look Once), within the object detection landscape. YOLO employed a network that is capable of mitigating the need for the complex pipelines associated with earlier approaches, through analyzing the entire image during training and testing in one go, yielding fewer background errors. (Redmon et al., 2016) However, YOLO had its own limitations, such as the challenge to localize small objects accurately.

Wang et al. (2022) focused on enhancing the training process through optimized modules and introduced YOLOv7. In comparison to its predecessors, YOLOv7 demonstrated improvements in accuracy and efficiency, having 75% fewer parameters and requiring 36% less computational resources (Wang et al., 2022). The researchers also found that YOLOv7 surpassed all known real-time object detectors in speed and accuracy, solidifying its position as a highly effective and efficient object detection solution (Wang et al., 2022).

In conclusion, while both YOLOv7 and Faster R-CNN present significant advancements in object detection, YOLOv7 seems to have the edge over object detection models in terms of speed and still obtains remarkable levels of accuracy. This makes it the more efficient solution for real-time object detection tasks, such as damage detection on the edge cases. Future developments in the field of object detection could further leverage the strengths of both approaches, pushing the boundaries of what is currently achievable.

2.3.4. Impact of Resolution on Performance

The importance of input resolution and quality on the performance of machine learning models, specifically for object detection, is increasingly recognized as a crucial aspect of achieving high-performance results. In their study, Wu et al. (2022) highlighted how variations in image resolution can substantially influence the efficacy of deep learning models in the diagnosis of breast cancer. Their research used grayscale ultrasound breast images from two Chinese hospitals, with resolutions of 224×224, 320×320, and 448×448 pixels, which are commonly used values due to their balance of performance and computational efficiency (Wu et al., 2022). The findings from this study underline that smaller resolution images, while requiring less computational time, may sacrifice critical information, thereby impacting the diagnosis outcomes. It was noted that different combinations of machine learning models and input image resolutions yielded diverse results, emphasizing the importance of finding the optimal pairings (Wu et al., 2022).

On the same theme, Talebi and Milanfar (2021) focused on the potential impacts of image size on training accuracy. They argued that the input images' resizing to a relatively small resolution has been treated as an afterthought in many machine learning applications, despite its potential to influence the overall performance of the trained models. Their research indicated that the commonly used image resizers could be replaced with learned resizers to improve performance. Interestingly, it was found that the replacement of these classical resizers with learned ones did not necessarily enhance the visual quality of the downscaled images, but they did improve task performance (Talebi and Milanfar, 2021). This demonstrates that a balance must be found between image quality and computational efficiency, which, in turn, could affect object detection performance.

The significance of image resolution was further underscored by Thambawita et al. (2021), who found that image resolution had a substantial impact on the performance of convolutional neural network (CNN)-based image classification in gastrointestinal endoscopy. They used a dataset comprising 10,662 images of 23 different findings to assess the performance of two models at different image resolutions. The findings of this study revealed that higher image resolution generally led to better performance (Thambawita et al., 2021). They suggested that the reduction of image resolution might lead to the loss of critical details, such as fine vessels and other patterns of the findings, which are important for accurate classification. Their research further demonstrated that upscaling from lower resolution images resulted in a more significant performance loss than downscaling from higher resolution images. This calls for high-resolution image collection in deep learning context, given that downscaling is easier than upscaling to the original resolution with the tools available at the time of the study (Thambawita et al., 2021).

Examining the impact of low resolution on image recognition, Koziarski and Cyganek (2018) discovered that even relatively mild decreases in image resolution could significantly decrease classification accuracy of deep neural networks. They observed that the performance decline was particularly noticeable for low-resolution levels and that super-resolution techniques could partially mitigate the negative impact but were still far from achieving results comparable to undistorted data (Koziarski and Cyganek, 2018). The authors acknowledged the ongoing research in super-resolution methods, suggesting that future improvements may further reduce the negative effects of low resolution on classification accuracy (Koziarski and Cyganek, 2018).

Dodge and Karam (2016) assessed the efficacy of cutting-edge deep neural network models in the realm of image classification, particularly when faced with different quality distortions. The results revealed these networks' vulnerability to distortions such as blur and noise, but they demonstrated resilience when dealing with compression artifacts and modifications in contrast (Dodge and Karam, 2016). Given that image quality is frequently compromised in real-world scenarios, the authors brought forward the importance of engineering deep neural networks that can better withstand quality distortions.

Overall, resolution impact on detection performance has been a focus of numerous studies within the medical field, but lacks attention in industrial applications. The studies explored so far highlight the significant impact of image resolution and quality on the efficiency of machine learning models. The choice of image resolution and the consideration of quality distortions are crucial factors in achieving optimal results, particularly in object detection within medical imaging applications. Further research on input resolution within specific use cases of computer vision, such as wind turbine inspection and other relevant industrial applications, is necessary to achieve more robust and invariant models.

2.4. Summary and Gaps in Literature

Wind turbine blades are central components in the operation of wind energy systems. Their maintenance is an area of increasing focus within the academic and industrial sectors, given the accelerating adoption of wind as a viable and competitive alternative source of energy. The steady increase in the number of turbines coming in operations, particularly those in remote or challenging locations such as offshore wind farms, has amplified the inherent complexities involved in maintaining these essential equipment. This rapidly expanding scheme of turbine operations demands emphasis on robust and efficient maintenance protocols for turbine blades, a factor critical to overall sustainability of operations and performance.

Currently, the focus is on predictive maintenance approaches through advanced sensor technologies and data-driven decision-making systems. The collection and utilization of real-time data offers a path to preemptive maintenance strategies that can significantly increase the operational lifespan of wind turbine blades. However, despite the promising potential of such approaches, the competitive nature of the wind energy industry may impose constraints on the kind collaboration necessary for large-scale reliability enhancements. This shared commitment to reliability is an essential aspect of ensuring the long-term sustainability of wind energy operations across the global energy industry.

Despite advancements in technology, inspections of leading-edge erosion on wind turbine blades remain largely reliant on traditional, labor-intensive methods. Inspectors typically have to physically climb towers to conduct visual checks, a process that is both time-consuming and inhibits potential safety risks. The potential for leveraging UAVs to enhance and simplify this process has been explored, however, their use has yet to be widely adopted or standardized, limiting their current impact on inspection practices.

Simultaneously, the progressive evolution of computer vision and object detection technologies has drawn considerable attention within the wind energy industry. These advancements lead to opportunities for enhancing turbine inspections. They can potentially drive down operational costs but also mitigate the human safety risks associated with manual inspections. By automating image analysis and damage detection, these emerging technologies can substantially streamline and upgrade the inspection process.

However, there's a lack of standardization in the tools required to implement this technology. The devices used in inspections, irrespective of UAV usage, currently produce variable resolutions. As object detection depends heavily on visual data, the size and quality of inspection images have the potential to influence the performance of models that detect and categorize erosion damage levels on wind turbine blades.

While other sectors, such as healthcare and medical technology, have investigated the impact of image resolution on machine learning performance, this critical factor remains comparatively under-researched in the realm of wind energy. This suggests a gap in the existing body of literature, emphasizing the need for research efforts to drive the industry towards the successful and effective deployment of these emerging technologies. As the world strives to utilize the power of wind more efficiently and sustainably, such research initiatives will be effective in shaping the future of wind turbine maintenance and management.

3. Methodology

This study employs a research design that applies the principles image processing techniques to the field of industrial asset management. The foundation of the research is based on the experimental method, where artificially created leading edge erosion (LEE) damages, created in a controlled environment, on a wind turbine blade section will be imaged, analyzed, and categorized using object detection capabilities of the YOLOv7 computer vision algorithm.

The experimental design includes the examination of the impact of image resolution variances obtained from different imaging devices on the accuracy of the computer vision algorithm. To effectively analyze this, the research is designed to compare the performance of the algorithm across three distinct types of imaging devices: (1) Ryze Tello, a mini drone available for educational purposes with an embedded lens, (2) a smartphone, which currently is a common tool in documenting damages in manned inspections due to its availability, and (3) Panasonic Lumix DMC-GF6, a consumer-grade digital camera. These devices were chosen due to their ability to represent a range of potential resolution and image quality outputs altogether.

The research seeks to explore new insights and understandings into how image resolution can impact the effectiveness of a computer vision algorithm in the context of wind turbine blade inspections. The outcomes of this research could potentially help to inform decisions around the imaging systems to enhance the effectiveness and efficiency of wind turbine blade inspections, contributing to both theoretical and practical advancements in the field.

3.1. Creation of Artificial Erosion

The first step of the methodology involved the acquisition of a decommissioned wind turbine blade, provided by the industry partner. This ensured an authentic experimental subject, embodying real-world conditions. A 2.5-meter section from the blade's tip was selected, as this area traditionally faces the highest wind velocities. Following the acquisition, the blade section was relocated to an area within the university grounds. This provided an appropriate environment for implementing controlled damage to the blade in a safe and practical manner.

To prepare the blade section for the creation of artificial LEE, it underwent an initial cleaning process. This stage was vital to ensure that the artificial damages created would be as accurate and representative as possible, removing any residues that could have interfered with the erosion simulation or detection process.

Artificial erosion was then implemented onto the blade's leading edge at three distinct locations, each corresponding to a different stage of LEE severity. This design choice ensured the study spanned across the spectrum of erosion, thus providing a comprehensive assessment of the capabilities of the computer vision algorithm under different damage conditions.

The first level of inflicted damage was primarily cosmetic, emulating the initial stages of erosion buildup. This level was minimal and did not expose the structural material beneath the surface layer of the blade.

The second level of damage was more advanced, with a deeper erosion that allowed for the partial visibility of the structural material beneath. This stage represented a more advanced state of LEE, which often leads to the need for repair or replacement actions in real-world conditions.

Finally, the third level of damage was the most severe, with the structural material of the blade completely exposed and a measurable depth to the erosion. This stage served to simulate the extreme cases of LEE, where the functionality and safety of the blade are compromised, and immediate maintenance action is necessary.

In essence, these three levels of artificial damage created on the leading edge of the blade provided a broad range of scenarios for testing the computer vision algorithm's ability to detect and categorize varying degrees of LEE. This was critical for measuring the impact of image resolution on the effectiveness of these algorithms, serving as the foundation of this study's experiment.

3.2. Data Collection

The data collection process was initialized with the imaging of the artificially eroded sections on the wind turbine blade. Three different imaging devices were selected to achieve this aim:

- The Ryze Tello drone, providing images with a resolution of 960 x 720
- A personal smartphone, capturing images with a resolution of 1536 x 2048
- A high-definition digital camera, producing images with a resolution of 4592 x 3448

In order to standardize the photography process and ensure the images were comparable across different erosion categories and devices, a structured process was carefully devised and followed throughout the image acquisition activity. This protocol required each erosion category to be photographed from three specified distances: 50cm, 100cm, and 150cm. At each distance, five images were taken while facing the leading edge directly (0 degrees), then again at an angle of 45 degrees, and finally at an angle of -45 degrees relative to the blade's leading edge. This was repeated for each damage instance located on the blade, and with each unique imaging device. Figure 3.2 visualizes the imaging layout of the process. This systematic approach provided a comprehensive visual dataset of each erosion category from multiple perspectives, resulting in an initial pool of 405 images.

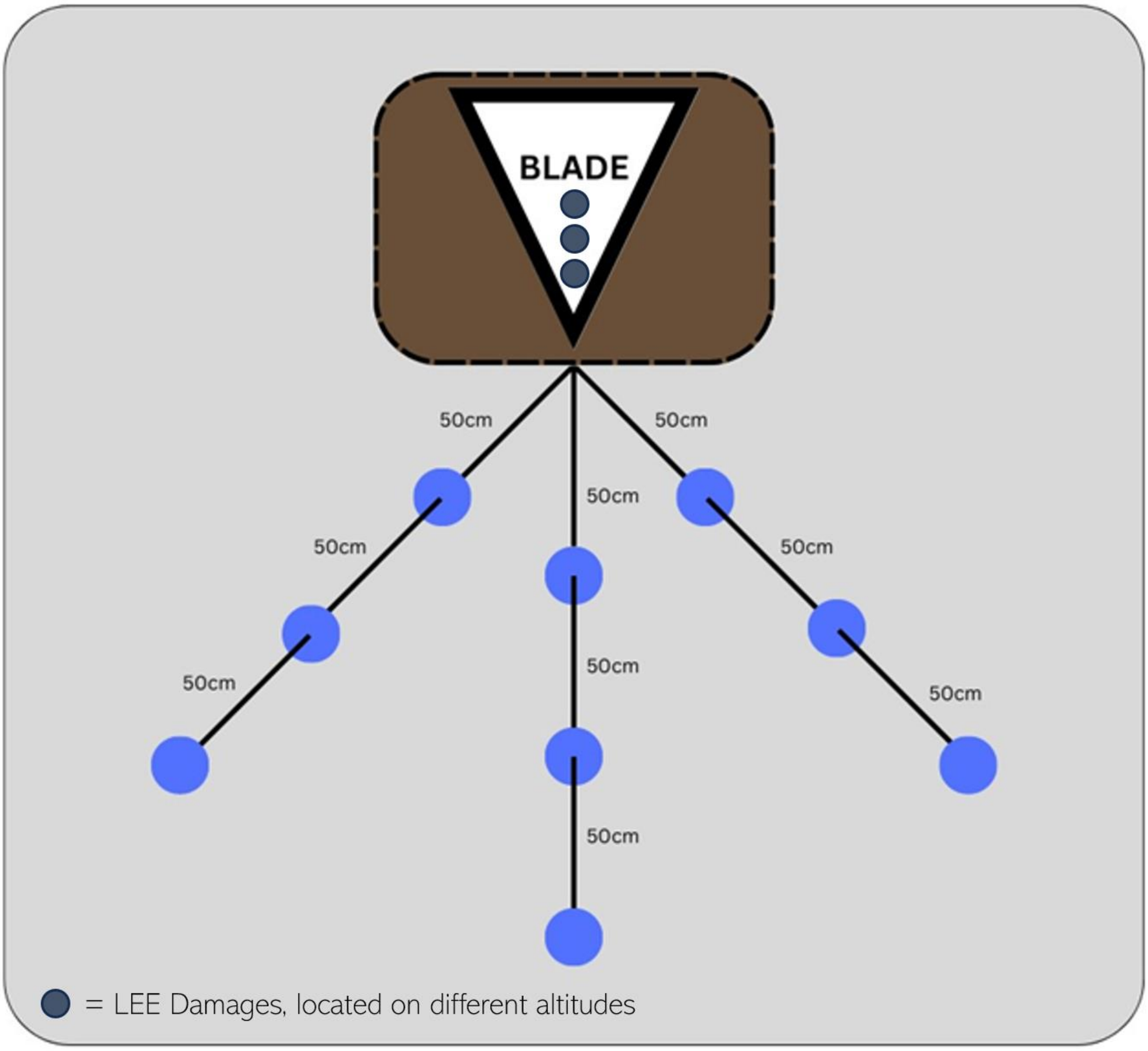

Figure 3.2 – Image Collection Layout

The capturing of images through the smartphone and the digital camera was conducted manually, however the capabilities of the drone allowed for a “mission” to be programmed for this process. The mission was coded in Python and sequentially commanded the drone to assume its position on image capture points and commence capturing pictures of the damages. The initial mission included one complete cycle for the drone to capture images at all dedicated points, however, this process consumed excessive power from the already limited capacity of the drone’s battery, therefore was deemed infeasible. Following this discovery, the mission was adjusted so that in a single execution, the drone would only capture images at points lying on individual angle lines. This modification required the drone to be placed at 50 centimeters distance from the blade’s leading edge as the takeoff position, and the mission to be executed thrice. The modified mission code can be found in Appendix A.

Following the image acquisition phase, a detailed selection was conducted to refine the dataset. From every combination of device, distance, and angle for each erosion category, one representative image was chosen out of the five available. This selection was based on factors such as clarity, focus, and how well they represented the damage stage in question. By the conclusion of this selection process, the dataset was reduced to 81 unique images that provided a comprehensive representation of the various categories of leading-edge erosion.

In addition to the images collected from the artificially damaged wind turbine blade, the dataset was further supplemented with real-world images obtained from an industrial partner. These images, taken during actual manned inspections of wind turbines, were carefully reviewed and a selection of these, representing various stages of leading-edge erosion, was added to the dataset. The inclusion of these real-world images provided an additional layer of validity to the study, offering insight into natural erosion on top of artificially created patterns. This combined dataset, representative of both simulated and actual conditions, formed the basis for the subsequent data analysis and machine learning model training stages of this research.

3.3. Image Preprocessing, Augmentation and Dataset Generation

This stage concentrated on preparing the images for successful machine learning applications. The images were annotated on Roboflow, a platform suitable for such operations, which involved marking every occurrence of damage visible on each image, categorizing them based on their levels of severity. This annotation process is essential for machine learning tasks, as it ensures precise algorithm training and accurate computer vision in later stages.

After thorough annotation, the dataset was separated into three distinct subsets for training, validation, and testing purposes. The division ensured the allocation of almost all the images provided by the industrial partner to the training and validation sets. Furthermore, 7 images from each device category were chosen for the testing process, 8 were assigned for validation, and 12 were set apart for training. This distribution aimed to ensure a comprehensive and thorough evaluation during each phase.

Training subsets tend to be the largest subset of datasets within computer vision applications and is used to train the model. The training set helps the model to learn the patterns, features, and relationships between inputs, image features, and the desired outputs, object labels and locations. The validation set is used to evaluate the model's performance during the training phase and to fine-tune model parameters. It acts as an “internal testing” set as the model trains over each epoch. Once the model's training and validation is complete it is then evaluated on the test set. This is a completely separate set of data that the model hasn't seen during its training or validation phases. The performance on the test set gives an unbiased estimate of the final model's performance and generalizability to unseen data.

The main difference between the validation and test sets is when and how they're used within the process. The validation set is used during training to make adjustments to the model's parameters and prevent overfitting, while the test set is used after training to provide an unbiased evaluation of the model's performance on previously unseen images.

In any computer vision application, the quantity of images with unique details provided for training is proportional to the performance of the model. Hence, to optimize the efficiency of this study's model, image augmentation and preprocessing was integrated into the methodology using Roboflow once more. This ensured that the training set was adequately robust, aiding in the development of a more accurate and efficient damage detection model.

The process of preprocessing specifically targeted the resizing of images, with the final dimensions of all being standardized to 640x640 pixels. This standardization enables the model to extract information from the images with increased effectiveness without losing source characteristics, therefore decreasing the computational load and considerably enhancing the training speed.

Simultaneously, the augmentation process included various operations on images aimed at artificially expanding the dataset. These included:

- Horizontal and vertical image flipping,
- 90-degree rotations both clockwise and counter-clockwise,
- Brightness adjustment within a range of 0% and +15%,
- Gaussian blur up to 1.5 pixels,
- and the inclusion of random noise up to 2% of the pixels.

Such augmentation methods help the machine learning model to better generalize by simulating a variety of possible scenarios it might face during the deployment phase.

The integration of these comprehensive preprocessing and augmentation steps resulted in a substantial increase in the overall dataset size, bringing it up to a total of 557 images, with a majority of 480 being allocated to training. The rest were utilized for validation and testing, with 50 images dedicated to the former and 27 to the latter. These steps ensured the creation of a well-rounded, high-quality, and extensive image dataset to be utilized in the following stages of the research.

3.4. Model Training

This section involves the essential process of feeding the improved and finalized dataset into the YOLOv7 object detection model, training it for the identification of different categories of damages in the test images. The objective was to train the model to efficiently place bounding boxes around the detected areas of damage in these images, as well as ensuring smooth categorization of the degree of damage.

The training was conducted on Google Colab, chosen for its superior GPU capabilities, a critical component in machine learning operations. Google Colab's available infrastructure allows for the use of high-performance GPUs such as the NVIDIA A100 with 40GB memory, which was utilized in this study. This GPU was selected for its proven efficiency, though it should be noted that alternative GPUs could also have been employed, based on available resources and specific project requirements.

The chosen GPU's capabilities informed the selection of certain adjustable training parameters, such as the batch size and the number of epochs. Batch size refers to the number of training examples utilized in one iteration, meanwhile, an epoch is a complete pass through the entire training dataset (Brownlee, 2018). These parameters can be adjusted to optimize model performance, balancing the speed of training against the final accuracy of the model. As you increase the number of epochs, the model has more opportunities to learn from the data, but this also increases the risk of overfitting. Essentially, overfitting is when the model becomes too well-adjusted to the training data. If the model is allowed to train for too many epochs, it may become "too specialized", performing poorly on the validation or new input (Ying, 2019). Therefore, careful tuning of the number of epochs is necessary. Enough epochs should be allowed for the model to learn from the data sufficiently, but not so many that it starts to overfit. For this study, through trial and error, the chosen final number of epochs reflects a balance between allowing the model to adequately learn from the data while avoiding overfitting to ensure it generalizes well.

To accelerate the training process and enhance the final performance, the model training began using a "checkpoint". This strategy means that the training process wasn't initiated from scratch; rather, it was built upon a pre-existing, successful iteration. MS COCO, which stands for Microsoft Common Objects in Context, is an extensive dataset used for tasks such as object detection, segmentation, and others (Lin et al., 2014). It's commonly utilized as a base for training models that are designed for object detection. By using this checkpoint, the model was able to leverage prior learnings, focusing on refining its ability to detect and categorize the specific wind turbine blade erosion damage as presented in the dataset. The complete code ran on Google Colab can be found in Appendix B.

3.5. Performance Analysis & Reparameterization

The performance of the trained YOLOv7 model was represented visually using a confusion matrix. A confusion matrix is a specific table layout that enables easy visualization of the performance of an algorithm, typically a machine learning model. In the context of a binary or multiclass classification problem such as the one in this study, each column of the matrix corresponds to instances of an actual class, and each row corresponds to instances in a predicted class. Thus, the confusion matrix presents a comprehensive view of how well the machine learning model has performed with respect to the annotated images, or the ground truth, thereby capturing outcomes.

True positives and true negatives represent the model's correct predictions, whereas false positives and false negatives reflect incorrect predictions. Ideally, a perfectly performing model would have all true positives and true negatives, meaning the model correctly identified all the damages according to their categories. On the other hand, the presence of any false positives and false negatives indicates the areas where the model has failed to identify the damage correctly. A clear advantage of using a confusion matrix is that it not only presents the errors, false positives and negatives, made by a model, but also shows the types of errors, allowing for targeted improvement.

The training, by default, outputs valuable insights into the model's performance. The output contained various performance metrics graphically plotted per epoch; including Precision, Recall, mAP@0.5, and mAP@0.5:0.95. Precision refers to the proportion of correctly predicted positive observations out of the total predicted positives. Recall is the proportion of correctly predicted positive observations out of the actual positives. Mean Average Precision (mAP) is a popular metric in measuring the accuracy of object detectors like YOLOv7. mAP is calculated by taking the mean of average precision scores for each class. Two critical values in this metric are mAP@0.5 and mAP@0.5:0.95, which specifically involve the concept of Intersection over Union (IoU).

IoU is an evaluation metric used to measure the overlap between the predicted bounding box (by the object detection model) and the ground truth bounding box. It is the ratio of the area of overlap and the area of union of the two bounding boxes. A higher IoU indicates a more accurate detection and is therefore preferred. The term mAP@0.5 refers to the scenario where the model is evaluated at a single IoU threshold of 0.5. This means that if the overlap between the predicted bounding box and the ground truth bounding box is at least 50%, the detection is considered a "true positive"; otherwise, it is considered a "false positive". On the other hand, mAP@0.5:0.95 means the model is evaluated at multiple IoU thresholds, from 0.5 to 0.95, in steps of 0.05. Here, average precision is calculated at each step and the mean of these values is reported. This provides a more robust metric, reflecting the model's performance across different levels of overlap and detection difficulty.

Observing the mAP values' progress over time allowed the identification of a suitable epoch number to prevent overfitting. Subsequently, separate training sessions were conducted, adjusting the number of epochs, to determine the optimal point that maximized model performance while minimizing the risk of overfitting. Ultimately, a final selection on the number of epochs was made to proceed into the further stages of the study. As the batch size parameter primarily concerns the total training time of the model, the selected size of 64 was not modified. It's worth noting that this selection was compatible only with the NVIDIA A100 GPU or higher models on Google Colab.

In addition to the previous output graphs, the training results also included a graph for the resulting F1 value, the harmonic mean of precision and recall, for the training against different confidence values, allowing for the optimal confidence value to be identified for the following stages of testing and detection. This phase of performance analysis and reparameterization allowed for an iterative improvement in the model, enabling the fine-tuning of the number of epochs to optimize performance, while ensuring the applicability and reliability of the model in real-world scenarios.

3.6. Model Testing

Upon the completion of the training phase, the next step was to test the trained model. This is a relatively simpler yet critical step, as it evaluates the model's performance and capacity to accurately detect and categorize the damages. The testing process was conducted in the Google Colab environment due to the powerful computational resources it offers, similar to the training phase, as well as to contain the computational part of the study to a single environment for better control and customization.

The primary aim of this stage was to input the images that were specifically allocated to the test subset into the trained YOLOv7 model. This is a critical phase as the performance of the model on this set of data helps to assess the effectiveness of the training process, providing insight into the model's ability to generalize its learning to new, unseen data.

The testing process essentially allowed the trained model to scan the input images and use its learnt parameters to identify possible areas of damage. Upon detection, the model drew bounding boxes around these areas, categorizing the damage as per the classification it had learnt during training. Not only did the model categorize the damage, but it also provided a corresponding confidence value for each detection. This value represents the model's level of certainty regarding the damage category of the detected area, hence providing a quantitative measure of the model's detection accuracy.

An important parameter that was set during the testing phase was the minimum confidence threshold. This threshold determined the confidence value below which a detection would not be accepted as valid. Consequently, any detections that had a confidence value lower than this pre-specified threshold were not marked with bounding boxes in the output images. This thresholding serves to maintain the quality of detections, ensuring that only those detections that the model is reasonably confident about are considered valid and presented in the final output. It is, therefore, a crucial aspect of maintaining the accuracy and reliability of the model's damage detection performance.

Overall, the testing phase provided a practical application of the trained model on new data, serving as a crucial assessment of its real-world usability and performance.

3.7. Imaging Device Evaluation

The final stage of Imaging Device Evaluation was a critical point in the research process, where the efficacy of three distinct sources of images was thoroughly evaluated. An evaluative framework was developed for this purpose, using a point system to assess the success rate of damage detection for each imaging device.

For each photograph captured, a potential detection scenario involved the algorithm identifying a damage with a confidence value between 0 and 1. For instance, in a photograph containing two discernible damages, regardless of their categories, the maximum overall confidence value achievable would be the sum of the maximum confidence values per damage, amounting to 2 in this scenario. This maximum value constituted the upper bound of the rate to be calculated.

For the actual value present on an image instance, confidence values of true positive detections were summed. However, exceptions were incorporated into the scoring system to address false detections. In the event of the algorithm detecting a damage but categorizing it incorrectly, the confidence value assigned to that detection was halved before being added to the total. It's important to note that this approach was taken because, regardless of the misclassification, a correct detection can still provide valuable insights to industrial professionals. This information can be crucial in maintenance planning for the asset, emphasizing the significance of correct detection despite potential categorization errors. Moreover, in instances where the algorithm incorrectly detected nonexistent damages, such as misinterpreting background objects as damages, the confidence value was deducted from the total. Finally, this total was divided by the maximum theoretical confidence achievable on the image based on the ground truth, forming the final "success" rate for every image.

Each test image, therefore, received a success rate based on this calculated score, effectively quantifying the accuracy of damage detection per image. Subsequently, these images were grouped according to their source, namely the drone, the smartphone, and the digital camera, and the average success rate for each group was computed. This resulted in an overall success rate for each imaging device, thus providing a comparison of their respective performances in the context of damage detection.

This evaluative approach allowed for a thorough assessment of the three imaging devices, providing crucial insights into their relative efficacy in capturing high-quality images in terms of effective damage detection.

4. Results and Analysis

In this chapter, the results and findings of the research are presented and evaluated. Each stage of the methodology detailed in Chapter 3 has produced quantifiable outcomes, which will be presented here in a structured manner. This structured presentation is designed to reflect against the stages of the methodology, thereby providing a direct correlation between the applied methods and the obtained results. In essence, this chapter's objective is to present and assess the findings, following the steps of image preprocessing and augmentation, training of the YOLOv7 object detection algorithm, testing of the trained model, performance analysis and reparameterization, and the evaluation of imaging devices. The subsections of the chapter will correspond to the steps in the methodology, presenting the related results, followed by a thorough analysis and interpretation of these results.

4.1. Creation of Artificial Erosion

In accordance with the experimental procedures outlined in the methodology, the section of the wind turbine blade was intentionally subjected to damage in three separate locations along its leading edge. The purpose was to simulate distinct categories of leading edge erosion that could occur in a real-world scenario. The types of damage ranged from the initial stages of cosmetic erosion, characterized by minor surface wear and tear, to the severe end of the spectrum where the structural material of the blade is completely exposed, necessitating urgent intervention. The damages inflicted on the blade are visually represented in Figure 4.1.1.

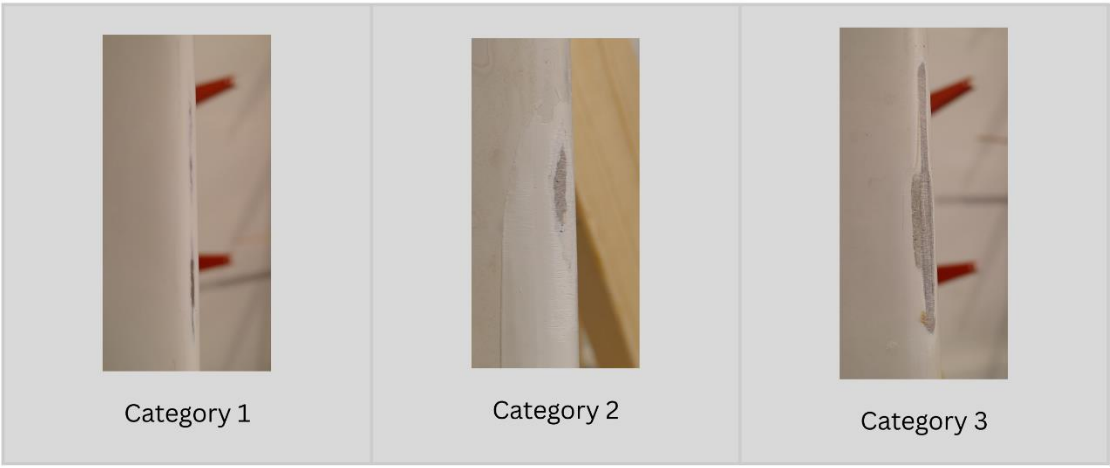

Figure 4.1.1 – Artificial LEE on Blade Section

The task of quantifying the distinct characteristics of each erosion category is notably complex and poses a significant challenge, which is also reflected in industry practices and academic literature. The scanner's capability to deliver reliable, quantitative results was found to be more effective in the instances of severe erosion damage, specifically for Category 3. In this case, the 3D scanner was able to accurately capture and quantify the depth of artificial erosion. Figure 4.1.2 shows reconstructed model of the blade where Category 3 erosion damage is present, captured by the scanner.

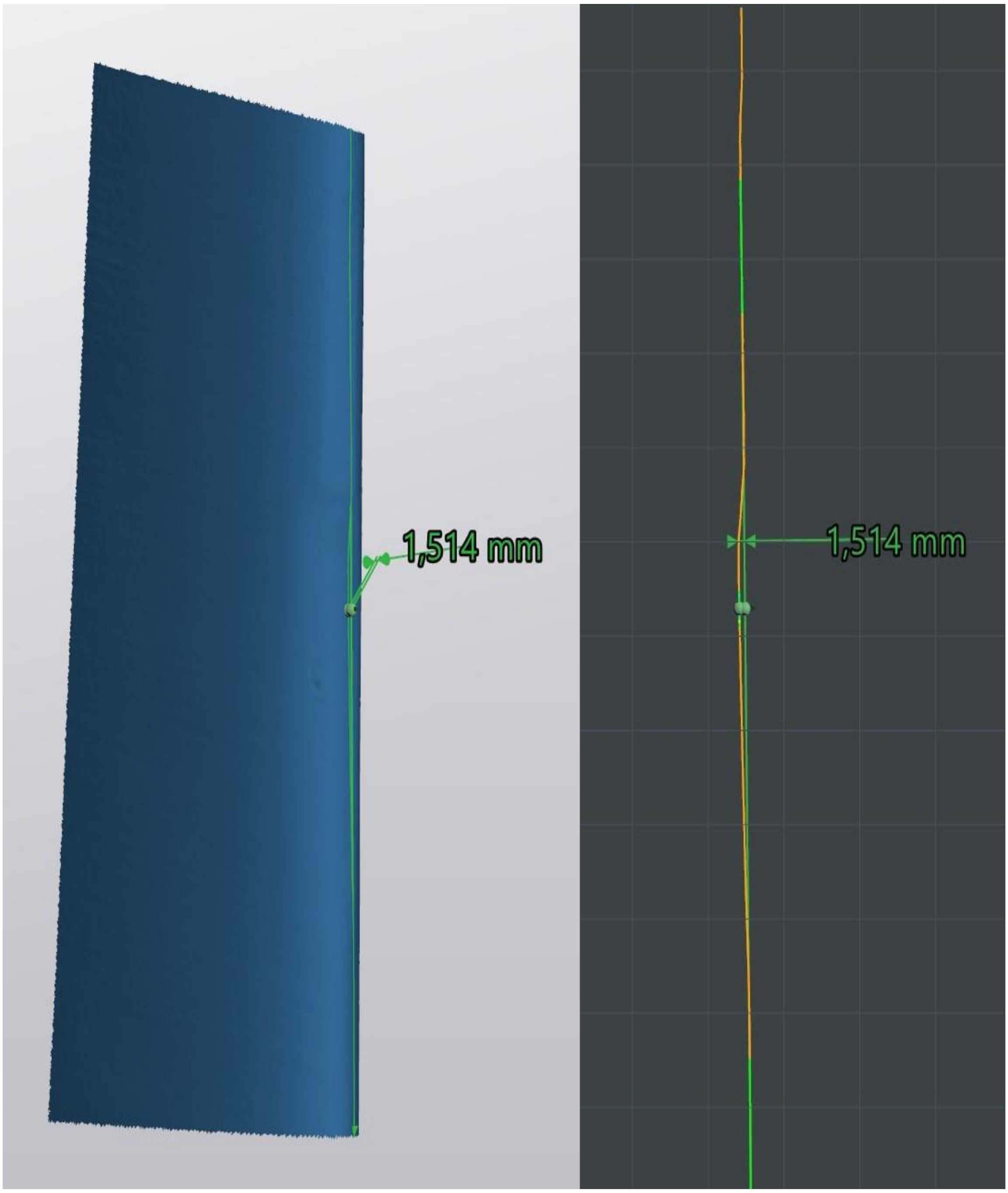

Figure 4.1.2 – 3D Model of Artificial Category 3 Erosion Damage on Blade

Although prior categories of damage cannot be accurately quantified regarding depth, Visual inspections and the nature of machine learning applications considered in this research primarily deals with the qualitative characteristics of the damage to categorize.

While the depth of the less severe erosion categories (Category 1 and 2) could not be accurately quantified using the industrial 3D scanner, this does not provide a significant obstacle for the further stages of the study. The nature of visual inspections and the machine learning applications considered in this study primarily focus on recognizing the qualitative features of the erosion damage for categorization, rather than relying heavily on the exact depth measurements. Visual inspections are designed to identify and categorize the damage based on its appearance and observable characteristics. These could include the shape, size, and pattern of the erosion on the blade's surface.

Moreover, the machine learning algorithms employed in this study are particularly adept at identifying patterns, nuances, and deviations in the data. They work by extracting features from the input images and learning to associate these features with particular categories of damage. This implies that even if the exact depth of erosion cannot be quantified for the less severe categories, the algorithms can still learn to recognize these damage categories based on other distinct, qualitative features visible in the images. While the quantification of damage depth is beneficial for certain analytical purposes, its absence does not delay the study's progress or its primary objective, to assess the impact of image resolution captured in industrial inspections.

4.2. Image Preprocessing, Augmentation and Dataset Generation

The image preprocessing and augmentation phase played a critical role in enhancing the dataset used in this study. Before any preprocessing and augmentation steps, the initial dataset consisted of 173 images, with 310 annotations across 3 damage categories. The average number of annotations per image was 1.8. The average size of the images was 15.68 megapixels, ranging from 0.25 megapixels to 16.06 megapixels. The median image ratio was 3006x3448, representing a tall image orientation.

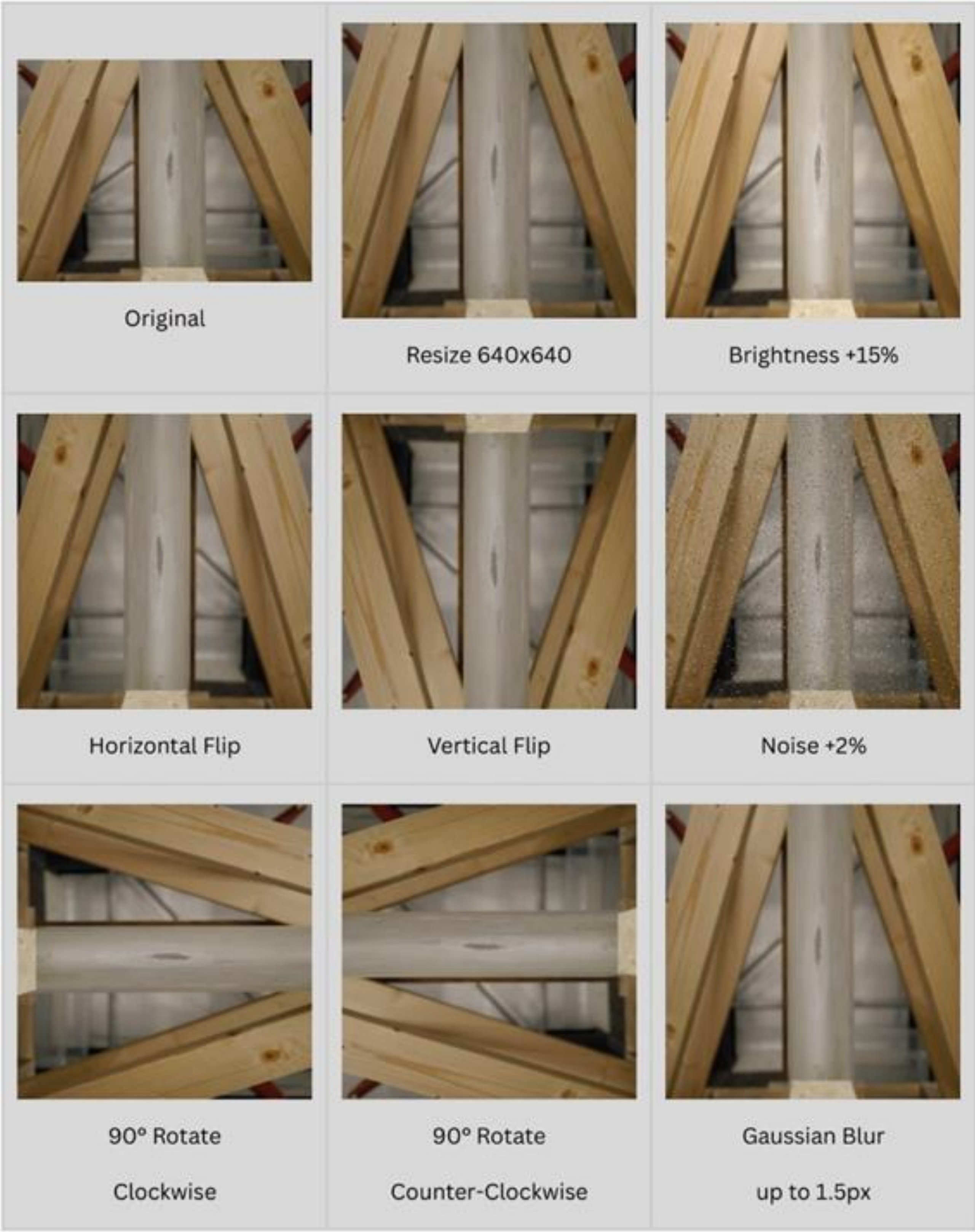

Figure 4.2.1 – Augmentation Outpus

This initial dataset, while substantial, needed further enhancement to optimally train the YOLOv7 model. The preprocessing and augmentation pipeline addressed this need by allowing for an increase in the size of the dataset and contributing to the overall diversity of the images.

The specific steps involved in this process, and their impacts, are analyzed in this section.

1) **Resizing of Images:** The initial step in the preprocessing was to resize the images to a uniform dimension of 640x640. Ensuring uniformity in image size is critical in machine learning models, as inconsistent image sizes can lead to irregularities in the learning process. It is important to note that original images were stretched to achieve the desired quality, rather than being cropped. The standard dimension of 640x640 pixels provided an optimal balance between image quality and computational efficiency. The difference between an original image and its resized version can be seen in Figure 4.2.1.

2) **Augmentation Techniques:** The augmentation process involved the application of multiple techniques to the resized images, aimed at artificially expanding the dataset. The techniques used were horizontal and vertical flipping, 90-degree clockwise and counter-clockwise rotation, brightness adjustment between 0% and +15%, Gaussian blur up to 1.5px, and the addition of random noise up to 2% of pixels. Examples of these augmentations can be seen in Figure 4.2.1.

3) **Increased Dataset Size:** As a direct result of the augmentation techniques, the dataset size increased from 81 unique images to a more substantial dataset of 557 images. This increase expanded the training data available to the YOLOv7 model, which generally leads to better model performance. Figure 4.2.2 provides a visualization of the dataset size before and after the augmentation process.

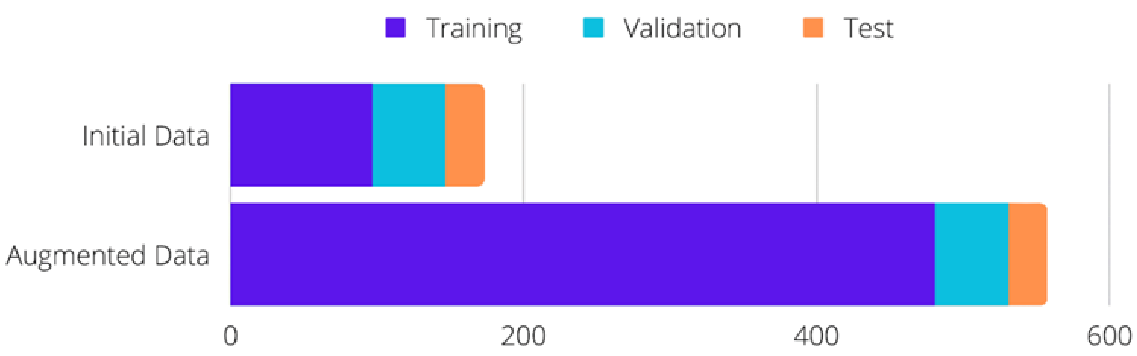

Figure 4.2.2 – Dataset Size and Distribution

4) **Allocations to Training, Validation, and Testing:** Following the augmentation process, the images were allocated to the training, validation, and testing subsets. The training subset received the most significant portion with 480 images, while the validation and testing subsets received 50 and 27 images, respectively. The allocation of the dataset to these subsets was done to ensure the optimal performance of the YOLOv7 model. The proportion of images allocated to each subset is illustrated in Figure 4.2.2.

This comprehensive approach to image preprocessing and augmentation demonstrates the importance of careful dataset construction in the successful application of machine learning models to real-world scenarios.

4.3. Model Training

The evaluation of the model training phase began with an analysis of the training outputs, focusing on the evolution of various key metrics over the sequence of training epochs. For context, an epoch in machine learning refers to one cycle through the entire training dataset. During each epoch, the model learns and updates its parameters to minimize the difference between the predicted and actual outputs. The key performance measures considered in this evaluation include precision, recall, mAP@0.5, mAP@0.5:0.95, and F1 values.

Precision is a metric that represents the proportion of true positive instances among all instances that the model has predicted as positive. A higher precision indicates a model that produces fewer false positives. In Figure 4.3.1, the trend of precision throughout the training process can be observed. The decrease in variance between consecutive precision values displays the learning process of the custom model.

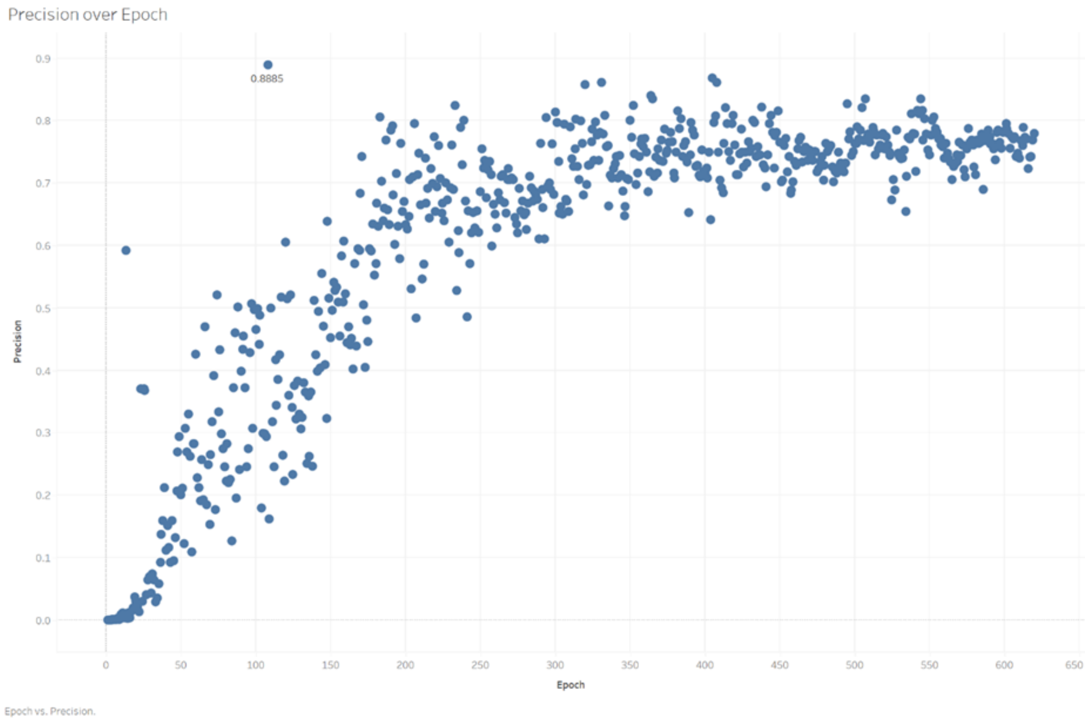

Figure 4.3.1 – Precision over Training

Similarly, recall, also known as sensitivity or true positive rate, is the portion of actual positive instances that the model has accurately predicted as positive. Higher recall means the model can identify more true positives, thereby producing fewer false negatives. Figure 4.3.2 presents the evolution of recall over the course of the training epochs. Much like precision, a decrease in variance is also present in the evolution of recall value, indicating successive learning.

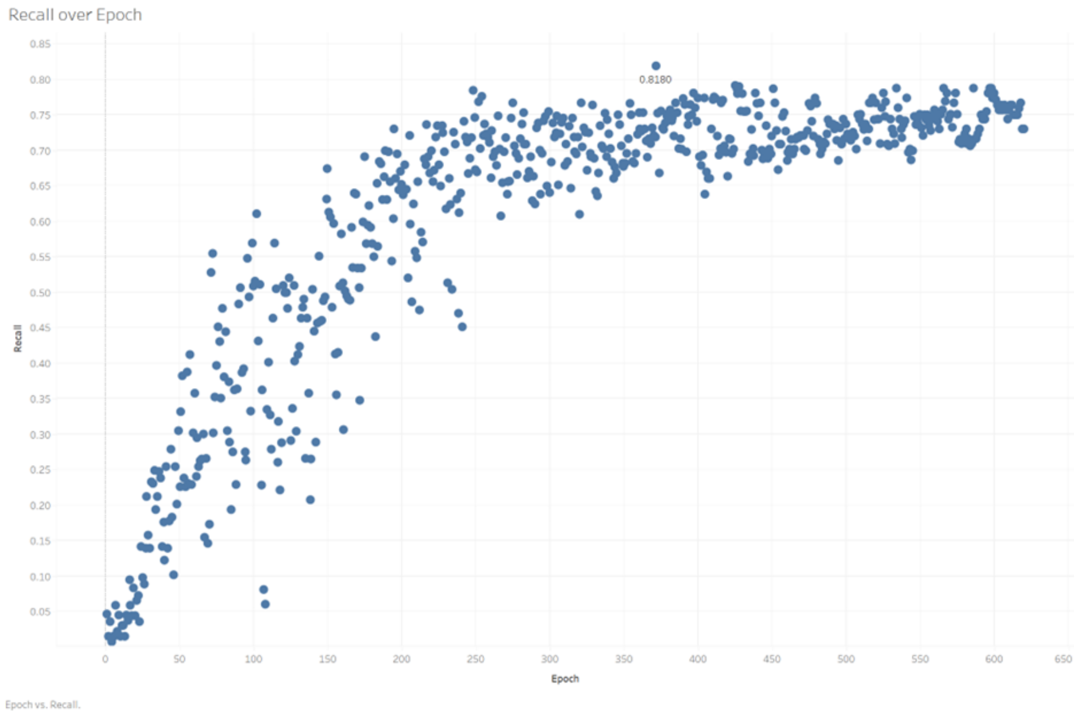

Figure 4.3.2 – Recall over Training

Next, the Mean Average Precision (mAP) is inspected, both at an Intersection over Union (IoU) threshold of 0.5 (mAP@0.5) and over a range of IoU values from 0.5 to 0.95 (mAP@0.5:0.95). Figure 4.3.3 displays the evolution of mAP over the course of 620 training epochs.

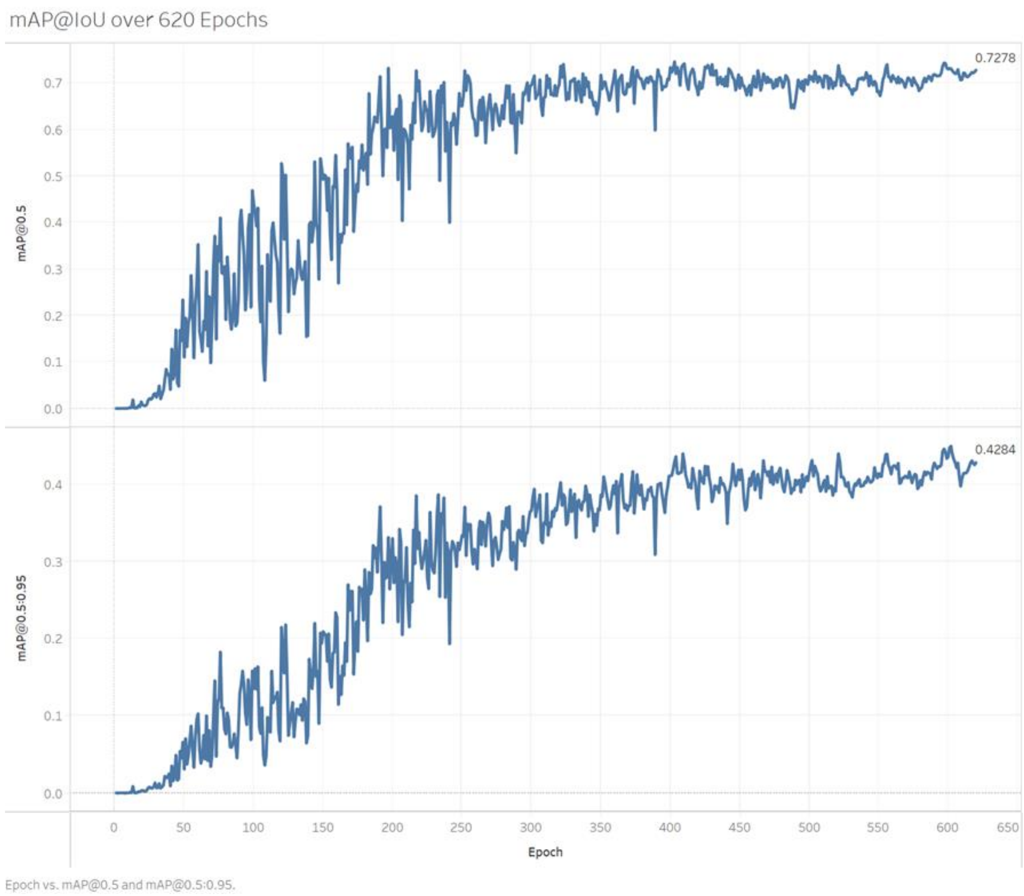

Figure 4.3.3 – mAP over Training

The mAP is a popular metric in measuring the accuracy of object detectors like YOLOv7. Specifically, mAP@0.5 means the model is evaluated at a single IoU threshold of 0.5, and mAP@0.5:0.95 means the model is evaluated at multiple IoU thresholds from 0.5 to 0.95. Recalling the definition, Intersection over Union (IoU) is a measure used in object detection to determine the accuracy of the bounding boxes predicted by the model. It is essentially the ratio of the area of overlap and the area of union of the predicted and actual bounding boxes.

After a thorough iterative approach, which will be further detailed in section 4.3 Performance Analysis & Reparameterization, it was decided to train the model for 620 epochs, using an NVIDIA A100 GPU on Google Colab, which took roughly 1.5 hours to conclude. This number was decided based on the observation that mAP@0.5:0.95 converged into its optimal value around this mark, and any further training could risk overfitting. The mAP@0.5:0.95 was prioritized over mAP@0.5 as it provides a more comprehensive evaluation, taking into account the mAP over a range of IoU values between 0.5 and 0.95. The model achieved mAP@0.5 and mAP@0.5:0.95 of 0.7278 and 0.4284, respectively, at the end of 620 epochs, with their peak values of 0.7458 and 0.4497 achieved at epochs 403 and 602. This confirms that average precision over lower IoU values converges into optimum more quickly, as detections with higher IoU values require further training to improve precision.

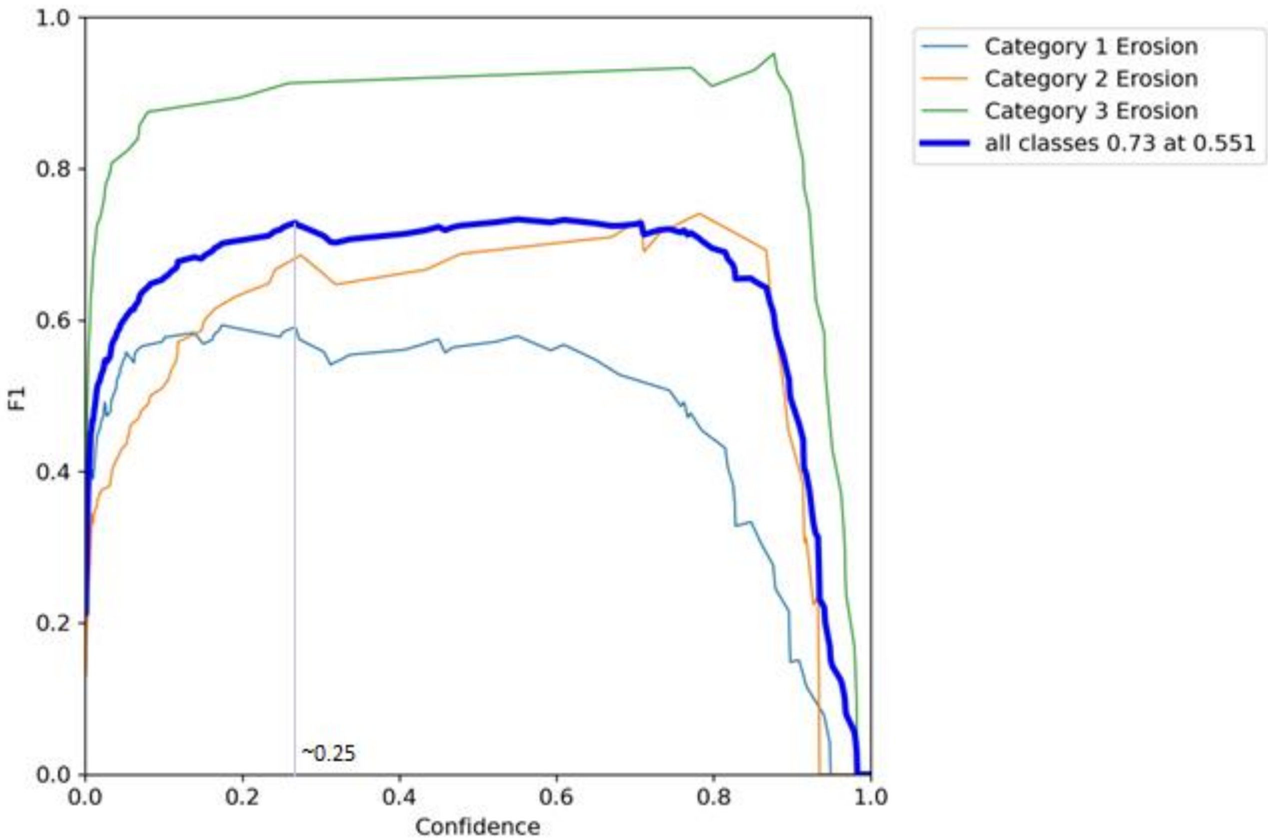

Figure 4.3.4 – F1 over Confidence Values

The F1 value, which is the harmonic mean of precision and recall, also plays an important role in the training outputs. Figure 4.3.4 shows the F1 value plotted over different confidence values. The term "confidence" in this context refers to the probability assigned by the model to the prediction of the object class and the bounding box. A higher confidence score implies that the model is more certain about the class of the object and the location of the bounding box. The best F1 value across all categories was observed around a confidence value of 0.25. As a result, this confidence value was used in the subsequent detection and test results evaluations.

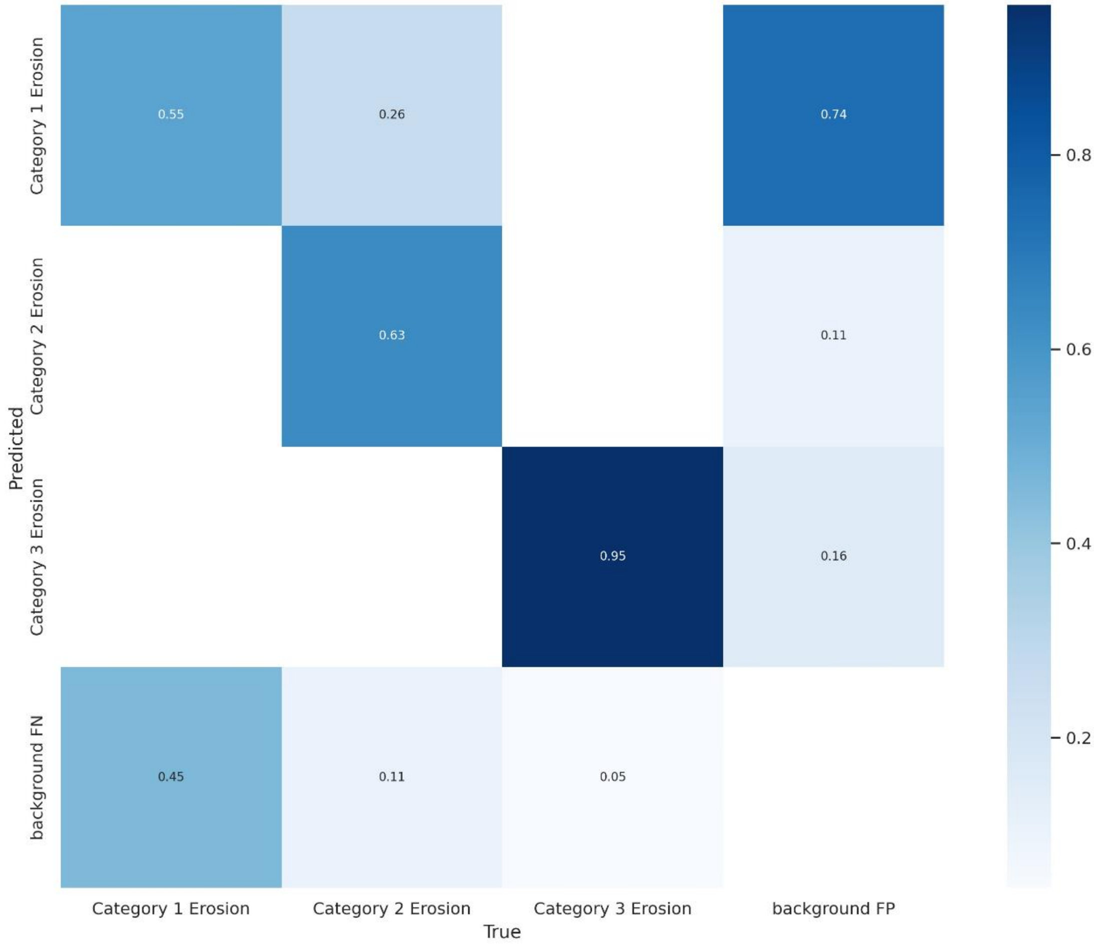

Figure 4.3.5 – Confusion Matrix (Training)

The confusion matrix is another essential output of the training phase. It is a representation that illustrates the performance of the model by comparing the actual and predicted damage categories, as well as false positives (FPs) and false negatives (FNs). This matrix allows us to see where the model is most and least “confused” in its predictions. It is noticeable from the confusion matrix, provided within the outputs of the model and displayed in Figure 4.3.5, that the model demonstrated less confusion when detecting Category 3 Erosion but struggled more with Category 1 Erosion. This finding is understandable, as Category 3 Erosions tend to be larger and more distinct due to exposed blade material, making them easier for the model to identify. Conversely, Category 1 Erosions, being less severe, smaller in size, and more challenging to detect as the distance increases, understandably resulted in greater confusion for the model. Meanwhile, Category 2 Erosion detection results provided a midpoint between the extremes.

These training outputs and the analysis therein set the foundation for further model testing and refinement as discussed in the following sections.

4.4. Performance Analysis & Reparameterization

Performance analysis and reparameterization efforts, within the scope of this study, were primarily concentrated on achieving a balance between the successful training of the model and the risk of overfitting. The primary training performance metric under observation was the mAP@0.5:0.95 score across different numbers of training epochs, namely 220, 300, 500, 620, 1000, and 1500 epochs. Graphical representations of mAP over epochs, displayed in Figure 4.4, for these respective training instances reveal noteworthy patterns and insights, providing a valuable visual aid in understanding the model's performance evolution.

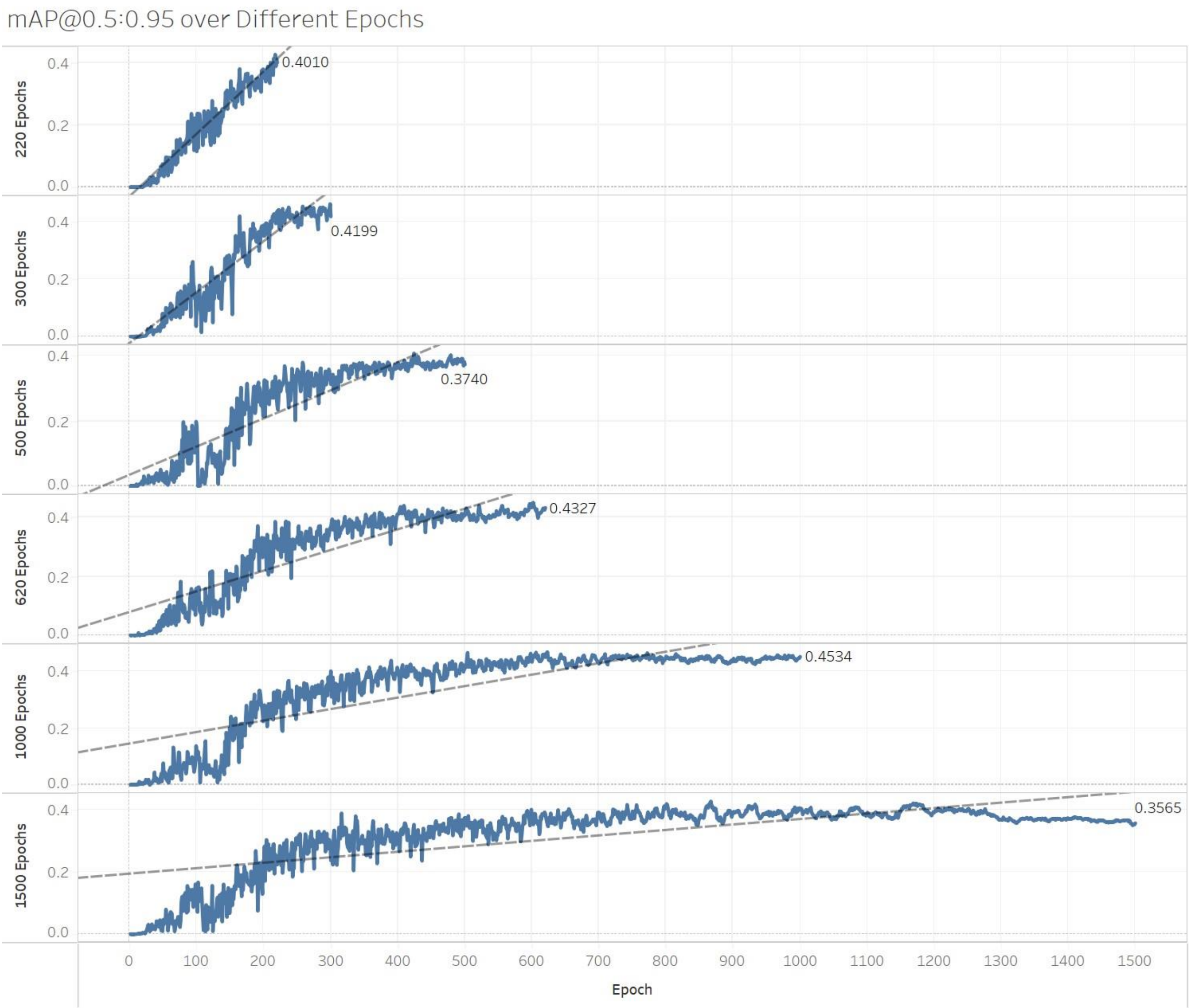

Figure 4.4 - mAP over Different Training Instances

Upon a detailed review of these varied training instances, it can be seen that each instance of training follows a similar pattern, but is actually unique in detail. This makes it quite complex to pinpoint the exact epoch that results in optimal training, rather it is more practical to approximate this epoch based on the results of each training. The aim was, therefore, to identify the mAP@0.5:0.95 value to which the training sessions converged. Once identified, the training was supposed to be terminated to prevent the model from training too long at this convergence point, which increases the risk overfitting.

Through the evaluation of the various training instances, it was found that the mAP@0.5:0.95 score, to which the trainings converged, ranged approximately between 0.4 and 0.45. The maximum concluding mAP recorded among the training instances was 0.4534, achieved at 1000 epochs. However, a notable deceleration in the increase of mAP over epochs was observed at around 620 epochs. The final mAP score reached at the end of training for 620 epochs was 0.4327. This 620-epoch instance offered an ideal compromise between training thoroughness and overfitting risk, thereby serving as the chosen model configuration for the remaining phases of the research.

4.5. Model Testing

Following the model training, the testing phase was carried out to evaluate the trained model's performance on a separate set of images. Each image was processed by the model, which then identified potential areas of damage based on its training. Bounding boxes were drawn around these regions, with the model assigning a category of damage and a confidence value to each detected area. A sample of detections made by the model are shown in Figure 4.5.1, additional output images can also be found in Appendix C. The confidence value signifies the model's certainty regarding the type of damage detected and serves as a quantifiable measure of its precision.

The minimum confidence threshold, a critical parameter that was identified during the testing phase, plays a crucial role in validation of the detections. Only detections with confidence values surpassing this threshold were deemed valid and marked with bounding boxes in the images. To optimize the balance between precision and recall, this threshold was set to 0.25, as per the best F1 score achieved during training.

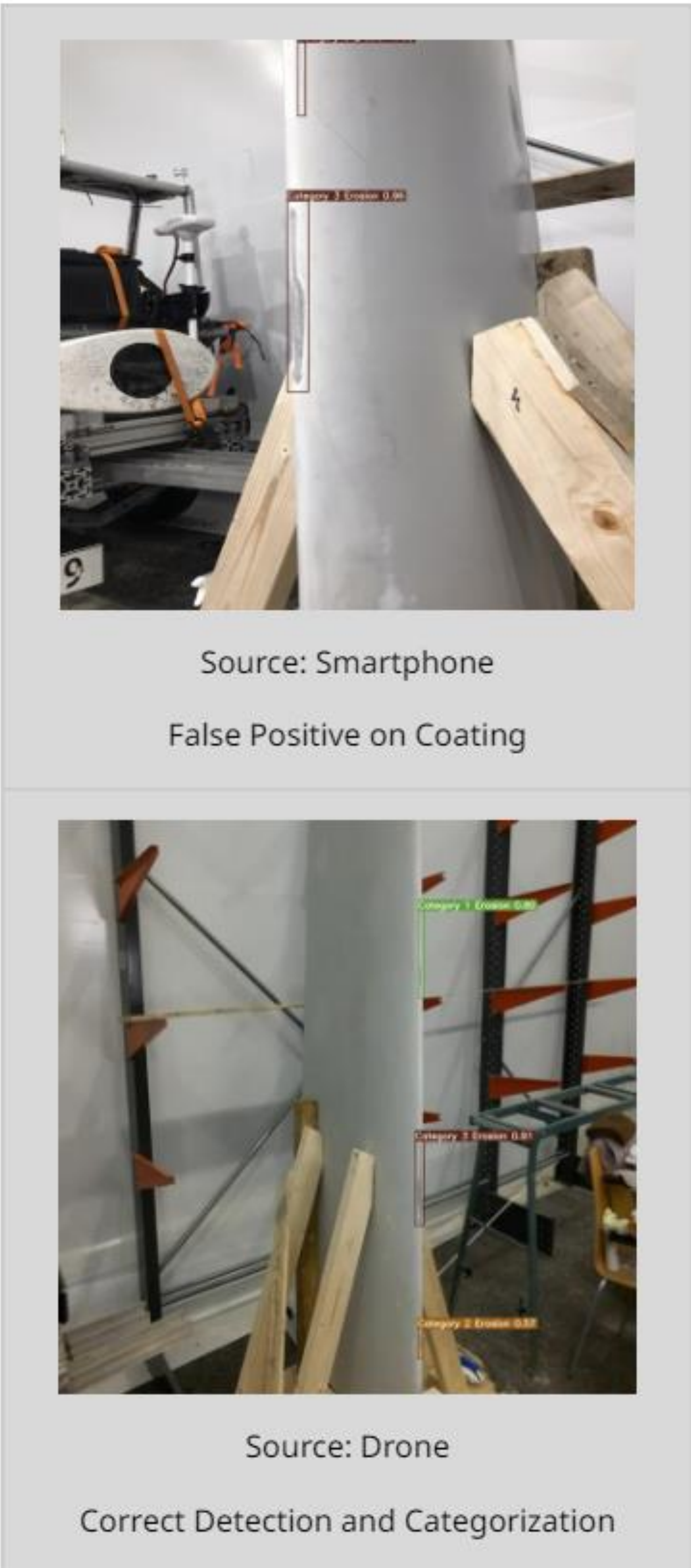

Figure 4.5.1 – Model Detections

An essential component of the test results once more is the confusion matrix displayed in Figure 4.5.2, a graphical representation illustrating the model's performance against the ground truth for the test set. Upon examination, it becomes evident that the model performed exceptionally well in detecting Category 3 Erosion damages. In fact, every instance of actual Category 3 damage in the test set was correctly identified by the model. However, among the total detections made, a minor portion of Category 3 detections were false positives, with the model mistakenly identifying background elements as damage.

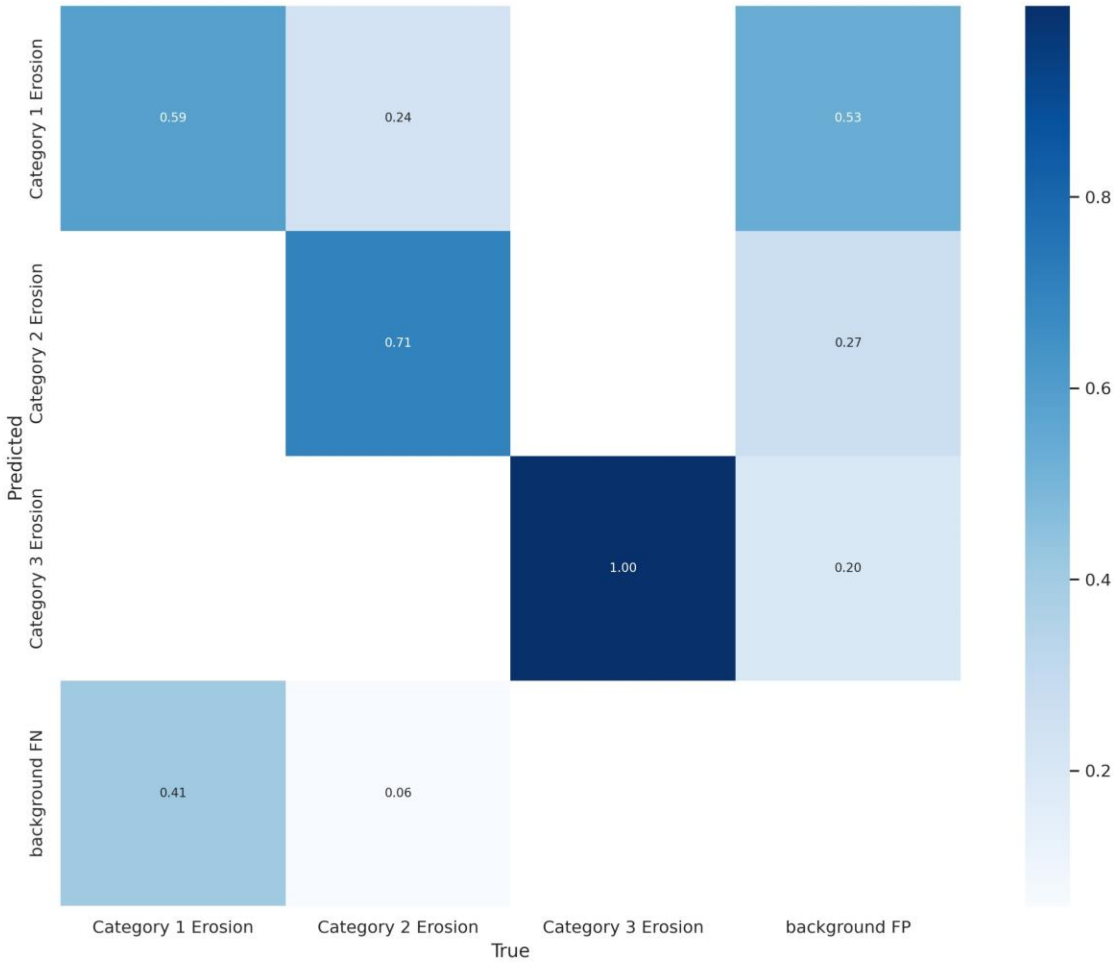

Figure 4.5.2 – Confusion Matrix (Testing)

This superior detection rate for Category 3 damages can be attributed to the unique visual characteristics of these damages, which makes them easier to recognize. On the other hand, the performance was somewhat less precise for Category 2 Erosion damages. Although the model accurately classified the majority of Category 2 instances, it can tend to incorrectly classify a quarter of the instances as Category 1 Erosion. A small fraction of Category 2 Erosions were also mislabeled as background objects.

Turning the attention to Category 1 Erosion damages, the model demonstrated a satisfactory degree of detection and categorization accuracy. Despite being the most challenging to detect due to their smaller size and lesser impact on the turbine's performance, the detection of Category 1 damages is arguably the most critical aspect of this study. The early detection of these damages can drive proactive maintenance planning, which in turn prevents the development of these damages to more severe categories. Therefore, while the overall detection rate for Category 1 Erosion is the lowest among the three categories, it remains at an acceptable level, thus making the model a viable tool for maintenance activities in most cases. An analysis of false positives reveals that most errors occur with Category 1 Erosion. Approximately half of the instances, background objects were categorized as Category 1 Erosion, with the remainder divided between Categories 2 and 3.

On a broader perspective, the testing outcomes correlate closely with the results obtained during the training phase. For instance, the model's superior performance in identifying Category 3 Erosion damages and the greater degree of difficulty it encounters when detecting Category 1 Erosion damages are aspects that persist between both stages. This consistent pattern of performance reinforces the model's strengths and highlights areas where further fine-tuning may be necessary.

This level of consistency is a promising indicator of the model's ability to generalize and reliably detect and categorize damages in wind turbine blades, regardless of the dataset it is applied to. This adaptability is also crucial as it affirms the model's utility in a real-world context, where input data can greatly vary. This is a significant achievement, as the primary goal of any machine learning model is to not only perform well on training data but also maintain that performance level on new, untrained data.

Therefore, the close alignment of the testing results with the training outputs underscores the model's robustness and its potential as a valuable tool in the ongoing maintenance and monitoring of wind turbines. These findings suggest that the model can potentially aid in the early detection of blade erosion, paving the way for timely maintenance activities and, in turn, potentially contributing to more efficient and sustainable wind energy production.

4.6. Imaging Device Evaluation

The implementation of the evaluative framework detailed in the methodology section yielded specific success rates for each of the three imaging devices. The results show that the digital camera was the most successful with a rate of 84%, followed by the smartphone at 78%, and the drone at 75%. It's crucial to clarify that these percentages do not represent the proportion of damages correctly identified; instead, they signify the confidence level of the detections produced by each imaging device. In other words, a success rate of 75% for the drone does not suggest that it misses 1 out of every 4 damage present. Rather, it indicates that the confidence levels associated with its detections tend to be lower than those of the other devices. Figure 4.6.1 provides a comprehensive representation of these detailed percentages and corresponding visual data.

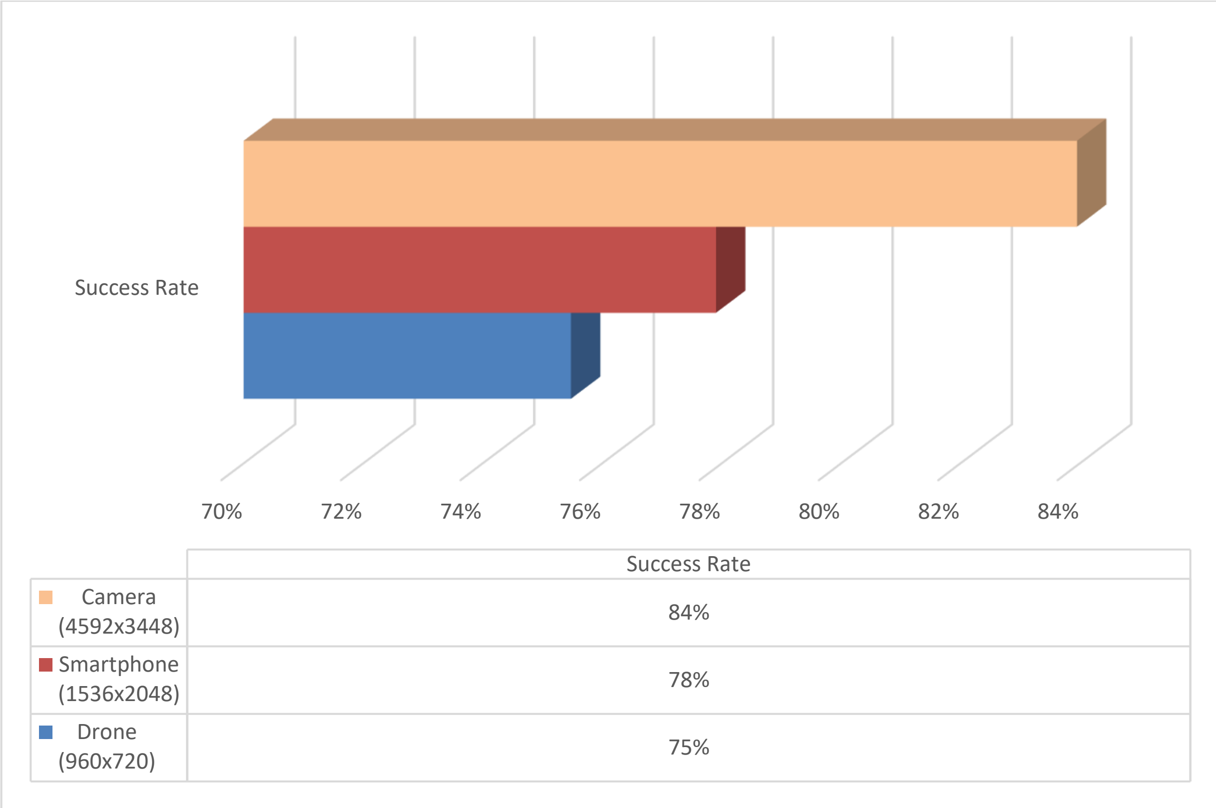

Figure 4.6.1 – Success Rate of Devices

To put things into perspective, the resolutions of the images from the drone, smartphone, and digital camera were 960x720, 1536x2048, and 4592x3448 pixels, respectively. In order to lay the foundations for an appropriate comparison, diagonal resolution values were used in visuals, instead of length and width resolution values of images. When these values are compared directly with the respective success rates, a clear pattern emerges, suggesting that the resolution of the input images indeed impacts the performance of the custom model. This correlation can be attributed to the fact that higher resolution images are able to represent more detailed and unique characteristics of each damage category, thereby enabling the model to make more confident detections. Indeed, a higher resolution tends to yield sharper images, which in turn allows for more precise and accurate damage detections. This is supported by Figure 4.6.2, illustrating the relationship between the diagonal resolution of each device and their respective success rates.

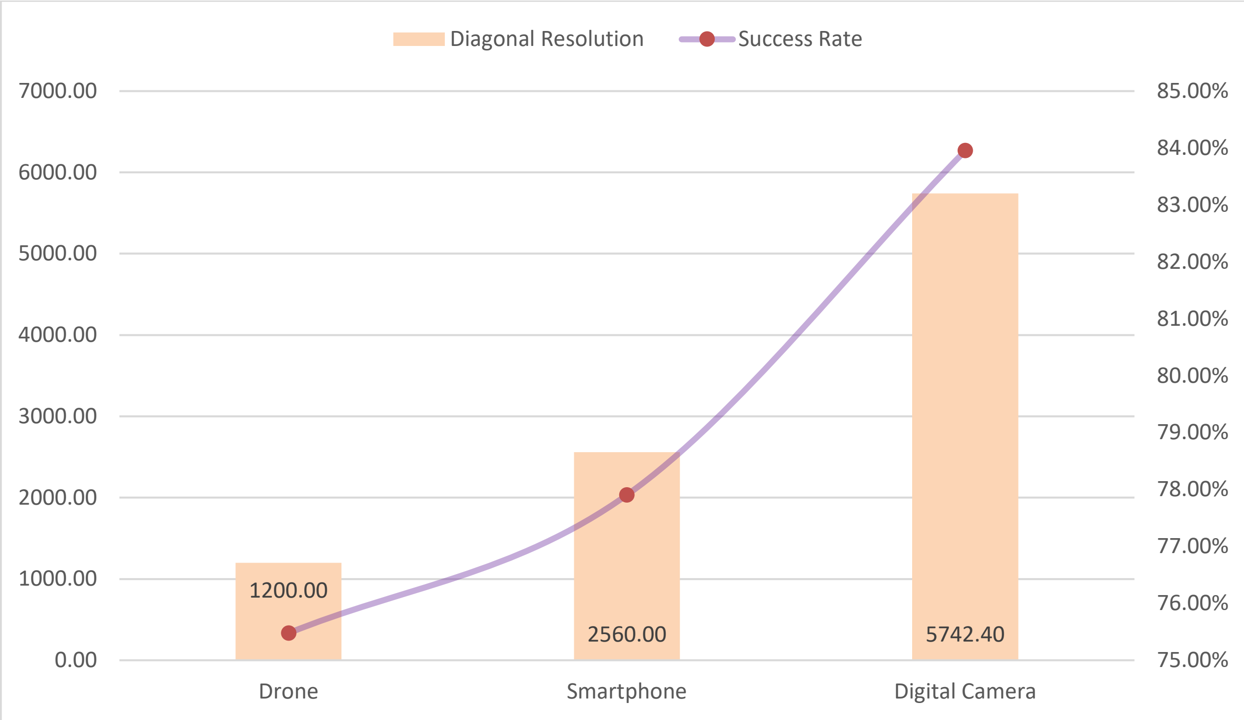

Figure 4.6.2 – Resolution and Success Rate Relation

In conclusion, while all three devices achieved considerable success rates, the digital camera emerged as the most effective imaging device for damage detection, likely due to its superior resolution capabilities. This evaluation and the resulting findings contribute a valuable perspective on the role of imaging devices in damage detection and categorization for wind turbine blades, as well as providing guidance for future studies and industrial applications.

5. Discussion

This chapter of the research is dedicated to a comprehensive discussion on the interpretation of the study's findings, their potential implications for the industry, the key takeaways, and how future studies can build upon this work. This discussion provides an opportunity to touch upon the aspects that could influence the future direction of research in this field.

As the research evolves, it becomes clear that the resolution of the input images can indeed have a noticeable impact on the confidence with which detections are made by the YOLOv7 object detection model when it is trained on custom data. One of the key findings of this study is that the model's performance is at its peak when dealing with Category 3 damages, which are the most advanced and severe types of damage. These are also the largest in terms of size and are highly distinguishable in color due to the often exposure of the blade's structural material. While this high level of performance in detecting Category 3 damages is indeed encouraging, it is, however, to some extent expected due to the distinct visual characteristics that this category typically displays.

On the other side of the spectrum, it is observed that while the model's performance in detecting Category 1 and 2 damages is acceptable, there exists a degree of potential for further optimization and fine-tuning. These initial categories of damage, which are relatively smaller in size and more challenging to detect based on turbine performance metrics, could greatly benefit from an improved detection model. As the accumulation and progression of leading-edge erosion typically begin with these less noticeable damages, the model's ability to correctly identify them can provide significant value.

Currently, the model has the potential provide beneficial insights and contribute to resource-efficient maintenance planning when supervised. However, its readiness for unsupervised operation is still uncertain due to the potential risk of generating false alarms. This characteristic brings into focus the importance of Category 1 and 2 damages in terms of early detection and preventive measures.

To enhance the model's performance in identifying these initial stages of damage, expanding the dataset to cover a greater variety of unique examples from these categories could be a potential solution. It is understood that, due to confidentiality concerns within the wind energy industry, there is limited focus and prioritization towards compiling a standardized dataset among operators of wind farms or parties overseeing maintenance activities. Yet, the insights gained from this study hopefully should incentivize these industry players to consider developing their own image datasets, varying in scale from turbine-specific to site or location-specific, as well as damage category-specific.

The potential of machine learning techniques in visual inspections is undeniable. If datasets of appropriate quality were readily available, it could significantly reduce the human labor required for manual annotation of existing images and dataset generation, thereby boosting the efficiency and effectiveness of these inspections. This shift towards leveraging artificial intelligence could help not only in damage detection but also in streamlining maintenance strategies, making a significant impact on the industry as a whole.

6. Conclusion and Future Work

Overall, this research highlights the stages of training a custom object detection model, suitable for use in detecting and categorizing LEE damages, including testing its capabilities on a training set comprised of images from different sources. This allowed for a comparison between the imaging devices, thereby evaluating the model's performance on different input resolutions.

The evaluation resulted in the discovery and affirmation that detection attempts on different classes being made with greater confidence on images with higher resolution, by the custom model. The outputs of the experiments brought forward a potential correlation between input image resolution and detection confidence on multi-class object detection, in the context of wind turbine maintenance. Ultimately, the findings highlight the importance and impact of imaging hardware selection on inspections, as well as encouraging industry professionals to construct datasets of appropriate quality to be machine learning-ready.

Looking ahead, there are several potentials for future research that could build upon the findings of this study. The first concerns the exploration of other computer vision algorithms. While this research was focused on the use of the main configuration of YOLOv7 object detection model due to its superior performance within the time constraints of the study, examining other configurations of YOLOv7 or exploring different architectures like Faster R-CNN might provide additional insight onto the effect of input resolution on damage detection and categorization. Within each model itself, exploring the optimal parameter selection regarding the number of epochs or confidence threshold could also further enhance the existing practices.

Additionally, the method employed in this study for gathering damage information, namely inspection photography, could be further enhanced with the inclusion of recent advancements in Light Detection and Ranging (LiDAR) technology. LiDAR is a remote sensing method that uses light in the form of a laser to measure distances between the source and objects, creating precise, three-dimensional information about shapes and surface characteristics. By utilizing LiDAR sensors, it may be possible to obtain the dimensional properties of each unique damage. This information could offer a deeper understanding of natural erosion patterns, improving predictive modeling on erosion and strategic maintenance planning.

A final promising field for future research lies on the use of artificial intelligence to enhance image quality. Recent advancements in this field have led to the development and distribution of image upscalers that can artificially increase the resolution of images to ultra-high values. It would be worthwhile to examine the effects of such artificial upscaling on input images. If successful, this could reduce the need for high-resolution inspection devices, further cutting hardware costs. Instead, an optimal imaging mode could be identified and used in combination with artificial upscaling to form a complete object detection input pipeline.

In conclusion, the findings of this study offer valuable insights into the role of image resolution in damage detection and categorization for wind turbine blades, through utilizing machine learning technology. It provides a strong foundation for future research, displaying new possibilities for the application of artificial intelligence in the field of wind energy maintenance and management. This work, finally, serves as an incentive and encouragement for industry stakeholders to leverage the power of machine learning and consider creating their own tailored image datasets with tools that provide higher resolution, thus leading the way for more efficient and effective visual inspections. Such advancements would introduce opportunities for more streamlined and efficient operational procedures in the field of wind turbine maintenance, ultimately contributing towards a more sustainable future.

References

Anisimov, A. G., Beukema, R., Hwang, J., Nijssen, R., & Groves, R. M. (2021). *AIRTuB: towards automated inspection of leading-edge erosion of wind turbine blades by shape analysis*. 29. https://doi.org/10.1117/12.2592291

Artigao, E., Martín-Martínez, S., Honrubia-Escribano, A., & Gómez-Lázaro, E. (2018). Wind turbine reliability: A comprehensive review towards effective condition monitoring development. In *Applied Energy* (Vol. 228). https://doi.org/10.1016/j.apenergy.2018.07.037

Brownlee, J. (2018). What is the Difference Between a Batch and an Epoch in a Neural Network? *Machine Learning Mastery*, *20*.

Carroll, J., McDonald, A., & McMillan, D. (2016). Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind Energy*, *19*(6), 1107–1119. https://doi.org/10.1002/we.1887

Ciang, C. C., Lee, J. R., & Bang, H. J. (2008). Structural health monitoring for a wind turbine system: A review of damage detection methods. *Measurement Science and Technology*, *19*(12). https://doi.org/10.1088/0957-0233/19/12/122001

Dimitrov, N. (2018). Risk-based approach for rational categorization of damage observations from wind turbine blade inspections. *Journal of Physics: Conference Series*, *1037*(4). https://doi.org/10.1088/1742-6596/1037/4/042021

Dodge, S., & Karam, L. (2016). Understanding how image quality affects deep neural networks. *2016 8th International Conference on Quality of Multimedia Experience, QoMEX 2016*. https://doi.org/10.1109/QoMEX.2016.7498955

Du, Y., Zhou, S., Jing, X., Peng, Y., Wu, H., & Kwok, N. (2020). Damage detection techniques for wind turbine blades: A review. In *Mechanical Systems and Signal Processing* (Vol. 141). https://doi.org/10.1016/j.ymssp.2019.106445

Duthé, G., Abdallah, I., Barber, S., & Chatzi, E. (2021). Modeling and monitoring erosion of the leading edge of wind turbine blades. *Energies*, *14*(21). https://doi.org/10.3390/en14217262

García Márquez, F. P., Tobias, A. M., Pinar Pérez, J. M., & Papaelias, M. (2012). Condition monitoring of wind turbines: Techniques and methods. *Renewable Energy*, *46*. https://doi.org/10.1016/j.renene.2012.03.003

Herring, R., Dyer, K., Martin, F., & Ward, C. (2019a). The increasing importance of leading edge erosion and a review of existing protection solutions. In *Renewable and Sustainable Energy Reviews* (Vol. 115). Elsevier Ltd. https://doi.org/10.1016/j.rser.2019.109382

Keegan, M. H., Nash, D. H., & Stack, M. M. (2013). On erosion issues associated with the leading edge of wind turbine blades. In *Journal of Physics D: Applied Physics* (Vol. 46, Issue 38). https://doi.org/10.1088/0022-3727/46/38/383001

Koziarski, M., & Cyganek, B. (2018). Impact of low resolution on image recognition with deep neural networks: An experimental study. *International Journal of Applied Mathematics and Computer Science*, *28*(4). https://doi.org/10.2478/amcs-2018-0056

Law, H., & Koutsos, V. (2020). Leading edge erosion of wind turbines: Effect of solid airborne particles and rain on operational wind farms. *Wind Energy*, *23*(10). https://doi.org/10.1002/we.2540

Lee, D., Kim, J., & Lee, D. (2019). Robust Concrete Crack Detection Using Deep Learning-Based Semantic Segmentation. *International Journal of Aeronautical and Space Sciences*, *20*(1). https://doi.org/10.1007/s42405-018-0120-5

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2014). *Microsoft COCO: Common Objects in Context*. http://arxiv.org/abs/1405.0312

Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., & Pietikäinen, M. (2020). Deep Learning for Generic Object Detection: A Survey. *International Journal of Computer Vision*, *128*(2). https://doi.org/10.1007/s11263-019-01247-4

López, J. C., Kolios, A., Wang, L., & Chiachio, M. (2023). A wind turbine blade leading edge rain erosion computational framework. *Renewable Energy*, *203*, 131–141. https://doi.org/10.1016/j.renene.2022.12.050

McGugan, M., & Mishnaevsky, L. (2020). Damage mechanism based approach to the structural health monitoring of wind turbine blades. *Coatings*, *10*(12). https://doi.org/10.3390/coatings10121223

Mishnaevsky, L. (2019). Repair of wind turbine blades: Review of methods and related computational mechanics problems. In *Renewable Energy* (Vol. 140). https://doi.org/10.1016/j.renene.2019.03.113

Mishnaevsky, L. (2022). Root Causes and Mechanisms of Failure of Wind Turbine Blades: Overview. In *Materials* (Vol. 15, Issue 9). MDPI. https://doi.org/10.3390/ma15092959

Mishnaevsky, L., Hasager, C. B., Bak, C., Tilg, A. M., Bech, J. I., Doagou Rad, S., & Fæster, S. (2021). Leading edge erosion of wind turbine blades: Understanding, prevention and protection. *Renewable Energy*, *169*, 953–969. https://doi.org/10.1016/J.RENENE.2021.01.044

Perez, L., & Wang, J. (2017). *The Effectiveness of Data Augmentation in Image Classification using Deep Learning*. http://arxiv.org/abs/1712.04621

Pinar Pérez, J. M., García Márquez, F. P., Tobias, A., & Papaelias, M. (2013). Wind turbine reliability analysis. In *Renewable and Sustainable Energy Reviews* (Vol. 23, pp. 463–472). Elsevier Ltd. https://doi.org/10.1016/j.rser.2013.03.018

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*. https://doi.org/10.1109/CVPR.2016.91

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6). https://doi.org/10.1109/TPAMI.2016.2577031

Sabottke, C. F., & Spieler, B. M. (2020). The effect of image resolution on deep learning in radiography. *Radiology: Artificial Intelligence*, *2*(1). https://doi.org/10.1148/ryai.2019190015

Sareen, A., Sapre, C. A., & Selig, M. S. (2014). Effects of leading edge erosion on wind turbine blade performance. *Wind Energy*, *17*(10). https://doi.org/10.1002/we.1649

Shihavuddin, A. S. M., Chen, X., Fedorov, V., Christensen, A. N., Riis, N. A. B., Branner, K., Dahl, A. B., & Paulsen, R. R. (2019). Wind turbine surface damage detection by deep learning aided drone inspection analysis. *Energies*, *12*(4). https://doi.org/10.3390/en12040676

Stokkeland, M., Klausen, K., & Johansen, T. A. (2015). Autonomous visual navigation of Unmanned Aerial Vehicle for wind turbine inspection. *2015 International Conference on Unmanned Aircraft Systems, ICUAS 2015*. https://doi.org/10.1109/ICUAS.2015.7152389

Talebi, H., & Milanfar, P. (2021). Learning to Resize Images for Computer Vision Tasks. *Proceedings of the IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV48922.2021.00055

Tchakoua, P., Wamkeue, R., Ouhrouche, M., Slaoui-Hasnaoui, F., Tameghe, T. A., & Ekemb, G. (2014). Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges. In *Energies* (Vol. 7, Issue 4). https://doi.org/10.3390/en7042595

Thambawita, V., Strümke, I., Hicks, S. A., Halvorsen, P., Parasa, S., & Riegler, M. A. (2021). Impact of image resolution on deep learning performance in endoscopy image classification: An experimental study using a large dataset of endoscopic images. *Diagnostics*, *11*(12). https://doi.org/10.3390/diagnostics11122183

Verma, A. S., Jiang, Z., Ren, Z., & Teuwen, J. J. E. (2020). Leading edge erosion of wind turbine blades: Effects of environmental parameters on impact velocities and erosion damage rate. *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE*, *9*. https://doi.org/10.1115/OMAE2020-18173

Walford, C. a. (2006). Wind turbine reliability: understanding and minimizing wind turbine operation and maintenance costs. *Energy*, *March*.

Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). *YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors*. http://arxiv.org/abs/2207.02696

Wang, L., Yang, J., Huang, C., & Luo, X. (2022). An Improved U-Net Model for Segmenting Wind Turbines from UAV-Taken Images. *IEEE Sensors Letters*, *6*(7). https://doi.org/10.1109/LSENS.2022.3184521

Wang, L., & Zhang, Z. (2017). Automatic Detection of Wind Turbine Blade Surface Cracks Based on UAV-Taken Images. *IEEE Transactions on Industrial Electronics*, *64*(9), 7293–7309. https://doi.org/10.1109/TIE.2017.2682037

Wang, L., Zhang, Z., & Luo, X. (2019). A Two-Stage Data-Driven Approach for Image-Based Wind Turbine Blade Crack Inspections. *IEEE/ASME Transactions on Mechatronics*, *24*(3). https://doi.org/10.1109/TMECH.2019.2908233

Wu, H., Ye, X., Jiang, Y., Tian, H., Yang, K., Cui, C., Shi, S., Liu, Y., Huang, S., Chen, J., Xu, J., & Dong, F. (2022). A Comparative Study of Multiple Deep Learning Models Based on Multi-Input Resolution for Breast Ultrasound Images. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/fonc.2022.869421

Xu, D., Wen, C., & Liu, J. (2019). Wind turbine blade surface inspection based on deep learning and UAV-taken images. *Journal of Renewable and Sustainable Energy*, *11*(5). https://doi.org/10.1063/1.5113532

Xu, X., Zhao, M., Shi, P., Ren, R., He, X., Wei, X., & Yang, H. (2022). Crack Detection and Comparison Study Based on Faster R-CNN and Mask R-CNN. *Sensors*, *22*(3). https://doi.org/10.3390/s22031215

Ye, X. W., Dong, C. Z., & Liu, T. (2016). A Review of Machine Vision-Based Structural Health Monitoring: Methodologies and Applications. *Journal of Sensors*, *2016*. https://doi.org/10.1155/2016/7103039

Ying, X. (2019). An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*, *1168*(2). https://doi.org/10.1088/1742-6596/1168/2/022022

Zhang, D., Watson, R., Dobie, G., MacLeod, C., Khan, A., & Pierce, G. (2020). Quantifying impacts on remote photogrammetric inspection using unmanned aerial vehicles. *Engineering Structures*, *209*. https://doi.org/10.1016/j.engstruct.2019.109940

Zoph, B., Cubuk, E. D., Ghiasi, G., Lin, T. Y., Shlens, J., & Le, Q. V. (2020). Learning Data Augmentation Strategies for Object Detection. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12372 LNCS*. https://doi.org/10.1007/978-3-030-58583-9_34

Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object Detection in 20 Years: A Survey. *Proceedings of the IEEE*. https://doi.org/10.1109/JPROC.2023.3238524

Appendices

Appendix A – Drone Image Capture Mission

```
# Import requirements

import time
from djitellopy import Tello
import cv2
import math


# Define function for image capture

def save_image(frame, file_path):
    # Save the image to a file
    cv2.imwrite(file_path, frame)
    print(f"Image saved to {file_path}")


# Set mission parameters

init_dist = 50                                          # Takeoff point distance from blade
altitudes = [20, 30, 70]                                # Altitudes of damage locations
distances = [init_dist, init_dist*2, init_dist*3]       # Distances for loop
num_photos = 5                                          # Number of images to capture
wait = 2                                                # Seconds between actions to prevent errors
name = "blade"                                          # Part of saved image name

# Connect to the drone

try:
    tello = Tello()

    # Connect to the drone
    tello.connect()
    time.sleep(wait)
    print(f"Battery percentage: {tello.get_battery()}%")

except:
    print("ERROR: Could not connect to Tello.")
```

```
# Initiate mission

time.sleep(wait)

# Take off and reach the desired altitude
print(f"Takeoff...")
tello.takeoff()
time.sleep(wait)


for alt in altitudes:

    tello.move_up(altitudes)
    time.sleep(wait)

    for distance in distances:

        # Image
        for i in range(num_photos-1):
            # Capture an image
            tello.streamon()
            frame = tello.get_frame_read().frame
            time.sleep(wait-1)
            tello.streamoff()
            time.sleep(wait-1)

            # Save the captured image
            save_image(frame, f"photo_{name}_{distance}cm_{i+1}.jpg")
            print(f"Image saved as: photo_{name}_{distance}cm_{i+1}.jpg.")
            time.sleep(wait-1)



        # Move to next position
        if distance != distances[-1]:
            tello.move_back(dist_incr)
            time.sleep(wait)
            print(f"Drone in position.")
        else:
                tello.move_forward((len(distances)-1)*dist_incr)
                time.sleep(wait)

# Land
print(f"Landing...")
tello.land()
print(f"Battery percentage: {tello.get_battery()}%")
print(f"Mission completed in {tello.get_flight_time()} time units.")
```

Appendix B – YOLOv7 Custom Training, Testing and Export

```
Download YOLOv7 and install requirements #
*****************************************

!git clone https://github.com/WongKinYiu/yolov7
%cd yolov7
!pip install -r requirements.txt


# Export and paste connection training-validation-test dataset from Roboflow in "YOLOv7 PyTorch" format
# *****************************************************************************************************

!pip install roboflow

from roboflow import Roboflow
rf = Roboflow(api_key="**********************")
project = rf.workspace("*****").project("*************")
dataset = project.version(**).download("******")


# Download YOLOv7 MS COCO starting checkpoint
# *******************************************

%cd /content/yolov7
!wget https://github.com/WongKinYiu/yolov7/releases/download/v0.1/yolov7_training.pt


# Begin training on COCO checkpoint (set batch size 2^n based on GPU power, adjust number of epochs for training duration)
# ************************************************************************************************************************

%cd /content/yolov7
!python train.py --batch 64 --epochs 300 --data {dataset.location}/data.yaml --weights 'yolov7_training.pt' --device 0


# Run detection on test images (adjust confidence based on F1 vs Confidence graph on training output)
# ***************************************************************************************************

%cd /content/yolov7
!python detect.py --weights runs/train/exp/weights/best.pt --conf 0.1 --source {dataset.location}/test/images
!python test.py --weights runs/train/exp/weights/best.pt --conf 0.1 --data {dataset.location}/data.yaml --device 0  --batch 64


# Zip training weights and results
# ********************************

!zip -r export.zip runs/detect
!zip -r export.zip runs/test
!zip -r export.zip runs/train/exp/weights/best.pt
!zip export.zip runs/train/exp/*


# Download zipped results
# ************************

from google.colab import files
files.download("/content/yolov7/export.zip")
```

Appendix C – Certain Model Outputs

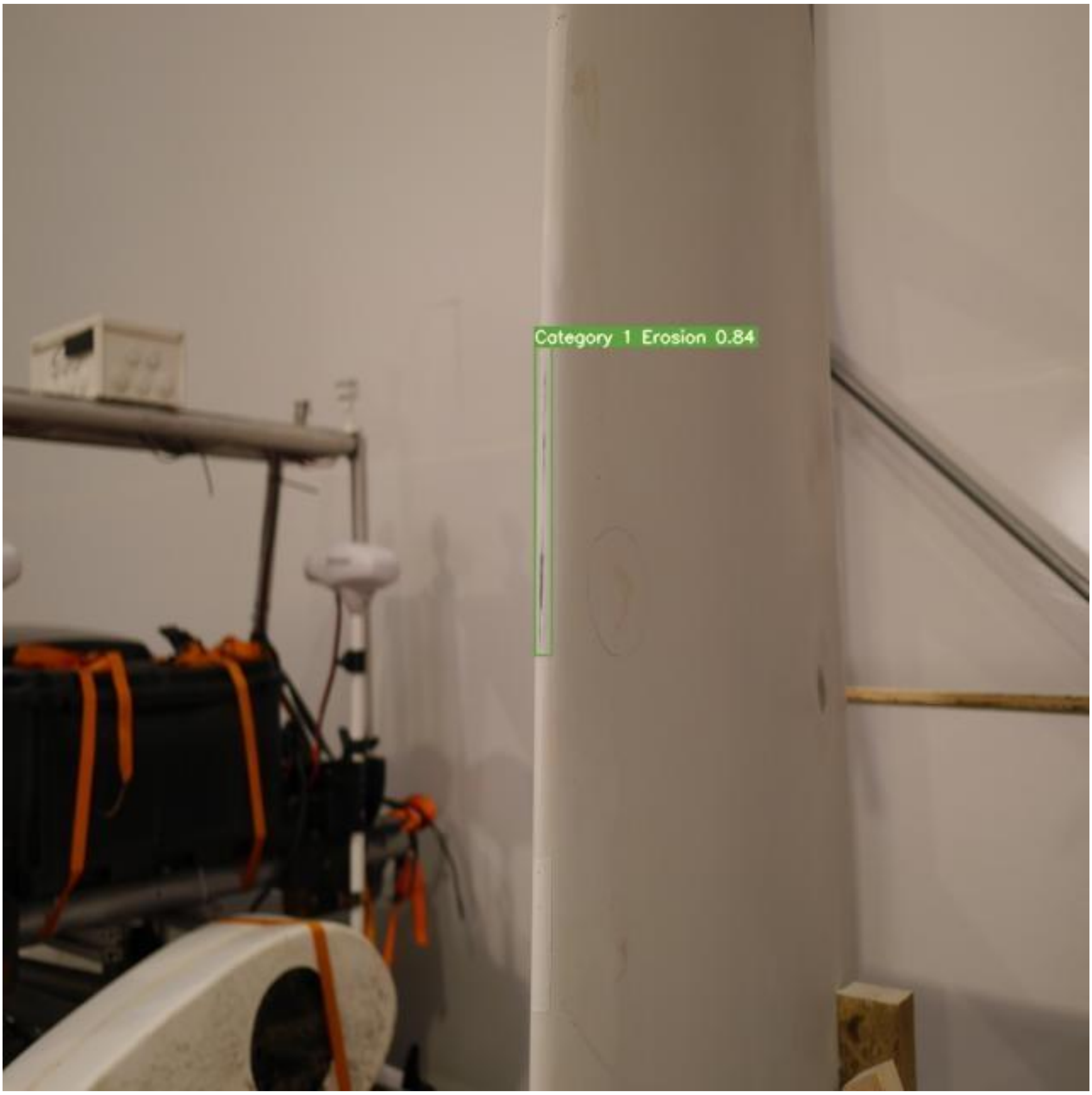

Image Device: Digital Camera

Image Device: Smartphone

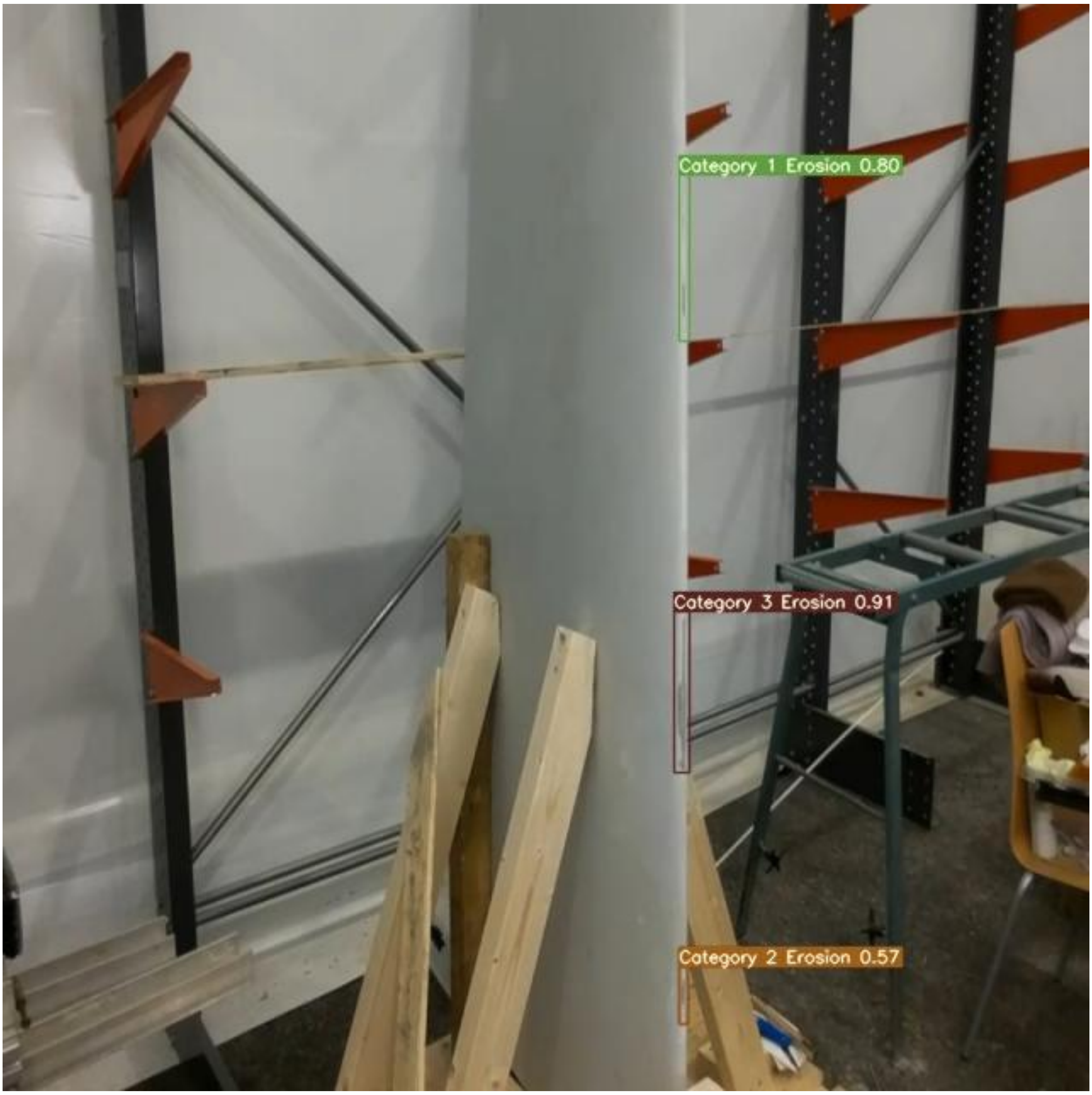

Image Device: Ryze Tello (Drone)

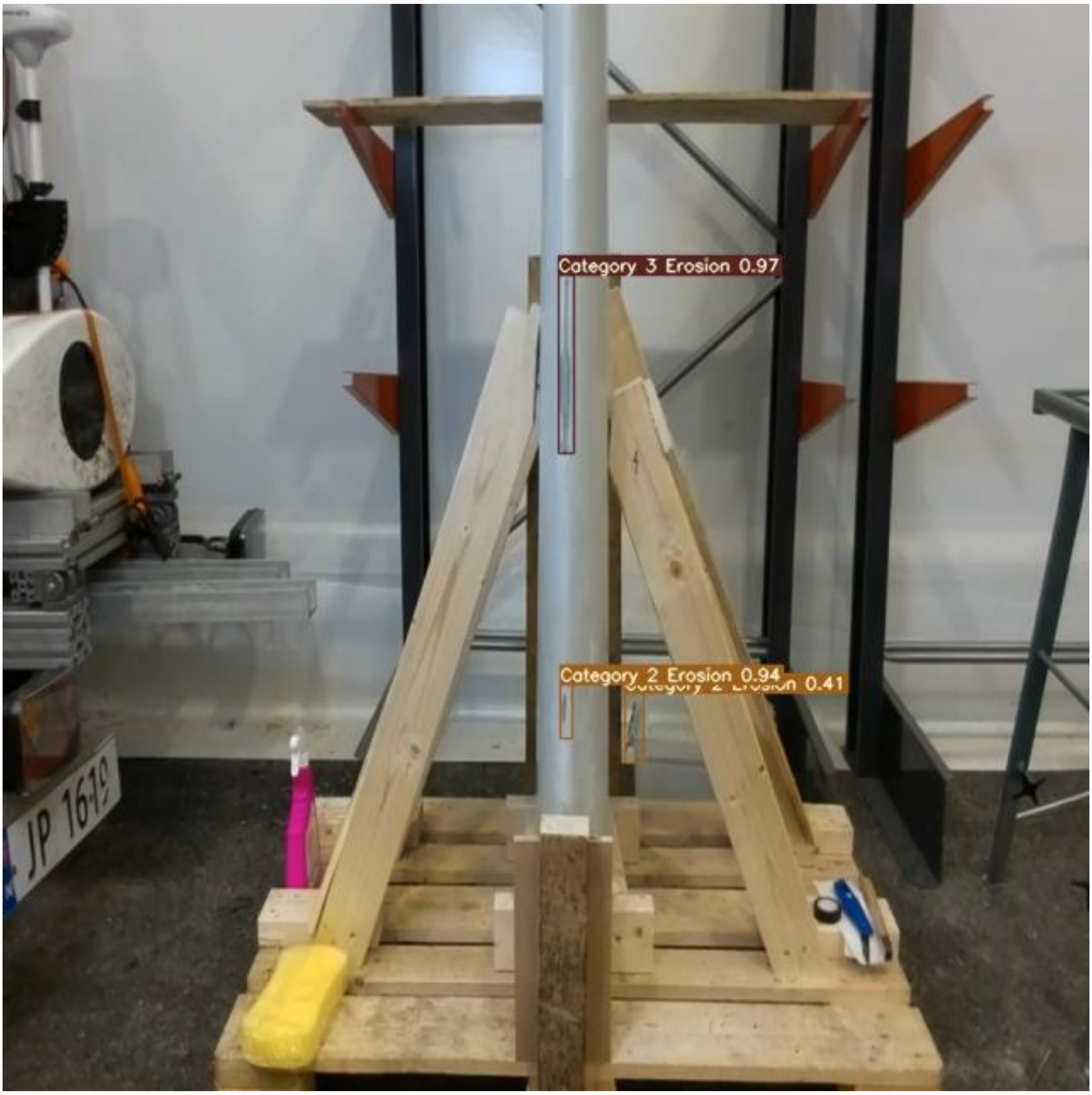

Image Device: Ryze Tello (Drone)
False Positive in Background

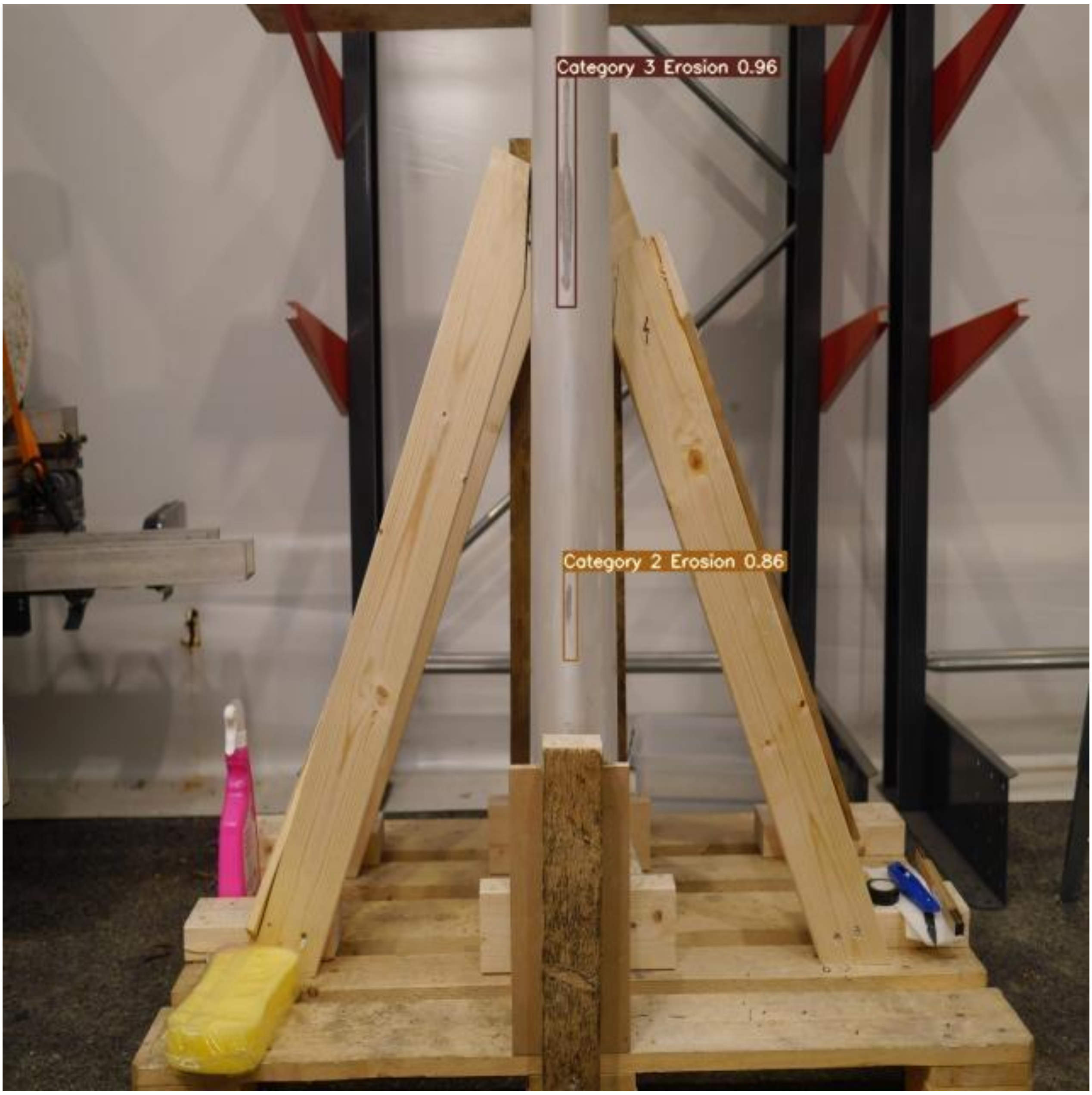

Image Device: Digital Camera