Abordagens de Deep Learning para Segmentação de Defeitos em Materiais Compósitos usando Termografia Infravermelha

Iago Garcia Vargas

Universidade Federal de Uberlândia
Faculdade de Computação
Programa de Pós-Graduação em Ciência da Computação

Uberlândia
2024

Iago Garcia Vargas

Abordagens de Deep Learning para Segmentação de Defeitos em Materiais Compósitos usando Termografia Infravermelha

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Henrique Coelho Fernandes

Uberlândia
2024

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

V297 Vargas, Iago Garcia, 1997-
2024 Abordagens de Deep Learning para Segmentação de Defeitos em Materiais Compósitos usando Termografia Infravermelha [recurso eletrônico] / Iago Garcia Vargas. - 2024.

Orientador: Henrique Coelho Fernandes.
Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: http://doi.org/10.14393/ufu.di.2024.813
Inclui bibliografia.
Inclui ilustrações.

1. Computação. I. Fernandes, Henrique Coelho,1986-, (Orient.). II. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. III. Título.

CDU: 681.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Computação
Av. João Naves de Ávila, 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgfacom@ufu.br

ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Dissertação, 46/2024, PPGCO				
Data:	18 de dezembro de 2024	Hora de início:	13:30	Hora de encerramento:	15:10
Matrícula do Discente:	12312CCP013				
Nome do Discente:	Iago Garcia Vargas				
Título do Trabalho:	Deep Learning Approaches for Defect Segmentation on Composite Materials using Infrared Thermography				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	"Inteligência artificial para a caracterização de defeitos em compósitos através de termografia infravermelha", projeto CNPq número 407140/2021-2				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Bruno Augusto Nassif Travençolo FACOM/UFU , Matheus Pereira Porto - Escola de Engenharia - UFMG e Henrique Coelho Fernandes - Cranfield University - UK, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Henrique Coelho Fernandes - Milton Keynes/Inglaterra e Matheus Pereira Porto - Belo Horizonte/MG . Os outros membros da banca e o aluno participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Henrique Coelho Fernandes, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação da Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir ao candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação

interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

Documento assinado eletronicamente por **Bruno Augusto Nassif Travençolo**, **Professor(a) do Magistério Superior**, em 18/12/2024, às 16:05, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Henrique Coelho Fernandes**, **Professor(a) do Magistério Superior**, em 18/12/2024, às 16:08, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Matheus Pereira Porto**, **Usuário Externo**, em 24/12/2024, às 06:43, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5946574** e o código CRC **0E2914CD**.

Referência: Processo nº 23117.085470/2024-17 SEI nº 5946574

Aos futuros pesquisadores que um dia abrirão estas páginas em busca de respostas ou inspiração. Que nunca percam a curiosidade e a coragem de seguir em frente, mesmo quando os caminhos parecerem incertos.

Agradecimentos

Com profunda gratidão inicio estes agradecimentos, primeiramente a Deus, por me conceder força e orientação em cada etapa desta jornada.

Expresso minha sincera apreciação à Universidade Federal de Uberlândia, em especial à Faculdade de Computação, incluindo os membros do corpo docente, a administração e toda a liderança, por me oferecerem esta oportunidade inestimável de crescimento pessoal, profissional e acadêmico.

Sou especialmente grato ao Prof. Dr. Henrique Fernandes Coelho, cujo mentorado, paciência e conhecimento não só moldaram este trabalho, mas também contribuíram significativamente para o meu desenvolvimento pessoal e profissional.

Agradecimentos especiais à minha família: minha mãe, meu pai e meu irmão. Vocês foram minha base firme nos momentos de incerteza, minha alegria nos momentos de celebração e meu conforto diante dos desafios. O apoio inabalável e amor incondicional me sustentaram ao longo desta jornada.

A todos vocês, meus mais sinceros agradecimentos. Vocês são uma parte essencial desta conquista.

"The mind that opens to a new idea never returns to its original size."
(Albert Einstein)

Resumo

A termografia infravermelha é amplamente utilizada para a detecção de defeitos em materiais compósitos. No entanto, identificar com confiança essas falhas apresenta desafios significativos devido à complexidade das propriedades térmicas dos compósitos e à variabilidade das condições de inspeção. Nesta pesquisa, propomos uma nova abordagem baseada em deep learning para a detecção de defeitos em materiais compósitos utilizando termografia infravermelha.

O modelo proposto integra duas arquiteturas de rede neural: a rede neural espacial DeepLabv3 e a rede neural temporal Bidirectional Long Short-Term Memory (BiLSTM). A combinação dessas arquiteturas permite a análise eficiente das características espaciais e temporais das imagens térmicas, melhorando a identificação de defeitos. A DeepLabv3 é utilizada para segmentar e destacar áreas de interesse nas imagens térmicas, enquanto a BiLSTM é responsável por analisar a evolução temporal das temperaturas nessas áreas, proporcionando uma visão mais completa do comportamento térmico dos defeitos ao longo do tempo. Esta abordagem permite uma detecção mais precisa e robusta em comparação com métodos tradicionais. Experimentos foram conduzidos utilizando um conjunto de dados reais composto por imagens térmicas de Carbon-fiber Reinforced Polymer (CFRP) submetidos a diferentes condições de teste. Os resultados demonstram que o método proposto neste trabalho, que combina DeepLabv3 e BiLSTM, melhora significativamente a detecção de defeitos, quando comparado com técnicas tradicionais.

Assim, esta pesquisa contribui para o avanço das técnicas de detecção de defeitos em materiais compósitos, utilizando a termografia infravermelha e deep learning, demonstrando o potencial das arquiteturas de rede neural para aplicações em inspeção e controle de qualidade.

Palavras-chave: Termografia Infravermelha. Segmentação. Aprendizado Profundo. Materiais Compósitos. DeepLabv3. BiLSTM.

Deep Learning Approaches for Defect Segmentation on Composite Materials using Infrared Thermography

Iago Garcia Vargas

Universidade Federal de Uberlândia
Faculdade de Computação
Programa de Pós-Graduação em Ciência da Computação

Uberlândia
2024

Abstract

Infrared Thermography (IRT) is widely used for detecting defects in composite materials. However, accurately identifying these faults presents significant challenges due to the complexity of the thermal properties of composites and the variability of inspection conditions. In this research, we propose a new approach based on deep learning for detecting defects in composite materials using infrared thermography.

The proposed model integrates two neural network architectures: the DeepLabv3 spatial neural network and the BiLSTM temporal neural network. The combination of these architectures allows for efficient analysis of the spatial and temporal characteristics of thermal images, improving the accuracy of defect identification. DeepLabv3 is used to segment and highlight areas of interest in thermal images, while BiLSTM is responsible for analyzing the temporal evolution of temperatures in these areas, providing a more comprehensive view of the thermal behavior of defects over time. This approach allows for more precise and robust detection compared to traditional methods. Experiments were conducted using a real dataset composed of thermal images of Carbon-fiber reinforced polymer CFRP subjected to different test conditions. The results demonstrate that the combined use of DeepLabv3 and BiLSTM significantly improves defect detection accuracy, outperforming traditional techniques.

Thus, this research contributes to the advancement of defect detection techniques in composite materials using infrared thermography and deep learning, demonstrating the potential of neural network architectures for applications in inspection and quality control.

Keywords: Infrared Thermography. Segmentation. Deep Learning. Composite Materials. DeepLabv3. BiLSTM..

List of Figures

List of Tables

Acronyms list

AI Artificial Intelligence

ASPP Atrous Spatial Pyramid Pooling

ADAM Adaptive Moment Estimation

BiLSTM Bidirectional Long Short-Term Memory

CFRP Carbon-fiber Reinforced Polymer

CNN Convolutional Neural Network

DL Deep Learning

FEM Finite Element Method

GRU Gated Recurrent Unit

GPU Graphics Processing Unit

IRT Infrared Thermography

IoU Intersection over Union

LWIR Long-Wave InfraRed

LIT Laser Infrared Thermography

ML Machine Learning

MWIR Mid-Wave InfraRed

MLP Multi Layer Perceptron

MAE Mean Absolute error

NDT Non-Destructive Testing

NN Neural Network

NIR Near-infrared

PPT Pulsed Phase Thermography

PT Pulsed Thermography

POD Probability of Detection

PCA Principal Components Analysis

RNNs Recurrent Neural Networks

RNN Recurrent Neural Network

ROI Region of Interest

SWIR Short-Wave InfraRed

LSTM Long Short-Term Memory

TCN Temporal Convolutional Network

THz TeraHertz

TSR thermography signal reconstruction

VGG-Unet Visual Geometry Group-Unet

Contents

Chapter 1

Introduction

Composite materials are important structural materials that have been widely used in aerospace, wind turbine blades, etc. (ALSAMHI et al., 2021). Defects generated during the manufacturing and handling of laminated composites can pose a potential safety risk to the device and result in significant losses (DAS; SAHU; PARHI, 2021). To address these challenges, Non-Destructive Testing (NDT) methods have emerged as essential tools for quality assurance in manufacturing and for ensuring operational safety.

The rise of composites as a unique class of materials occurred in the mid-20th century, with the deliberate design of multiphase composites such as CFRP (GAO et al., 2022). Despite advances, the damage mechanisms in composite materials remain unrecognized. Defects such as delaminations, disbonds, and voids can arise during manufacturing or develop throughout the lifecycle of the structure. Consequently, numerous defect detection techniques have been developed over the past years to identify the presence, location, and extent of damage (WANG et al., 2022).

Among the various NDT approaches, IRT has proven to be a promising and efficient method for defect detection in composite materials. IRT operates on the principle that defective regions display different thermal patterns compared to intact areas when subjected to thermal excitation. By capturing and analyzing IRT images, this method enables the identification of defects without altering the material under investigation (KHODAYAR; SOJASI; MALDAGUE, 2016). Despite their advantages, the complex geometries and distinct properties of composite materials present challenges, necessitating the development of advanced analytical and pattern extraction techniques (NIENDORF; RAEYMAEKERS, 2021).

In recent years, the integration of Artificial Intelligence (AI) techniques into NDT processes has gained traction. AI, particularly through Deep Learning (DL) approaches, has demonstrated capabilities in processing large datasets and extracting features with minimal human intervention (VANRULLEN; KANAI, 2021). DL techniques, which include Neural Network (NN) with multiple layers, aim to emulate the human brain by leveraging data inputs, weights, and biases to perform accurate classification and pattern

recognition.

Despite advances in the field, existing approaches to defect detection face several challenges. One primary issue is the need for segmentation of defect regions within composite materials, which is a crucial step in defect identification and classification. Conventional methods often rely on manual segmentation or simplistic models, which may lack robustness when applied to complex geometries and varying defect patterns (WANG et al., 2022).

To address these limitations, this study proposes a novel method that combines DeepLabv3 and BiLSTM networks. DeepLabv3, a state-of-the-art semantic segmentation model, excels at extracting spatial features and delineating boundaries. Complementarily, BiLSTM networks enhance temporal feature extraction by processing information in both forward and backward directions, allowing improved pattern recognition and classification. This hybrid approach aims to improve the accuracy and efficiency of defect detection in composite materials, offering a more reliable and automated solution.

The proposed model leverages DL techniques to perform the segmentation task, addressing the shortcomings of traditional approaches. By integrating spatial and temporal analysis, the method improves the robustness of defect detection against variations in geometry and defect patterns. An overview of this proposed solution is presented in Figure 1.

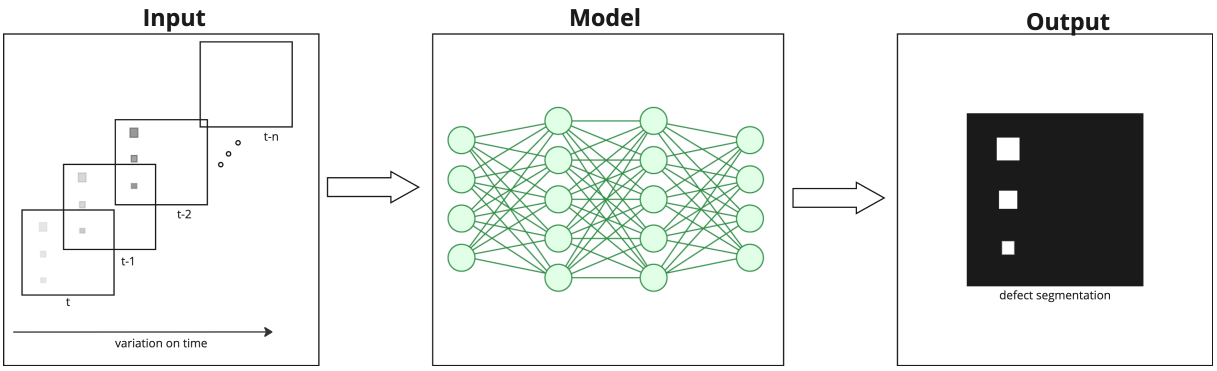


Figura 1 – Overall structure of the proposed solution. Source: Own work.

1.1 Motivation

The advent of composite materials has transformed industries such as aerospace and energy by providing high strength-to-weight ratios and improving durability (HSISSOU et al., 2021). However, their complex structures introduce challenges, including manufacturing defects and damage during service life, which can compromise safety and performance (DAS; SAHU; PARHI, 2021).

NDT techniques play an important role in ensuring the structural integrity of composites by allowing defect detection without damaging the material (GUPTA et al., 2021).

IRT has gained popularity as an NDT method due to its ability to detect subsurface defects through thermal patterns. Despite its advantages, IRT still faces limitations in dealing with complex geometries and overlapping defect patterns (KHODAYAR; SOJASI; MALDAGUE, 2016).

The increasing availability of sensor data has led to the integration of AI techniques, particularly DL, in NDT processes. DL methods, such as NN, excel at processing large datasets and automating defect detection with high accuracy (VANRULLEN; KANAI, 2021). However, conventional methods such as traditional image processing algorithms or manual segmentation present limitations in terms of detecting defects with high confidence.

In this context, this work aims to contribute to the scenario of seeking more robust alternatives to detect defects in composite materials using DL techniques. To this end, this research proposes a new hybrid approach combining DeepLabv3 and BiLSTM networks. DeepLabv3 effectively performs spatial segmentation, while BiLSTM enhances temporal pattern recognition. This integration enables precise and automated defect detection, improving scalability and accuracy in composite material inspections.

1.2 Objective and Challenges

1.2.1 General Objectives

To develop a model capable of detecting and segmenting defects in CFRP generated by IRT using both temporal and spatial features.

1.2.2 Specific Objectives

The specific objectives to be achieved are:

- ❏ To review the state-of-the-art solutions for deep learning (DL) applied to defect detection and segmentation in CFRP.
- ❏ To study IRT to understand its properties and importance as a NDT technique.
- ❏ To analyze the properties and characteristics of CFRP materials.
- ❏ To explore various DL architectures to design an effective solution.
- ❏ Implement and test Temporal Convolutional Network (TCN) for CFRP defect segmentation.
- ❏ Compare different DL architectures, such as U-Net, DeepLabv3, and BiLSTM, for defect segmentation in CFRP.

- ❏ To develop a model capable of detecting and segmenting defects in CFRP using data generated through IRT.
- ❏ Evaluate the model developed to ensure its generalization across different datasets.
- ❏ Compare the performance of the proposed method against state-of-the-art solutions.

1.3 Hypothesis

This research hypothesizes that **advanced DL architectures can effectively analyze and assess the structural integrity of CFRP by interpreting data obtained through IRT to detect defects**. Furthermore, it is expected that **combining TCN and Convolutional Neural Network (CNN) will improve the performance of defect segmentation, particularly in CFRP**.

The goal of this study is to investigate the potential of DL architectures applied to IRT data for defect detection in composite materials and to analyze the importance of combining temporal and spatial information in these models to enhance segmentation. By evaluating the performance of these methods, this research aims to provide insights into their feasibility and effectiveness for structural health monitoring of CFRP.

1.4 Contribution

The project presents several contributions. Firstly, it provides a comprehensive literature review on composite materials and their applications using IRT, along with DL architectures and methods. Secondly, the project will contribute to the generation of datasets on composite materials with defects for use in NDT. Third, create and present DL models capable of detecting and segmenting defects in composite materials using IRT data. Fourth, propose a new method combining BiLSTM network with DeepLabv3 for the process of IRT image segmentation, which will improve the ability to capture temporal dependencies in IRT data and enhance the segmentation performance. Lastly, reproducibility of our findings given the fact that all the data will be available under public repository and the code used in this work is documented at github repository[1] and its readme can be seen in Appendix A;

1.5 Dissertation Organization

The dissertation is organized into six chapters, each addressing aspects of the research. The chapters are structured as follows:

[1] <https://github.com/iagogv/master-research>

In Chapter 2, the theoretical concepts that will be used in the solution are presented. The chapter begins with an explanation of **composite materials**, emphasizing their properties and significance. This is followed by an analysis of the principles and applications of **IRT** within the research context. Next, the chapter delves into **image processing**, with a particular focus on **adaptive thresholding** technique. The architectures of **U-Net** and **DeepLabv3** are explained in detail, along with an investigation of the **BiLSTM** model and its integration with other techniques. Additionally, essential concepts of NN are discussed, including **activation functions**, **loss functions**, **training**, and **validation**.

In Chapter 3, a review of the existing literature and research related to the topics addressed in this dissertation is provided. It presents an analysis of previous studies and identifies gaps that this research aims to fill.

The Chapter 4 details the methodology employed in the research. It covers several sections: **Data Collection**, which discusses the methods and protocols for data acquisition; **Data Annotation**, which describes the techniques and tools used for annotating the dataset; **Proposed Method**, which presents the proposed method to improve defect segmentation on CFRP using temporal and spatial informations; **Hardware and Software Setup**, which provides details on the computational resources and software environments utilized; **Training and Validation**, which explains the processes for training and validating the models; **Evaluation Metrics**, which outlines the metrics used to evaluate model performance.

Chapter 5 presents the experimental setup, results, and analysis. It includes a detailed account of the training process for the U-Net model, a description of the training process for the DeepLabv3 model, and an explanation of the proposed method using DeepLabv3 and BiLSTM. The chapter also provides a comprehensive evaluation of the experimental results and concludes with an interpretation and discussion of the results.

The Chapter 6 summarizes the key contributions of the research and suggests directions for future work. It highlights the main contributions and significance of the research findings and proposes potential areas for further research and development.

Chapter 2

Fundamentals

In this chapter, we examine the fundamental concepts essential for a comprehensive understanding of this research. For this purpose, the following sections are organized as follows:

- ❏ **Section 2.1:** Provides an in-depth discussion of composite materials and CFRP.
- ❏ **Section 2.2:** Explains IRT as a method for NDT.
- ❏ **Section 2.3:** Describes the adaptive threshold technique, a traditional image processing method employed in this study.
- ❏ **Section 2.4:** Details the U-Net architecture and its role in the proposed approach.
- ❏ **Section 2.5:** Discusses the DeepLabv3 architecture and its application in this work.
- ❏ **Section 2.6:** Explores the BiLSTM architecture and its functionality.
- ❏ **Section 2.7:** Explain the activation function, which adds non-linearity to the model.
- ❏ **Section 2.8:** Explains the loss function, which quantifies the model's prediction error and measures how close the predictions are to the expected outputs.
- ❏ **Section 2.9:** Describes the methodology for training a neural network.
- ❏ **Section 2.10:** Highlights the importance of validation sets and presents a strategy for cross-validation.

2.1 Composite Materials

Composite materials represent a class of engineered substances, formed through the combination of two or more constituent materials that possess significantly different physical or chemical properties. The synthesis of these materials results in a composite

whose overall characteristics are distinct and often superior to those of the individual components, thereby enhancing the material's performance and broadening its range of applications (CLYNE; HULL, 2019).

A composite material generally comprises two primary phases: the matrix and the reinforcement. The matrix, which can be a polymer, metal, or ceramic, serves as the binding medium for the reinforcement phase. This reinforcement phase typically consists of fibers, particles, or flakes. The resulting structure allows the composite to leverage the advantageous properties of both components. For example, CFRP integrate the ductility and toughness of the matrix with the high tensile strength of the fibers, creating a material that is simultaneously strong and lightweight (CHUNG, 2010).

In this study, a set of CFRP was utilized as the primary composite material. The selection of CFRP was based on its advantageous properties, including high tensile strength, lightweight characteristics, and resistance to external factors.

2.1.1 Carbon Fiber Reinforced Polymers

CFRP are advanced composite materials mostly used in aerospace, automotive, and structural engineering applications due to their high strength-to-weight ratio, corrosion resistance, and excellent mechanical properties. CFRP consists of carbon fibers embedded in a polymer matrix, typically epoxy, which provides rigidity and durability as shown in Figure 2.

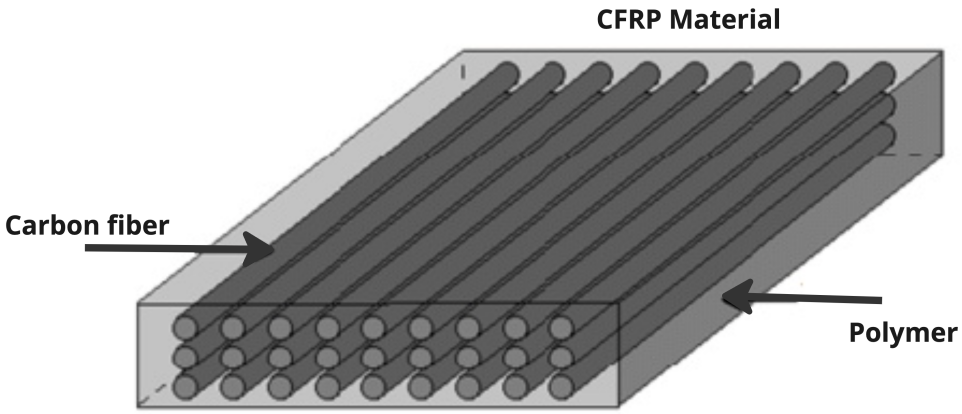


Figura 2 – Structure of a CFRP Composite Material. Source: Own work.

The alignment of carbon fibers within the polymer matrix can be tailored to meet specific mechanical demands, such as tensile strength, stiffness, and impact resistance. However, CFRP materials are susceptible to defects, as shown in Figure 3. These defects can include delamination, fiber breakage, and matrix cracking, which can compromise their performance (CLYNE; HULL, 2019). These defects are commonly inspected using NDT techniques such as IRT, enabling early detection and maintenance of structural integrity.

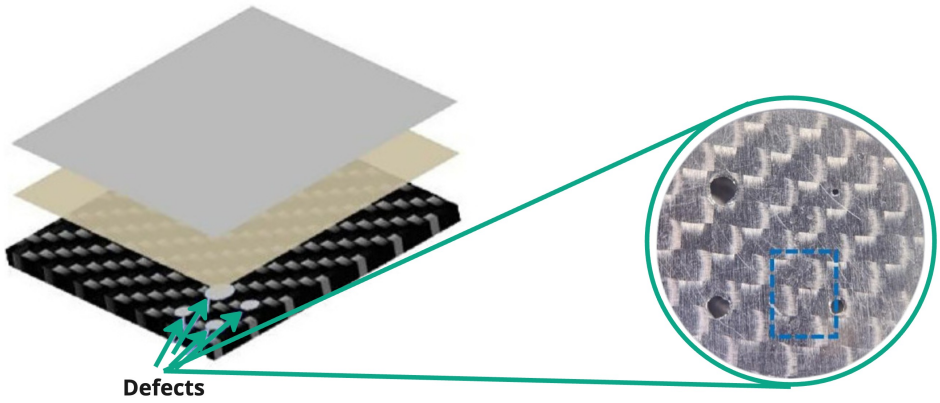


Figura 3 – Representation of Defects in a CFRP Composite Material. Source: Own work.

2.2 Infrared Thermography as a NDT method

The IRT has emerged as a powerful and versatile NDT method, playing a crucial role in various industries in inspecting and evaluating the integrity of materials and structures. It utilizes the principle of capturing and analyzing infrared radiation emitted by objects to detect and assess hidden flaws, defects, or irregularities without causing any damage (QU; JIANG; ZHANG, 2020).

IRT growing popularity is attributed to its non-invasive and non-contact nature, which allows for rapid and efficient inspections. The method involves the use of infrared cameras or thermographic devices to detect variations in surface temperatures. When materials experience changes in their internal conditions, such as the presence of defects, delaminations, or voids, these anomalies alter the surface temperatures, creating distinguishable thermal patterns (LEI, 2018).

IRT refers to the region of the spectrum that lies between visible light and microwaves. It covers wavelengths ranging from approximately 0.7 μm to 1000 μm, as seen in Figure 4. This region is commonly divided into five parts: Near-infrared (NIR) from 0.74 μm to 1 μm; Short-Wave InfraRed (SWIR) from 1 μm to 3 μm; Mid-Wave InfraRed (MWIR) from 3 μm to 5 μm; Long-Wave InfraRed (LWIR) from 8 μm to 14 μm; and TeraHertz (THz) from 100 μm to 1000 μm.

The NIR region lies just beyond the visible light spectrum and extends from approximately 0.74 to 1 micrometer (μm) in wavelength. NIR radiation is commonly used in applications such as remote sensing, spectroscopy, and optical communication. It is also used in agriculture, food analysis, and medical imaging. The SWIR region spans from 1 to 3 micrometers (μm) in wavelength. SWIR radiation is used in various fields, including night vision devices, moisture detection, and material inspection. It is parti-

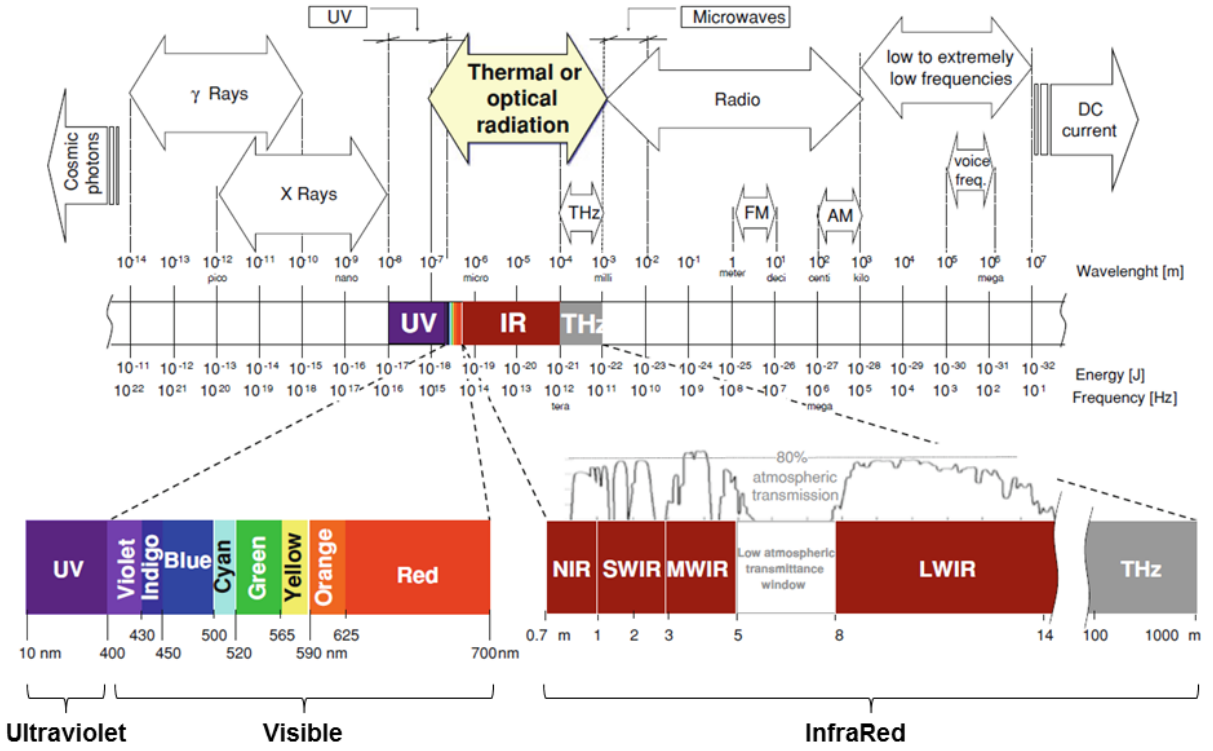


Figura 4 – EM Spectrum. Source: Adapted from (CASTANEDO, 2005)

cularly useful for imaging through haze, fog, and smoke, making it valuable in certain military and surveillance applications. The MWIR region covers wavelengths from 3 to 5 micrometers (μm). MWIR radiation is often utilized in thermal imaging cameras, where it can detect the heat emitted by objects at elevated temperatures. It is commonly used for industrial inspections, monitoring of manufacturing processes, and in certain defense and security applications. The LWIR region ranges from 8 to 14 micrometers (μm) in wavelength. LWIR radiation is widely used in thermal imaging systems to detect heat signatures emitted by objects, making it valuable in applications such as surveillance, firefighting, and medical thermography. It is also commonly used in night vision equipment. The THz region covers a wide range of wavelengths, from 100 to 1000 micrometers (μm). THz radiation is still an emerging field, with applications in spectroscopy, security screening, and imaging. It has unique properties that enable NDT and imaging through certain materials, such as clothing or paper, making it useful in security and inspection scenarios (BUTCHER, 2016).

2.2.1 Active thermography

Active infrared thermography is a NDT technique that leverages the principles of thermodynamics and infrared imaging to, among other things, evaluate the thermal properties of materials. In this process, a heat source is applied to the surface of the material being inspected, resulting in a localized temperature increase. The material's thermal response

is subsequently recorded using an infrared camera, which detects the emitted infrared radiation (KOSOVA; ALTAY; ÜNVER, 2024). This entire process is shown in Figure 5. By analyzing the thermal patterns captured by the camera, it is possible to infer valuable information about the material's internal structure, including the presence of defects and anomalies.

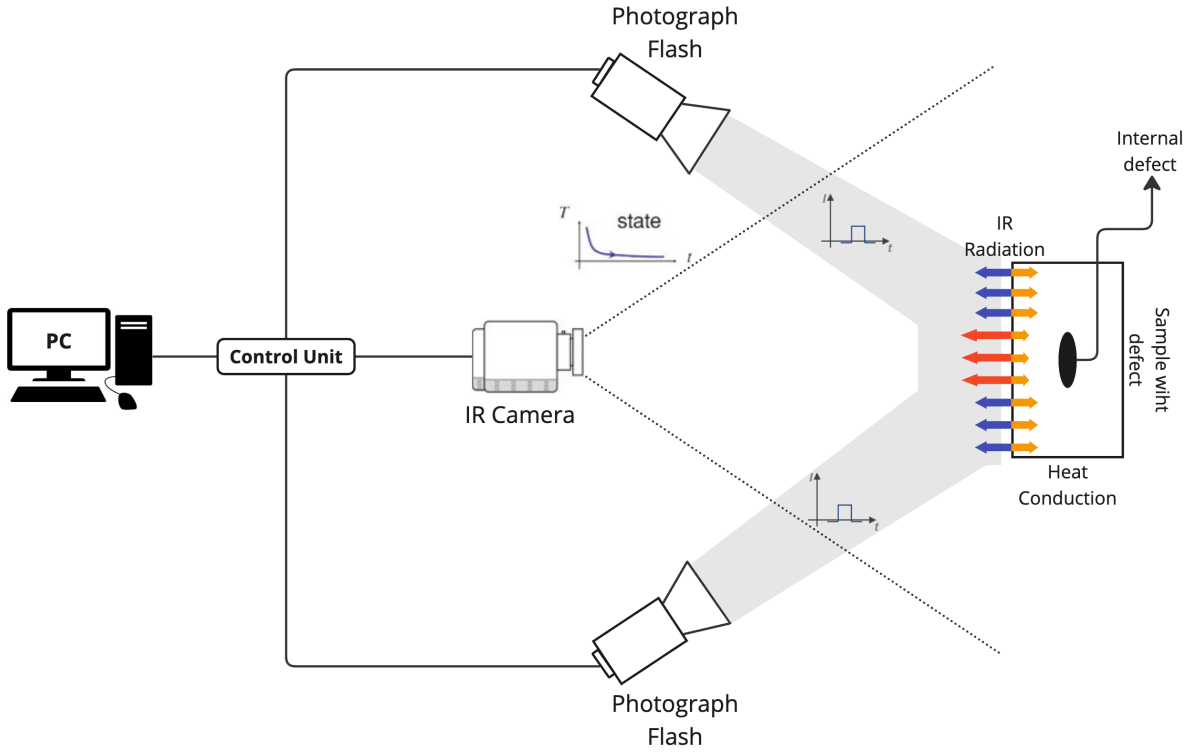


Figura 5 – Active thermography. Source: Adapted from (KOSOVA; ALTAY; ÜNVER, 2024)

2.3 Adaptative Thresholding

Image processing is an essential technique in the field of signal processing, focusing on performing operations on images to enhance their quality or extract valuable information. At its core, image processing involves a series of techniques and procedures aimed at improving the visual aspects of images or drawing out specific characteristics and parameters from them. Whether it is sharpening an image, removing noise, or detecting edges, the goal is to make the image more useful and informative. This can be particularly crucial in areas such as medical imaging, where clarity and detail can significantly impact diagnostics, or in automated systems such as self-driving cars, where precise image interpretation is vital for safety and functionality (NIBLACK, 1985). In this research, a specific image processing technique known as adaptive thresholding was employed to segment defects in CFRP.

Adaptive thresholding is a technique in image processing that helps convert a grayscale image into a binary image, which is basically an image that only has two colors: black and white. This method is particularly handy when dealing with images that have uneven lighting, where applying a single threshold value across the entire image would not work well (BRADLEY; ROTH, 2007). Instead of using a global threshold, adaptive thresholding calculates the threshold for each pixel based on the local characteristics of the surrounding pixels. Essentially, the image is divided into smaller regions and a unique threshold value is determined for each of these regions. This approach allows for better handling of images with varying lighting conditions.

The general form of the adaptive threshold is represent by:

$$T(x, y) = f(I(x, y), \mathcal{N}(x, y)) \tag{1}$$

Where:

- $I(x, y)$: The intensity or value of the input data at position (x, y).
- $\mathcal{N}(x, y)$: A neighborhood or region surrounding (x, y).
- $f(\cdot)$: A function that computes the threshold based on the local properties of $\mathcal{N}(x, y)$.

Adaptive thresholding techniques are mainly divided into two categories: mean-based and Gaussian-based methods. This study focuses on mean-based thresholding, which calculates the threshold for each pixel by averaging the pixel values within its local neighborhood. This approach is particularly effective in yielding accurate results, especially in images with varying lighting conditions (SINGH et al., 2012). The threshold is determined using Equation (2).

$$T(x, y) = \mu_{\mathcal{N}} - C \tag{2}$$

Where:

- $\mu_{\mathcal{N}}$: Mean of the values in the neighborhood $\mathcal{N}(x, y)$.
- C: A constant used to fine-tune the threshold.

2.4 U-Net

U-Net is a type of CNN designed specifically for biomedical image segmentation but has proven effective in other domains, including defect segmentation in infrared images of CFRP (ANBUDEVI; SUGANTHI, 2022).

The architecture of U-Net, as shown in Figure 6, consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. This

design allows U-Net to learn high-resolution features from relatively small datasets, making it suitable for tasks where annotated data is limited. The following subsections will detail some the components of this architecture.

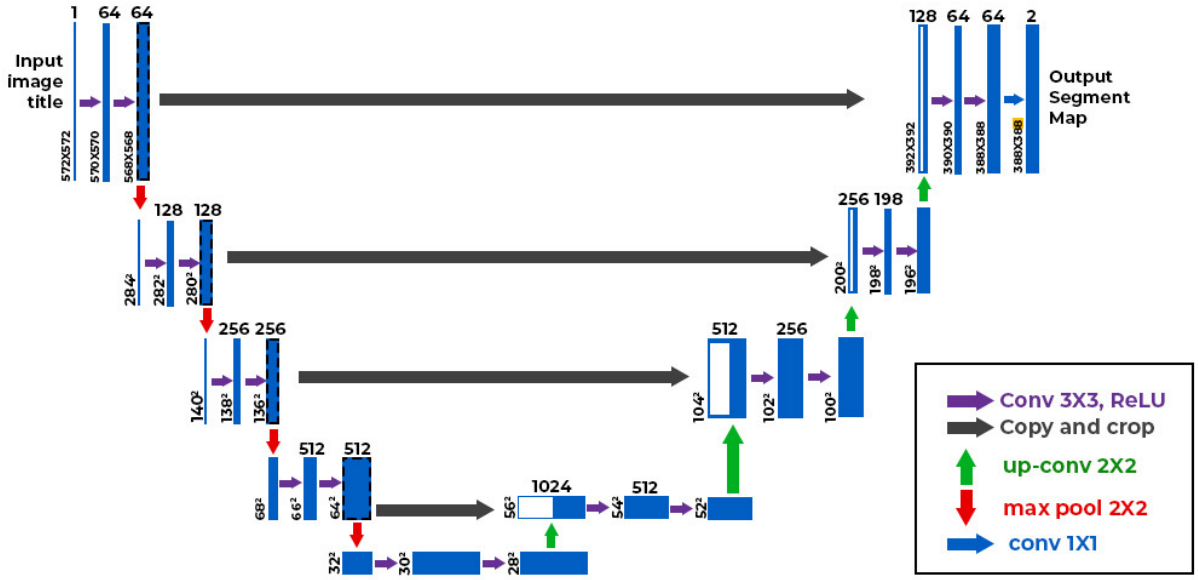


Figura 6 – U-Net Architecture. Source: Adapted from (RONNEBERGER; FISCHER; BROX, 2015)

2.4.1 Contracting Path

In the contracting path, the network applies a series of convolutions and max-pooling operations to progressively reduce the spatial dimensions while capturing increasingly complex features. It uses convolutional layers, where for a 2D input image X, a convolution layer applies a kernel W and bias b:

$$Y = \text{ReLU}(W * X + b), \tag{3}$$

where $*$ denotes the convolution operation.

This component also uses down-sampling (pooling). After convolution, max-pooling reduces spatial dimensions. This process can be explained by:

$$Y_{\text{pooled}}(i, j) = \max_{m,n} Y(2i + m, 2j + n), \tag{4}$$

where (m, n) represents the pooling window.

2.4.2 Expansive Path

The expansive path uses transposed convolutions to upsample the feature maps, restoring the original spatial dimensions. Skip connections between corresponding layers in the contracting and expansive paths enable the network to combine high-level semantic

information with low-level detailed information, enhancing the accuracy of segmentation (RONNEBERGER; FISCHER; BROX, 2015). The expansive path involves three key components: the transposed convolution, the upsampling, and convolutional layers.

Transposed convolutions are used to increase the spatial resolution by a factor of 2:

$$Y_{\text{upsampled}} = W^{\top} * X + b, \tag{5}$$

where $W^{\top}$ is the transpose of the convolution kernel, X is the input feature map, and b is the bias term.

The upsampled feature map is concatenated with the corresponding feature map from the encoder to retain fine-grained information. This process is represented as:

$$F_{\text{skip}} = \text{Concat}(F_{\text{encoder}}, F_{\text{decoder}}), \tag{6}$$

where F_{encoder} and F_{decoder} represent the feature maps from the encoder and decoder paths, respectively.

Finally, convolutional layers are applied to further refine the upsampled feature map using standard convolution operations, as described in Equation (3).

2.4.3 Output Layer

The output is typically a 1×1 convolution layer that maps the final feature map F to the desired number of classes C for pixel-wise classification:

$$P(x, y) = \text{Softmax}(W_{\text{out}} * F(x, y) + b_{\text{out}}), \tag{7}$$

where $P(x, y)$ is the probability of each pixel belonging to a class.

2.5 DeepLabv3

DeepLabv3 is another powerful DL model used for semantic image segmentation, leveraging atrous (dilated) convolutions to effectively capture multi-scale contextual information while preserving spatial resolution. This property is particularly beneficial for detecting and segmenting defects in IRT images of CFRP, where accurate localization and context are critical (CHEN et al., 2017).

The innovation in DeepLabv3 lies in its Atrous Spatial Pyramid Pooling (ASPP) module, which employs atrous convolutions with multiple dilation rates in parallel to extract features at different scales. This section details the theoretical concepts behind atrous convolutions and ASPP.

2.5.1 Atrous Convolutions

Atrous convolutions, also known as dilated convolutions, introduce a dilation rate r that defines the spacing between kernel elements. This allows the convolution to cover a larger receptive field without increasing the number of parameters or computational complexity. For a 2D input feature map F, the output of an atrous convolution is given by:

$$Y[i,j] = \sum_{m=1}^{M} \sum_{n=1}^{N} W[m,n] \cdot F[i + r \cdot m, j + r \cdot n], \tag{8}$$

where:

- $Y[i,j]$ is the output feature at spatial location (i,j),
- $W[m,n]$ represents the convolutional kernel weights,
- $F[i,j]$ is the input feature map,
- r is the dilation rate, and
- $M \times N$ is the kernel size.

Atrous convolutions preserve the spatial resolution of the feature map, which is crucial for dense prediction tasks like semantic segmentation, while effectively expanding the receptive field to incorporate broader contextual information.

2.5.2 Atrous Spatial Pyramid Pooling (ASPP)

The ASPP module in DeepLabv3 is designed to capture multi-scale context by applying atrous convolutions with varying dilation rates in parallel. This enables the model to extract features at different resolutions and scales. Mathematically, the ASPP module can be represented as:

$$F_{\text{ASPP}} = \text{Concat}\left[\text{Conv}_{r_1}(F), \text{Conv}_{r_2}(F), \ldots, \text{Conv}_{r_N}(F), \text{GAP}(F)\right], \tag{9}$$

where:

- $\text{Conv}_{r_i}(F)$ denotes an atrous convolution with dilation rate r_i applied to the input feature map F,
- $\text{GAP}(F)$ represents global average pooling, given by:

$$\text{GAP}(F) = \frac{1}{|F|} \sum_{i,j} F[i,j], \tag{10}$$

- Concat denotes concatenation along the channel dimension.

The outputs from the parallel atrous convolutions and the global average pooling are concatenated and passed through a 1×1 convolution to reduce dimensionality and merge the multi-scale information:

$$F_{\text{output}} = \text{Conv}_{1\times1}\left(F_{\text{ASPP}}\right). \tag{11}$$

2.5.3 DeepLabv3 Architecture

The overall architecture of DeepLabv3 as shown in Figure 7, begins with an input image that is processed through an initial convolution and pooling layer, reducing its spatial dimensions. This is followed by a backbone (e.g., ResNet or Xception) that extracts features through several convolutional blocks. Atrous convolutions are introduced in the later blocks of the backbone to maintain spatial resolution while expanding the receptive field.

The ASPP module is applied to the extracted feature maps, aggregating multi-scale context through parallel atrous convolutions and global pooling. The output is merged and upsampled using bilinear interpolation to restore the spatial dimensions of the input image, ensuring precise pixel-wise segmentation.

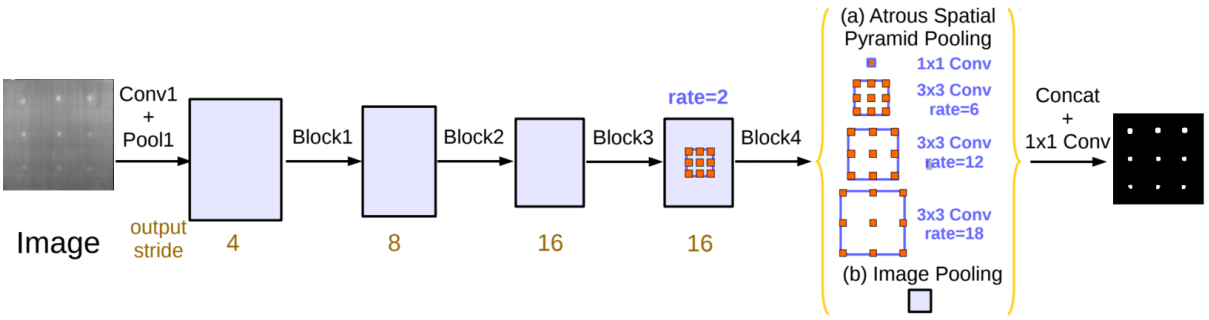


Figura 7 – DeepLabv3 Architecture. Source: Adapted from (CHEN et al., 2017)

2.6 BiLSTM

While CNNs excel at capturing spatial features, Recurrent Neural Networks (RNNs) are adept at modeling temporal dependencies in sequential data. BiLSTM networks, a variant of RNNs, are particularly effective in capturing long-range dependencies by processing data in both forward and backward directions. In the context of defect segmentation in IRT images of CFRP, BiLSTM networks can analyze temporal sequences of images captured over time (ZHANG; YIN, 2021). This is especially useful when monitoring the evolution of thermal patterns, providing additional information on the presence and progression of defects.

2.6.1 LSTM

To understand BiLSTM, it is essential to first grasp Long Short-Term Memory (LSTM) networks. LSTM was introduced as a variant of RNNs designed to overcome the vanishing gradient problem, which makes standard RNNs ineffective at learning long-term dependencies (HOCHREITER; SCHMIDHUBER, 1997). As shown in Figure 8, LSTM introduces a memory cell and three gates (input, forget, and output gates) that regulate the flow of information.

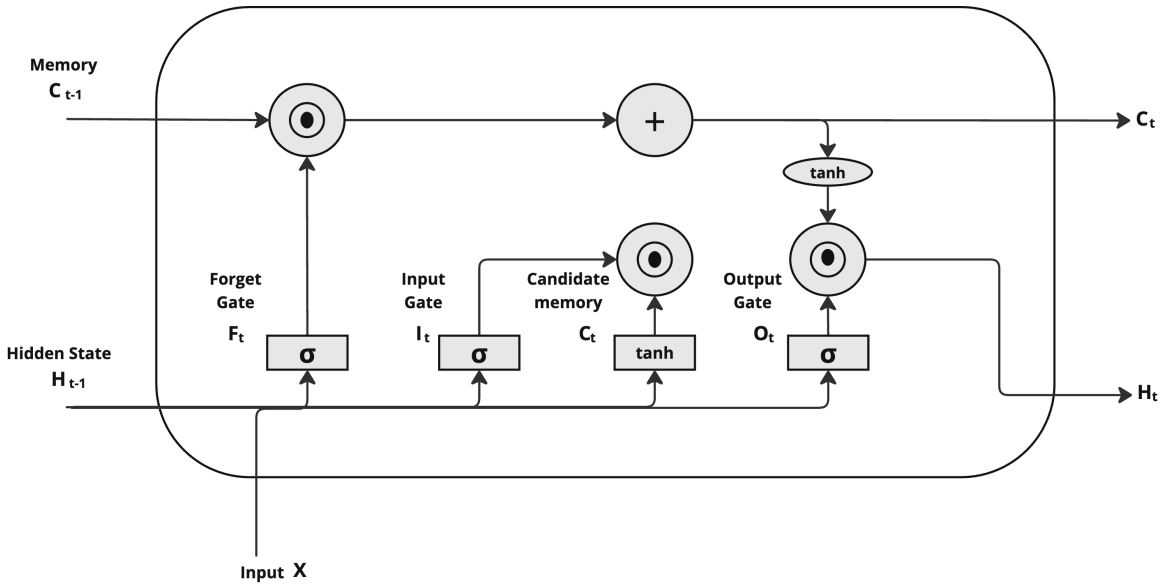


Figura 8 – LSTM Architecture. Source: Adapted from (HOCHREITER; SCHMIDHUBER, 1997)

The **input gate** controls how much of the new input to add to the cell state. It is computed as follows:

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right), \tag{12}$$

where x_t is the input at time t, h_{t-1} is the previous hidden state, and σ denotes the sigmoid activation function.

The **forget gate** decides how much of the previous cell state to retain:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right). \tag{13}$$

The **output gate** determines the output based on the cell state:

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right). \tag{14}$$

The **cell state** is updated by combining the forget gate, input gate, and a candidate cell state:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right), \tag{15}$$

where $\odot$ represents element-wise multiplication.

Finally, the **hidden state** is updated based on the output gate and the cell state:

$$h_t = o_t \odot \tanh(c_t). \tag{16}$$

These components work together to enable LSTM networks to learn long-term dependencies effectively, making them a cornerstone for many sequential modeling tasks.

2.6.2 BiLSTM Architecture

A Bidirectional LSTM was introduced by (SCHUSTER; PALIWAL, 1997). Figure 9 presents the overall architecture, which consists of two LSTM layers: one that processes the sequence from start to end (forward direction) and another that processes the sequence from end to start (backward direction). The outputs of these two layers are combined, typically by concatenation, enabling the model to capture both past and future context within the sequence.

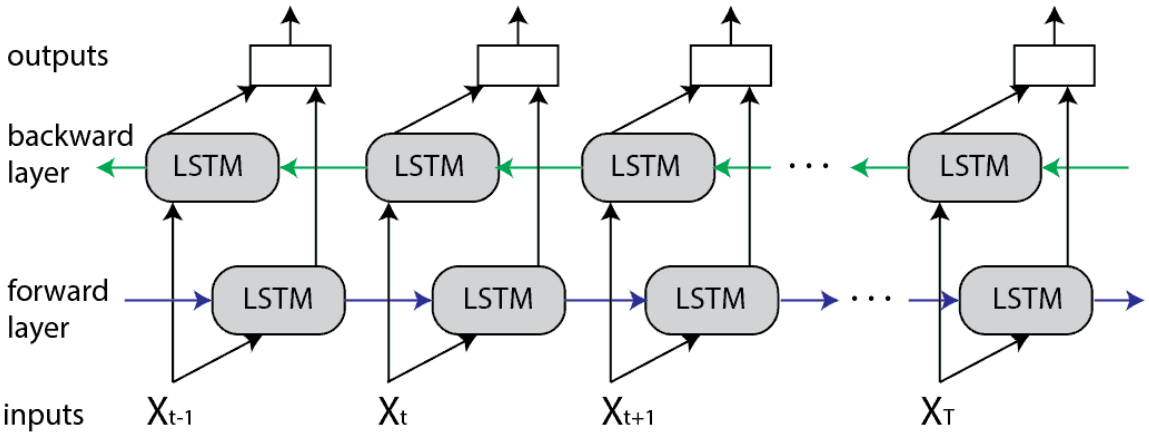


Figura 9 – BiLSTM Architecture. Source: Adapted from (IHIANLE et al., 2020)

The **forward LSTM layer** processes the input sequence $X = \{x_1, x_2, \ldots, x_T\}$ in its natural order. The hidden state at time step t is computed as:

$$\overrightarrow{h_t} = \text{LSTM}_f\left(x_t, \overrightarrow{h_{t-1}}\right). \tag{17}$$

The **backward LSTM layer** processes the input sequence in reverse order, and the hidden state at time step t is given by:

$$\overleftarrow{h_t} = \text{LSTM}_b\left(x_t, \overleftarrow{h_{t+1}}\right). \tag{18}$$

The **combined output** of the BiLSTM is obtained by concatenating the hidden states from the forward and backward layers:

$$h_t = \left[\overrightarrow{h_t}; \overleftarrow{h_t}\right], \tag{19}$$

where $[\cdot;\cdot]$ denotes concatenation.

This architecture allows BiLSTM models to leverage information from both the past and future context within the sequence, making them particularly effective for tasks involving sequential or temporal data.

2.7 Activation Functions

Activation functions play a critical role in NN by introducing nonlinearity, enabling the network to learn complex mappings from inputs to outputs. In the context of the architectures discussed in sections 2.5, 2.4 and 2.6, activation functions are indispensable for both extracting meaningful spatial features in convolutional layers and modeling temporal dependencies in recurrent layers. This section reviews some widely used activation functions and their relevance to the models explored earlier.

2.7.1 Rectified Linear Unit (ReLU)

The ReLU activation function introduced by (GLOROT; BORDES; BENGIO, 2011) is one of the most commonly used in DL models, particularly in CNN like DeepLabv3. It is defined as:

$$\mathrm{ReLU}(x) = \max(0, x). \tag{20}$$

ReLU introduces sparsity into the network by zeroing out negative values, helping mitigate the vanishing gradient problem and accelerates training. Its simplicity and computational efficiency make it an ideal choice for models like DeepLabv3 that require processing large feature maps.

2.7.2 Sigmoid Function

The sigmoid function introduced by (RUMELHART; HINTON; WILLIAMS, 1986) maps inputs to a range between 0 and 1, making it suitable for probabilistic interpretations in binary classification tasks. It is defined as:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \tag{21}$$

Although the sigmoid function has been widely used in earlier NN, it suffers from vanishing gradient issues, making it less suitable for deeper architectures like those used in DeepLabv3. However, it remains relevant in certain gating mechanisms, such as those in LSTM and BiLSTM, where it helps control the flow of information.

2.7.3 Hyperbolic Tangent (Tanh)

The Tanh function is another activation function used in recurrent architectures like BiLSTM. It is defined as:

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \tag{22}$$

and maps inputs to the range $[-1, 1]$. Unlike the sigmoid function, Tanh is zero-centered, which often leads to faster convergence during training. In the context of BiLSTM, Tanh is frequently used within the memory cell computations to provide smooth gradients and balance updates.

2.7.4 Softmax Function

The Softmax function is commonly used in the final layer of a network for multi-class classification tasks. It transforms the raw network outputs (logits) into probabilities that sum to 1. For a vector of logits $z = [z_1, z_2, \ldots, z_K]$, the Softmax function is given by:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}, \quad \forall i \in \{1, 2, \ldots, K\}. \tag{23}$$

In models like DeepLabv3, Softmax is used to compute class probabilities for each pixel in the segmentation map.

2.7.5 Relevance to DeepLabv3 and BiLSTM

In DeepLabv3, ReLU is predominantly used in convolutional layers to handle feature extraction efficiently. Softmax or sigmoid is employed in the final layers to produce segmentation maps with class probabilities. In BiLSTM, sigmoid and Tanh play a key role in regulating the flow of information through gates and ensuring smooth gradients during training. These activation functions collectively enable the seamless integration of spatial and temporal features in hybrid models.

2.8 Loss Function

The loss function is a critical component in training deep learning models, as it quantifies the discrepancy between the model's predictions and the true labels. By providing a measure of prediction error, the loss function guides the optimization process, enabling the model to iteratively adjust its parameters to minimize this error. In the context of the architectures discussed earlier, such as DeepLabv3 and BiLSTM, the choice of a suitable loss function is crucial to achieve high performance in tasks such as semantic segmentation and sequence modeling.

For classification tasks, including pixel-wise classification in semantic segmentation, cross-entropy loss is commonly used. It is defined as:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{i,k}\log \hat{y}_{i,k}, \tag{24}$$

where:

- N is the number of samples (or pixels, in the case of semantic segmentation),
- K is the number of classes,
- $y_{i,k}$ is the ground truth label (one-hot encoded) for class k at sample i,
- $\hat{y}_{i,k}$ is the predicted probability for class k at sample i.

Cross-entropy loss penalizes incorrect predictions based on the confidence of the model, encouraging accurate and confident classifications. In DeepLabv3, cross-entropy loss is typically applied to pixel-wise predictions to train the model for semantic segmentation tasks.

2.9 Training

Training a NN involves optimizing its parameters to minimize the loss function, ensuring that the model accurately maps the inputs to desired outputs. This section outlines the steps and strategies employed in training the discussed architectures, such as DeepLabv3 and BiLSTM, for their respective tasks.

2.9.1 Dataset Preparation

The first step in training a neural network is to prepare the dataset. This involves gathering a diverse and representative dataset that reflects the variability of the task (e.g., IRT images of CFRP defects or sequential data for temporal analysis). The preprocessing is then applied, which typically includes normalization or standardization to ensure consistent input scales. For image data, resizing or cropping images to a uniform size compatible with the network architecture is essential.

Data augmentation can improve the robustness of the dataset by applying transformations such as rotation, flipping, or noise addition. This is particularly important for tasks with limited training data. The dataset should be divided into training, validation, and test sets, ensuring no overlap between these subsets to prevent data leakage.

2.9.2 Initialization of Parameters

Network parameters, such as weights and biases, must be initialized before training begins. Common initialization techniques include **Xavier Initialization** described by (GLOROT; BENGIO, 2010), which ensures the variance of weights remains consistent across layers, helping to prevent vanishing or exploding gradients. Another popular method introduced by (HE et al., 2015) is **He Initialization**, optimized for ReLU activation functions, which improves convergence in deep architectures by appropriately scaling weights. Additionally, for architectures like DeepLabv3, **Pretrained Models** are often utilized. These pretrained weights, such as those from ResNet backbones, can accelerate convergence and enhance performance.

2.9.3 Optimization Algorithm

Training requires an optimization algorithm to minimize the loss function by updating the network parameters. Two commonly used optimization algorithms are **Stochastic Gradient Descent (SGD)** and the **Adam Optimizer**.

Stochastic Gradient Descent (SGD) updates parameters based on a small batch of data. The update rule is defined as:

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta_t} \mathcal{L}, \tag{25}$$

where θ_t represents the parameters at iteration t, η is the learning rate, and $\nabla_{\theta_t} \mathcal{L}$ is the gradient of the loss function.

The **Adam Optimizer** combines momentum and adaptive learning rates to achieve efficient convergence. The algorithm is governed by the following equations:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \tag{26}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t}, \quad \theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}, \tag{27}$$

where g_t is the gradient at iteration t, and β_1, β_2, and ϵ are hyperparameters.

2.9.4 Batch Size and Epochs

The training process iteratively updates parameters based on subsets of the dataset (batches) for a fixed number of epochs. The **batch size** determines the number of samples used per gradient update. Smaller batch sizes reduce memory usage but introduce more noise into the gradient estimates, which can help avoid local minima. The **epochs** specify the number of complete passes through the training dataset, which ensures the model adequately learns from the data.

2.9.5 Regularization Techniques

Regularization techniques help prevent overfitting by constraining the model's complexity. One popular method is **dropout**, which randomly sets a fraction of neurons to zero during training, reducing co-adaptation among neurons and improving generalization. Another approach is L_2 **regularization**, which penalizes large weight values by adding a term to the loss function:

$$\mathcal{L}_{\text{regularized}} = \mathcal{L} + \lambda \|\theta\|_2^2, \tag{28}$$

where λ is the regularization coefficient. This technique helps ensure that the weights remain small, contributing to model stability and reduced overfitting.

2.9.6 Validation and Early Stopping

Validation and early stopping are essential for evaluating model performance and preventing overfitting. **Validation** involves monitoring the loss in a separate validation set during training, providing insights into the model's generalization capabilities. **Early stopping** halts the training process when the validation loss stops improving, preventing the model from overfitting to the training data.

2.9.7 Performance Metrics

To evaluate model performance beyond the loss function, metrics such as accuracy, precision, recall, and Intersection over Union (IoU) are commonly used. These metrics provide a more comprehensive view of the model's effectiveness, particularly for tasks like image segmentation, where IoU is critical for assessing performance.

By following this methodology, the architectures discussed in Sections 2.5 and 2.6 can be effectively trained to achieve high performance on their respective tasks.

2.10 Validation and Cross-Validation

Validation sets play a crucial role in training neural networks by providing an unbiased evaluation of the model's performance on unseen data. This evaluation helps to monitor overfitting and guides critical decisions such as hyperparameter tuning and early stopping.

During training, the model learns patterns from the training data, but its ability to generalize to unseen data is assessed using a separate validation set. Key benefits of using a validation set include:

1. **Hyperparameter Optimization:** Validation helps to select optimal hyperparameters such as learning rate, batch size, and regularization coefficients.

2. **Model Comparison:** Among different architectures or configurations, the validation set helps identify the best-performing model.

3. **Overfitting Detection:** A widening gap between training and validation performance indicates overfitting, necessitating strategies like regularization or early stopping.

2.10.1 Cross-Validation Strategy

Cross-validation is an effective technique to maximize the use of available data, particularly in scenarios with limited datasets. It involves partitioning the data into multiple subsets (folds) and rotating the validation and training sets across these folds.

The most common approach is k**-fold cross-validation**, where the dataset is divided into k equally sized folds. The model is trained k times, each time using a different fold as the validation set and the remaining $k-1$ folds as the training set. The final performance metric is obtained by averaging the results across all k iterations.

Mathematically, for a performance metric M, the final metric is computed as:

$$M_{\text{final}} = \frac{1}{k} \sum_{i=1}^{k} M_i, \tag{29}$$

where M_i represents the metric computed during the ith fold.

Cross-validation provides a robust estimate of model performance by mitigating the variance caused by data partitioning. It is particularly valuable for evaluating models in situations where the dataset size is small or imbalanced.

Chapter 3

Literature Review

This chapter presents the state of the art of using deep learning techniques in the context of NDT and IRT. For this, it first presents works using similar technology and their results. Finally, it presents the contributions of this work to CFRP defect segmentation using IRT.

3.1 Related Works

This Section presents the literature review and briefly discusses correlated works and their results. The study presented by (MALDAGUE; LARGOUËT; COUTURIER, 1998) is the first attempt to extract quantitative information using Machine Learning (ML) from the results of NDT evaluation using Pulsed Phase Thermography (PPT). The proposed approach is based on NN, which are well known for their ability to deal with complex nonlinear problems with access to noisy partial data. In the article, a thermal model is presented and some sections describe the correlation between phase and frequency, NN, and experimental data on aluminum and plastic materials. The paper concludes with encouraging results.

A three-dimensional thermal model for inhomogeneous materials such as CFRP is presented by (DARABI; MALDAGUE, 2002). Two backpropagation NN as flaw detector and depth estimator are then presented, and their result was compared with the analytical solution based on Duhamel's theorem. A Multi Layer Perceptron (MLP) network was trained with simulated data containing defects with different shapes, and the network classified 96.8% of the pixels using the simulated dataset. The trained model was also evaluated on experimental dataset and the network was unable to detect and reveal all the known defects.

In the study presented by (BENITEZ et al., 2006), an application of NN and thermography signal reconstruction (TSR) was proposed to characterize defects in composite materials. TSR is a technique used to process and filter thermograph data in order to find patterns in thermograph signals; it can be used, for example, to segment defects in composite ma-

terials. The paper presents the application of a NN to map between the TSR coefficients and defect depths mathematical spaces. They trained and validated several MLP architectures with TSR coefficients extracted from thermogram sequences of a CFRP sample and proposed a NN architecture. According to the authors, it was observed that the defects within 3 mm were not correctly classified. In addition, false defects detection occur especially in those areas surrounding the smallest defects (less than 3 mm). For this test, 80% of defective pixels and 96% of non-defective pixels were correctly classified.

The work carried out by (SAEED; ZARKANI; OMAR, 2019) aims to perform a comprehensive analysis of a NN when used as a post processor to quantify defect depths from CFRP thermograms. The authors used Pulsed Thermography (PT) methodology to inspect the materials and generate the thermograms. A pipeline based on filters was applied in the thermal data to remove noises such as reflections and residual heat. An in depth study is carried out to compare and analyze the effect of different parameters on the accuracy and robustness of the NN result using two sets of data, a original dataset create by PT, and a cleaned dataset which received a treatment pipeline to remove thermal noise such as the reflection and the contribution from the surroundings and the residual heat from the pulse source. The best result achieved in this study was an accuracy of 92% using the cleaned dataset.

A hybrid of spatial and temporal DL architecture for automatic thermography defect detection is proposed by (LUO et al., 2019). The integration of cross-network learning strategies can minimize uneven lighting and enhance the defect detection rate. The paper presents the influence of different algorithms in enhancing detection contrast from the spatial and the temporal domain, respectively. Different samples of CFRP were inspected using optical pulsed thermography and the produced image sequences were used to train the model with the proposed method. Six different DL algorithms were performed on the dataset, namely U-Net, VGG-U-Net, Segnet, VGG-Segnet, FCN8, and 3-layer-LSTM. The results show that the Visual Geometry Group-Unet (VGG-Unet) cross learning structure can significantly improve the contrast between defective and non defective regions, which achieved the Probability of Detection (POD) average value compared with the others spatial models, which was 66.7%.

An approach to determine the defect depth in composite materials, more precisely, in CFRP of aviation structure is presents by (WANG et al., 2020). The authors used a Laser Infrared Thermography (LIT) system to inspect the material and built a dataset to train and validate the developed model. An architecture of LSTM-Recurrent Neural Network (RNN) is used to train the proposed model. The purpose of the paper is to compare the results using the proposed methodology with other two traditional architectures, RNN and CNN, concluding that they achieved better results using the proposed methodology. The model achieved a recall of 95% while the traditional algorithms, RNN (one hidden layer and containing 128 nodes) and CNN (64 convolutional kernel), achieved a recall of

84% and 80%, respectively.

A sophisticated methodology called Gated Recurrent Unit (GRU) was used in the paper presented in 2020 by (FANG; MALDAGUE, 2020). The authors developed a deep neural network for defect depth estimation in composite materials through PT. They used a synthetic database composed of 270 thermal images of CFRP with known defects generated by Finite Element Method (FEM). The synthetic database was built with a variety of defects with different controlled depth and geometries. To train the model, they used a strategy of splitting the data in 80.0% of it for training and 20.0% of it for testing. The implementation was done on Graphics Processing Unit (GPU) using a machine with 64 GB of RAM. They used Adaptive Moment Estimation (ADAM) as an optimizer and a learning rate of 0.0001. The GRU developed is capable of quantifying automatically the depth of the defects precisely, with a Mean Absolute error (MAE) ranging from -0.11 mm to 0.11 mm for the normalized database, and the results provided an accuracy of 95.0%.

The work developed in 2021 by (XU; HU, 2021) presents a study of defect depth estimation based on GRU. The work proposes to solve the problem of insufficient precision in the estimation of the depth of defects. Active IRT was applied to obtain raw thermal sequences of the surface temperature field distribution of the inspected sample. Before training the GRU model, Principal Components Analysis (PCA) was used to reduce the dimension of the raw sequence and eliminate the correlation of the raw data sets. Then, the GRU model was used to automatically estimate the defect's depth. The results indicate that PCA-processed datasets outperform raw temperature datasets in model learning when evaluating defect depth characteristics. According to the authors, the recall achieved by the GRU network was over 95.0% under the condition of 1 mm depth defect.

Another common application in the field of image segmentation was described in 2021 by (FANG; IBARRA-CASTANEDO; MALDAGUE, 2021) in the paper entitled "Automatic Defects Segmentation and Identification by Deep Learning Algorithm with Pulsed Thermography: Synthetic and Experimental Data". The authors used DL to segment defects on composite materials. According to the authors, one of the most expensive phases in a DL project is the acquisition/construction of a good dataset to train the algorithm, especially when we work with infrared images. Fortunately, the industry does not produce many faulty samples. If one wants to build a database with balanced individuals (sound samples vs. defective samples) it will have a difficult time doing that with samples produced by the industry. In this way, they build an image dataset using a fusion approach between synthetic images generated by FEM and experimental images. The main goal of this paper was to demonstrate the possibility of training a model even with a limited dataset. According to the results presented by the authors, the trained CNN achieved more than 98% of accuracy and a precision of 68% on the validation dataset. However, they did not present a harmonic metric such as F1-score.

The detection of subsurface defects is one of the hardest tasks in the industry, the paper presented by (LEMA et al., 2022) addresses the challenging problem of subsurface defect detection using a combination of active thermography and DL. The main contribution of this work is to compare object detection rather than semantic segmentation or classification. The experiments were performed in a dataset obtained by applying active thermography to a single carbon fiber solid laminate sheet; this type of material consists of several carbon fiber layers, each layer with fibers oriented in a specific direction, stacked and bonded together with a resin. The defects in the laminate have different sizes: 12 x 12 mm, 7 x 7 mm, and 5 x 5 mm, and different depths: 0.63 mm, 1.46 mm, and 2.08 mm. They used a mobilenetv2 architecture to perform the classification model, a U-Net to perform segmentation and Yolov5 to perform the detection model. According to the authors, the detection model achieved the best result on the test dataset compared with the others. The Yolov5 architecture achieved a precision of 97.0%, a recall of 95.0% and a F1-score of 96.0%.

3.2 Work Contribution

Considering the criteria discussed so far, it is evident that the related works presented numerous strengths along with certain limitations, as outlined in Table 1. Considering the limitations primarily, the main original contributions of this work can be summarized as follows:

- ❏ A novel architecture is proposed that integrates both temporal and spatial features, enhancing the model's ability to capture complex patterns. Unlike most studies that use standalone models with spatial or temporal data, this approach leverages both data types to improve performance.

- ❏ The results obtained are equivalent or superior (in most cases analyzed) to the state-of-the-art methods for the dataset evaluated.

- ❏ The F1-score is used as the primary evaluation metric, ensuring high precision and high recall in performance assessment. This metric is important because it balances the trade-off between precision and recall, providing a more comprehensive evaluation of model performance, especially in cases where data imbalance are critical.

- ❏ The combination of temporal and spatial features demonstrates the potential to achieve the best results, as evidenced by the superior performance metrics.

Tabela 1 – Summary of Related Works

Work	**Spatial Feature**	**Temporal Feature**	**Used Metric**	**Result**
(DARABI; MALDAGUE, 2002)	Yes	No	Pixel-wise Accuracy	96.8%
(BENITEZ et al., 2006)	Yes	No	Accuracy	94.4%
(SAEED; ZARKANI; OMAR, 2019)	Yes	No	Accuracy	92.0%
(LUO et al., 2019)	Yes	Yes	Average POD value	66.7%
(WANG et al., 2020)	No	Yes	Accuracy	95.0%
(FANG; MALDAGUE, 2020)	Yes	No	Accuracy	95.0%
(XU; HU, 2021)	Yes	No	Recall	95.0%
(FANG; IBARRA-CASTANEDO; MALDAGUE, 2021)	Yes	No	Precision	68.0%
(LEMA et al., 2022)	Yes	No	F1-score	96.0%
This Work	**Yes**	**Yes**	**F1-score**	**96.29**%

Chapter 4

Methodology

This chapter describes the methodologies employed in this research to achieve the objectives outlined in previous chapters. It begins by detailing the experimental setup, including the data acquisition process, the equipments and setup under which the experiments were conducted. This is followed by a thorough explanation of the data annotation process. The proposed method combining DeepLabv3 and BiLSTM is described. In addition, the training process is described, focusing on the algorithms and models used to process the data in the context of defect segmentation using IRT. Finally, the validation and evaluation strategies implemented to assess the effectiveness and accuracy of the proposed methodologies are discussed. The methodology used in this study is also illustrated in Figure 10, providing a clear overview of the process and the key steps involved. This chapter serves as a bridge between the theoretical concepts introduced earlier and the practical implementation discussed in subsequent chapters.

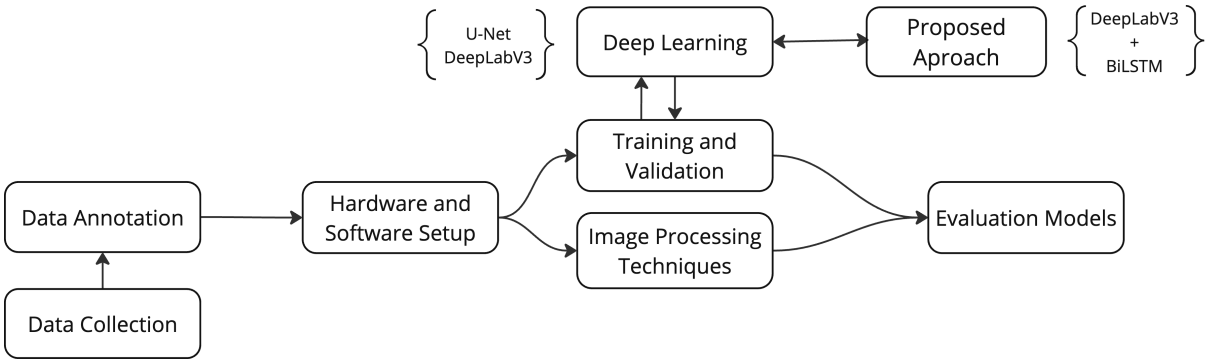


Figura 10 – Research Methodology. Source: Own work.

4.1 Data Collection

Data acquisition is a crucial component of any data project. In this research, a dataset composed of CFRP laminates examined using active IRT was utilized. The dataset was created in 2016 by (FERNANDES et al., 2016). The techniques employed, including

PT, were instrumental in gathering the data. This section outlines the procedures and equipment used to collect the data necessary for analysis.

In order to test the proposed approaches, a PT experiment was conducted for this work. A flat of CFRP laminate, with nine artificial inserts manufactured with carbon / PEEK (polyether ether ketone) APC-2 / AS4, which contained a volume fraction of 61% fiber with a $[0_2/90_2]_6$ setup.

In Table 2, the thermal properties of APC-2/AS4 are listed. Moreover, the positions of the defects inserted in the sample are shown in Figure 11. The nominal depths of the defects are 0.13 mm for D1, 0.26 mm for D2, and 0.39 mm for D3. Artificial Kapton tape inserts, with nominal sizes of 2 × 2 mm, 3 × 3 mm, and 4 × 4 mm, were placed at these specific depths (D1, D2, and D3) and in different layers during the molding process.

Tabela 2 – Thermal properties of APC-2/AS4 (AGEORGES et al., 1998).

	Property		Value	
k:	Thermal conductivity	Longitudinal	5.65	[W/mK]
		Transverse	0.35	[W/mK]
C_p:	Specific heat$^{(a)}$		1310	[J/kgK]
ρ:	Density		1584	[kg/m^3]

$^{(a)}$*at constant pressure.*

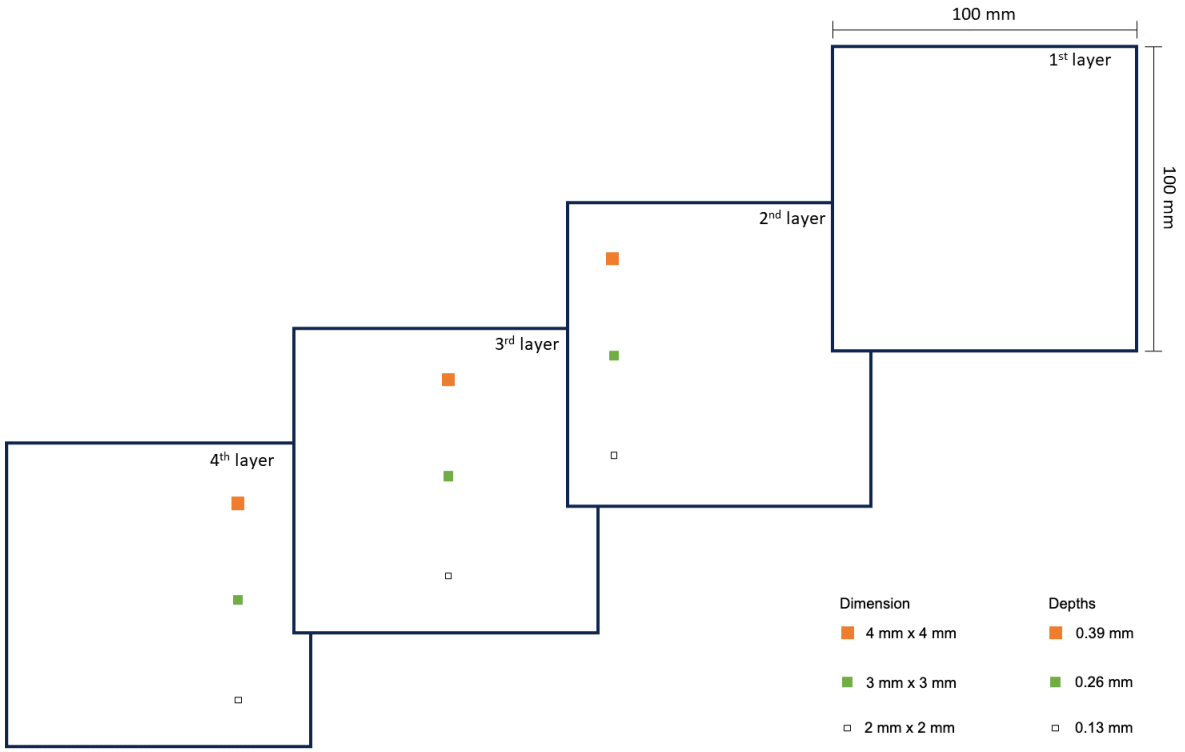


Figura 11 – Position of the artificial inserts on the layers of the laminate. Source: Adapted from (FERNANDES et al., 2016).

The experimental setup utilized the pulsed thermography technique, where the specimen's surface is subjected to a heat pulse using a high-power source, such as photographic

flashes. A heat pulse can be thought of as the combination of several periodic waves at different frequencies and amplitudes. After the thermal front comes into contact with the specimen's surface, it propagates from the surface through the specimen. As time elapses, the surface temperature decreases uniformly for a defect-free region (e.g., a sound area) (SHEPARD, 2001; BALAGEAS, 2012).

On the other hand, subsurface discontinuities (e.g., porosity, delaminations, disbonds, fiber breakage, and inclusions) act as barriers to heat flow, producing abnormal temperature patterns on the surface of the sample. This abnormal temperature distribution can be detected using an infrared (IR) camera.

The one-dimensional (1D) solution of the Fourier equation for the propagation of a Dirac heat pulse an ideal waveform defined as an intense unit-area pulse of such brief duration that no measuring equipment can distinguish it from even shorter pulses (BRACEWELL; KAHN, 1966) in a semi-infinite isotropic solid by conduction is given by:

$$T(z,t) = T_0 + \frac{Q}{\sqrt{k\rho C_p \pi t}} \exp\left(-\frac{z^2}{4\alpha t}\right), \tag{30}$$

where:

- Q is the energy absorbed by the surface ($\mathrm{J/m^2}$),
- T_0 is the initial temperature (K),
- α is the thermal diffusivity ($\mathrm{m^2/s}$),
- t is the time (s),
- k is the thermal conductivity ($\mathrm{W\ m^{-1}K^{-1}}$),
- C_p is the specific heat ($\mathrm{J\ kg^{-1}K^{-1}}$),
- ρ is the density ($\mathrm{kg/m^3}$), and
- z is the depth (m).

A Dirac heat pulse consists of periodic waves at all frequencies and amplitudes. Although reproducing such a waveform in practice is not feasible, a heat pulse from a powerful source, such as a photographic flash with an approximately square shape can be used. In this case, the signal consists of periodic waves at several (but not all) frequencies. The shorter the pulse, the broader the range of frequencies.

The configuration used for the experiment is shown in Figure 12. The camera used, indicated by number 1 in Figure 12, was a FLIR Phoenix, InSb, 3–5 µm, with a resolution of 640 × 512 pixels. The camera frame rate was 55 Hz, with a 5 ms photographic pulse, producing 6.4 kJ of energy per flash for sample heating. The recording was done in reflection mode. The camera was placed approximately 1.5 m from the sample, indicated

by number 3 in Figure 12, and the flashes, indicated by number 2, were placed about 1.0 m from the sample.

Figura 12 – Laboratory setup at Laval University. Source: Adapted from (FERNANDES et al., 2016).

The images were acquired for 25 seconds, resulting in a sequence containing 1385 images comprised sequences of thermograms that recorded temperature variations across the sample surface over time. These thermograms were stored as 2D matrices, with spatial coordinates (x, y) and temporal data (t) representing the thermal response.

The raw thermal data collected from the MWIR camera were processed using PCA and Pulsed Phase Thermography (PPT). PCA decomposed the data into orthogonal statistical modes, enhancing defect visibility by isolating significant thermal variations. PPT transformed the data into the frequency domain, enabling quantitative analysis of defect depths based on phase contrast.

Figure 13 illustrates the temporal evolution of a surface defect captured at different time steps, showing variations in its appearance over time. Analyzing such temporal changes is important for defect segmentation, as it enables early detection, tracks progression, and improves the segmentation models.

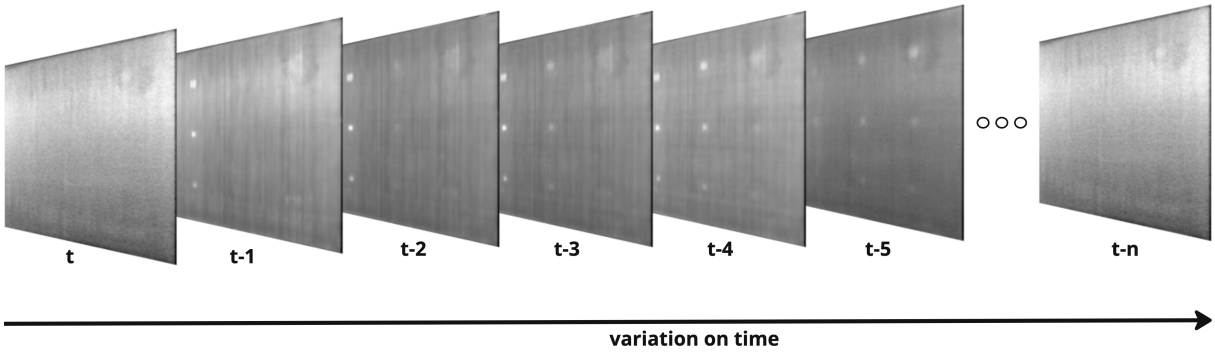


Figura 13 – Temporal Evolution of Surface Defects. Source: Own work.

4.2 Data Annotation

For this study, all images described in Section 4.1 were annotated by a student and validated by an expert to create the dataset used to train the model. The annotation process was done using the VGG Image Annotator, a web-based annotation tool developed by the Visual Geometry Group at the University of Oxford (DUTTA; GUPTA; ZISSERMANN, 2016). This tool was chosen for its simplicity, ease of use, and flexibility in handling image annotation tasks.

The annotation process involved four key steps:

1. **Image Upload and Preparation:** All thermal images were first uploaded to the tool. The images were pre-processed to ensure good resolution and format across the dataset, enabling uniformity in the annotation.

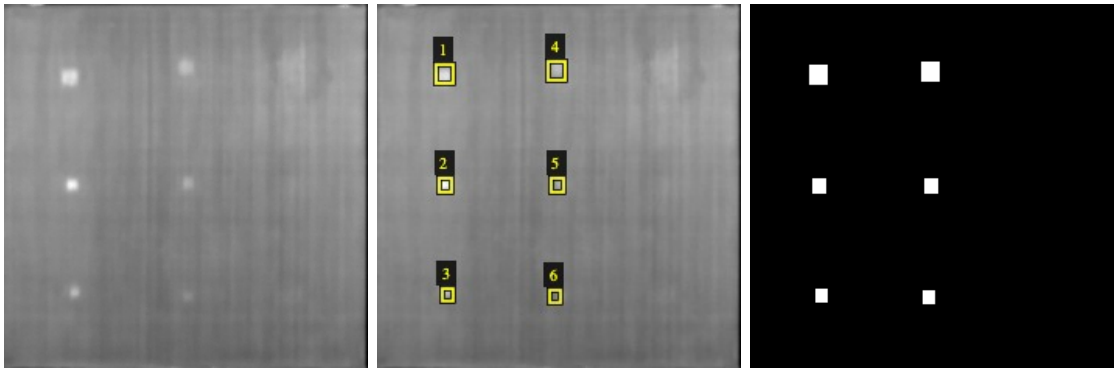
2. **Region Annotation:** For each image, Region of Interest (ROI) were manually annotated. These regions corresponded to specific defects within the CFRP materials.

3. **Labeling:** Each annotated region was assigned a label corresponding to the defect.

4. **Quality Control:** After the initial round of annotations, a quality control process was implemented. This involved cross-validation by multiple annotators to ensure the consistency and accuracy of the annotations. Discrepancies were resolved through discussion, and annotations were revised as necessary to maintain a high standard of data integrity.

Once the annotation process was completed, the annotated regions were compiled to create the ground truth dataset. This ground truth serves as a reference for training and validating the results of the models created in this study. The annotated regions and the corresponding labels were exported from the tool in JSON format, which provided a structured and easily interpretable representation of the data. Figure 14 presents a

comparison of the original image, the annotated image, and the corresponding ground truth representation.



Original Image Annotated Image Ground Truth Image

Figura 14 – Comparison of Annotated Image and Corresponding Ground Truth Representation. Source: Own work.

4.3 Proposed Method (DeepLabv3+BiLSTM)

Figure 15 presents the model proposed to solve the problems described in this work. This model addresses the challenges of accurately segmenting CFRP materials using IRT data. This method introduces a hybrid architecture that combines convolutional feature extraction with temporal sequence modeling. By leveraging temporal information in sequential thermal images and extracting spatial features at multiple scales, the model produces segmentation maps to identify defects. The architecture processes input data in three main steps: **feature extraction**, **temporal modeling**, and **decoding**, which collectively ensure segmentation performance.

4.3.1 Architecture

As shown in Figure 15, the proposed architecture comprises three main components: the input representation, a feature extraction module that incorporates temporal modeling through a combination of DeepLabv3 and BiLSTM, and the segmentation output.

4.3.1.1 Model Input

The model takes as input a sequence of thermal frames:

$$X = \{X_t, X_{t-1}, \ldots, X_{t-T+1}\}, \tag{31}$$

where $X_t \in \mathbb{R}^{H \times W}$ represents the thermal image captured at time t, with height H and width W. The sequence length T captures temporal variations in the thermal response of the material. These frames represent the evolution of thermal patterns that

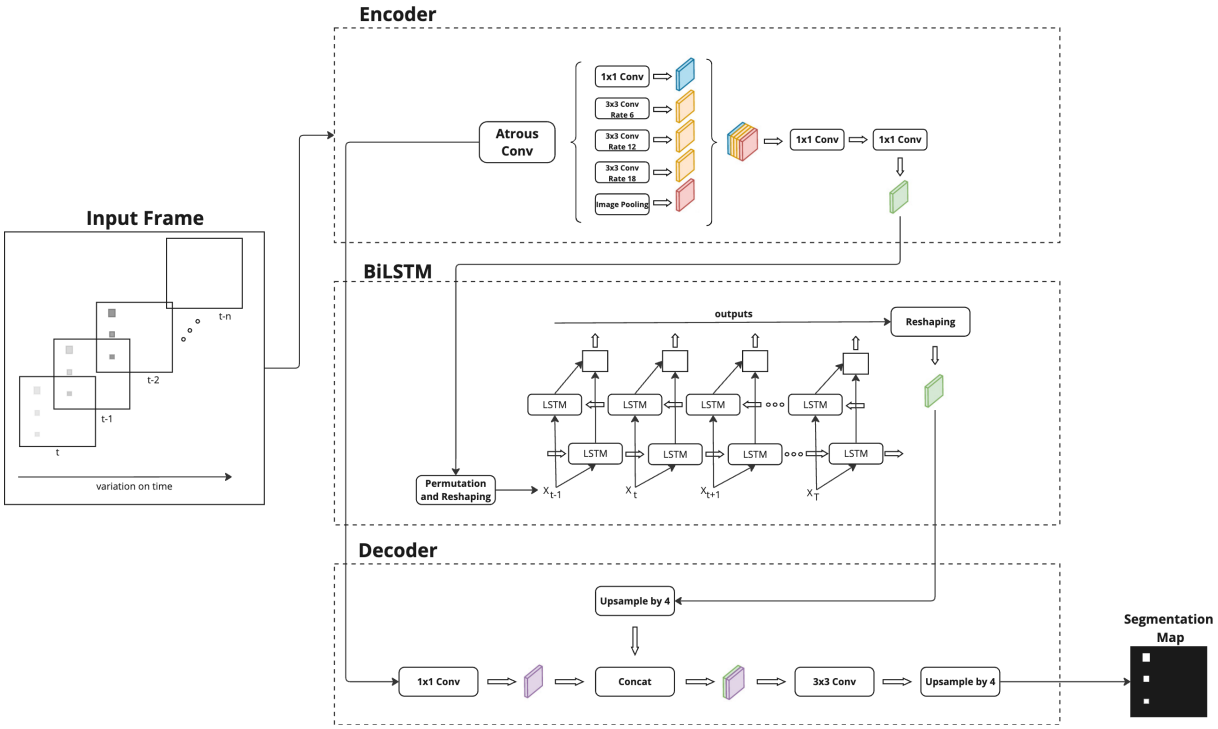


Figura 15 – Overview of the proposed method. Source: Own work.

reveal potential defects in the CFRP. The input is normalized and passed through a preprocessing layer to ensure consistency across samples.

4.3.1.2 Feature Extraction

The feature extraction process is carried out by a hybrid DeepLabv3-BiLSTM module. The encoder utilizes a combination of atrous convolutions with different dilation rates to capture spatial features at multiple scales. These convolutions ensure that the model can extract both fine-grained details and broader contextual features, which are critical for detecting small and large defects. The resulting feature maps are transformed into a compact latent representation, denoted as

$$F \in \mathbb{R}^{T \times C \times H' \times W'}, \tag{32}$$

where C is the number of channels and H', W' are the reduced spatial dimensions.

The temporal relationships among frames are modeled using a BiLSTM network. Let F_t represent the feature vector in time t. The BiLSTM processes the sequence $\{F_t\}$ both forward and backward, generating an output H_t for each time step:

$$H_t = \text{BiLSTM}(F_t) = \overrightarrow{h_t} + \overleftarrow{h_t}, \tag{33}$$

where $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are the forward and backward hidden states, respectively. The BiLSTM enables the model to capture temporal dependencies across the input sequence, enhancing its ability to detect time-dependent thermal variations.

4.3.1.3 Model Output

The features from the BiLSTM are passed to the decoder, which reconstructs a segmentation map $Y \in \mathbb{R}^{H \times W}$. The decoder combines upsampling and convolutional operations to produce high-resolution outputs aligned with the input resolution. The segmentation map represents areas of interest that are defects, by assigning a probability score to each pixel, which represents the likelihood of belonging to a defective region.

Mathematically, the segmentation map is defined as:

$$Y_{i,j} = \sigma(f_{\text{decoder}}(H)), \tag{34}$$

where σ is the softmax activation function and f_{decoder} represents the series of operations in the decoder. The resulting map is thresholded to create a binary segmentation output, which delineates the boundaries of the defect.

4.3.2 Benefits of Combining Temporal and Spatial Information

The proposed architecture leverages both temporal and spatial information, offering advantages for the segmentation of CFRP materials using IRT data. By integrating these two dimensions, the model addresses the limitations of approaches that rely solely on either spatial features or temporal sequences.

Spatial information, extracted through the encoder's multi-scale convolutional layers, captures both fine-grained details and larger contextual patterns. This is particularly important for identifying defects in CFRP materials, which can vary in size and shape, ranging from small localized cracks to broader irregularities. The spatial feature extraction process enables the model to localize and show these patterns within individual thermal frames effectively.

Temporal information, generated by the BiLSTM network, captures dynamic changes in thermal signals over time. In IRT data, defects often manifest as temporal variations in heat dissipation, which may not be visible in a single frame. By processing a sequence of frames, the BiLSTM learns to identify these subtle temporal patterns, allowing the model to detect defects that would otherwise be overlooked through spatial analysis alone.

Combining temporal and spatial information provides a more holistic understanding of the input data. This dual approach enables the model to detect defects of various sizes and shapes across the frames, producing robust segmentation results even in the presence of noise in the input data. This combination of spatial and temporal dimensions ensures that the proposed architecture achieves high accuracy in CFRP defect analysis.

4.4 Hardware and software setup

Using Google Colaboratory, a cloud-based platform that offers access to powerful computational resources, the models were trained and validated. In particular, the deep learning models that were used in the tests allowed for efficient processing and shorter training periods using the GPU accelerator option that was available in Colaboratory.

Python was used to create and train the models, making use of well-known DL packages like TensorFlow and Keras. TensorFlow was chosen because of its strong support for DL operations and GPU resource compliance with Google Colaboratory. With its high-level API, Keras made it easier to quickly prototype and experiment with the various NN topologies that were previously discussed.

The software environment included additional libraries such as NumPy and Pandas for data manipulation, Matplotlib for visualization, and Scikit-learn for preprocessing and evaluation metrics. The setup provided an efficient environment for model development, experimentation, and validation, ensuring that the computational demands of the models were adequately met by the available hardware in Google colaboratory.

The hardware resources provided by Colaboratory included access to an NVIDIA GPU; specifically, a Tesla T4 was used for these experiments. This GPU offer a substantial amount of computational power, with GPU RAM of 16 GB. This GPU memory was crucial for handling large-scale matrix operations required by the DL models, especially during backpropagation and when processing large batches of data.

4.5 Training and validation

Figure 16 describes the steps required to train and validate the architecture presented in Section 4.3. The first step is preprocessing the dataset to prepare them for training and inference. The dataset were divided into training and validation sets using 70% for training, 20% for validation, and 10% for the final test. This split was chosen to ensure that the models could be trained on a sufficiently large portion of the data while maintaining a representative subset for validation and test.

Several tests were run to compare different model architectures. First, we put into practice a U-Net model, which is a well-established architecture for image segmentation applications. Next, we evaluated the DeepLabv3 model, which is well-known for its robust semantic segmentation performance. We also performed additional experiments in our proposed method presented in Section 4.3.

The performance of each model was evaluated on the validation and test set, with metrics such as Intersection over Union (IoU), and F1-score being recorded for comparative analysis. The results of these experiments are detailed in Section 5.

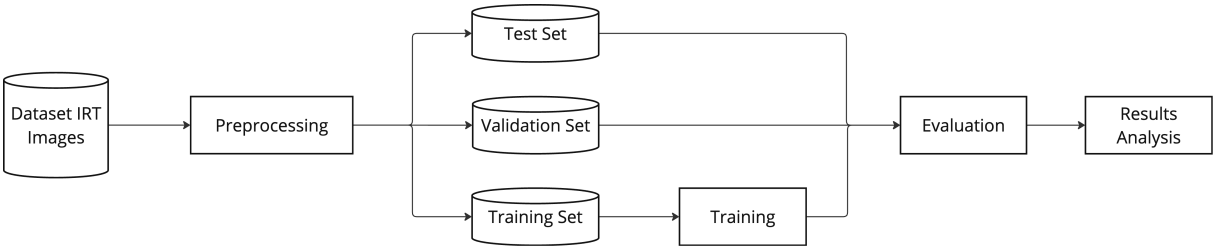


Figura 16 – Overview of the training and evaluation processes. Source: Own work.

4.5.1 Preprocessing

To ensure consistency and improve the performance of the proposed model, a preprocessing step is applied to the input data. Each thermal image and its corresponding mask undergo a series of transformations to prepare them for training and inference.

The input images are first converted to grayscale format, ensuring a standardized representation of the color channels. Masks, representing the ground truth segmentation, are converted to grayscale to simplify the segmentation task. Both images and masks are resized to a fixed resolution (H, W) to maintain uniform input dimensions across the dataset.

For images, transformations include normalization to standardize pixel intensities and conversion to tensors to make the data compatible with the deep learning framework. Specifically, the pixel values are normalized with a mean of $[0.485, 0.456, 0.406]$ and a standard deviation of $[0.229, 0.224, 0.225]$, which are commonly used for pre-trained models. Masks are also resized and converted to tensors, with an additional channel dimension added to ensure compatibility with the network's input format.

4.5.2 Training Parameters

The selection of training parameters is crucial to achieving optimal model performance. The Adam optimizer was chosen because of its adaptive learning rate mechanism and computational efficiency. Adam is well-suited for segmentation tasks as it combines the advantages of RMSProp and momentum optimization. By adjusting learning rates for each parameter, it ensures faster convergence while being robust to sparse gradients. The learning rate was set to 1×10^{-4}, balancing convergence speed and stability.

The `BCEWithLogitsLoss` was selected as the criterion to train the model. This loss function is ideal for binary segmentation tasks, as it combines the binary cross-entropy loss with a sigmoid activation. It directly handles unnormalized outputs (logits) and computes the loss in a numerically stable way.

The training process was initially configured to run for 60 epochs; however, early stopping was applied based on the improvement of the IoU metric across epochs to prevent

overfitting. Monitoring these metrics during training ensured that the model's performance consistently improved without excessive training.

A batch size of 4 was selected considering the hardware limitations and the need for effective gradient updates. Smaller batch sizes often lead to noisier gradient estimates but allow for greater hardware utilization in memory-constrained environments.

To ensure that the model trains efficiently, the data loaders were configured with shuffle for the training dataset. Shuffling improves generalization by preventing the model from learning spurious patterns in data order. Additionally, 2 works were used to parallelize data loading, speeding up the training process by preloading data during GPU computations.

4.5.3 Evaluation metrics

For this work, F1-score and IoU were used as evaluation metrics. These metrics provide a comprehensive assessment of the model's performance in terms of accuracy, precision, and recall, particularly in scenarios involving imbalanced datasets and object detection tasks.

The F1-score is a harmonic mean of precision and recall, balancing the two metrics to provide a single performance indicator. Precision, defined in Equation (35), measures the proportion of true positive predictions among all positive predictions made by the model:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{35}$$

where TP represents true positives and FP denotes false positives.

Recall, defined in Equation (36), evaluates the proportion of true positive predictions out of all actual positive instances:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{36}$$

where FN represents false negatives.

The F1-score incorporates both precision and recall into a single metric, making it especially useful when classes are imbalanced (OBI, 2023). It is calculated as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{37}$$

IoU is a common metric for evaluating the performance of segmentation tasks. It quantifies the overlap between the predicted and actual regions of interest (RAINIO; TEUHO; KLÉN, 2024). Mathematically, IoU is defined in Equation (37):

$$IoU = \frac{\text{Intersection Area}}{\text{Union Area}} \tag{38}$$

where the numerator represents the area of intersection between the predicted and actual regions, and the denominator represents the total area covered by both regions.

The closer the IoU is to 1, the more accurate the model's detection. Conversely, an IoU close to 0 indicates poor localization and low accuracy in predicting object locations.

Together, the F1-score and IoU provide a robust framework for evaluating model performance, capturing aspects of classification accuracy and spatial alignment to segmentation tasks.

Chapter 5

Experiments and Results

This chapter presents the experimental results, focusing on evaluating the effectiveness of defect segmentation in CFRP materials using IRT. To achieve this, a series of experiments was conducted to test various segmentation algorithms, including traditional methods such as adaptive thresholding and DL algorithms like U-Net, DeepLabv3, and the proposed method, which combines DeepLabv3 and BiLSTM, as discussed in Section 4.3.

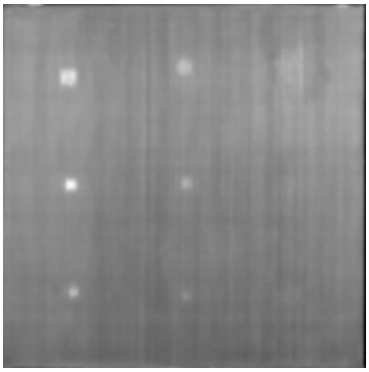
The experiments were structured into multiple phases. First, the data was obtained, as described in Sections 4.1 and 4.2. Subsequently, the training phase was conducted, as detailed in Section 4.5. During this phase, various AI models were trained and validated using the prepared dataset. Hyperparameters were adjusted, and evaluation metrics such as F1-score and IoU, as described in Section 4.5.3, were employed to compare the performance of the models. The training processes and results are presented in a structured manner, facilitating easy comparison between the algorithms.

Finally, the implications of the results are discussed, including the limitations observed during the experiments and potential improvements for future research. This chapter provides a foundation for the final discussion and conclusion, where the overall success of the research and its contributions to the field are assessed.

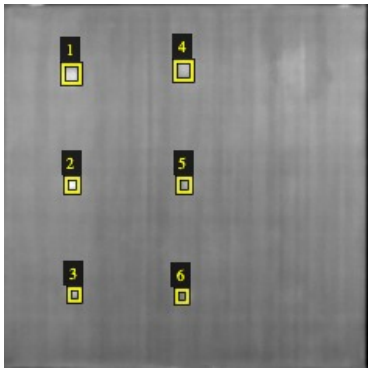
5.1 Experimental Results

In this research, some experiments were conducted to evaluate the effectiveness of the image processing technique and DL models for defect segmentation in CFRP with artificial delaminations. Four segmentation methods, as described in Chapter 4, were employed: adaptive thresholding, the U-Net model, the DeepLabv3 model, and the proposed hybrid approach integrating DeepLabv3 with a BiLSTM network. The segmentation results for a representative image, displaying six defects, are presented in Figure 17. These defects include two with dimensions of 4x4 mm, two with dimensions of 3x3 mm, and two with dimensions of 2x2 mm. Figure 17(a) shows the original raw image, extract from time t =

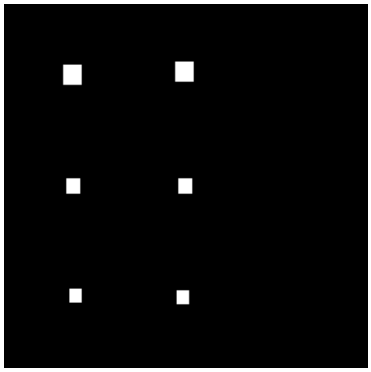
yy s from the pulsed thermography sequence, Figure 17(b) shows the manually annotated defects, Figure 17(c) shows the corresponding segmentation mask, Figure 17(d) illustrates the results of adaptive thresholding, Figure 17(e) shows the segmentation using the U-Net model, Figure 17(f) shows the results from DeepLabv3, and Figure 17(g) shows the segmentation achieved using our proposed approach combining DeepLabv3 with BiLSTM.



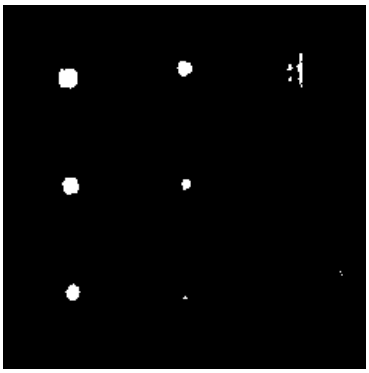
(a) Original Image



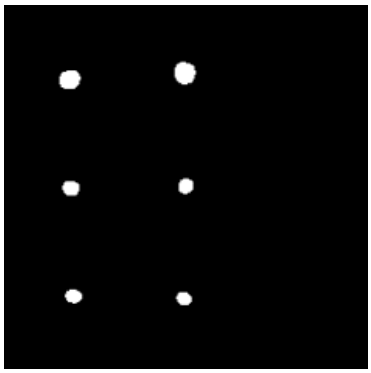
(b) Annotated Image



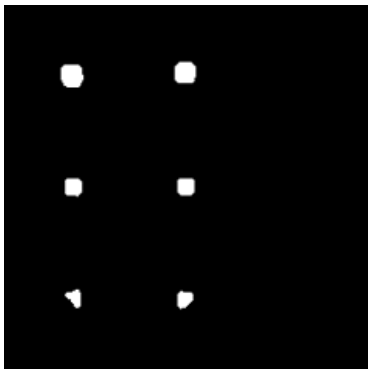
(c) Segmentation Mask



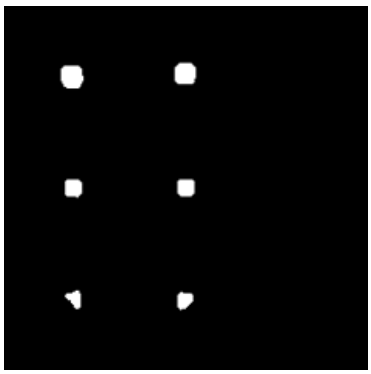
(d) Adaptive Threshold



(e) U-Net



(f) DeepLabv3



(g) Proposed Method (DeepLabv3+BiLSTM)

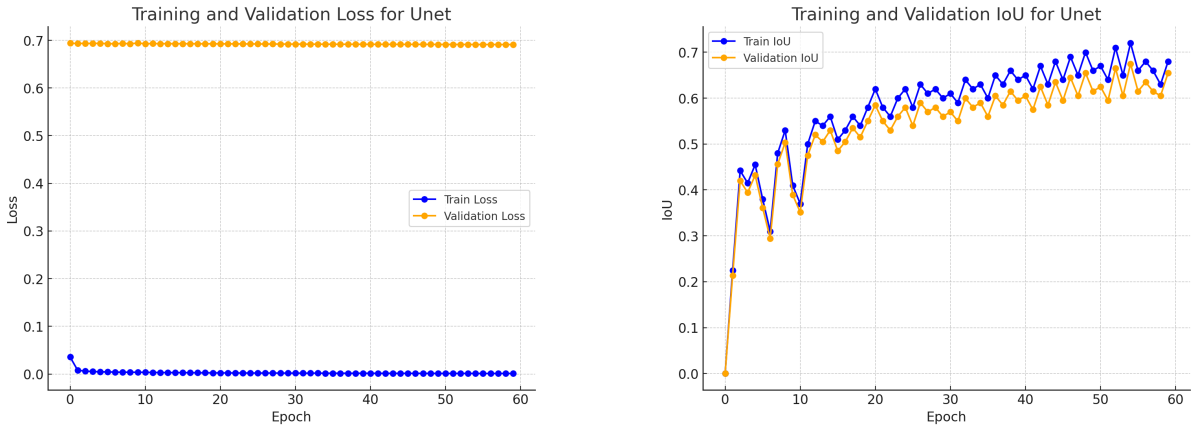
Figura 17 – Comparison of the different methods applied. Source: Own work.

The mask for segmentation shown in Figure 17(c) is a binary mask created from manual annotations of the defects. Each pixel in the mask is labeled by an expert as either a defect (value 1) or non-defect (value 0). This mask serves as a ground truth

reference for evaluating the segmentation performance of different methods. The manual annotation process involves marking the exact locations and boundaries of defects in infrared images, ensuring that the segmentation models have accurate targets for training and validation.

Initially, an adaptive thresholding method was employed to segment the infrared images, resulting in an average F1-score of 0.7519 and an IoU metric of 0.2375. While this approach was computationally efficient, the outcomes shows that its segmentation results was limited.

Subsequently, the U-Net model was trained to develop a better robust segmentation approach. The results demonstrated a modest improvement over the initial approach, with the model achieving an IoU of 0.6141 and an F1-score of 0.9142. This performance, compared to adaptive thresholding, suggests that a deep learning-based approach is more effective in capturing the complexities associated with the thermal data. Although the model outperformed in terms of IoU and F1-score, it exhibited signs of overfitting, as illustrated in Figure 18(a), where the training loss decreases rapidly and stabilizes at a very low value close to zero. This suggests that the model is performing exceptionally well on the training data, achieving minimal loss. However, the validation loss remains constant and does not decrease over epochs, staying at a relatively high value, indicating that the model is not improving on the validation data. Figure 18(b) shows the training and validation IoU over epochs. The training IoU starts low but rapidly increases within the first 10 epochs, then continues to improve more gradually and it stabilizes with some flutuations. The validation IoU also starts low, increases quickly in the first few epochs, and then follows a similar trend to the training IoU.



(a) Training and validation loss

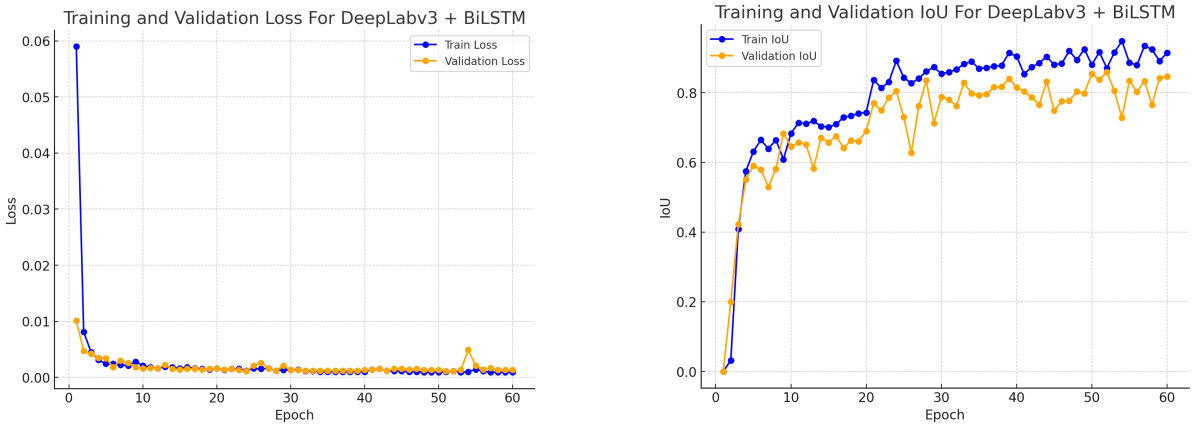
(b) Training and validation IoU

Figura 18 – Training and validation process for U-Net model. Source: Own work.

Subsequently, a DeepLabv3 neural network was implemented. This technique demonstrated a good improvement compared to previously tested approaches, achieving an F1-score of 0.9406 and an IoU of 0.7419. The high performance of DeepLabv3 shows its

ability to segment areas of interest in complex thermographic images.

Figure 19(a) displays the training and validation loss, starting relatively high and decreasing rapidly in the first few epochs. It stabilizes around a very low value close to zero, indicating that the model is learning the training data well. The validation loss follows a similar pattern to the training loss, starting high and then decreasing rapidly. It also stabilizes at a low value, almost matching the training loss. Both training and validation losses decrease very quickly within the first 10 epochs and then remain low and stable. This suggests that the model is minimizing loss and achieving good performance on both the training and validation datasets. Figure 19(b) shows the IoU for training and validation, both of which improve over time, suggesting that the model is learning and getting better at the task of image segmentation.



(a) Training and validation loss

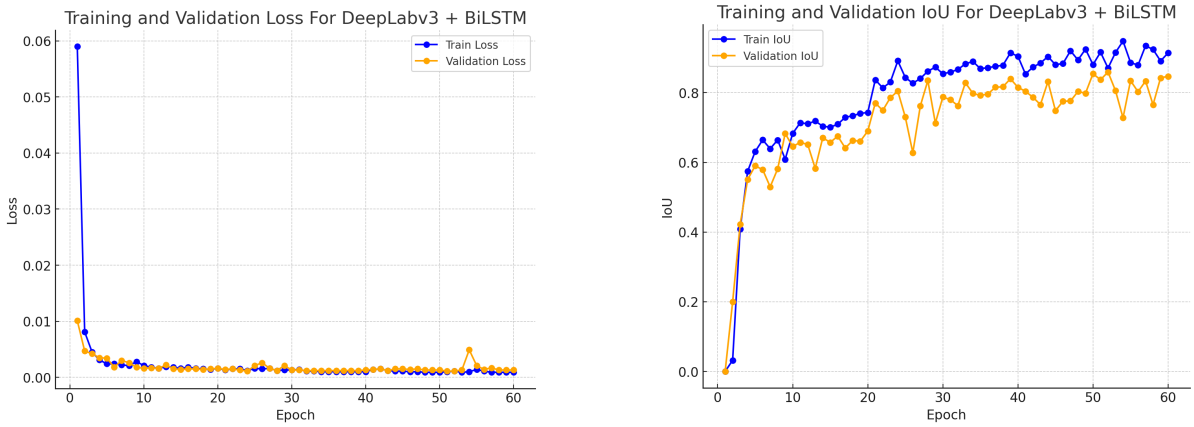
(b) Training and validation IoU

Figura 19 – Training and validation process for DeepLabv3 model. Source: Own work.

In the final phase of the experiments, the proposed approach, which integrates the DeepLabv3 neural network with a BiLSTM network, was tested. The introduction of BiLSTM is an innovative approach to segmenting this kind of material and aims to enhance the model's performance by incorporating temporal features, leading to a significant improvement in the results. Together, these techniques achieved an F1-score of 0.96298 and an IoU of 0.83123.

The performance evaluation of the DeepLabv3 model combined with a BiLSTM layer demonstrates good learning and generalization capabilities, as evidenced by Figure 20 with a convergence of training and validation IoU metrics around 0.83. The model's rapid reduction in loss during the initial epochs, with both training and validation losses stabilizing near zero, suggests effective learning.

Although the IoU and loss over epochs initially suggested satisfactory results without indications of overfitting, a 5-fold cross-validation was performed to thoroughly evaluate the generalizability of the model. Cross-validation analysis revealed that, despite initial observations indicating no overfitting, the model demonstrated limited generalization

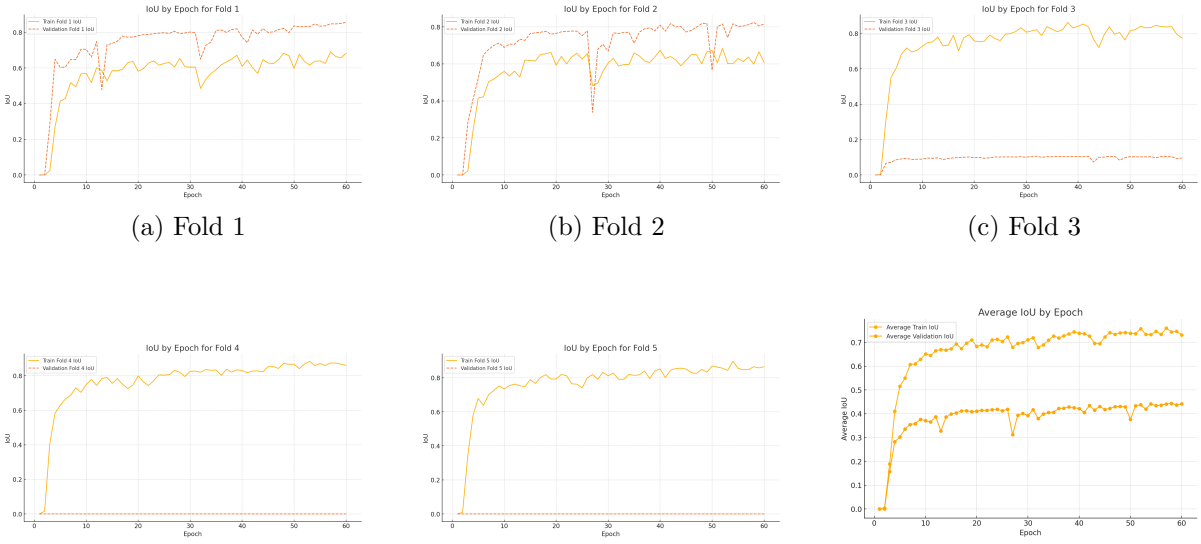


(a) Training and validation loss

(b) Training and validation IoU

Figura 20 – Training and validation process for DeepLabv3 + BiLSTM model. Source: Own work.

performance. This outcome is shown in Figure 21, which presents the IoU values for each fold and the computed average IoU for the folds in all epochs. Although the model demonstrated effective learning in Folds 1 and 2, its performance in Folds 3, 4, and 5 underscored significant challenges in the learning process, with validation IoU scores near to zero, indicating difficulty in generalizing when applied to unknown data.



(a) Fold 1

(b) Fold 2

(c) Fold 3

(d) Fold 4

(e) Fold 5

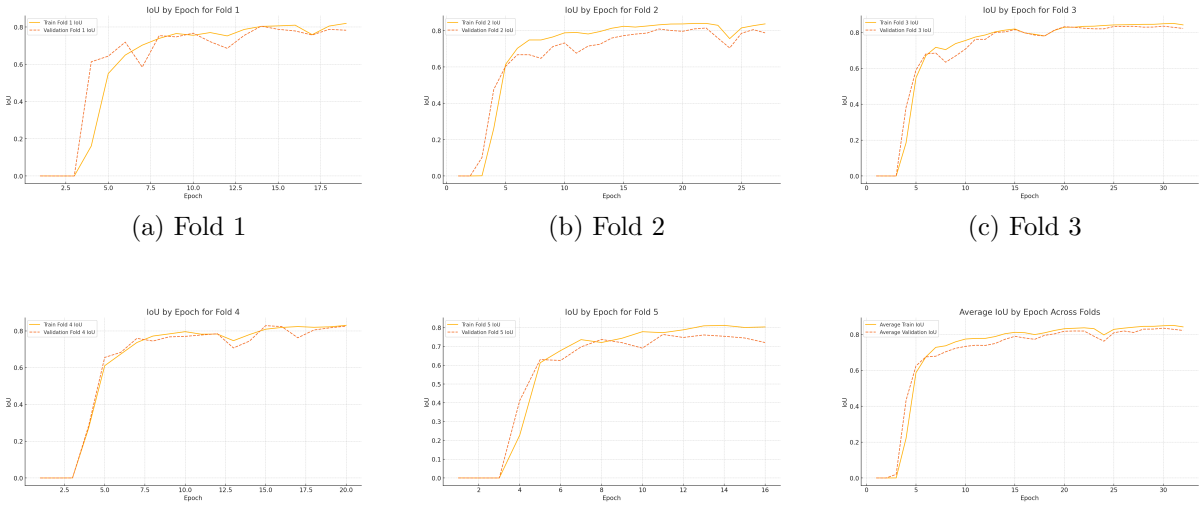
(f) Average for the 5 folds

Figura 21 – IoU for train and validation. Source: Own work.

It is possible to see in Figure 21(f) the average IoU across the 5 folds, where the validation average remains around 0.43, influenced by poor performance in folds 3, 4, and 5.

As shown, it was observed that the model lacked sufficient generalization capability, leading to the implementation of an early stopping strategy. This strategy involved stopping model iterations after three consecutive epochs without improvement in the va-

lidation IoU. In addition, dropout layers were incorporated before and after the BiLSTM layer to enhance regularization. After retraining the model using 5-fold cross-validation with these strategies in place, a significant improvement in the results was observed. Figure 22 shows the IoU for training and validation for each fold, demonstrating that the model achieved better learning performance in all training sets. This process ultimately resulted in a final average IoU of 0.83, as shown in Figure 22(f).



(a) Fold 1 (b) Fold 2 (c) Fold 3

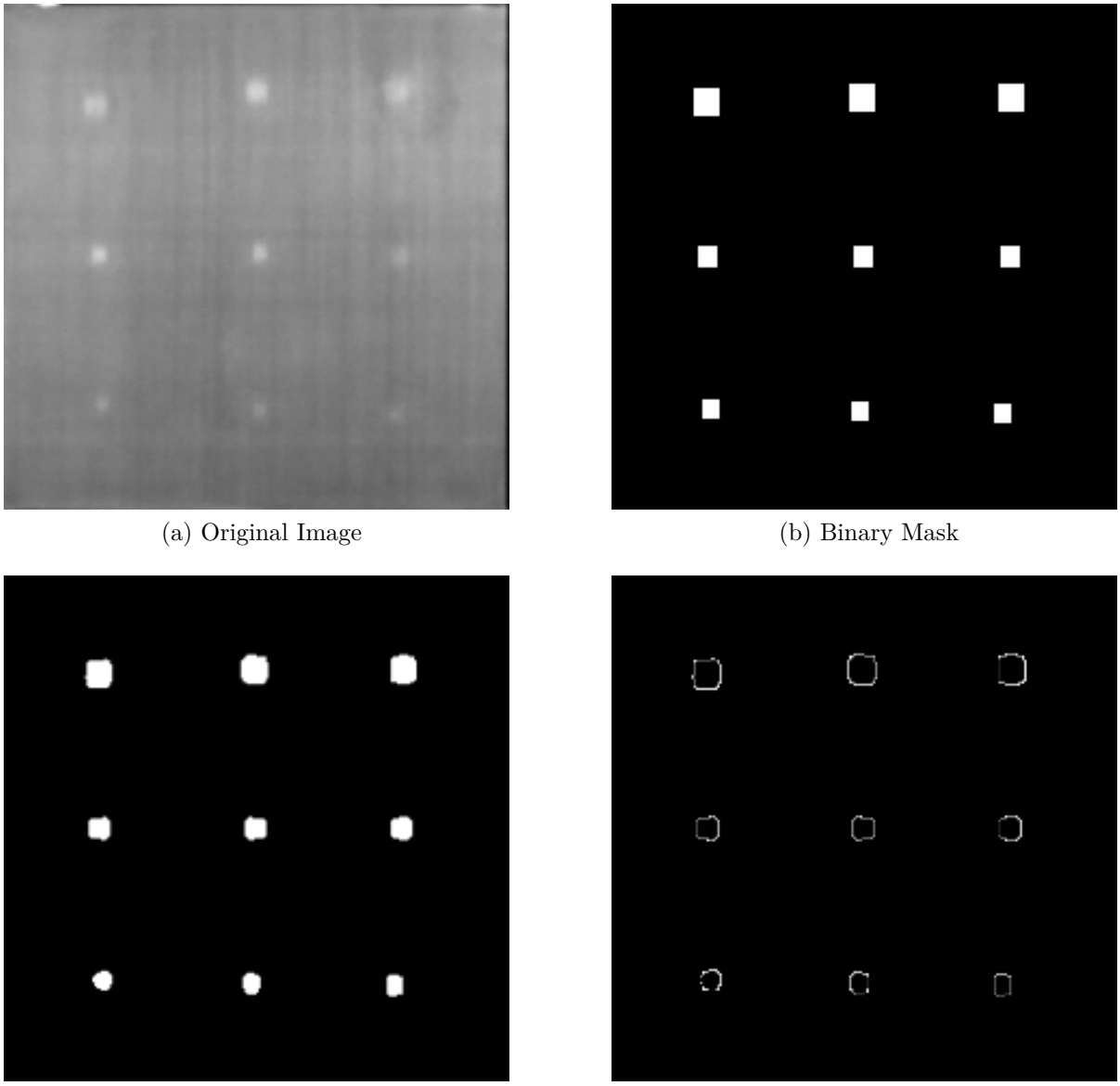
(d) Fold 4 (e) Fold 5 (f) Average for the 5 folds

Figura 22 – IoU for training and validation after the application of strategies. Source: Own work.

The performance of the proposed method after applying early stopping and dropout strategies can be seen in more detail in Figure 23. Figure 23(b) represents a mask of a CFRP sample annotated by an expert, while Figure 23(c) shows the predicted image from the developed model. Analyzing the results, it is possible to note small differences between the binary mask and the predicted image, Figure 23(d) presents the prediction differences.

These errors can be attributed to various factors, such as variations in object boundaries or the model's difficulty in capturing certain details in the image. The addition of the BiLSTM layer was designed to improve the capture of contextual information over a sequence of images, allowing the model to take into account not only spatial features but also temporal or sequential relationships between images. However, in situations where the boundary between classes is ambiguous or indistinct, it is difficult for the model to generate a perfectly accurate segmentation.

From the perspective of the IoU metric, which measures the overlap between the prediction and the ground truth, the identified error was relatively low. This indicates that overall, the model successfully segmented most of the objects in the image, with only small areas of disagreement. This performance suggests that our proposed method using Deeplabv3 with BiLSTM model holds promising potential for segmentation applications,



(a) Original Image

(b) Binary Mask

(c) Predicted by DeepLabv3+BiLSTM model

(d) Errors between mask and predicted

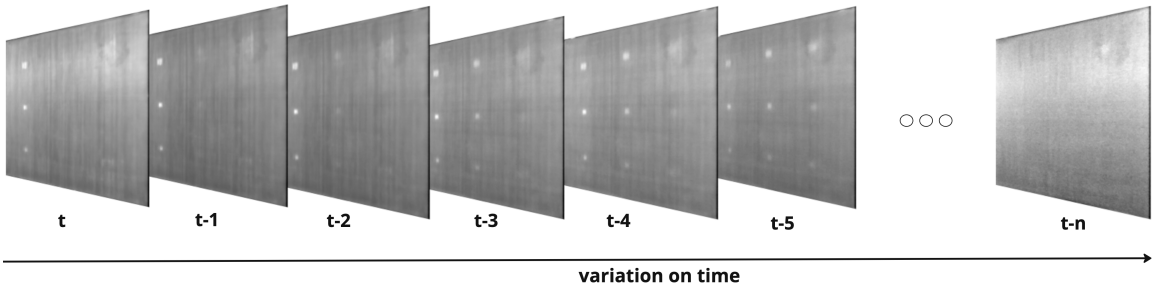
Figura 23 – Difference between binary mask and prediction. Source: Own work.

though some adjustments could be made to further minimize errors at the boundaries and improve generalization in other scenarios.

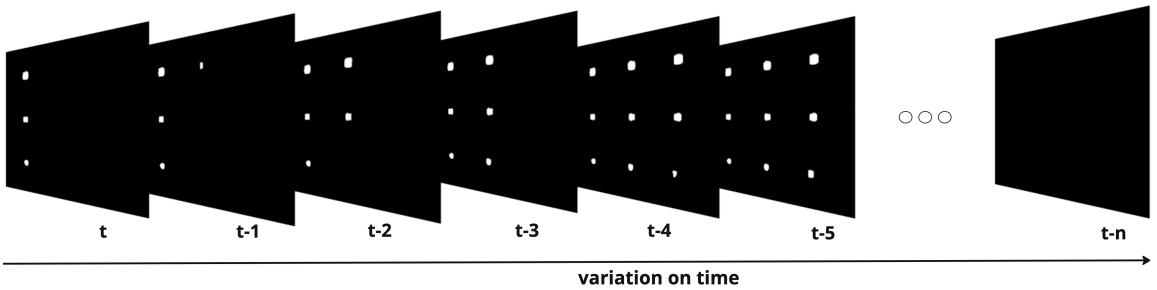
Figure 24 illustrates a sequence of frames showing variations over time. Figure 24(a) displays the sample of the original images captured over time, while Figure 24(b) shows the corresponding predicted images, highlighting detected defects in the same temporal sequence.

5.2 Results Discussion

Based on the results achieved in this work, as presented in detail in Table 3, we can conclude that the performance of the adaptive thresholding method in this study, which



(a) Original images over time.



(b) Predicted images over time.

Figura 24 – Comparison of original and predicted images over time.

achieved an F1-score of 0.7519 and an IoU of 0.2375, reflects its limitations in handling the complexities of CFRP defect detection. While adaptive thresholding offers computational simplicity, its low performance represents the need for more robust approaches. This aligns with findings from similar studies, where traditional methods such as thresholding often fall short in capturing defect boundaries due to noise and irregular thermal contrasts.

The U-Net architecture, commonly applied in defect segmentation, demonstrated significant improvements, achieving an F1-score of 0.9142 and an IoU of 0.6141. These results surpass those reported by (PEDRAYES et al., 2022), where U-Net achieved an F1-score of 0.745 and an IoU of 0.593 for step-heating thermography images. This improvement can probabily be attributed to differences in training methodologies and the custom dataset used in this study. However, the signs of overfitting in the U-Net results, as indicated by a low training loss and a stagnating validation loss, point to potential model optimization issues that could be addressed in future research.

DeepLabv3, a more advanced deep learning model, showed further performance gains, with an F1-score of 0.9406 and an IoU of 0.7419. These metrics outperform the results from (PEDRAYES et al., 2022), where DeepLabv3 achieved an F1-score of 0.776 and an IoU of 0.629. This suggests that DeepLabv3 is better suited for defect segmentation in thermographic data due to its ability to handle complex image features, such as variations in thermal contrasts and noise. The model's training process exhibited good convergence between training and validation losses, indicating effective generalization without overfitting.

The introduction of the BiLSTM layer in conjunction with DeepLabv3 brought about

further improvements, with the model achieving an F1-score of 0.96298 and an IoU of 0.83123. The proposed method combining a CNN with a RNN architecture enabled the model to capture both spatial and temporal features, significantly improving segmentation accuracy.

Tabela 3 – Comparison of Results for Different Models

Model	**F1-score**	**IoU**	**Notes**
Adaptive Thresholding	0.7519	0.2375	Low performance; struggles with complex defect detection.
U-Net	0.9142	0.6141	Improved performance; slight overfitting observed.
DeepLabv3	0.9406	0.7419	Significant improvement; good generalization without overfitting.
Our proposed method	0.96298	0.83123	Best performance; effectively captures spatial and temporal features.

The results suggest that integrating temporal features through a BiLSTM layer can significantly improve segmentation accuracy in thermographic datasets. The success of this approach demonstrates the potential for further research in hybrid architectures that combine CNNs with RNNs or other advanced models. While the proposed approach performed well compared to similar studies, there remains room for improvement, particularly in addressing optimization hyperparameters to further enhance IoU scores.

Chapter 6

Conclusions and Future Works

This chapter provides an overall summary of the work conducted throughout this dissertation, revisits the objectives outlined at the start of the research, and discusses how these objectives were met. We also review the main contributions of this study, propose potential directions for future work, and outline the bibliographic production generated from the research.

6.1 Conclusions

In order to enhance the reliability and efficiency of defect detection in CFRP materials, advanced segmentation techniques using IRT play a crucial role, enabling the identification and characterization of structural anomalies. One of the main challenges in defect segmentation is accurately determinate between true defects and noise in thermographic data, especially when defects exhibit irregular shapes or faint thermal contrasts.

Traditional algorithms for defect detection in CFRP materials face certain limitations, primarily because these techniques do not leverage temporal information, making it challenging to achieve high-confidence defect identification. These limitations underscore the need for an architecture that can effectively integrate both spatial and temporal features to improve segmentation accuracy in thermographic data. To address this issue, this work proposes a hybrid approach that combines the spatial feature extraction capabilities of CNN with the temporal features of BiLSTM networks.

Thus, all of the objectives listed in Section 1.2 were achieved through the steps listed in this document. The techniques and concepts presented in Chapter 2 have been studied in order to elaborate the final architecture. The state-of-the-art solutions for defect segmentation in CFRP were also examined, focusing mainly on DL techniques, as presented in Chapter 3. Finally, a novel hybrid model that combines the spatial features of DeepLabv3 with the temporal features generated by a BiLSTM network was developed using the methodology presented in Chapter 4.

The results obtained for the proposed solution architecture, presented in Chapter 5, demonstrate its superiority over other state-of-the-art methods, as shown in Table 1. The hybrid model achieved an F1-score of 0.96298 and an IoU of 0.83123, as detailed in Table 3, highlighting the effectiveness of combining spatial and temporal features for defect detection in CFRP materials. Additionally, the findings discussed in Section 5.1 validate the model's generalization capability through cross-validation across the dataset. Strategies such as early stopping and dropout layers further enhanced the model's performance by mitigating overfitting and improving accuracy.

Despite the promising outcomes, several challenges were encountered during the research process. The high noise levels and variability inherent in IRT data, often influenced by environmental conditions and equipment sensitivity, posed difficulties in achieving consistent and reliable detection, particularly for subtle defects. Furthermore, the limited availability of annotated datasets for training DL models constrained the model's capacity to learn complex patterns effectively.

In conclusion, this study contributes to a novel hybrid approach to defect segmentation in CFRP materials, integrating spatial and temporal features to achieve state-of-the-art performance. While challenges related to IRT data variability and dataset scarcity remain, the findings pave the way for further advancements in defect detection techniques, including the exploration of domain adaptation and synthetic data generation to address data limitations. This work not only highlights the potential of DL in industrial applications but also establishes a robust foundation for future research in this domain.

6.2 Future Work

While the proposed method achieved promising results, there are several areas for improvement:

- **Addressing overfitting in U-Net:** One potential shortcoming observed was the overfitting behavior in the U-Net model, where the validation loss stagnated despite a low training loss. Future work could explore regularization techniques such as dropout or data augmentation to improve the generalization of the model.

- **Further optimization of the hybrid model:** Although the DeepLabv3+BiLSTM model performed well, there is still a place for optimization. Fine-tuning hyperparameters, experimenting with different recurrent layers (e.g., GRU), or exploring attention mechanisms could further improve segmentation accuracy.

- **Expanding the dataset:** The current dataset used in this research, while effective, could be expanded to include more diverse types of CFRP defects. A broader dataset would improve the robustness and generalizability of the model for industrial applications.

- ❑ **Real-world implementation:** Future research could focus on deploying the proposed segmentation model in real-time industrial environments, exploring how it performs with live IRT data in non-laboratory conditions. This would provide valuable insights into the practical applicability of the method.

Furthermore, our work opens up new directions for research in hybrid architectures for defect detection. The projects could build on this by combining different DL architectures or even exploring multimodal approaches that integrate IRT with other nondestructive evaluation methods.

6.3 Contributions in Bibliographic Production

This research has contributed to the academic community through the following bibliographic productions:

- ❑ **Conference presentation:** A paper detailing the performance of different segmentation methods for IRT defect detection using IRT, presented at the *XIX Workshop de Visão Computacional*[1].
- ❑ **Journal publications:**
  - A manuscript was accepted for publication in *RITA - Revista de Informática Teórica e Aplicada* (Qualis B3). (VARGAS; FERNANDES, 2024a)
  - A manuscript describing the development and validation of the hybrid DeepLabv3 + BiLSTM model was accepted for publication in *Nondestructive Testing and Evaluation* (Qualis A2). (VARGAS; FERNANDES, 2024b)
- ❑ **Dataset publication:** The annotated dataset used in this research, along with the ground truth segmentation masks, will be publicly available, allowing other researchers to build upon our work.

[1] This paper was invited for publication in a special issue of the Revista de Informática Teórica e Aplicada (RITA).

References

AGEORGES, C. et al. Characteristics of resistance welding of lap shear coupons. part i: Heat transfer. **Composites Part A: Applied Science and Manufacturing**, Elsevier, v. 29, n. 8, p. 899–909, 1998. Disponível em: <https://doi.org/10.1016/S1359-835X(98)00022-0>.

ALSAMHI, S. H. et al. Green internet of things using uavs in b5g networks: A review of applications and strategies. **Ad Hoc Networks**, Elsevier, v. 117, p. 102505, 2021. Disponível em: <https://doi.org/10.1016/j.adhoc.2021.102505>.

ANBUDEVI, M. K. A.; SUGANTHI, K. Review of semantic segmentation of medical images using modified architectures of unet. **Diagnostics**, MDPI, v. 12, n. 12, p. 3064, 2022. Disponível em: <https://doi.org/10.3390/diagnostics12123064>.

BALAGEAS, D. L. Defense and illustration of time-resolved pulsed thermography for nde. **Quantitative InfraRed Thermography Journal**, Taylor & Francis, v. 9, n. 1, p. 3–32, 2012. Disponível em: <https://doi.org/10.1080/17686733.2012.676902>.

BENITEZ, H. et al. Defect quantification with thermographic signal reconstruction and artificial neural networks. In: **Proceedings of 8th conference on quantitative infrared thermography, Padova, Italy**. [s.n.], 2006. v. 2, p. 6. Disponível em: <https://doi.org/10.21611/qirt.2006.010>.

BRACEWELL, R.; KAHN, P. B. The fourier transform and its applications. **American Journal of Physics**, American Association of Physics Teachers, v. 34, n. 8, p. 712–712, 1966. Disponível em: <https://doi.org/10.1119/1.1973431>.

BRADLEY, D.; ROTH, G. Adaptive thresholding using the integral image. **Journal of Graphics Tools**, Taylor & Francis, v. 12, n. 2, p. 13–21, 2007. Disponível em: <https://doi.org/10.1080/2151237X.2007.10129236>.

BUTCHER, G. **Tour of the electromagnetic spectrum**. [S.l.]: Government Printing Office, 2016.

CASTANEDO, C. I. Quantitative subsurface defect evaluation by pulsed phase thermography: depth retrieval with the phase. 2005.

CHEN, L.-C. et al. Rethinking atrous convolution for semantic image segmentation. **arXiv preprint arXiv:1706.05587**, 2017.

CHUNG, D. D. **Composite materials: science and applications**. Springer Science & Business Media, 2010. Disponível em: <https://doi.org/10.1007/978-1-84882-831-5>.

CLYNE, T. W.; HULL, D. **An introduction to composite materials**. Cambridge University Press, 2019. Disponível em: <https://doi.org/10.1017/9781139050586>.

DARABI, A.; MALDAGUE, X. Neural network based defect detection and depth estimation in tnde. **NDT & E International**, Elsevier, v. 35, n. 3, p. 165–175, 2002. Disponível em: <https://doi.org/10.1016/S0963-8695(01)00041-X>.

DAS, M.; SAHU, S.; PARHI, D. Composite materials and their damage detection using ai techniques for aerospace application: A brief review. **Materials Today: Proceedings**, Elsevier, v. 44, p. 955–960, 2021. Disponível em: <https://doi.org/10.1016/j.matpr.2020.11.005>.

DUTTA, A.; GUPTA, A.; ZISSERMANN, A. **VGG image annotator (VIA)**. 2016.

FANG, Q.; IBARRA-CASTANEDO, C.; MALDAGUE, X. Automatic defects segmentation and identification by deep learning algorithm with pulsed thermography: Synthetic and experimental data. **Big Data and Cognitive Computing**, MDPI, v. 5, n. 1, p. 9, 2021. Disponível em: <https://doi.org/10.3390/bdcc5010009>.

FANG, Q.; MALDAGUE, X. A method of defect depth estimation for simulated infrared thermography data with deep learning. **Applied Sciences**, MDPI, v. 10, n. 19, p. 6819, 2020. Disponível em: <https://doi.org/10.3390/app10196819>.

FERNANDES, H. et al. Carbon fiber composites inspection and defect characterization using active infrared thermography: numerical simulations and experimental results. **Applied optics**, Optica Publishing Group, v. 55, n. 34, p. D46–D53, 2016.

GAO, T. et al. Carbon fiber reinforced polymer in drilling: From damage mechanisms to suppression. **Composite Structures**, Elsevier, v. 286, p. 115232, 2022. Disponível em: <https://doi.org/10.1016/j.compstruct.2022.115232>.

GLOROT, X.; BENGIO, Y. Understanding the difficulty of training deep feedforward neural networks. In: JMLR WORKSHOP AND CONFERENCE PROCEEDINGS. **Proceedings of the thirteenth international conference on artificial intelligence and statistics**. [S.l.], 2010. p. 249–256.

GLOROT, X.; BORDES, A.; BENGIO, Y. Deep sparse rectifier neural networks. In: JMLR WORKSHOP AND CONFERENCE PROCEEDINGS. **Proceedings of the fourteenth international conference on artificial intelligence and statistics**. [S.l.], 2011. p. 315–323.

GUPTA, R. et al. A review of sensing technologies for non-destructive evaluation of structural composite materials. **Journal of Composites Science**, MDPI, v. 5, n. 12, p. 319, 2021. Disponível em: <https://doi.org/10.3390/jcs5120319>.

HE, K. et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: **Proceedings of the IEEE International Conference on Computer Vision (ICCV)**. [s.n.], 2015. p. 1026–1034. Disponível em: <https://doi.org/10.1109/ICCV.2015.123>.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, MIT Press, v. 9, n. 8, p. 1735–1780, 1997. Disponível em: <https://doi.org/10.1162/neco.1997.9.8.1735>.

HSISSOU, R. et al. Polymer composite materials: A comprehensive review. **Composite Structures**, Elsevier, v. 262, p. 113640, 2021. Disponível em: <https://doi.org/10.1016/j.compstruct.2021.113640>.

IHIANLE, I. K. et al. A deep learning approach for human activities recognition from multimodal sensing devices. **IEEE Access**, IEEE, v. 8, p. 179028–179038, 2020. Disponível em: <https://doi.org/10.1109/ACCESS.2020.3027979>.

KHODAYAR, F.; SOJASI, S.; MALDAGUE, X. Infrared thermography and ndt: 2050 horizon. **Quantitative InfraRed Thermography Journal**, Taylor & Francis, v. 13, n. 2, p. 210–231, 2016. Disponível em: <https://doi.org/10.1080/17686733.2016.1200265>.

KOSOVA, F.; ALTAY, Ö.; ÜNVER, H. Ö. Structural health monitoring in aviation: a comprehensive review and future directions for machine learning. **Nondestructive Testing and Evaluation**, Taylor & Francis, p. 1–60, 2024. Disponível em: <https://doi.org/10.1080/10589759.2024.2350575>.

LEI, L. **Cold food chain: Infrared thermography applied to the evaluation of insulation anomalies in refrigerated vehicles for the transport of food & Exploration of cold approach in infrared thermography for Non-Destructive Testing**. Tese (Doutorado) — Université Laval, 2018.

LEMA, D. G. et al. Automated detection of subsurface defects using active thermography and deep learning object detectors. **IEEE Transactions on Instrumentation and Measurement**, IEEE, v. 71, p. 1–13, 2022. Disponível em: <https://doi.org/10.1109/TIM.2022.3169484>.

LUO, Q. et al. Temporal and spatial deep learning network for infrared thermal defect detection. **NDT & E International**, Elsevier, v. 108, p. 102164, 2019. Disponível em: <https://doi.org/10.1016/j.ndteint.2019.102164>.

MALDAGUE, X.; LARGOUËT, Y.; COUTURIER, J.-P. A study of defect depth using neural networks in pulsed phase thermography: modelling, noise, experiments. **Revue générale de thermique**, Elsevier, v. 37, n. 8, p. 704–717, 1998. Disponível em: <https://doi.org/10.1016/S0035-3159(98)80048-2>.

NIBLACK, W. **An introduction to digital image processing**. [S.l.]: Strandberg Publishing Company, 1985.

NIENDORF, K.; RAEYMAEKERS, B. Additive manufacturing of polymer matrix composite materials with aligned or organized filler material: a review. **Advanced Engineering Materials**, Wiley Online Library, v. 23, n. 4, p. 2001002, 2021. Disponível em: <https://doi.org/10.1002/adem.202001002>.

OBI, J. C. A comparative study of several classification metrics and their performances on data. **World Journal of Advanced Engineering Technology and Sciences**, World Journal of Advanced Engineering Technology and Sciences, v. 8, n. 1, p. 308–314, 2023. Disponível em: <https://doi.org/10.30574/wjaets.2023.8.1.0054>.

PEDRAYES, O. D. et al. Semantic segmentation for non-destructive testing with step-heating thermography for composite laminates. **Measurement**, Elsevier, v. 200, p. 111653, 2022. Disponível em: <https://doi.org/10.1016/j.measurement.2022.111653>.

QU, Z.; JIANG, P.; ZHANG, W. Development and application of infrared thermography non-destructive testing techniques. **Sensors**, MDPI, v. 20, n. 14, p. 3851, 2020. Disponível em: <https://doi.org/10.3390/s20143851>.

RAINIO, O.; TEUHO, J.; KLÉN, R. Evaluation metrics and statistical tests for machine learning. **Scientific Reports**, Nature Publishing Group UK London, v. 14, n. 1, p. 6086, 2024. Disponível em: <https://doi.org/10.1038/s41598-024-56706-x>.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. **Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III**. 2015. p. 234–241. Disponível em: <https://doi.org/10.1007/978-3-319-24574-4_28>.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, Nature Publishing Group UK London, v. 323, n. 6088, p. 533–536, 1986. Disponível em: <https://doi.org/10.1038/323533a0>.

SAEED, N.; ZARKANI, H. A.; OMAR, M. A. Sensitivity and robustness of neural networks for defect-depth estimation in cfrp composites. **Journal of Nondestructive Evaluation**, Springer, v. 38, p. 1–10, 2019. Disponível em: <https://doi.org/10.1007/s10921-019-0607-4>.

SCHUSTER, M.; PALIWAL, K. K. Bidirectional recurrent neural networks. **IEEE Transactions on Signal Processing**, IEEE, v. 45, n. 11, p. 2673–2681, 1997. Disponível em: <https://doi.org/10.1109/78.650093>.

SHEPARD, S. M. Advances in pulsed thermography. In: SPIE. **Thermosense XXIII**. [S.l.], 2001. v. 4360, p. 511–515.

SINGH, O. I. et al. Local contrast and mean based thresholding technique in image binarization. **International Journal of Computer Applications**, v. 51, n. 6, p. 5–10, 2012. Disponível em: <https://doi.org/10.5120/8044-1362>.

VANRULLEN, R.; KANAI, R. Deep learning and the global workspace theory. **Trends in Neurosciences**, Elsevier, v. 44, n. 9, p. 692–704, 2021. Disponível em: <https://doi.org/10.1016/j.tins.2021.04.005>.

VARGAS, I. G.; FERNANDES, H. C. Composite material defect segmentation using deep learning models and infrared thermography. **RITA - Revista de Informática Teórica e Aplicada**, 2024. Accepted for publication.

_____. Spatial and temporal deep learning algorithms for defect segmentation in infrared thermographic imaging of carbon fiber–reinforced polymers. **Nondestructive Testing and Evaluation**, 2024. Accepted for publication.

WANG, Q. et al. Defect depth determination in laser infrared thermography based on lstm-rnn. **IEEE Access**, IEEE, v. 8, p. 153385–153393, 2020. Disponível em: <https://doi.org/10.1109/ACCESS.2020.3018116>.

WANG, Z. et al. Application of carbon dots and their composite materials for the detection and removal of radioactive ions: A review. **Chemosphere**, Elsevier, v. 287, p. 132313, 2022. Disponível em: <https://doi.org/10.1016/j.chemosphere.2021.132313>.

XU, L.; HU, J. A method of defect depth recognition in active infrared thermography based on gru networks. **Applied Sciences**, MDPI, v. 11, n. 14, p. 6387, 2021. Disponível em: <https://doi.org/10.3390/app11146387>.

ZHANG, P.; YIN, Z.-Y. A novel deep learning-based modelling strategy from image of particles to mechanical properties for granular materials with cnn and bilstm. **Computer Methods in Applied Mechanics and Engineering**, Elsevier, v. 382, p. 113858, 2021. Disponível em: <https://doi.org/10.1016/j.cma.2021.113858>.

Appendices

APPENDIX A

Readme of the source project

This repository hosts the research and development work for my master's thesis, which focuses on the innovative application of defect detection in composite materials. The objective of the thesis is to explore and develop advanced methodologies for accurately identifying, characterizing, and quantifying defects within composite materials using cutting-edge technologies and analytical techniques. This includes the integration of machine learning algorithms, image processing techniques, and non-destructive testing methods to enhance the reliability and efficiency of defect detection processes. The outcomes of this research aim to contribute to the improvement of material integrity and performance in critical applications, such as aerospace, automotive, and renewable energy sectors.

Repository Structure

The repository is organized into the following directories:

- **data**: Contains datasets used for training and testing the defect detection models.
- **docs**: Includes documentation related to the project, such as research papers, reports, and presentations.
- **literature**: Contains literature reviews, relevant research articles, and references.
- **notebooks**: Jupyter notebooks with experiments, data analysis, and visualizations.
- **proposals**: Research proposals, project plans, and funding applications.
- **src**: Source code for the developed methodologies, including machine learning models, image processing algorithms, and non-destructive testing techniques.

How to Use

1. **Clone the repository**:

```
git clone https://github.com/iagogv/master-research.git
cd master-research
```

2. **Install the required dependencies**:

```
pip install -r requirements.txt
```

3. **Explore the notebooks**: Navigate to the `notebooks` directory and open the Jupyter notebooks to review the experiments and analyses.

4. **Run the code**: The source code is located in the `src` directory. Follow the instructions in the README files within each subdirectory to run the models and algorithms.

Contributions

Contributions are welcome! If you have suggestions for improvements or find any issues, please feel free to submit a pull request or open an issue.

License

This project is licensed under the MIT License.

Contact

For any questions or further information, please contact:

Iago Vargas
iagogarcia@ufu.br
Federal University of Uberlandia