RESEARCH PROGRESS IN MECHANICAL AND MANUFACTURING ENGINEERING
e-ISSN: 2773-4765

RPMME

Vol. 7 No. 1 (2026) 140-149
https://publisher.uthm.edu.my/periodicals/index.php/rpmme

Application of Artificial Intelligence Models for Enhancing Non-Destructive Inspection of Aircraft Wheel Tie Bolt Component Through Fluorescent Penetrant Inspection

M. Afiq Norwahidi[1], M.F Mahmod[1*]

[1] *Faculty of Mechanical and Manufacturing Engineering*
Universiti Tun Hussein Onn Malaysia, Batu Pahat, 86400, Malaysia

*Corresponding Author: mfaisal@uthm.edu.my
DOI: https://doi.org/10.30880/rpmme.2026.07.01.015

Article Info

Received: 10 December 2025
Accepted: 9 March 2026
Available online: 28 February 2026

Keywords

Non Destructive Testing (NDT), Fluorescent Penetrant Inspection (FPI), Artificial Intelligence (AI), Ultraviolet (UV)

Abstract

Non-destructive testing (NDT) is crucial for maintaining the structural integrity of aircraft components, particularly wheel tie bolts that are prone to surface defects like cracks due to stress and environmental factors. Fluorescent Penetrant Inspection (FPI), with its high sensitivity compared to Dye Penetrant Testing (DPT), is commonly used in the aviation industry to detect such anomalies. This study presents an application of artificial intelligence (AI) image detection inspection method. FPI images of the aircraft wheel tie bolt of a narrow body type of aircraft were trained and analyzed. The defects were identified and inspected using machine learning. The findings present an application, which can improve the aerospace Non-Destructive testing procedures with higher detection consistency on human inspection.

1. Introduction

Aviation has significantly influenced global development by transforming how people travel, trade, and connect across vast distances. It is a complex, technology driven industry that supports a wide range of sectors such as tourism, logistics, and international commerce, while also generating extensive employment opportunities worldwide [1]. Aircraft systems consist of several critical components, including the fuselage, wings, empennage, and landing gear. The landing gear system comprising essential parts such as the wheel hub and aircraft wheel tie bolts ensures structural stability during ground operations and absorbs high-impact forces during landing and takeoff [2]. In line with this, the objective of this study is to enhance the detection and analysis of surface defects, specifically cracks in aircraft wheel tie bolts, through the application of Artificial Intelligence techniques to Fluorescent Penetrant Inspection data.

Fluorescent Penetrant Inspection (FPI) is a highly effective Non-Destructive Testing (NDT) technique commonly used to detect surface-breaking defects in metals and other non-porous materials on aircraft component [3]. Surface cleaning, penetrant application, dwell time, excess penetrant removal, and developer application to bring out the flaws for improved visibility are all steps in the methodical procedure [4]. The fluorescent dye accentuates any imperfections under ultraviolet (UV) light [5], which enables inspectors to find minute cracks or discontinuities that are not visible to the naked eye.

Defect learning-based detection methods are broadly categorized into three categories: supervised, semi-supervised, and unsupervised. Large number of different types of defects in a single defect detection task, where defect diversity substantially contributes to the problem complexity of defect detection and makes it difficult to develop a generalized solution [6]. Each of them has some pros and cons. Because certain defects are rare, obtaining such data especially those labeled as defects is usually expensive and time-consuming. Stronger is the ability provided by object detection models to detect and identify multiple defect regions from a single image, thereby addressing severe weaknesses of classification-based methods. Generative adversarial networks and convolutional neural network and stabilizes the training process more than the first one [6]. The training process of generators and discriminators needs to be well balanced in order to avoid the unstable process for improving the diversity of samples and convergence possibility.

2. Methodology

This research follows a structured methodology as illustrated in the figure 2.1 that begins with the sample preparation phase, involving both real and artificially simulated defects on aircraft wheel tie bolt components to ensure diverse and representative data. The prepared samples then undergo fluorescent penetrant inspection (FPI) to reveal surface-breaking defects for further analysis. The captured FPI images are subjected to image pre-processing, which includes data augmentation techniques [7] such as flipping, zooming, and rotation, along with grayscale conversion [8] to enhance dataset diversity and simplify image features. Following pre-processing, the post-processing phase is carried out, consisting of the annotation of defect areas using CVAT and pixel size calibration to standardize all images to a uniform resolution of 416 × 416 pixels. The final stage involves the training and validation setup of the artificial intelligence model, including the design of a graphical user interface (GUI) to allow for practical implementation and testing of defect detection capabilities.

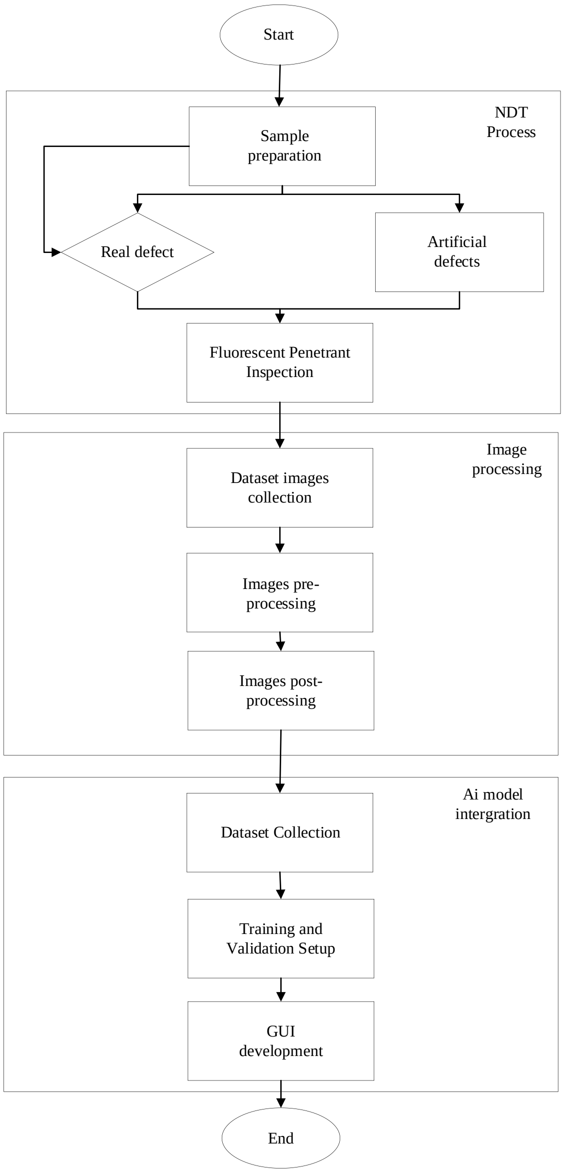


Fig. 2.1 *Process flowchart methodology*

3. Results and Discussion

This section presents the results of each main stage in the project. Section 3.1 shows the outcome of the fluorescent penetrant inspection (FPI), where both real and artificial cracks on aircraft wheel tie bolts were detected. In Section 3.2, the dataset was prepared and expanded using data augmentation techniques like flipping, rotation, zooming, and grayscale conversion to improve model performance. Section 3.3 covers image preprocessing and post-processing, including pixel calibration and manual annotation using CVAT to label cracks. These steps were essential in building an accurate and reliable machine learning model for defect detection.

3.1 Fluorescent Penetrant Inspection on Aircraft Tie Bolt

FPI process involved cleaning the surface, applying penetrant, removing excess penetrant, applying developer, and UV-A light inspection. Cleaning was accomplished with Magnaflux SKC-S followed by ZL-27A fluorescent penetrant. After dwell time and cleaning, entrapped penetrant was pulled out with SKD-S2 developer. Upon illumination under UV-A light, surface cracks fluoresced and revealed defects in the aircraft wheel tie bolt.

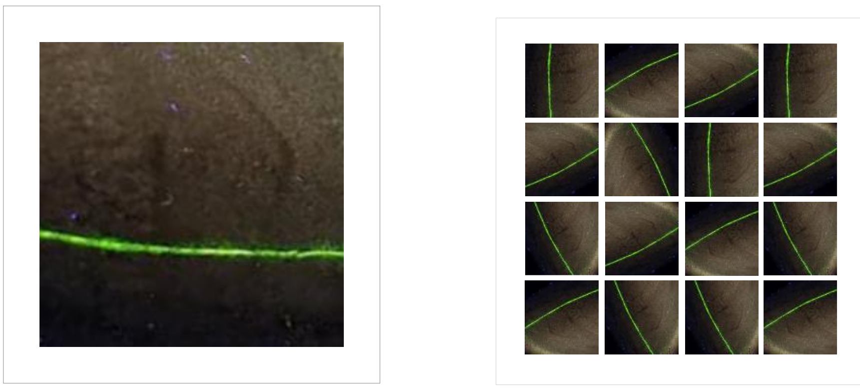
(a) (b)

Fig. 3.1 *Detection of defects using fluorescent penetrant inspection (a) Before FPI; (b) After FPI*

Based on the results, the FPI process successfully revealed visible indications of surface-breaking cracks on the aircraft wheel tie bolt, confirming the presence of both real and artificial defects.

3.2 Dataset Preparation and Augmentation

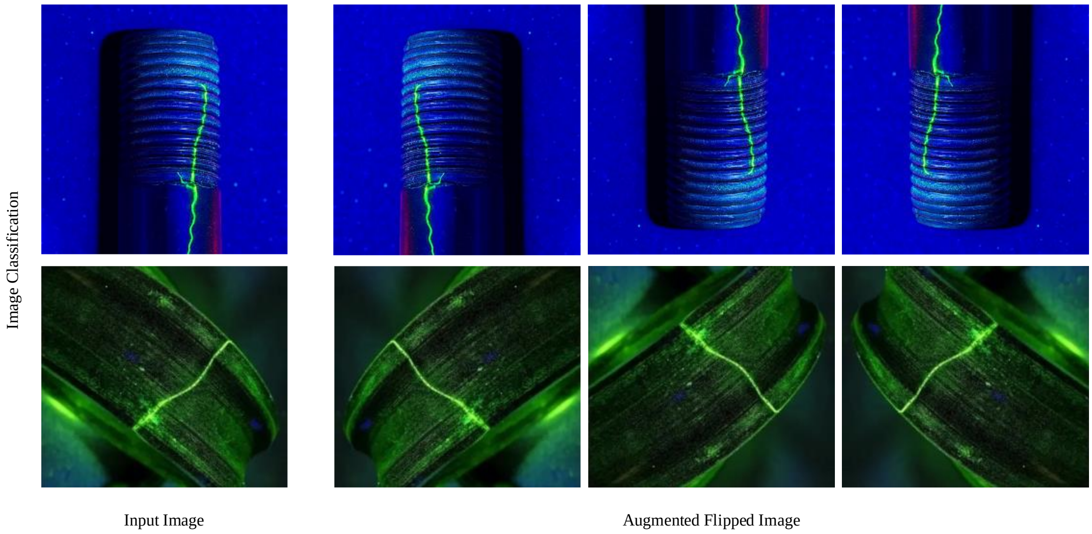
FPI photos of aircraft wheel tie bolts were collected to capture surface breaking cracks for assessment. Flipping, rotation, and zooming data augmentation methods were applied to increase dataset diversity and model efficiency. In addition, grayscale conversion was utilized to reduce color complexity and improve crack contrast. These augmentations preserved essential defect characteristics, facilitating more effective and efficient training of the machine learning model.



(a) (b)

Fig. 3.2 *Augmented Fluorescent penetrant inspection image showing rotation at +20° and -20°. (a) Original dataset; (b) Augmented rotated dataset*

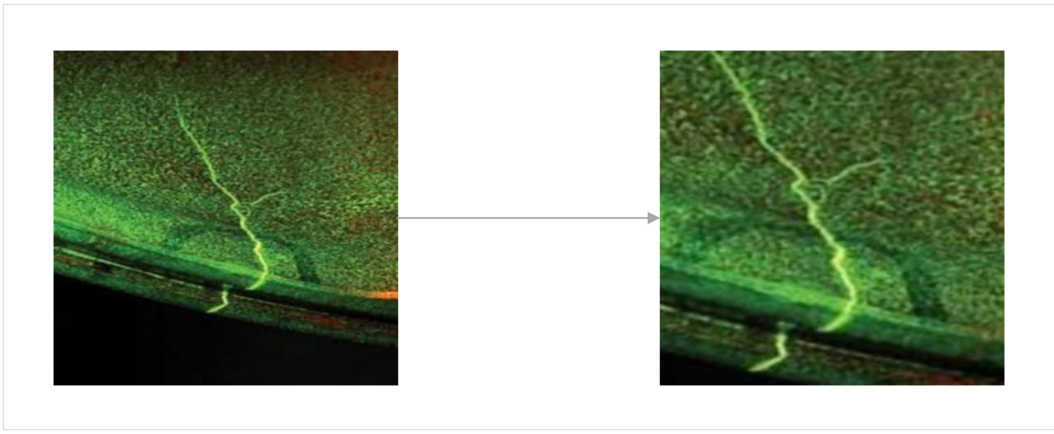
The enhanced fluorescent penetrant inspection (FPI) image used with rotations of +20° and -20°, which is used to mimic other angles of view by inspections, is shown in Figure 3.2. An augmentation process like this enables the machine learning model to generalize better by exposing it to many different orientations of surface cracks and, thus, enhance its real world defect detection.



(a) **(b)** **(c)** **(d)**

Fig. 3.3 *Flipped augmented result of fluorescent penetrant inspection image. (a) original image; (b) flipped horizontally; (c) flipped vertically; (d) horizontal and vertical flipping*

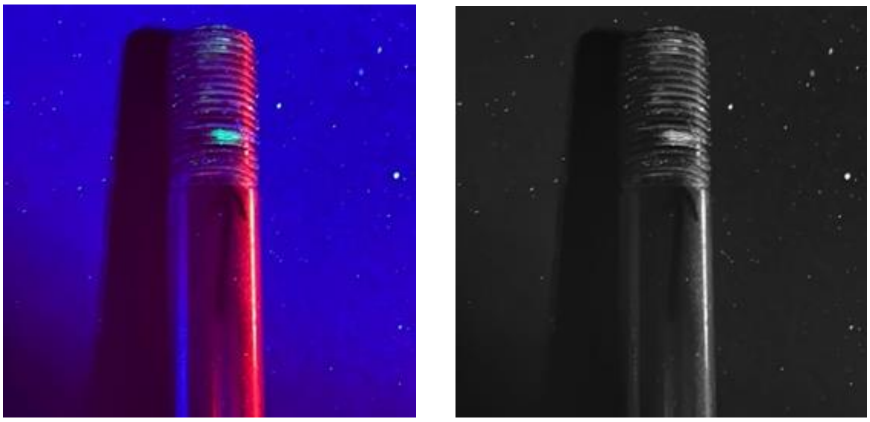
The flipped augmented results as illustrated in figure 3.3 demonstrate how machine learning can optimize dataset variability through orientation transformation. The visibility and crack structure are unaffected even when flipping in both directions and horizontally and vertically. This demonstrates that flipping shows the model to a variety of visual inputs while maintaining significant features in defects. This makes the trained model more consistent and flexible in identifying cracks from different perspectives in practical inspection tasks.



(a) **(b)**

Fig. 3.4 *Zoomed Augmented Result of Fluorescent Penetrant Inspection Image. (a) original image dataset; (b) zoomed image dataset*

The results of the zooming augmentation as illustrated in figure 3.4 show how the machine learning models capacity to identify the surface flaws can be improved by using closer inspection views. Without changing the original structure of the defect, key features like crack edges and intensity contrasts are made more noticeable by slightly enlarging the image. As a result, subtle features that might be harder to see at normal scale can be better learned by the machine learning model.



(a) **(b)**

Fig. 3.5 *Result of grayscale conversion image dataset. (a) Original RGB Image; (b) Grayscale Converted Image*

The grayscale conversion process as shown in figure 3.5 simplifies the image by removing color information while maintaining the critical structural details of surface cracks. This reduction in complexity allows the machine learning model to focus on intensity differences and contrast, which are essential for detecting defects. In order to ensure that the model continues to learn from precise images of crack patterns, the converted images preserve the essential visual cues. During training, this step also helps to increase computational efficiency.

3.3 Visual Data Preprocessing and Annotation

This section describes the preparation of FPI images for machine learning, a critical step to ensure the model receives clean, consistent, and meaningful input. All images were first resized to a standardized 416 × 416 pixel resolution to maintain uniformity and compatibility with the model's input layer. Following this, CVAT (Computer Vision Annotation Tool) was used to annotate surface cracks on selected samples, resulting in 300 annotated images from the full dataset. These preprocessing and annotation steps are essential for training a reliable supervised learning model by providing accurate, labeled data and reducing inconsistencies in image size and quality.

3.3.1 Image Pre-processing

Image preprocessing involved preparing the raw fluorescent penetrant inspection (FPI) images to ensure consistency and compatibility with machine learning requirements. This step included resizing and standardizing the images to create a uniform dataset, reduce variability, and improve the efficiency and accuracy of model training.

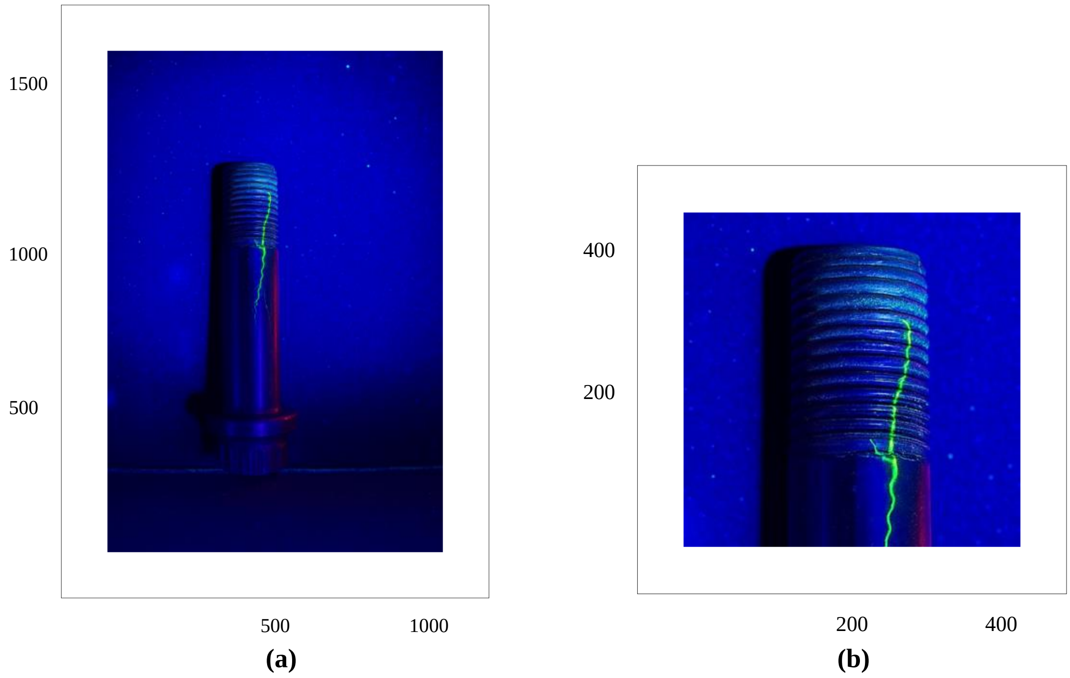


Fig. 3.6 *Image Resizing Result for Pixel Calibration. (a) uncalibrated image; (b) calibrated image*

Based on one of the sample original dataset image were resized from 1024 × 1536 pixels to 416 × 416 pixels as shown in figure 3.6. This resizing process reduced the total number of pixels per image from 1,572,864 to 173,056, which equates to an 88.95% reduction in image size. As part of the image preprocessing phase, a pixel calibration process was conducted to standardize all dataset images for compatibility with the machine learning model.

Table. 3.1 *Dataset Image Standardization to 416 × 416 Pixels*

Uncalibrated	1024 X 1536	PNG
Calibrated	416 X 416	PNG

The original uncalibrated images had a resolution of 1024 × 1536 pixels, amounting to a total of 1,572,864 pixels per image. These were resized to a uniform dimension of 416 × 416 pixels, reducing each image to 173,056 pixels. This resizing process significantly decreased the image size by 1,399,808 pixels, which translates to an approximate 88.95% reduction in pixel count.

3.3.2 Image Post-processing

In this section the annotation process was conducted as illustrated in figure 3.7 using Computer vision Annotation Tool (CVAT), the selected images from the dataset were manually labeled to identify visible surface cracks. Based on my project, a total of 300 annotated images were annotate out of 3,500 dataset images, focusing on representative samples with clear defect indications.

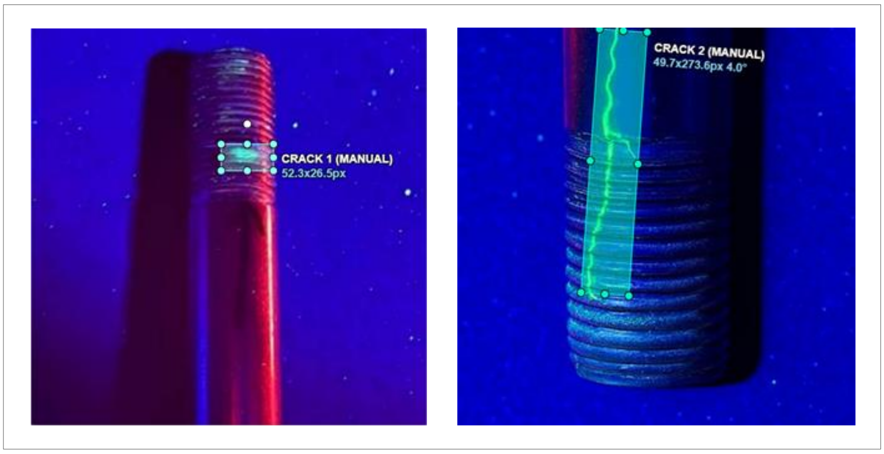


Fig. 3.7 *Annotated image result using CVAT*

3.4 Training and Validation

In this study, two machine learning models, YOLOv8m and YOLOv8s, were used for crack defect detection on fluorescent penetrant inspection (FPI) images. These models were applied to evaluate and compare their performance in identifying surface-breaking cracks based on the annotated dataset.

3.4.1 Model Learning

The model training results of both YOLOv8m and YOLOv8s models for crack defect detection are summarized in Table 3.2 to compare their performance in terms of accuracy, loss values, and processing time.

Table 3.2 *Model Training Metrics of YOLOv8m and YOLOv8s*

Testing	Model	
	Yolo v8m	Yolo v8s
Epoch	100	100
Ground truth	269	80
Predicted box	271	49
Box loss	0.6991	0.9473
Train loss	0.5129	0.9473
DFL loss	1.121	1.248
Processing time	2.007 hr	1.004 hr

As shown in Table 3.2, YOLOv8m demonstrates lower train box loss values and higher detection performance, while YOLOv8s offers a faster training time with lower detection sensitivity.

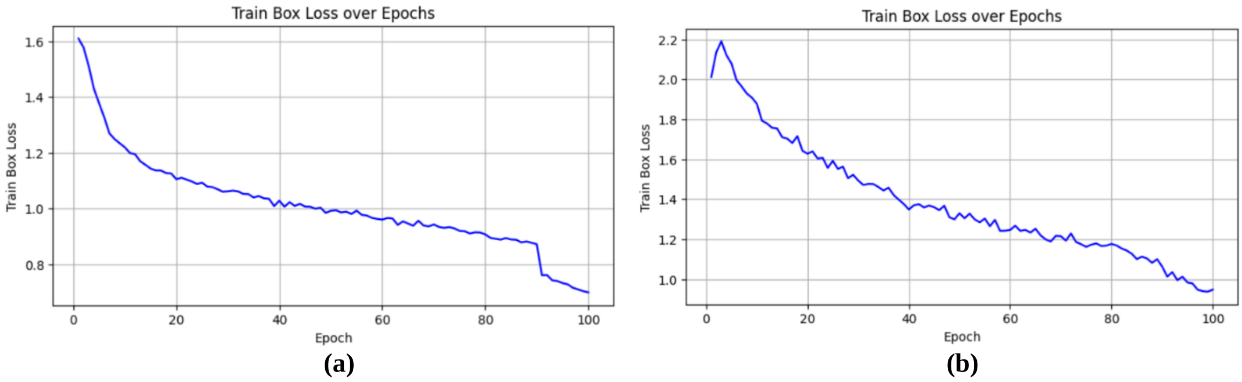


Fig. 3.8 *Model Training Loss Curve Comparison Between YOLOv8m and YOLOv8s. (a) Yolo v8m; (b) Yolo v8s*

The model training involved comparing two YOLOv8 models, YOLOv8m and YOLOv8s, for crack defect detection. Both models were trained for 100 epochs. However, their performance differed in terms of prediction accuracy and training efficiency. YOLOv8m showed superior performance, with a lower box loss (0.6991 vs. 0.9473) and train loss (0.5129 vs. 0.9473), resulting in better overall model learning. It also achieved a higher number of predicted boxes (271) compared to YOLOv8s (49), demonstrating better sensitivity in detecting cracks defect. Additionally, the DFL (Distribution Focal Loss) was slightly lower for YOLOv8m (1.121 vs. 1.248), supporting its improved prediction confidence. However, this came at the cost of longer training time YOLOv8m took around 2.007 hours, meanwhile the YOLOv8s completed training in 1.004 hours. These results suggest that while YOLOv8s is faster, YOLOv8m provides higher accuracy and better detection capability, making it more suitable for applications where precision in crack identification is critical.

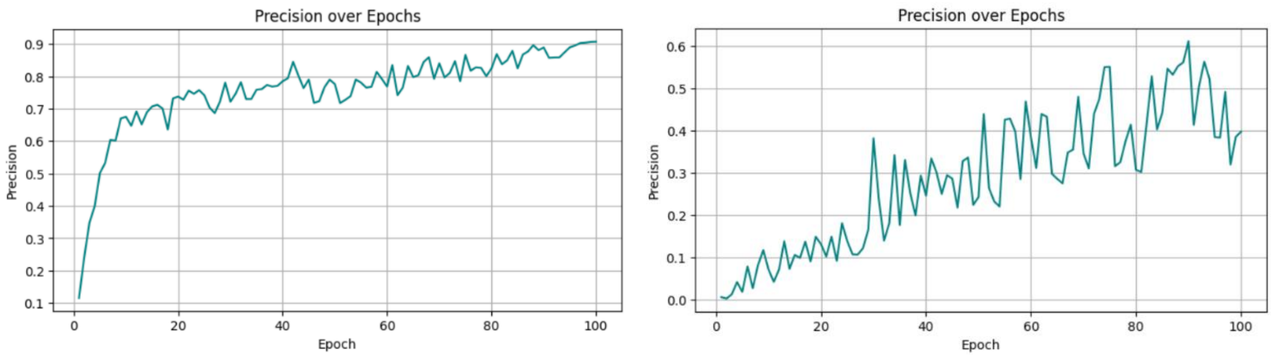
3.4.2 Model Testing

The model training results of both YOLOv8m and YOLOv8s models for crack defect detection are summarized in Table 3.3 to compare their performance in the testing phase, specifically in terms of precision, recall, and mean average precision (mAP).

Table 3.3 *Model Testing Metrics of YOLOv8m and YOLOv8s*

Testing	Model	
	Yolo v8m	Yolo v8s
Precision	0.907	0.397
Recall	0.871	0.327
Mean Average Precision (mAP)	0.884	0.302

Table 3.3 shows a visual comparison of the precision, recall, and mean average precision (mAP) achieved by YOLOv8m and YOLOv8s during the testing phase. The values indicate each models performance in detecting crack defects based on the annotated dataset.



(a) **(b)**

Fig. 3.9 *Model Testing precision over epoch Comparison Between YOLOv8m and YOLOv8s. (a) Yolo v8m; (b) Yolo v8s*

The evaluation results shows that the YOLOv8m significantly outperforms YOLOv8s through all key performance metrics. YOLOv8m has achieved a precision of 0.907, showing a higher rate of success crack predictions among all positive predictions, compared to only 0.397 in YOLOv8s. Similarly, YOLOv8m recorded a recall of 0.871, meaning it successfully detected the majority of actual cracks, while YOLOv8s showed much lower recall at 0.327, missing some of the defects. The mean average precision (mAP), which reflects overall detection accuracy, further highlights this gap 0.884 for YOLOv8m versus only 0.302 for YOLOv8s. These results shows that YOLOv8m is more reliable for crack detection tasks in fluorescent penetrant inspection images, making it the preferred model for practical deployment in defect identification.

3.5 Graphic User Interface results

This section presents the final result of the graphical user interface (GUI) developed for the crack defect detection system using the trained YOLOv8 model. It basically allows users to upload an image, start the detection process, and view bounding boxes overlaid on detected cracks. This website interface serves as a practical implementation tool, bridging the gap between model development and real world usability.

(a) **(b)**

Fig. 3.10 *Graphic user interface result (a) Yolo v8s; (b) Yolo v8m*

Both YOLOv8m and YOLOv8s models into a functional interface to perform real-time crack detection on FPI images. During testing, YOLOv8m consistently delivered more accurate and reliable detection results, displaying precise bounding boxes and higher confidence scores within the GUI. In contrast, YOLOv8s showed noticeably lower detection accuracy, often missing fine cracks or producing incomplete annotations. While YOLOv8s runs faster and requires less computational power, YOLOv8m proved to be more suitable for practical deployment, especially in scenarios where detection accuracy is critical.

4. Conclusion

The research was able to efficiently demonstrate the application of artificial intelligence in the enhancement of crack defect detection through the use of fluorescent penetrant inspection (FPI) on aircraft wheel tie bolts. The research began by simulating actual defect aircraft component, following which the FPI process was then performed in order to make surface breaking cracks detectable. The resulting images were further preprocessed via pixel calibration for a uniformed dataset and augmented afterwards with the techniques such as flipping, rotation, zooming, and grayscale to increase data diversification and model insensitivity.

All of the individual images in the post-processing phase were manually labeled using CVAT, with 300 annotated images selected from a total dataset of 3,000 to provide a representative and high quality training subset for the machine learning model by creating labeled data for supervised learning. YOLOv8m and YOLOv8s models were trained and evaluated using the dataset that have been gathered. Based on performance metrics like precision, recall, and mean average precision (mAP), YOLOv8m consistently outperformed YOLOv8s across all accuracy-related metrics. However, this improved performance came with a trade off in training time, as YOLOv8m required approximately 100% more training time compared to YOLOv8s, which completed the training process nearly 50% faster.

Finally, a graphical user interface (GUI) was created to integrate the trained models into a practical tool for real time crack detection. The GUI helps the importation of FPI images and viewing of detection results in real time, showing that the practicality of AI image based detection systems in the aerospace industry. Overall, this research highlights the potential of combining image processing and machine learning to help non-destructive testing (NDT) processes.

Acknowledgement

Communication of this research is made possible through monetary assistance by Universiti Tun Hussein Onn Malaysia and the UTHM Publisher's Office via Publication Fund E15216.

References

[1] Ragan, P. (1997). Aviation English: An Introduction. *Journal of Aviation/Aerospace Education & Research*. https://doi.org/10.15394/jaaer.1997.1189

[2] Rohit, B., Gowda, P. D., Nagaraju, N. B., & Meghana, S. (2019). Life and Failure of Aircraft Wheels - A Review. *IOP Conference Series: Materials Science and Engineering*, *520*(1). https://doi.org/10.1088/1757-899X/520/1/012002

[3] Shipway, N. (2019). Automated Defect Detection for Fluorescent Penetrant Inspection using Machine Learning.

[4] Jose, J. K. (2016). Analysis of Aircraft Wheel Hub Assembly Using NDT Techniques. *International Advanced Research Journal in Science, Engineering and Technology*, *3*(5). https://doi.org/10.17148/IARJSET.2016.3504

[5] Xu, B., Feng, R., Zhao, T., Zhou, L., Ju, Y., Chen, J., Zhang, Q., & Dong, L. (2024). Fluorinated polyimide with enhanced ultraviolet-resistant radiation via a self-sacrificial strategy. *European Polymer Journal*, *202*. https://doi.org/10.1016/j.eurpolymj.2023.112627

[6] Le, X., Mei, J., Zhang, H., Zhou, B., & Xi, J. (2020). A learning-based approach for surface defect detection using small image datasets. *Neurocomputing*, *408*, 112–120. https://doi.org/10.1016/j.neucom.2019.09.107

[7] Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, *3*(1), 91–99. https://doi.org/10.1016/j.gltp.2022.04.020

[8] Saravanan, C. (2010). Color image to grayscale image conversion. *2010 2nd International Conference on Computer Engineering and Applications, ICCEA 2010*, *2*, 196–199. https://doi.org/10.1109/ICCEA.2010.192