

Article

Advanced Defect Detection on Curved Aeronautical Surfaces Through Infrared Imaging and Deep Learning

Leith Bounenni ¹, Mohamed Arbane ¹, Clemente Ibarra-Castanedo ^{1,*} , Yacine Yaddaden ² ,
Sreedhar Unnikrishnakurup ³ , Andrew Ngo Chun Yong ³ and Xavier Maldague ^{1,*} 

¹ Department of Electrical and Computer Engineering, Université Laval, 1065, av. de la Médecine, Québec, QC G1V 0A6, Canada; leith.bounenni.1@ulaval.ca (L.B.); mohamed.arbane.1@ulaval.ca (M.A.)

² University of Quebec at Rimouski, Rimouski, QC G5L-3A1, Canada; yacine_yaddaden@uqar.ca

³ Institute of Materials Research & Engineering (IMRE), Agency for Science, Technology and Research (A*STAR), 2 Fusionopolis Way, 08-03 Innovis, Singapore 138634, Singapore; sreedharun@imre.a-star.edu.sg (S.U.); ngocya@imre.a-star.edu.sg (A.N.C.Y.)

* Correspondence: clemente.ibarra-castanedo@gel.ulaval.ca (C.I.-C.); xavier.maldague@gel.ulaval.ca (X.M.)

Abstract: Detecting defects on aerospace surfaces is critical to ensure safety and maintain the integrity of aircraft structures. Traditional methods often need more precision and efficiency for effective defect detection. This paper proposes an innovative approach that leverages deep learning and infrared imaging techniques to detect defects with high precision. The core contribution of our work lies in accurately detecting the size and depth of defects. Our method involves segmenting the size of the defect and calculating its centre to determine its depth. We achieve a more comprehensive and precise assessment of defects by integrating deep learning with infrared imaging based on the U-net model for segmentation and the CNN model for classification. The proposed model was rigorously tested on both a simulation dataset and an experimental dataset, demonstrating its robustness and effectiveness in accurately identifying and assessing defects on aerospace surfaces. The results indicate significant improvements in detection accuracy and computational efficiency, showing advancements over state-of-the-art methods and paving the way for enhanced maintenance protocols in the aerospace industry.

Keywords: deep learning; infrared imaging; aeronautical surfaces; defect detection



Citation: Bounenni, L.; Arbane, M.; Ibarra-Castanedo, C.; Yaddaden, Y.; Unnikrishnakurup, S.; Yong, A.N.C.; Maldague, X. Advanced Defect Detection on Curved Aeronautical Surfaces Through Infrared Imaging and Deep Learning. *NDT* **2024**, *2*, 519–531. <https://doi.org/10.3390/ndt2040032>

Academic Editor: Fabio Tosti

Received: 30 September 2024

Revised: 3 November 2024

Accepted: 12 November 2024

Published: 2 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Defects in aerospace surfaces present a severe and multifaceted problem that directly impacts aircraft safety, reliability, and operational efficiency. These defects, microscopic cracks, corrosion, delamination, or other structural anomalies can lead to catastrophic failures if undetected and unaddressed [1]. The high-stress environment that aircraft endure, including extreme temperature variations, pressure changes, and mechanical loads, exacerbates the risk of defect propagation. Such defects can compromise critical structural components, leading to potential in-flight failures that endanger passenger and crew lives [2]. Additionally, the financial implications are substantial, as undetected defects can result in unexpected maintenance, costly repairs, and prolonged downtime, disrupting airline schedules and reducing fleet availability. The aerospace industry must continuously strive to improve defect detection methods to ensure aircraft integrity, prevent accidents, and maintain rigorous safety standards [3].

Detecting defects in aerospace surfaces has traditionally relied on various non-destructive testing methods, including ultrasound and infrared imaging. Ultrasound detection involves using high-frequency sound waves that penetrate materials to identify internal flaws based on the reflection or transmission of these waves. This method effectively identifies subsurface defects and provides detailed information about the material's internal structure [4]. However, ultrasound testing requires direct contact with the surface, often

necessitating coupling media, such as gels or water, to facilitate the transmission of sound waves. Additionally, interpreting ultrasound signals can be complex, requiring skilled technicians and sophisticated equipment to analyse the data accurately. The process can be time-consuming and is typically limited to point-by-point inspections, which may need to be more efficient for large-scale assessments [5].

In contrast, infrared imaging offers distinct advantages over traditional ultrasound techniques. Infrared detection leverages thermal variations on the surface of materials to identify defects. This method is highly advantageous because it is contactless and non-invasive, allowing for the inspection of surfaces without the need for coupling media or physical contact [6]. Infrared systems can rapidly scan large areas, making them suitable for comprehensive surface inspections. This technology excels in detecting defects caused by variations in thermal conductivity or heat retention, such as delaminations, cracks, and corrosion. Infrared imaging can visualise defects in real time, providing immediate feedback and facilitating quicker decision-making processes. The potential for automated, high-speed inspections further enhances efficiency, reducing the time and labour required for thorough evaluations [5].

Moreover, infrared imaging can be integrated with advanced data processing algorithms and deep learning techniques to enhance defect detection accuracy [7]. These systems can be trained to recognize specific thermal signatures associated with different types of defects, improving the reliability of inspections. The use of infrared technology in conjunction with deep learning models allows for the automatic identification and classification of defects, minimizing human error and increasing the consistency of inspections. The ability to perform continuous, real-time monitoring of aerospace surfaces with infrared systems also supports proactive maintenance strategies, enabling the early detection of defects before they become critical issues. Overall, the advantages of infrared imaging make it a powerful and efficient tool for enhancing the accuracy and effectiveness of defect detection in the aerospace industry, providing significant improvements over traditional ultrasound methods.

In order to construct a deep learning model that can identify different sorts of defects in composite samples using infrared images as input and classify them according to their depths in the material, we needed a large dataset with a variety of defect types and depths. However, only three actual samples—a planar, a trapezoidal, and a curved sample—were available to us in our lab, and each sample had fifteen square-shaped defects of varying sizes and depths. To increase our dataset, we had two options: either generate the necessary data using simulation, which would save time and money, or create fresh physical samples. We opted for the simulation solution and used the COMSOL simulation program available in our lab.

In this study, we employed pulse phase thermography (PPT) to enhance defect visualization on curved aeronautical surfaces, allowing for precise thermal imaging of structural anomalies. We utilised a U-Net model to segment and localise the defects within the infrared images, accurately identifying their size and position. Following this, a convolutional neural network (CNN) was applied to classify the defect depth based on the temperature at the defect centre, dividing them into distinct classes. To validate our models, we first created a simulated dataset featuring different defect types and sizes, enabling us to rigorously test and fine-tune our approach before applying it to real-world data. The structure of this paper begins with a comprehensive literature review in Section 2, which examines existing methods and highlights the gaps addressed in this study. Section 3 details the proposed methodology, including dataset creation, defect segmentation, and classification techniques. Section 4 analyses the experimental results to evaluate the model's performance. Finally, this paper is concluded in Section 5, discussing the findings and potential future work.

2. Literature Review

Tur et al. [8] study how an aerospace cylindrical structure component, including a hole, responds to cyclic mechanical loads. It uses COMSOL Multiphysics for simulation and the

Taguchi method for optimisation. While this method provides a good understanding of both fatigue causes and breakdown trends, it has a number of constraints compared to our study. The simulation used in this article focuses on examining the fatigue performance of a single type of structural component (a cylindrical component with a hole) with defined cyclic mechanical pressures. Meanwhile, our article discusses a wider variety of composite models, such as planar, trapezoidal, and many curved models with different curvatures. It includes various internal defects, like square, ellipsoidal, and star shapes. Their article studies the impact of a singular stress concentrator (a hole) on the fatigue life of the component, while our article examines various internal defects in different models to aid in the identification of the different defects in every composite model. Also, the goals behind the simulations differ significantly. The provided article focuses on improving the longevity of a component. In contrast, our article concentrates on generating a dataset and training a deep learning model to identify and classify flaws in infrared images, focusing on practical uses in non-destructive testing.

Ziang Wei et al. [9] propose that combining deep learning with infrared thermography can effectively detect and segment impact damage in curved carbon fibre-reinforced polymer (CFRP) specimens. They highlight that the deep neural networks used in their study could accurately identify damage regions in both mid-wave and long-wave infrared images, with F1-scores of 92.74% and 87.39%, respectively. This demonstrates the method's potential for non-destructive testing of advanced materials.

R. Marani et al. [10] introduce a novel method for detecting subsurface defects in composite materials using pulsed thermography. This method models the thermal decay on laminate surfaces, induced by a brief heat pulse, with an exponential model characterised by three parameters, estimated using the least squares method. These parameters are distinctive and resistant to noise, making them ideal features for training classifiers that label potential defects based on their depth. Experimental tests on a carbon fibre-reinforced polymer (CFRP) laminate with known inclusions demonstrated that the best results were achieved using a decision forest with 30 trees, resulting in standard and balanced accuracies of 99.47% and 86.9%, respectively, and precision and recall rates of 89.87% and 73.67%.

Numan Saeed et al. [11] propose an automated method for detecting defects in carbon fibre-reinforced polymer (CFRP) materials using thermographic imaging. They use convolutional neural networks (CNNs) with transfer learning to identify defects and estimate their depths. The method achieves high accuracy in detecting defects by fine-tuning pre-trained networks with a small dataset of thermographic images. The approach was validated experimentally on CFRP samples with known defects, demonstrating an accuracy of approximately 88% in-depth estimation.

Pengfei Pan et al. [12] propose integrating infrared thermography (IRT) with deep learning, using the Mask R-CNN neural network, to detect internal defects in fibre-reinforced polymer (FRP)-reinforced concrete structures. The study involved capturing and aligning image data from a dual RGB and thermal camera setup, annotated for semantic segmentation to train the deep learning model. This approach achieved high accuracy, with an average accuracy of 96.28%, specificity of 96.78%, precision of 96.42%, recall of 96.91%, and an F1-score of 96.78%.

Peigen Li et al. [13] proposed a method for detecting pavement defects using convolutional neural networks (CNNs). The authors utilised grey and depth images captured via a 3D pavement data collection system, which was then pre-processed and labelled. They developed two network architectures, a classic U-shaped and a double-headed structure, and integrated attention modules to improve defect detection accuracy. Four types of pavement defects (potholes, cracks, bleeding, and patches) were analysed, and training with combined grey and depth images resulted in a significant improvement in detection performance, with a global pixel accuracy (GPA) of 97.36% and a mean intersection over union (MIoU) of 80.28%.

Jha et al. [14] explore advancements in defect detection within high-tech industries, especially aerospace manufacturing, using computer vision and deep learning. They

categorise defect inspection methods into manual, traditional computer vision, and modern deep learning-based automatic optical inspection (AOI) methods. CNN algorithms, specifically for supervised (classification and segmentation) and unsupervised learning models, form the basis of modern AOI. The authors emphasize the benefits of pixel-level segmentation techniques for high precision but acknowledge challenges, such as labelling complexity and data imbalance.

3. Methodology

This section provides the methodology we propose for detecting defects on aerospace curved surfaces, employing a two-part approach involving localization and segmentation, followed by classification, as illustrated in the Figure 1.

The first step involves pre-processing the infrared thermal images by applying a smoothing algorithm to remove noise from the pixel-level temperature data. This is essential because noise in infrared images can obscure the subtle signals that indicate the presence of defects. Smoothing, such as the Savitzky–Golay [15] filter, ensures cleaner input for further analysis, enhancing the precision of defect detection.

Next, pulse phase thermography (PPT) is applied to analyse the thermal signals more effectively [16]. This step introduces a thermal pulse to the surface, and the subsequent cooling process is recorded. The thermal response is transformed into the frequency domain through fast Fourier transform (FFT) techniques. PPT helps to visualise the defects more clearly by highlighting variations in thermal conductivity and heat retention across the surface, making it easier to detect defects in the data.

Once the defects are highlighted, a U-Net model is used for semantic segmentation. This deep learning architecture is specifically designed for image segmentation tasks. It processes the infrared images and identifies the precise boundaries of the defects. Segmentation is crucial for localizing the defect regions on aeronautical surfaces, allowing for accurate identification of the defects.

Finally, the depth of the detected defects is classified using a 2D convolutional neural network (CNN). The CNN analyses the thermal data from the centre of each defect, classifying the defects into different depth categories. This classification is critical for determining the severity of the structural anomalies, enabling more informed maintenance and repair decisions in aerospace applications.

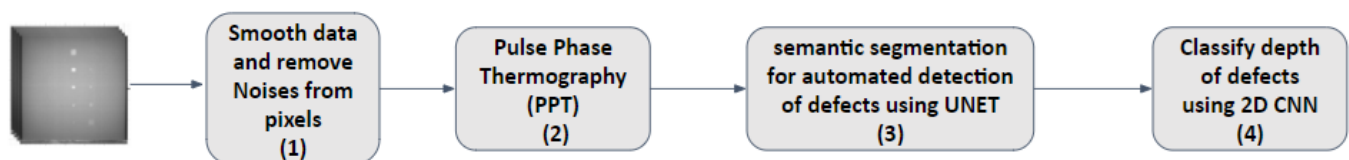


Figure 1. Overview of the proposed system.

4. Model Generation Using COMSOL and Dataset Creation

This section covers the different steps in configuring the COMSOL Multiphysics infrared thermography simulation. The main objective of this simulation is to produce data for training and testing a deep learning model that will identify and categorise defects in composite samples with high curvature.

4.1. COMSOL Simulation Setup

In this subsection, we will describe the steps used to create models in COMSOL and outline the process from model setup to final simulation to provide a good understanding of the modelling workflow.

4.1.1. Creating the Geometric Model

Different types of composite models were used for the simulation process, such as planar, trapezoidal, and various curved models with varying curvatures. Fifteen defects of

various sizes, forms, and depths, such as square, ellipsoidal, and star-shaped air cavities, were included in each of the samples.

- **Module Selection:** For this thermal simulation, the “Heat Transfer in Solids” COMSOL module was chosen.
- **CMaterial Definition:** The samples are composed of carbon fibre, with a surface emissivity of 0.85 and a thermal conductivity of $0.8 \text{ W/(m}\cdot\text{K)}$ [17].
- **Geometry Creation:** Every composite model was made to faithfully simulate the real parts, including the modelled defects.

4.1.2. Boundary Conditions Setup

Boundary conditions are essential for the simulation’s accuracy. The following parameters were established: two external radiation sources, each with a power of 500 W, were placed 30 cm from the piece to simulate the application of infrared thermography, as shown in Figure 2.

- **Initial Conditions:** The initial temperature of the sample was set to 293.15 K, equivalent to $20 \text{ }^\circ\text{C}$.
- **Heat Sources:** Two external radiation sources, with a power of 500 W for each, were placed 30 cm away from the borders of the sample to simulate the application of infrared thermography.

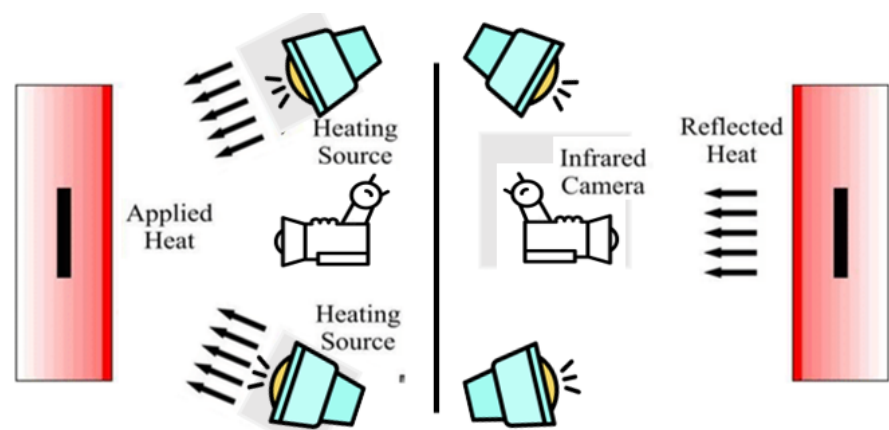


Figure 2. Experimental setup for thermography.

4.1.3. Meshing the Model

The mesh used in our simulation process was physics-controlled and set to normal, which ensured a good balance between accuracy and computational time.

4.1.4. Solving and Analysing the Results

With a time step of 0.033 s, the simulation was run in time-dependent mode from 0 to 10 s, allowing us to generate 300 infrared images for each sample.

- **Solving the Simulation:** The thermal variation over time was captured using a transient solver.
- **Analysis of the Results:** The results were analysed to detect different defects in the sample based on their form, depth, and size. The main tools used for this analysis were the infrared images produced by the simulation.

4.1.5. Data Export

The simulation’s data were exported for further use in training and testing the deep learning model.

- **Results Export:** Videos of thermography and sequences of infrared images were exported for a detailed analysis.
- **Documentation:** All simulation parameters and outcomes were recorded for future reference.

4.2. Some of the Models Created

As previously said, we have three actual carbon fibre samples in our lab (Figure 3), each 30 cm by 30 cm in size: a planar, a trapezoidal, and a curved sample. The following figures show that each sample has 15 square flaws of five different sizes and depths.

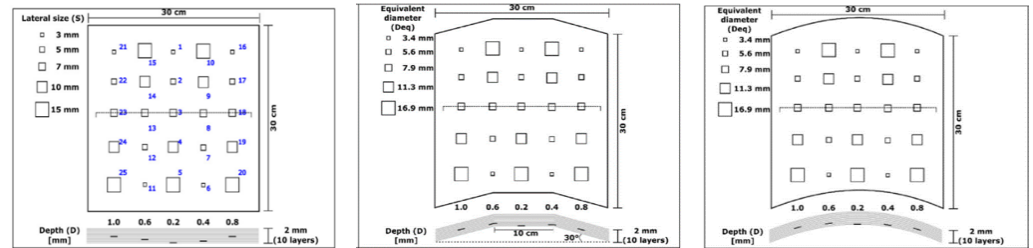


Figure 3. Experimental samples.

To evaluate the quality and similarity of the simulated data concerning the real-world data, we first built models that were precisely the same as the ones we had in our lab.

4.2.1. Planar Model

A composite sample with dimensions of 30 cm × 30 cm and a thickness of 2 mm serves as the planar model utilised in the simulation. This model has 15 square-shaped defects, all with different sizes and depths, and is made to look exactly like the original sample. A comprehensive dataset for identifying and evaluating thermal anomalies in a flat composite structure is provided by the defects dispersed throughout the surface and varying in size. There are five distinct dimensions and depths to the faults, as given in Figure 4.

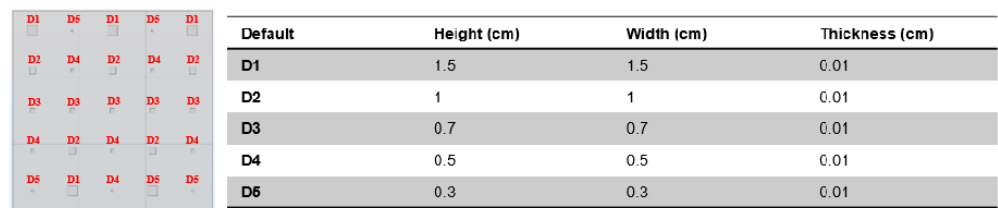


Figure 4. Planar model defaults' description.

Although the simulated and real-world data are not exactly the same, they do resemble one other. Because of this closeness, simulations are used to create new models with certainty (Figure 5).

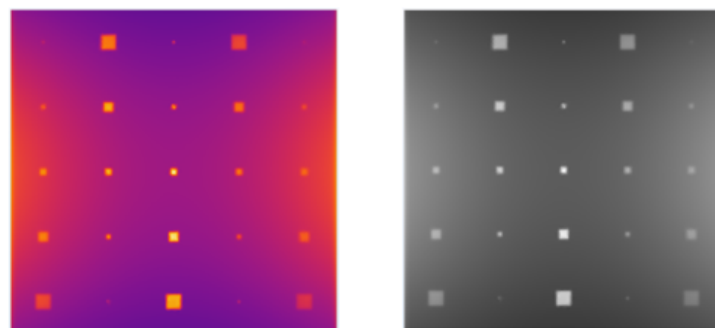


Figure 5. Infrared pics taken after 2.5 s.

4.2.2. Curved Models with Ellipsoidal Defects

After simulating the three models we had in our lab and confirming that the data from the simulation and the real world were similar, we created other models, primarily curved

ones, to increase the size of our dataset. We decided not to use square defects in this step to keep our model from overfitting to square patterns. Instead, we developed two curved models with defects that are ellipsoidal, as shown in Figure 6. Every model has 15 defects of five different sizes and five different depths, 30 cm by 30 cm in length. The distinction between the two models lies in the orientation of the defects: the first model has defects aligned horizontally and vertically. In contrast, the second model has defects directed at a 45-degree angle.

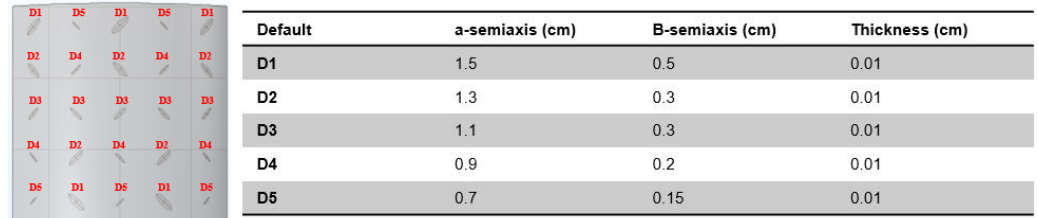


Figure 6. Curved model ellipsoidal defects description.

By applying infrared thermography to these two models, we were able to extract high-quality infrared images that clearly show the defects, as shown in Figures 7 and 8.

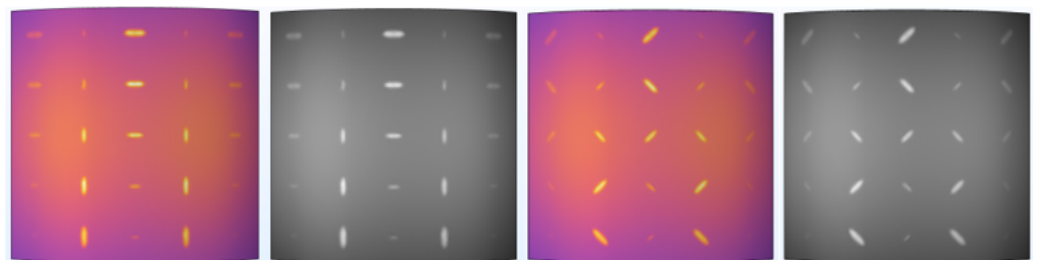


Figure 7. Infrared pics taken after 2.1 s.

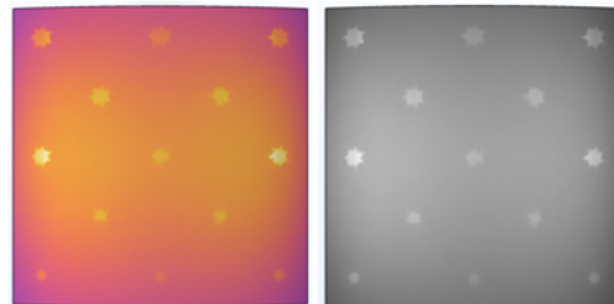


Figure 8. Infrared pictures taken after 2.2 s.

Afterwards, we developed a new curved model incorporating 15 ellipsoidal defects with a 45-degree orientation. However, in this instance, the defect depths were determined at random. This model will be utilised as a test model to assess how well our deep learning model can detect the depths of the defects.

4.2.3. Curved Models with Star Shaped Defects

We then moved on to the next step, which involved simulating a sample with defects that resembled real-world defects, like cracks. After successfully simulating curved models with ellipsoidal defects and receiving good simulation results, we created a sample simulation that contained 13 star-shaped flaws in 5 different depths and sizes, as shown in Figure 9.

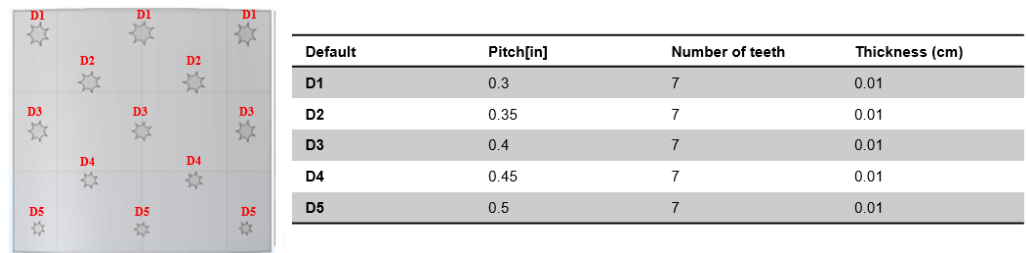


Figure 9. Curved model star-shaped defects' description.

By applying infrared thermography to this model, we could extract high-quality infrared images that clearly show the defects.

5. Deep Learning Model Architectures

This section will provide a detailed overview of the deep learning model architectures and the processes we employed in our study. We will explain the configurations and algorithms used, including the data pre-processing, segmentation, and classification techniques.

5.1. Data Pre-Processing Steps

In the pre-processing step, we utilised the Savitzky–Golay filter to smooth the temperature signals of the pixels in the infrared images. This filter effectively removes high-frequency noise while preserving the low-frequency signal, which is crucial for accurately detecting and characterizing defects. The Savitzky–Golay filter applies a polynomial smoothing technique, fitting successive sub-sets of adjacent data points with a low-degree polynomial using linear least squares. The equation for the Savitzky–Golay filter can be represented as

$$\hat{y}(i) = \sum_{j=-m}^m c_j y(i + j) \tag{1}$$

where $\hat{y}(i)$ is the smoothed value at position i , $y(i + j)$ represents the input signal values, and c_j represents the filter coefficients determined by the polynomial fit. This approach helps to enhance the quality of the input data, providing a cleaner signal for subsequent processing stages.

The second part of the pre-processing involves applying pulse phase infrared thermography (PPT) to enhance the visualisation of defects using frequency analysis. PPT is a non-destructive testing method that helps identify and characterise defects by analysing the thermal response of materials to a pulse of thermal energy. This technique enables us to highlight subtle variations in the temperature signals that may not be immediately apparent in the raw data. In the PPT process, we first apply a thermal pulse to the surface and then record the cooling process using infrared cameras. The temperature data are then transformed into the frequency domain using a fast Fourier transform (FFT), which helps identify the frequency components associated with different defects. The results of this process are illustrated in Figure 10. The period T and frequency F for the analysis are calculated using the following formulas: $T = \text{samplingrate} / \text{numberof frames}$ and $F = 1 / T$.

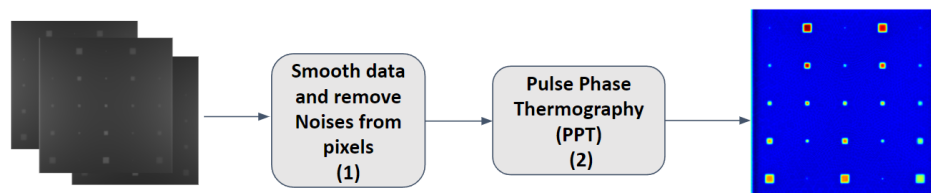


Figure 10. Frequency analysis of defects via pulse phase infrared thermography (PPT).

5.2. Image Segmentation Model for Defect Segmentation

After pre-processing, we developed a segmentation model based on the U-Net architecture [18]. Before applying the U-Net model, we implemented a subsystem to divide the PPT image into multiple smaller images. This step ensures that each segment of the PPT retains a reference to the original image, allowing us to identify the position of the first pixel on each of the four sides relative to the original image. This referencing is crucial for accurately mapping the segmented defects back to their original positions on the larger surface.

This subdivision process also serves as a data augmentation technique, increasing the diversity and quantity of images available for training the segmentation model. By augmenting the dataset, we enhance the model’s ability to generalise and improve its performance in identifying defects across various conditions.

Figure 11 illustrates the architecture of the proposed segmentation system, including this subdivision and referencing mechanism. This detailed architectural layout highlights the steps in segmenting the image data, ensuring accurate and effective defect detection.

The U-Net model was configured with a learning rate of 0.001, which was chosen to facilitate efficient and stable convergence during training. We utilised a batch size of 10 to balance the computational load and the ability of the model to generalise from the data. The model was trained over 20 epochs, allowing sufficient iterations for the network to learn the complex patterns associated with defect segmentation. To optimise the model’s performance, we employed the Adam optimiser [19], renowned for its adaptive learning rate capabilities and efficiency in handling sparse gradients. The loss function used was Binary Cross-Entropy, appropriate for the binary nature of the segmentation task, distinguishing between defect and non-defect regions. Additionally, a dropout rate of 0.5 was implemented to prevent overfitting, ensuring that the model maintains its generalization ability when exposed to new data. The Rectified Linear Unit (ReLU) activation function [20] was utilised throughout the network to introduce non-linearity, enabling the model to capture intricate features within the infrared images effectively.

The U-Net model was trained using 70% of the available data, while the remaining 30% was reserved for testing. This split ensures the model is exposed to diverse examples during training while retaining a separate dataset to evaluate its performance. We used Xavier initialisation [21], which helps maintain the scale of gradients in all network layers, thereby promoting stable training. All other layer parameters, such as the number of filters and kernel sizes, were configured as specified in the original U-Net model.

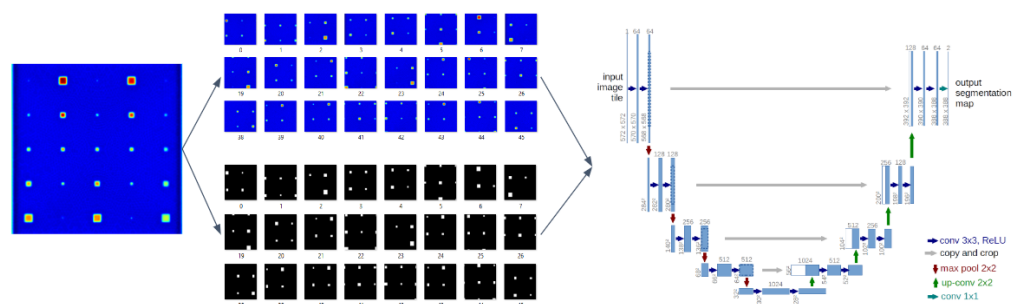


Figure 11. Image segmentation workflow using U-Net with PPT subdivision and mapping.

5.3. Image Classification Model for Depth Defect Classification

After segmentation, we can accurately determine the location of defects on the image, focusing mainly on the central region of each defect. This crucial area is identified by extracting a 3×3 pixel grid centred on the defect. For each detected defect, we generate an Excel file containing the addresses of the defects and their centres on the original image.

Using these addresses, we map the identified defects back to the raw dataset to extract temperature data over time for each defect. We then apply a fast Fourier transform (FFT) to

this temperature dataset to generate a spectrogram for each defect, capturing the frequency characteristics associated with the defects' thermal responses.

We utilise a 2D convolutional neural network (CNN) [22] to classify defects based on depth as shown in Figure 12. This model is trained to categorise defects into three classes based on their depth. Class 1 corresponds to defects with a depth range of 1.0 to 0.8, Class 2 covers depths from 0.7 to 0.4, and Class 3 includes defects between 0.3 and 0.1. This classification helps assess the defects' severity and potential impact, which is crucial for maintenance and repair decisions in aerospace applications.

The CNN model employed for classifying defect depths was meticulously configured to optimise performance. We set the learning rate to 0.001, a value selected to ensure gradual and stable updates to the network weights, thereby enhancing the model's ability to converge to an optimal solution without overshooting. A larger batch size of 10 was utilised to improve the statistical estimation of the gradient, promoting more robust learning across the dataset. The model underwent training for 20 epochs, and the Adam optimiser was chosen for its adaptive learning rate properties, which are beneficial in managing the complexities of the classification task. For the loss function, we implemented Categorical Cross-Entropy, suitable for multi-class classification scenarios where each defect is categorised into one of three depth classes. The network architecture comprised five layers, each contributing to the extraction of hierarchical features from the spectrograms. ReLU activation functions were applied to all hidden layers to introduce non-linearity, while a Softmax activation function was employed in the output layer to facilitate probability-based classification of defect depths.

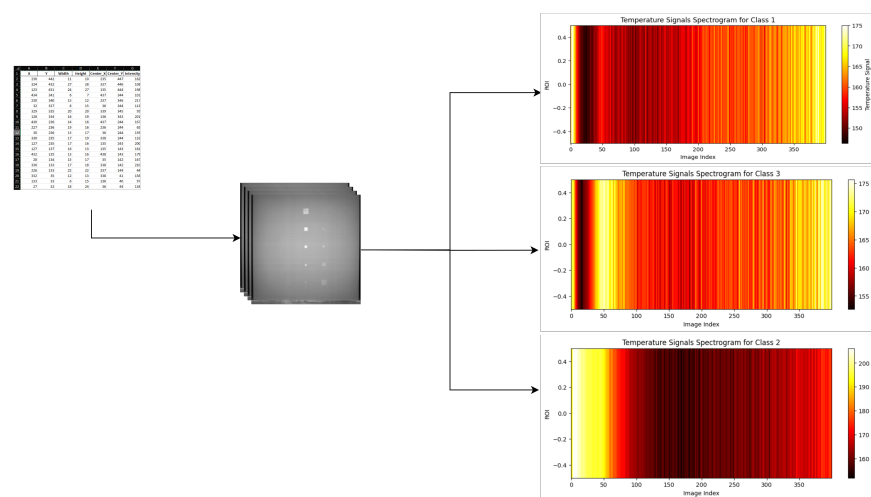


Figure 12. The proposed approach of 2D CNN classification.

6. Results and Discussion

Figure 13 presents the evaluation table results of different models and datasets used for centre detection and segmentation tasks. The models used are "Curved 1", "Curved 2", "Trapezoid", "M2", and "Curved 5", where Curved 1 and 2 are 70% used for training and 30% for testing for segmentation and classification, and other datasets models used only for testing. The results are reported for three processing methods: the raw dataset (unprocessed images), the Savitzky–Golay filter combined with PPT, and the U-NET segmentation model for the centre and classification results.




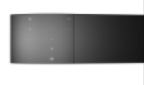
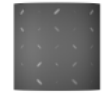
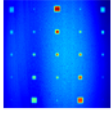
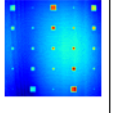
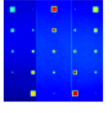
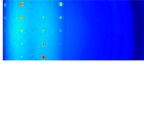
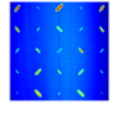
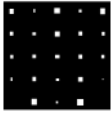
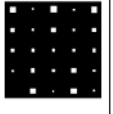
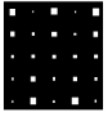

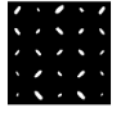
Model/	Curved 1	Curved 2	Trapezoid	M2	Curved 5
Raw Dataset					
Savgol filter + PPT					
Unet (segmentation)					
centre detection	0.99	0.99	0.99	0.99	0.99
Precision	0.91	0.93	0.87	0.92	0.93
Recall	0.90	0.94	0.87	0.92	0.88
F1-Score	0.90	0.93	0.87	0.92	0.90

Figure 13. Overview of the obtained results.

The metrics used to evaluate the model’s depth classification include centre detection, precision, recall, and F1-score. Centre detection measures the ability of the model to accurately detect the centre points in the images, with values close to 1 indicating high accuracy. Precision represents the ratio of true positive classification to the total number of depth defects. At the same time, recall indicates the ratio of true positive classification to the actual number of true depth classes. The F1-score provides a harmonic mean of precision and recall, offering a balanced measure of both.

The analysis shows that centre detection scores are consistently high (0.99) across all datasets, demonstrating excellent performance in identifying centre points. However, the precision and recall scores vary slightly across different datasets and models. The “Curved 2” and “Curved 5” datasets achieved the highest precision (0.93) with the classification model. In contrast, the “Trapezoid” dataset showed a relatively lower precision (0.87). Recall scores were generally high, with “Curved 2” achieving the highest score (0.94). The “Curved 5” dataset had the lowest recall (0.88). The F1-score, which balances precision and recall, ranged from 0.87 to 0.92 across datasets. The “Trapezoid” dataset had the lowest F1-score (0.87), reflecting the challenges seen in both precision and recall metrics. While the U-NET segmentation model generally maintains high accuracy in centre detection, there are variations in precision and recall, particularly with the “Trapezoid” dataset, which appears more challenging for the models.

As depicted in Figure 14, we applied the model to another dataset-based modelling dataset to a real dataset, specifically retraining it for the model “CRFP06 Curved Surface”. The process involved using a sequence of images (video) processed through a PPT method, resulting in a labelled dataset for training. This dataset was then used for defect segmentation using the U-NET model. The results of this application were promising, achieving an intersection over union (IoU) score of 0.86 for segmentation, which measures the accuracy of the segmentation in identifying the area of defects. Additionally, the system demonstrated high precision in identifying the centre of defects, with a detection accuracy of 0.99. These metrics indicate the model’s effectiveness in accurately detecting and segmenting defects on curved surfaces.

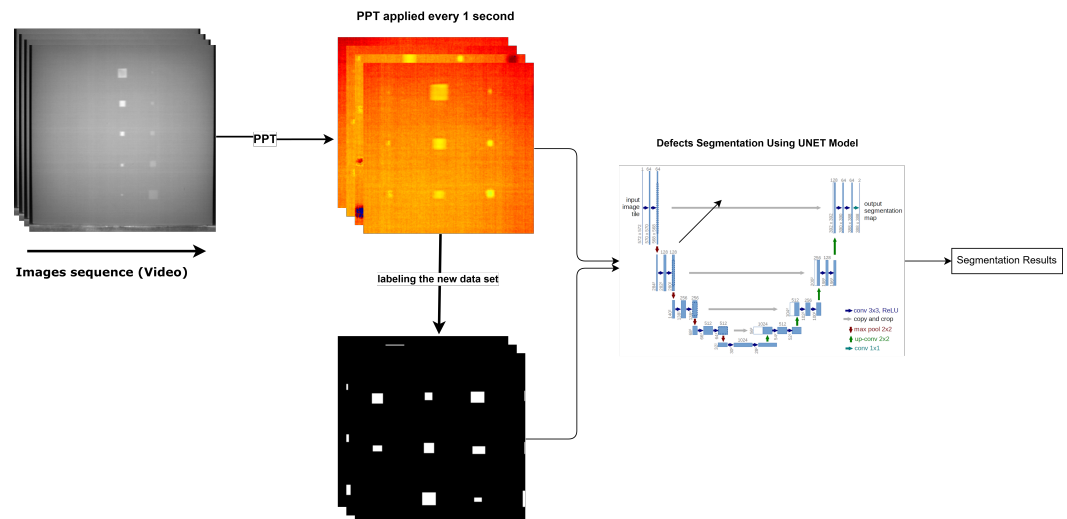


Figure 14. Proposed of segmentation model on the experimentation dataset.

7. Conclusions

In conclusion, this paper proposed a developing and validated deep learning model combined with infrared imaging to detect defects on aeronautical curved surfaces. Our approach demonstrated significant improvements in accuracy and efficiency compared to traditional non-destructive testing methods and state-of-art methods, offering a more reliable solution for aerospace maintenance. By integrating COMSOL simulations to augment our dataset, we addressed the challenges of limited real-world samples, enabling robust model training and validation. Although our research primarily utilised simulated data, the promising results pave the way for future applications in real-time and in-situ inspections, potentially enhancing safety protocols and reducing maintenance costs in the aerospace industry. Future work should validate the model with diverse materials and real-world conditions to solidify its applicability in operational environments. The limitation of our proposed approach is that it relies on detecting the centre of defects to estimate their depth. This approach may be less practical or even inadequate for defects that do not have a well-defined centre, such as irregular cracks or diffuse damage areas. As a result, the accuracy of depth estimation for such defects may be compromised. For our future work, we plan to address this limitation by developing methods that can accurately assess the depth of defects without relying solely on their centre, thereby improving the model's ability to handle a broader range of defect types and geometries.

Author Contributions: Conceptualization, L.B., M.A., C.I.C. and X.M.; methodology, L.B., M.A., C.I.C., S.U. and X.M.; software, M.A; validation, X.M., C.I.C., Y.Y., S.U. and A.N.C.Y.; resources, L.B., C.I.C. and X.M.; writing—original draft preparation, M.A. and L.B.; writing—review and editing, X.M., Y.Y., C.I.C., S.U. and A.N.C.Y.; supervision, X.M., S.U. and A.N.C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: Research work was supported by the Agency for Science, Technology and Research (A*STAR) through grants under its Polymer Matrix Composites Programme (Grant No. A19C9a0044). We also acknowledge the contribution of Ministère des Relations internationales et de la Francophonie of Quebec in the framework of the Quebec/Singapore scientific collaboration, project SPiiRALE.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Yang, B.; Huang, Y.; Cheng, L. Defect detection and evaluation of ultrasonic infrared thermography for aerospace CFRP composites. *Infrared Phys. Technol.* **2013**, *60*, 166–173. [[CrossRef](#)]
2. Pozzer, S.; El Refai, A.; López, F.; Ibarra-Castanedo, C.; Maldague, X. Passive infrared thermography for subsurface delamination detection in concrete infrastructure: Inference on minimum requirements. *Comput. Struct.* **2024**, *305*, 107529. [[CrossRef](#)]

3. Zhang, Z.; Liu, M.; Li, Q.; Ang, Y. Visualized characterization of diversified defects in thick aerospace composites using ultrasonic B-scan. *Compos. Commun.* **2020**, *22*, 100435. [[CrossRef](#)]
4. Pozzer, S.; Ramos, G.; Azar, E.R.; Osman, A.; El Refai, A.; López, F.; Ibarra-Castanedo, C.; Maldague, X. Enhancing concrete defect segmentation using multimodal data and Siamese Neural Networks. *Autom. Constr.* **2024**, *166*, 105594. [[CrossRef](#)]
5. Fang, Q.; Nguyen, B.D.; Castanedo, C.I.; Duan, Y.; Maldague, X., II. Automatic defect detection in infrared thermography by deep learning algorithm. In Proceedings of the Thermosense: Thermal Infrared Applications XLII, Online, 27 April–8 May 2020; Volume 11409, pp. 180–195.
6. Deng, L.; Zuo, H.; Wang, W.; Xiang, C.; Chu, H. Internal defect detection of structures based on infrared thermography and deep learning. *KSCE J. Civ. Eng.* **2023**, *27*, 1136–1149. [[CrossRef](#)]
7. Zhou, G.; Zhang, Z.; Yin, W.; Chen, H.; Wang, L.; Wang, D.; Ma, H. Surface defect detection of CFRP materials based on infrared thermography and Attention U-Net algorithm. *Nondestruct. Test. Eval.* **2024**, *39*, 238–257. [[CrossRef](#)]
8. Tur, E. Fatigue Analysis of an Aerospace Elastoplastic Structural Cylindrical Component with Hole under Cyclic Mechanical Load using COMSOL Multiphysics and Taguchi Method Optimization. *Bayburt Üniversitesi Fen Bilimleri Dergisi* **2023**, *6*, 151–171. [[CrossRef](#)]
9. Wei, Z.; Fernandes, H.; Herrmann, H.G.; Tarpani, J.R.; Osman, A. A deep learning method for the impact damage segmentation of curve-shaped cfrp specimens inspected by infrared thermography. *Sensors* **2021**, *21*, 395. [[CrossRef](#)] [[PubMed](#)]
10. Marani, R.; Palumbo, D.; Renò, V.; Galietti, U.; Stella, E.; D’Orazio, T. Modeling and classification of defects in CFRP laminates by thermal non-destructive testing. *Compos. Part B Eng.* **2018**, *135*, 129–141. [[CrossRef](#)]
11. Saeed, N.; King, N.; Said, Z.; Omar, M.A. Automatic defects detection in CFRP thermograms, using convolutional neural networks and transfer learning. *Infrared Phys. Technol.* **2019**, *102*, 103048. [[CrossRef](#)]
12. Pan, P.; Zhang, R.; Zhang, Y.; Li, H. Detecting Internal Defects in FRP-Reinforced Concrete Structures through the Integration of Infrared Thermography and Deep Learning. *Materials* **2024**, *17*, 3350. [[CrossRef](#)] [[PubMed](#)]
13. Li, P.; Zhou, B.; Wang, C.; Hu, G.; Yan, Y.; Guo, R.; Xia, H. CNN-based pavement defects detection using grey and depth images. *Autom. Constr.* **2024**, *158*, 105192. [[CrossRef](#)]
14. Jha, S.B.; Babiceanu, R.F. Deep CNN-based visual defect detection: Survey of current literature. *Comput. Ind.* **2023**, *148*, 103911. [[CrossRef](#)]
15. Zimmermann, B.; Kohler, A. Optimizing Savitzky–Golay parameters for improving spectral resolution and quantification in infrared spectroscopy. *Appl. Spectrosc.* **2013**, *67*, 892–902. [[CrossRef](#)] [[PubMed](#)]
16. Maldague, X.; Marinetti, S. Pulse phase infrared thermography. *J. Appl. Phys.* **1996**, *79*, 2694–2698. [[CrossRef](#)]
17. Adam, T.; Hashim, U. COMSOL multiphysics simulation in biomedical engineering. *Adv. Mater. Res.* **2014**, *832*, 511–516. [[CrossRef](#)]
18. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv* **2015**, <http://arxiv.org/abs/1505.04597>.
19. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2017**, <http://arxiv.org/abs/1412.6980>.
20. Hara, K.; Saito, D.; Shouno, H. Analysis of function of rectified linear unit used in deep learning. In Proceedings of the 2015 international joint conference on neural networks (IJCNN), Killarney, Ireland, 12–17 July 2015; pp. 1–8.
21. Kumar, S.K. On weight initialization in deep neural networks. *arXiv* **2017**, <http://arxiv.org/abs/1704.08863>.
22. Li, Z.; Yang, W.; Peng, S.; Liu, F. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *arXiv* **2020**, <http://arxiv.org/abs/2004.02806>. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.