applied sciences

Review

Machine Learning in Thermography Non-Destructive Testing: A Systematic Review

Shaoyang Peng *, Sri Addepalli † and Maryam Farsi †

Centre for Digital and Design Engineering, Cranfield University, Bedford MK43 0AL, UK; p.n.addepalli@cranfield.ac.uk (S.A.); maryam.farsi@cranfield.ac.uk (M.F.)
* Correspondence: shaoyang.peng.875@cranfield.ac.uk
† These authors contributed equally to this work.

Abstract

This paper reviews recent advances in machine learning (ML) algorithms to improve the postprocessing and interpretation of thermographic data in non-destructive testing (NDT). While traditional NDT methods (e.g., visual inspection, ultrasonic testing) each have their own advantages and limitations, thermographic techniques (e.g., pulsed thermography, laser thermography) have become valuable complementary tools, particularly in inspecting advanced materials such as carbon fiber-reinforced polymers (CFRPs) and superalloys. These techniques generate large volumes of thermal data, which can be challenging to analyze efficiently and accurately. This review focuses on how ML can accelerate defect detection and automated classification in thermographic NDT. We summarize currently popular algorithms and analyze the limitations of existing workflows. Furthermore, this structured analysis provides an in-depth understanding of how artificial intelligence can assist in processing NDT data, with the potential to enable more accurate defect detection and characterization in industrial applications.

Keywords: machine learning; neural network algorithm; non-destructive testing; systematic review

check for updates

Academic Editors: Evangelos Hristoforou and Victor Giurgiutiu

Received: 6 May 2025
Revised: 19 July 2025
Accepted: 22 August 2025
Published: 1 September 2025

Citation: Peng, S.; Addepalli, S.; Farsi, M. Machine Learning in Thermography Non-Destructive Testing: A Systematic Review. *Appl. Sci.* **2025**, *15*, 9624. https://doi.org/10.3390/app15179624

1. Introduction

The proliferation of advanced manufacturing techniques, such as additive manufacturing and fiber-reinforced composites, has reshaped key industries such as aerospace while simultaneously introducing new challenges for inspection and validation. Traditional visual inspections, once sufficient to identify surface defects, have proven inadequate to assess the internal integrity of increasingly complex structures. As a result, non-destructive testing (NDT) technologies, including ultrasonic testing (UT) and infrared thermography (IRT), have seen widespread adoption in detecting subsurface anomalies via changes in physical properties such as heat, sound, and magnetism [1].

Among these, thermography stands out for its non-contact, real-time imaging capabilities, large-area coverage, and minimal operational disruption. Thermographic inspection is generally classified as passive, relying on natural thermal emissions, or active, where external stimuli such as flash or laser pulses are applied to induce measurable thermal responses. Active thermography, including pulsed and laser-based methods, has shown strong performance in the evaluation of carbon fiber-reinforced polymer (CFRP) materials [2].

Ali et al. [3] have provided a comprehensive overview of the primary categories of active thermographic techniques currently used in NDT. Based on their summary, six main types of active infrared thermography are widely adopted in the literature, each with distinct operational principles and application scopes.

Each of these techniques is tailored to specific material types, defect depths, and inspection environments. The applicable detection targets and optimal use cases for each method are summarized in Figure 1, providing a comparative framework to guide the selection of technologies in thermography NDT.

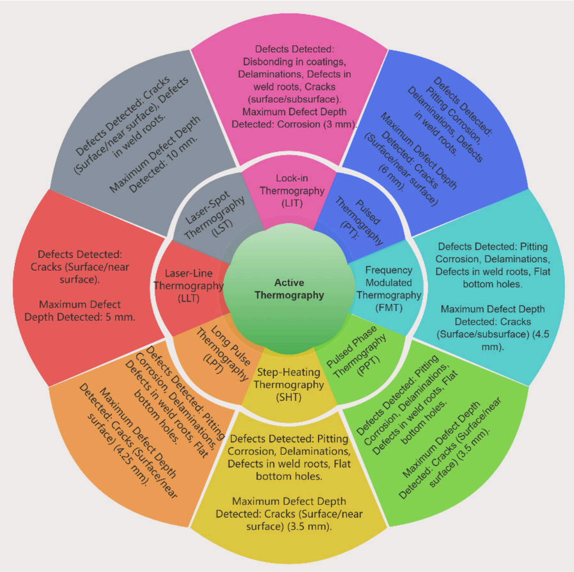

Figure 1. Overview of active IRT methods and empirical detectability guidelines (for sources, see Table 1).

Table 1. Common active IRT methods.

Terminology	**Principle**	**Representative ML Algorithm**
Lock-In Thermography (LIT) [4–6]	Periodically modulates the energy source (e.g., lamp or ultrasound) and analyzes each pixel's thermal response in the frequency domain to detect subsurface defects.	CNN [7,8]
Pulsed Thermography (PT) [9,10]	Applies a short heat pulse and monitors cooling to detect surface or shallow defects.	CNN [11,12]
Frequency-Modulated Thermography (FMT) [13]	Applies heating with a changing frequency to detect defects at different depths.	CNN [14]
Pulsed-Phase Thermography (PPT) [15]	Enhances defect contrast by applying phase analysis to data from PT.	SVM [16]
Step-Heating Thermography (SHT) [17]	Uses slow, steady heating to detect deeper or low-contrast defects.	SVM [18]
Long-Pulse Thermography (LPT) [19]	Applies long-duration heating for identification of defects with slow thermal responses.	SVM [20]
Laser-Line Thermography (LLT) [21–23]	Uses a narrow laser line for localized, high-resolution scanning.	Decision tree [24]
Laser-Spot Thermography (LST) [23,25,26]	Focuses a laser spot on a small area to detect fine or localized defects.	CNN [27]

Recent simulation-based studies have also reinforced the diagnostic value of thermal imaging. For example, Qiu [28] employed microCT-informed finite element models to analyze how defects, such as voids and delaminations, affect the thermal conductivity of FRP materials. The study revealed that defect-induced heat flow disturbances, coupled with the non-linear effects of moisture content, lead to localized thermal accumulation, supporting the value of thermography in structural assessment and highlighting the need for intelligent interpretation tools.

Coupled with robotics, thermographic systems now enable automated inspections, significantly improving efficiency [29]. However, as Bossi [30] observed, this automation also generates massive amounts of data, which are typically interpreted by trained human operators using specialized software—an approach that is labor-intensive and costly. There is a lack of clarity regarding the selection of a suitable machine learning (ML) algorithm that works with a specific type of geography data. To address this bottleneck, artificial intelligence (AI), particularly ML, offers a compelling solution by automating the interpretation of thermal data.

Recent ML-based studies—especially those using convolution neural networks (CNNs)—have demonstrated notable success in classifying thermal images, identifying defect types, and localizing damage. However, while many of these works highlight strong performance within their chosen models, there is a lack of comparative analysis between algorithm types, limiting the practical guidance available for researchers and practitioners.

This paper addresses this gap by presenting a systematic literature review (SLR) focused on the application of machine learning to thermographic NDT. Through a structured data collection and organization (SDCO) methodology, we evaluated how different ML approaches have been used to improve thermal image postprocessing, ranging from support vector machines (SVMs) to deep learning frameworks. We evaluated the effectiveness, limitations, and domain suitability of each method with the goal of providing clear evidence-based guidance for algorithm selection.

All thermographic modalities, whether passive or active, produce data that are amenable to ML-based analysis. This review examines the current landscape of ML integration in thermographic NDT, paying particular attention to its practical benefits, technical limitations, and implementation challenges. By identifying a critical gap in algorithm selection and application, we position this review as a resource for both researchers and practitioners seeking to deploy AI tools in thermal imaging-based inspection.

The remainder of this paper is split into four key sections: a thorough exploration of the adopted methodologies, a description of the data sources, the precision of ML algorithms for postprocessing in NDT, and a discussion of the nuances of the extant ML frameworks currently in use for similar applications. The concluding section summarizes the research achievements and inherent limitations and provides directions for future research endeavors.

2. Methods

A systematic literature review is described by Booth as an explicit and reproducible approach used to identify, evaluate, and synthesize the pre-existing work undertaken and documented by researchers, scholars, and practitioners [31]. This paper applies the SALSA framework, which consists of four primary steps: retrieval, assessment, synthesis, and analysis. Each of these steps is elaborated upon in the following subsections.

2.1. Step 1—Search

Defining the scope of the study is essential to formulate clear and answerable questions. This review adopts the SDCO (sample, design, comparison, outcome) framework (Table 2) to construct the purpose, objectives, and research questions.

Table 2. SDCO framework for search strategy.

Sample	Thermography data of metal or composite plates or structures with defects
Design	Feature extraction, dataset generation, and training design
Comparison	Different machine learning algorithms and neural network frameworks
Outcome	Classification accuracy of machine learning for postprocessing of thermography data and accuracy of defect size and depth prediction

2.1.1. Aim, Objectives, and Research Questions

Guided by the SDCO framework, this review seeks to explore the variations in the effectiveness of different machine learning algorithms in postprocessing thermography data. This encompasses tasks such as feature extraction, employing a neural network, ensuring computation precision, gauging processing times, and validating models. Beyond this, the review also elucidates both the inherent uncertainties present in thermal imaging data and those that arise during the application of the machine learning algorithm.

Central to our inquiry are the following research questions:

(a) In what ways does machine learning expedite the postprocessing of thermographic data in non-destructive testing?
(b) What factors should be prioritized when choosing the appropriate machine learning algorithm tailored to specific test data?

2.1.2. Search Strategy for Identification of Articles

To ensure the systematic nature of a review, it is imperative to have a well-defined strategy for the retrieval of the literature. Identified works subsequently undergo a filtering process dictated by both inclusion and exclusion criteria, which serve as critical determinants in the selection of studies for the review (refer to Table 3). Furthermore, the specific databases used for searching, the search terminologies employed, and the types of articles targeted are explicitly delineated in this review.

Table 3. Inclusion and exclusion criteria for search strategy.

SDCO	Inclusion Criteria	Exclusion Criteria
Sample	Thermography data of different materials (both metals and composites)	Other NDT data sources (like vibration)
Design	Dataset generation from different defect types or samples Feature extraction from different defect types or samples Experimental platform building based on thermography	Other defect test data sources Performance indices of infrared camera or other equipment
Comparison	Different machine learning algorithms and neural network frameworks	Other variables not related to machine learning algorithms and samples
Outcome	Performance of machine learning in postprocessing thermography data and accuracy of defect size and depth prediction	Studies that do not report relevant outcome measures

Based on these criteria, a literature review was conducted with the purpose of comparing references.

Data sources: The literature search used scientific indexing services such as Web of Science, Science Direct, and Scopus.

Keywords included “machine learning” AND “thermography” AND “non-destructive testing”.

2.2. Step 2—Appraisal

Approximately 100 English journal and conference papers were identified through our initial search through Scopus. Based on the inclusion and exclusion criteria, 67 were identified as distinct and directly relevant to this review. In cases where articles produced by the same authors presented similar methods and findings, they were treated as a singular representative article. Although there may have been inadvertent omissions due to search engine and platform limitations, the data derived from this search are broadly representative given the specific subjects under examination. As indicated in Table 2, the assessment framework encompassed the following:

- Data sources, data integrity, and data management;
- Exploration of methods and associated parameters (including preprocessing and ML);
- Sensitivity analysis related to parameters, data sources, and intrinsic uncertainties;
- Evaluation metrics tailored for ML efficacy;
- Established methods for validation and verification;
- Recognized biases within the data and the resultant conclusions.

2.3. Step 3—Synthesis

During the synthesis phase, pertinent information is extracted and classified from the articles under review. Taking cues from Booth's four elemental categories—publication aspects, research facets, distinguishing characteristics, and concluding elements—and combining these with pivotal information from works concerning machine learning and thermal imaging, our review predominantly utilized the "characteristics" metric. This allowed the categorization of the literature, highlighting the distinctive aspects of various documents within the identified literature pool. Specifically, in this review, two pivotal feature identifiers took precedence: the type of data and the machine learning methodology involved.

2.4. Step 4—Analysis

The analysis phase consists of three steps.

(i) Thematic analysis: This focuses on gauging the pertinence of each theme in relation to the articles assessed via qualitative data or quantifiable outcomes.
(ii) Discussion of the results: This involves a critical appraisal of the results, determining their alignment with the review's research questions. Such deliberations included in the Results section.
(iii) Drawing conclusions: This step revolves around articulating the primary takeaways in the Conclusions section, acknowledging the inherent limitations of the study and providing avenues for future research.

3. Results

In this section, we present and discuss the results in terms of the four subcategories: publishing elements, research elements, features, and conclusion elements.

Firstly, thorough quantitative evidence and explanation should be given regarding the process of study inclusion/exclusion (from start to finish), summarized in an appropriate flow diagram. Secondly, a summary of the study characteristics should be derived, informing the reader of the total number (sum of all studies), mean/median, age range, comorbidities, and other appropriate characteristics of the subjects considered. This information can be stored in a graph, but the key data must be present in the text.

3.1. Publication Elements

The distribution of the selected works by year of publication reveals a significant trend, which is shown in Figure 2. It is evident that research in the field of the automated processing of thermography data using ML algorithms has experienced exponential growth since 2020. Remarkably, the year 2021 alone saw 12 publications, which is almost equivalent to the total publications published between 2009 and 2019. This surge underscores the growing interest and focus in this research area over time. In particular, the period 2021 to 2023 contributed to most of these publications, and the increasing volume of research in recent years clearly indicates the increasing recognition of the potential and applicability of ML techniques in thermal imaging data analysis, reflecting both the rapid advancements in the field and its expanding significance in various applications. This trend not only demonstrates the evolving landscape of ML in thermography but also reaffirms the timeliness and importance of the current review in capturing and understanding these recent developments.

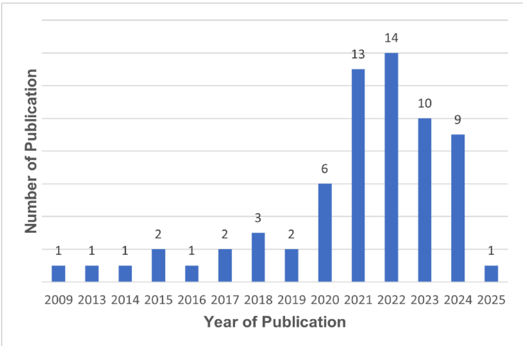

Figure 2. Distribution of the publications per year.

The distribution of the publications reviewed between journal and conference papers is shown in Figure 3. Of these, 4 out of 66 (6%) are conference papers, while 62 out of 66 (94%) are journal articles. In the compilation of the literature for this review, it was observed that conference papers typically offered less detailed explanations of the ML algorithms employed and a more superficial analysis of the data. Furthermore, access to some conference papers was restricted, which limited their availability for this review. Consequently, journal articles constitute most of the works reviewed. This preference for journals is attributed not only to their greater accessibility but also to their comprehensive and in-depth treatment of the given topic, including a detailed description of the methodologies and more rigorous data analysis. The list of journals and conferences where the reviewed articles have been published is shown in Tables 4 and 5.

Table 4. List of conferences.

Name of Conference	Ref.
2016 IEEE 8th International Conference on Intelligent Systems (IS)	[32]
Thermosense: Thermal Infrared Applications XLIII (2021)	[27]
12th CIRP Conference on Photonic Technologies [LANE 2022]	[33]
2018 6th International Renewable and Sustainable Energy Conference (IRSEC)	[34]

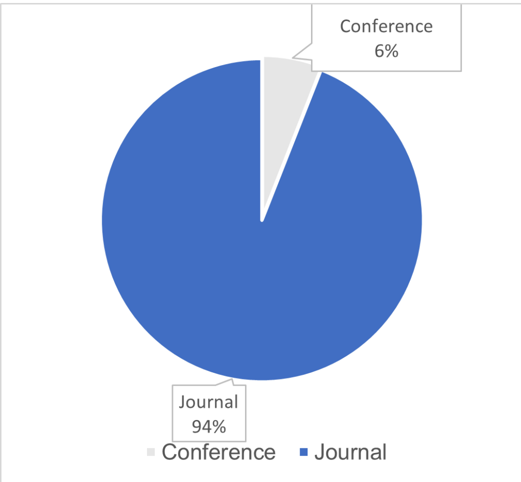

Figure 3. Distribution of the publications between journals and conferences.

Table 5. List of journals.

Name of Journal	Ref.
IEEE Geoscience and Remote Sensing Letters	[35]
IEEE Access	[36–40]
IEEE Transactions on Industrial Informatics	[41]
Wireless Personal Communications	[42]
Proceedings of the Institution of Mechanical Engineers	[43]
Advances in Engineering Software	[44]
Polymers	[45,46]
Sensors	[12,18,47–49]
Composites Science and Technology	[40]
Russian Journal of Nondestructive Testing	[50]
Energies	[51,52]
IEEE Sensors Journal	[53]
NDT&E International	[54,55]
Infrared Physics & Technology	[2,56–62]
Journal of Nondestructive Evaluation	[63]
Applied Sciences	[64]
Measurement	[18,65,66]
Engineering Applications of Artificial Intelligence	[67]
international Journal of Adhesion and Adhesives	[68]
Procedia Structural Integrity	[34]
Neural Computing and Applications	[69,70]
Remote Sensing	[35,71]
Smart Materials and Structures	[72]
ISA Transactions	[73]
JOM	[74]
Measurement Science and Technology	[1,45,75]
Computers, Materials, Continua	[76]
Forests	[77]
Composites Part B: Engineering	[78]
Composite Structures	[79]
Quantitative InfraRed Thermography Journal	[80,81]
Materials	[82]

3.2. Study Elements

In this review of 59 articles on the application of ML algorithms in thermography processing, the following classification is derived based on the different types of research objects.

- Composite Material Defect Detection: A significant portion, comprising 34 publications, focuses on defect detection in composite materials, such as carbon fiber-reinforced polymers (CFRPs), glass fiber-reinforced polymers (GFRPs), and various metal materials. This category includes studies that use thermography to identify structural weaknesses or inconsistencies in these materials.
- Electrical Equipment Defect Detection: Approximately nine articles are dedicated to the detection of electrical equipment defects. This field involves the use of thermal imaging to identify faults or anomalies in electrical components, contributing to preventive maintenance and safety in electrical systems.
- Additive Manufacturing Monitoring: Five publications delve into the monitoring of additive manufacturing processes. These studies explore the role of thermal imaging in ensuring the quality and precision during the fabrication stages of additive manufacturing.
- Plant Growth Monitoring: Two articles address the application of thermal imaging in monitoring plant growth. This unique application focuses on changes in specific components, such as the nitrogen content in plants, which can help to improve agricultural practices.
- Human Health Applications: In addition, there is one article that focuses on the use of thermal imaging to detect breast-related health issues in humans. This signifies the expansion of thermal imaging applications into the realm of medical diagnostics.

3.3. Characteristics

Machine learning is a method whereby human experts employ appropriate algorithms to enable machines to learn and adapt to human behavior, extracting information and knowledge directly from the given data. Once the machine completes its learning from these data, it can automatically evaluate and perform learning tasks without adhering to explicit rules. By applying various statistical and mathematical models to identify patterns and draw inferences from data, it achieves decision making with minimal human intervention. These algorithms improve their performance as they are exposed to more data over time. Fundamentally, ML algorithms are categorized into three types: supervised learning, where the algorithm learns from labeled training data; unsupervised learning, which involves finding hidden patterns in unlabeled data; and reinforcement learning, where an agent learns to make decisions by performing actions and receiving feedback.

In the articles reviewed, the primary objective of the machine learning algorithms employed is to accelerate the analysis of the given thermal imaging and ultrasonic data, while also enhancing the accuracy of the results. It should be noted that, for this review, the application of ML in the field of ultrasonic inspection has been considered due to similarities associated with NDT measurements and the limited literature on ML for thermography measurements.

A key distinguishing feature among the identified studies is the variety of ML algorithms used. This diversity reflects different approaches to processing and interpreting data, each with its unique strengths and limitations. In this review, based on the examination of the 62 journal articles identified, a comprehensive overview of the various methodologies applied by different researchers is obtained, as shown in Table A2. This table not only serves as a reference for the specific algorithms used but also highlights the breadth of ML applications in thermal imaging and ultrasonics. The comparison of these diverse algorithms provides valuable information about how different machine learning techniques can be optimized for specific types of data analysis, contributing to the field's understanding of the most effective strategies for processing thermal and ultrasonic data.

In addition to drawing conclusions, several recent studies have emphasized the importance of preprocessing strategies to enhance the quality of thermal data before feeding them into machine learning models. In active thermography, issues such as sensor noise, non-uniform heating, and background interference can significantly impair signal-to-noise ratios, reducing the effectiveness of subsequent ML-based defect characterization. To address these challenges, Dong et al. [83] proposed a two-stage convolutional neural network that jointly suppresses sensor noise and background variation in lock-in thermography data, enabling clearer signal interpretation and improved defect detection performance. Likewise, Matarrese et al. [84] explored how the lock-in excitation setup parameters (e.g., frequency, amplitude, phase delay) influence the ability of temporal convolutional neural networks to detect and characterize defects in CFRP materials. These studies underscore that the preprocessing quality is a critical determinant of downstream ML model performance and must be carefully considered when designing thermography-based NDT workflows. Incorporating such denoising and normalization techniques can significantly boost the robustness and generalization of ML algorithms across varied inspection scenarios.

3.4. Algorithm Types

Machine learning algorithms process thermographic data by extracting patterns from thermal images (2D spatial data) or thermal signal sequences (1D temporal data). The workflow typically involves data preprocessing (e.g., denoising, normalization), feature extraction (manual or automatic), and classification. Traditional models such as SVM and KNN rely on hand-crafted features, while deep learning models like CNNs learn feature representations directly from raw data. Some studies also employ dimensionality reduction techniques (e.g., PCA) to enhance the classification performance. These approaches enable the more accurate and efficient interpretation of thermal data for the detection and characterization of defects in NDT.

This section will provide a brief description of the different ML algorithms used in this review of the literature. A summary of the machine learning methods in the main papers reviewed is shown in Table A1 in Appendix A.

a. Artificial Neural Network (ANN)

The ANN is a computational framework modeled after the workings and architectures of biological neural networks. It consists of a series of interconnected units known as neurons, each capable of executing basic mathematical calculations to convert input data into output results. These interconnections have weights that are fine-tuned during training to map input variables to their corresponding outputs [49]. A standard ANN configuration comprises three primary layers: the input layer, which handles incoming data; one or multiple hidden layers, where data features are represented and transformed; and the output layer, which presents the final prediction or classification.

b. Multilayer Perceptron (MLP)

The MLP is a type of feedforward neural network that is notable for its straightforward computational structure, clearly specified parameter space, and consistent effectiveness across various tasks. Relative to other ANN designs, the MLP usually demands fewer parameters and adopts a more concise architecture to reach comparable accuracy. In the feature evaluation phase, each input is mapped to its specific node in the input layer. During the diagnostic phase, the input layer is set up to handle a vector composed of selected features, with each node representing one feature, enabling the model to learn from a systematically organized set of relevant inputs.

c. Deep Neural Network (DNN)

An ANN with several hidden layers is termed a deep neural network (DNN). These hidden layers enhance the model's ability to learn non-linear relationships via activation functions like sigmoid, tanh, or ReLU. Each neuron's full connection to adjacent layers ensures efficient feature propagation and transformation across the network.

d. K-Nearest Neighbors (KNN)

The KNN algorithm predicts the output for a given query point considering the outputs of its K-nearest training samples, usually identified via the Euclidean distance [40]. The parameter K specifies the number of neighbors to include in the prediction. Depending on the problem, KNN is employed for classification, selecting the label that appears the most frequently among neighbors, or regression, calculating the average of their values. Weighting can be performed uniformly or inversely proportional to the distance, allowing the model to prioritize closer samples during predictions.

e. Convolutional Neural Network (CNN)

Traditional machine learning techniques such as SVM and KNN require the manual extraction or statistical derivation of features based on expert insights. In contrast, deep learning models, especially convolutional neural networks, inherently learn pertinent features directly from unprocessed data, facilitating comprehensive defect detection with little human input. CNN architectures typically begin with one or more convolutional layers. These layers employ multiple adjustable filters that traverse the input feature maps to perform the weighted summation of local regions, allowing the extraction of localized spatial features [56]. The pooling layers are commonly placed after the convolutional layers to decrease the spatial dimensions of the feature maps. This involves scanning fixed-size windows across the feature maps and calculating statistical values like the maximum or average within each window. Following the feature extraction phase, the resulting features are flattened and passed through one or more fully connected layers for final classification [70]. CNNs have been extensively used for defect classification in both composite materials and metallic structures, especially in thermographic inspection systems. Their ability to discern spatial and texture-based patterns from thermal images facilitates the precise identification of typical defect types such as delamination, voids, cracks, and corrosion. This capability makes CNNs an invaluable tool for automated classification tasks in non-destructive evaluation across various material systems [12,52,64,70,85].

f. Support Vector Machine (SVM)

Traditional machine learning techniques such as SVM and KNN require the manual extraction or statistical derivation of features based on expert insights. In contrast, deep learning models, especially convolutional neural networks, inherently learn pertinent features directly from unprocessed thermal data, facilitating comprehensive defect detection with little human input. CNN architectures typically begin with one or more convolutional layers. These layers employ multiple adjustable filters that traverse the input feature maps to perform the weighted summation of local regions, allowing the extraction of localized spatial features [56]. The pooling layers are commonly placed after the convolutional layers to decrease the spatial dimensions of the feature maps. This involves scanning fixed-size windows across the feature maps and calculating statistical values like the maximum or average within each window. Following the feature extraction phase, the resulting features are flattened and passed through one or more fully connected layers for final classification [70]. CNNs have been extensively used for defect classification in both composite materials and metallic structures, especially in thermographic inspection systems. Their ability to discern spatial and texture-based patterns from thermal images facilitates the precise identification of typical defect types such as delamination, voids, cracks, and corrosion. This capability makes CNNs an invaluable tool for automated classification tasks in non-destructive evaluation across various material systems [12,52,64,70,85].

g. Generative Adversarial Networks (GANs)

Generative adversarial networks (GANs) are a class of deep generative models that consist of two competing neural networks: a generator and a discriminator. During training, the generator G learns to produce synthetic thermal images that mimic the true data distribution, while the discriminator D learns to distinguish between real and generated images [86]. In practice, G tries to create samples that closely resemble the target distribution, and D acts as a binary classifier to separate real infrared images from those produced by the GAN [86]. Through this adversarial, game-like training, the generator gradually improves until its outputs are difficult to distinguish from real thermographic data.

In thermographic NDT, GANs have been applied both to augment and enhance thermal image data and to automate defect detection. One common use is data augmentation: GANs learn the distribution of real infrared frames and generate new synthetic images to enrich limited NDT datasets. For example, Liu et al. proposed a spectral normalization GAN (SNGAN) to "learn the thermal image distribution, thereby generating virtual images to enrich the dataset" in composite inspections [46]. This approach (called GMLT) uses the GAN generator to create additional thermographic images that capture typical defect patterns and background variability. Another role of GANs is image enhancement: many thermographic NDT problems suffer from a low resolution or noise. GAN-based superresolution models (e.g., SRGAN variants) can upsample low-quality infrared frames. Cheng and Kersemans' Dual-IRT-GAN, for example, uses a generator to produce high-resolution thermograms from low-resolution input while simultaneously enhancing defect visibility with attention maps [78]. Finally, GANs can perform image-to-image translation for segmentation: a conditional GAN can directly map infrared images to defect masks. In the IRT-GAN framework, the generator takes multiple preprocessed thermal images (such as principal component or signal reconstruction frames) and fuses them using a multihead strategy to produce a single defect segmentation map [79]. In other words, the GAN learns an end-to-end transform from IR data to defect annotations, automating what is normally performed by an expert.

3.5. Feature Extraction

The ML algorithms used in the selected articles and the corresponding data sources and data types are compiled in Table A2 in Appendix B. Among them, 36.95% of the articles use digital signals and 73.05% use images. Most articles using image data for ML adopt the method of cropping and converting images into grayscale images to facilitate feature extraction. This may be because the first-hand data obtained through the camera are image data, assuming that the data are not processed at all. Some articles propose image denoising methods based on ML. Note that, in most cases, the image features are selected manually. Most articles using signals as ML algorithms use thermographic signal reconstruction (TSR), a polynomial fitting method, and principal component analysis (PCA) as data signal enhancement and denoising methods. The feature selection of signals is mostly based on the temperature decay curve of the selected experimental objects, such as metal plates.

3.5.1. Thermographic Signal Reconstruction (TSR)

Thermographic signal reconstruction involves adjusting the evolution of the temperature to an nth-degree polynomial, as shown in Equation (1) [87]:

$$T(t) = a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t^1 + a_0 \quad (1)$$

When the infrared camera captures pulsed thermographic data, polynomial fitting is used to filter out noise from the signal. The processed data are then converted to a logarithmic domain. This treatment compresses the thermal sequence of the original data. In addition, it is utilized for the analytical calculation of time derivatives. This has been shown to be one of the best methods of improving defect visualization [30], especially when using the second-order derivative.

3.5.2. Principal Component Analysis (PCA)

PCA is a method of projecting high-dimensional responses into a low-dimensional subspace through orthogonal projection using eigenvalue or singular value decomposition [57]. In thermography data processing, matrixing the original high-dimensional thermography data and using fewer orthogonal dimensions to represent the feature vectors and eigenvalues of the data covariance matrix allows one to identify the principal components. Finally, it is transformed into a smaller, low-dimensional dataset for better visualization.

Reducing the number of variables in a dataset naturally comes at the cost of reduced accuracy, but the key aspect of dimensionality reduction is to sacrifice a small amount of accuracy for simplicity. Because smaller datasets are easier to explore and visualize, the analysis of data points becomes easier and faster for machine learning algorithms, without dealing with irrelevant variables.

4. Discussion

Compared with other non-destructive testing technologies, thermography has advantages such as remote detection, the fast scanning of large areas, improved detection efficiency, and the ability to be applied to high-density materials. ML has also been proven to accelerate the analysis of data generated by thermal imaging with high accuracy. The combination of ML and IRT for non-destructive testing has broad application prospects. According to Shakeel's [29] research, ML algorithms, especially backpropagation neural networks combined with thermography technology, have the potential to improve NDT's accuracy, sensitivity, and specificity, particularly in medical detection. The SLR indicates that ML methods can be applied to most inspection cases and that there are no inspection-related constraints on the use of ML. The ML methods may need to be tuned (model structure, hyperparameters) to maximize their performance, but the diversity of data and methods in the literature seems to indicate the potential for the widespread use of ML for NDT, particularly showcasing these systems' potential in preventing equipment failure and system breakdowns [44]. However, research on the combination of thermography and ML technology for non-destructive testing still has certain limitations. These limitations are summarized in the following sections.

4.1. Significance of Using Machine Learning in the Field of NDT

The first-hand data provided by an IR camera are images consisting of spatial pixels containing temperature data. For engineers, the ideal outcome is to visually identify internal defects within the tested object directly from the images obtained. However, due to limitations in the performance of cameras and other conditions, detecting small-impact defects (with an impact energy of 3J) is more challenging compared to other methods, such as ultrasonic or vibration-based inspection [88]. Postprocessed thermal imaging data are necessary to help engineers to obtain accurate results.

The raw data generated by the IR camera are vast in volume, not only making manual processing time-consuming but also increasing the uncertainty. ML models can automatically identify and extract features from images and utilize high-performance computing resources to process large datasets in parallel. The goal of researchers is to employ anomaly detection algorithms so that machine learning models can automatically discover and report potential defects, without the need for engineers to manually inspect each image. This approach not only significantly shortens the inspection cycle but also enhances the reliability and consistency of the results, being of tremendous value in the industrial sector.

4.2. Lack of High-Quality Public and Unified Datasets

The dataset size/representation seems to be a limiting condition in most reviewed studies. The size of the dataset has a significant impact on the selection of ML algorithms. In the process of addressing problems using ML or DL methodologies, model training requires extensive datasets. However, in most current research within the relevant field, datasets are constructed by researchers through their own experiments [30,32,50] due to various issues such as ownership and confidentiality, as well as differences in the physical properties of different test subjects, and there is a lack of open-access datasets that could serve as benchmarks. Additionally, the constraints on experimental conditions mean that smaller datasets lead researchers to favor machine learning algorithms like SVM, which are less sensitive to the size of the dataset.

In the non-destructive testing of materials, such as CFRP or metal materials, whether the datasets are generated experimentally or based on finite element analysis, the predefined defect structures for testing are overly simplistic. The most applied defect structure consists of specifically shaped air gaps within the base structure. Some studies have also used cracks as a defect feature. However, the occurrence of defects is much more complex and varied. Datasets generated from singular defect features may reduce the confidence in the accuracy of the final algorithms.

4.3. Algorithm Performance

Variations in data analysis performance can be attributed to differences in the data used, in combination with the ML algorithms used. In Buongiorno's [64] study, modifications to the parameters or algorithms themselves resulted in data that demonstrated nearly identical accuracy for the same test datasets. However, deep learning algorithms required more learning time and did not always achieve an advantage in accuracy, especially in the case of small datasets [89,90]. The reported results indicated that there were no clear correlations between the methods used and the accuracy of the classification. The implication is that most methods are likely capable of good classification performance, with the results depending on the data used and model parameter tuning.

4.4. Deep Learning: Current Conditions

Deep learning, a relatively new and promising subfield of machine learning, trains on vast datasets and has surpassed the limitations of traditional ML algorithms, enabling more reliable and efficient detection in highly complex scenarios. However, the current research does not show that deep learning significantly outperforms traditional methods in defect prediction accuracy. In studies before 2019, the features applied to deep learning were manually selected, which could be one of the reasons for the decrease in algorithm performance [36]. Since 2019, researchers have increasingly used CNNs to extract high-dimensional data features, rather than employing PCA for feature extraction. Using CNNs for feature extraction has further enhanced the automation of the analysis process, and the extracted features are considered better suited for classifiers, although this process consumes more time [77]. Currently, algorithms based on large language models have the potential to allow the autonomous selection of features for ML, which could be one direction for the development of deep learning in the field of non-destructive thermography testing [91].

4.5. Uncertainty of Existing Research

Despite recent progress, several important aspects of machine learning applications in thermographic NDT remain underexplored in the current literature. These include challenges related to dataset size determination, class imbalance, and data bias, all of which can significantly affect model robustness and generalization. In addition, limited attention has been paid to feature selection strategies, hyperparameter tuning, and the documentation of model optimization workflows. The issues surrounding result reproducibility, the quantification of predictive uncertainty, and the use of probability of detection (POD) curves are also not adequately addressed in many studies. Furthermore, practical considerations, such as the software platforms used, the computational environment, and the required expertise for implementation and interpretation, are often omitted or only briefly mentioned. These gaps highlight the need for greater transparency and standardization. Fortunately, many of these concerns are receiving increased attention in the broader machine learning research community, and potential solutions are beginning to emerge in all domains [27].

4.6. Probability of Detection

The POD is also referred as the positive predictive probability. It defined as "the probability that a given damage in a component will be detected using a given inspection method". It is expressed as

$$POD = \frac{TP}{FN + TP} \tag{2}$$

FN = false negative;
TP = true positive. Probability of detection judgment criteria can be seen in Table 6.

Table 6. Probability of detection judgment criteria.

	True	**False**
Positive	An item has defects, and the NDT method detects it.	No flaw exists and the NDT method indicates a flaw.
Negative	No flaw exists, and the NDT method has no indication of a flaw.	An item has defects and the NDT method does not detect it.

The probability of detection (POD) curve is a widely used tool in NDT to quantify the relationship between defect detectability and defect characteristics, particularly size. It represents the likelihood of successfully identifying defects of varying severity, providing an intuitive way to assess and compare the effectiveness of different inspection approaches. In practical settings, it is generally acknowledged that no NDT technique can achieve absolute certainty in detecting all flaws, even those of considerable size. Therefore, a commonly accepted performance benchmark is a detection probability of 90% with statistical confidence of 95%. This standard indicates that the method should reliably detect 90% of relevant defects, with high confidence in the reliability of the estimate [92]. When applied to ML-based NDT analysis, POD curves serve as a valuable means of illustrating the predictive capacity of learning algorithms. They provide a statistical framework for the evaluation of model performance and can support validation by linking the algorithm output to practical defect detection expectations.

The general ML literature includes methods for data selection, model tuning, sensitivity analysis, performance evaluation metrics, and verification and validation. Open-source software is also available for many of these aspects. As a result, it is likely that these techniques are primarily applied to standardized image and time series data.

5. Conclusions

Our work comprehensively reviews the progress in applying ML techniques in the field of thermography NDT. The intersection of these two domains highlights the trends and opportunities associated with ML, leveraging its efficient data processing capabilities across various fields. The results were derived using a structured approach, reviewing over 60 related papers to thoroughly examine the current state of research in ML and deep learning algorithms for thermography NDT.

Convolutional neural networks (CNNs) and support vector machines (SVMs) are the main highlighted methods, accounting for 18% and 23%, respectively. This reflects their reliable performance in handling image and signal data, especially noting the excellent performance of SVMs when dealing with smaller datasets.

The type and source of data are key aspects that were elucidated in our analysis. Images and signals are dominant in analytical research. Observing the simultaneous use of images and signals in a single study highlights the potential benefits of multimodal data integration.

Demonstrating confidence in the results of ML algorithms will require careful attention to data selection. Representative, common datasets are likely to be necessary to increase the confidence in ML performance and allow easier comparison between methods and their verification and validation, as these are required to demonstrate the impact of ML on the reliability of NDT. Methods of quantifying confidence bounds or uncertainty in the predictions of the ML algorithm are also important to ensure valid classification results.

Sensitivity studies related to model parameters are likely to be important in improving the confidence in the reported results. The impact of noise in the data on the results and the uncertainty throughout the process must be a part of the assessment.

Moreover, it is crucial to acknowledge a significant limitation in the current research: the lack of a high-quality database hinders the training of high-quality ML algorithms. This limitation could cause algorithm selection to be restricted to those suitable for small datasets, ultimately leading to a lack of innovation. In addition, uncertainties in dataset optimization and parameter tuning are present in related studies. Therefore, the creation of a high-quality database becomes a major requirement to solve this problem. It is a good approach to combine finite element simulations with laboratory experiments to create a database with multiple defect modes and material properties.

Essentially, our synthesis of the research findings not only highlights the progress being made but also points out gaps in existing research. It suggests that diversification in data types and collection techniques will form the basis for the optimization of ML contributions to thermography data analysis.

Future work will focus on researching algorithms that offer higher performance and usability for integration with active thermographic NDT [89].

Exploring methods that can automate the entire thermographic NDT process may offer significant opportunities for the manufacturing, maintenance, and asset management fields in the future.

Author Contributions: S.P.: Conceptualization, Methodology, Data Curation, Writing—Original Draft Preparation. S.A.: Methodology, Writing—Reviewing and Editing. M.F.: Writing—Reviewing and Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the EPSRC platform grant (grant number EP/P027121/1).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed during this study. Data sharing is not applicable to this article.

Acknowledgments: This work was undertaken as part of a Ph.D. programme. Acknowledgments are extended to the Centre for Digital Engineering and Manufacturing, Cranfield University.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NDT	non-destructive testing
UT	ultrasonic testing
IRT	infrared testing
CFRP	carbon fiber-reinforced polymer
ML	machine learning
SLR	systematic literature review
AI	artificial intelligence
SDCO	systematic data collection and organization
ReLU	rectified linear unit
UD	uniform distribution of weights
KNN	K-nearest neighbors
DNN	deep neural network
MLP	multilayer perceptron
ANN	artificial neural network
CNN	convolutional Neural Network
SVM	support vector machine
TSR	thermographic signal reconstruction
PCA	principal component analysis

Appendix A

Table A1. Summary of machine learning methods in main papers reviewed.

Method	Features	Comments	Refs.
GAN	Comprises a generator and a discriminator that compete to improve performance.	Excellent for generating realistic images, useful in synthesizing or enhancing thermal images.	[38,45,46,51,78,79]
MLP	A basic form of neural network with multiple layers and backpropagation training.	Suitable for straightforward pattern recognition in thermal data but may struggle with complex spatial data.	[18,36,41,44,52,55,67,74]
ANN	Composed of interconnected nodes mimicking biological neural networks.	Versatile for various NDT tasks, offering flexibility in handling diverse thermal imaging data.	[27,34,42,44,49,72–74]
SVM	Efficient in high-dimensional spaces, focuses on finding the maximal margin.	Effective in classifying thermal image features, particularly in clear-cut defect recognition.	[18,27,34,43,52,53,55,57,62,66,93]
KNN	Unsupervised learning; simple, instance-based learning algorithm that classifies based on nearest training examples.	Useful for thermal imaging when patterns are well defined and consistent.	[40,62,64]
CNN	Using convolutional layers to efficiently capture spatial and hierarchical patterns within data.	Highly effective in feature extraction and pattern recognition in thermal imaging, especially in defect detection.	[12,33,38,39,48,50,52,63,68,71,82,85,94]

Table A1. *Cont.*

Method	Features	Comments	Refs.
Bi-LSTM	Processes data in both forward and backward directions, effectively capturing complex temporal sequences and dependencies in data.	Applicable for analysis of time series thermal imaging data, such as in continuous monitoring.	[67]
DNN	Contains multiple hidden layers for learning of complex data representations.	Capable of handling intricate thermal imaging data but requires substantial data and computational resources.	[50,65]
FCN	A non-parametric, probabilistic approach that models the distribution of possible outcomes, excelling in predicting continuous outputs with uncertainty estimation.	Performs well with small-scale thermal imaging datasets, providing uncertainty estimates with predictions.	[59]
RNN	A type of neural network designed for sequential data processing, characterized by their ability to	maintain internal memory of previous inputs, making them ideal for tasks involving temporal dependencies.	[56,60,61]
Hybrid algorithms	Use multiple distinct machine learning techniques, leveraging their individual strengths to enhance overall performance and address complex problems more effectively than single-method algorithms.	Provide improved performance in complex thermal imaging data analysis tasks but may be complex to implement and optimize.	[34,37,56,95]
SDL	An algorithmic approach where the model autonomously identifies learning tasks and strategies, adapting its learning process based on evolving data and objectives.	Has potential in scenarios requiring adaptive learning for thermal imaging analysis, an emerging area.	[75]
SK-Means	The K-means algorithm is a partitional clustering method that groups data by minimizing the squared distances between points and their nearest cluster centroids. The SK-means variant extends this approach by integrating a region segmentation step into the clustering process.	High performance in thermal imaging by incorporating spatial continuity into the standard K-means framework. This spatial constraint reduces sensitivity to noise, leading to more coherent and reliable defect localization.	[81]
U-Net	A convolutional neural network tailored for pixel-wise segmentation tasks. It combines a contracting path to extract the region surrounding each pixel and an expansive path for precise localization.	High performance with small datasets, but overlapping areas around adjacent pixels reduce training efficiency.	[80]

Appendix B

Table A2. Summary of feature extraction and data processing in main papers reviewed.

Ref.	Machine Learning Model	Experimental Object	Data Type	Features
[36]	MLBPNN	Brain	Image	Surface highlights of the areas of interest.
[41]	MLP	AL sheet, FS weld	Image	ROI was manually cropped from the recorded grayscale image.
[42]	ANN	Breast	Image	Fifteen features from each image. The features are entropy, mean, kurtosis, median, mode, contrast, standard deviation, correlation, variance, skewness, energy, homogeneity, dissimilarity, inverse difference moment, and area.
[32]	K-means unsupervised machine learning	CFRP plate	Image	Speeded-up robust features (SURF); each point of a blurred image is expressed in terms of an encoded word using 64 single-precision float variables and represented by its position in space, the scale level at which the gradient is the maximum (or minimum), and its dominant direction.
[43]	SVM	Electrical transformers	Image	A 2048-dimensional feature vector is formed for each image, obtained by 32 feature maps with 16 (4 × 4) region points.
[44]	SVM	CFRP plate	Image	Not mentioned
[45]	GAN	CFRP plate	Image	Total of 308 × 212 pixels were selected as the region of interest.
[18]	SVM MLP RF	FEM in AM	Signal	Not mentioned
[40]	KNN	CFRP	Signal	Not mentioned
[85]	CNN& DNN	CFRP plate	Signal	Thermal response at a spatial resolution of 315 × 317 pixels in all thermograms.
[52]	CNN-RF and SVM	High-voltage power transformer	Image	Features were learned using an innovative deep learning approach.
[74]	UL	AM metallic specimen	Image	Compression of thermography data cube (720 × 864 × 1200) matrix with 16-bit elements.
[51]	ANN MLP	Substation	Signal	The seven first-order features are mean, variance, standard deviation, skewness, kurtosis, energy, and entropy.
[53]	SVM CDT LDA	Induction motor	Signal	Feature values: mean, kurtosis, energy, standard deviation, entropy, and skewness.
[55]	SVM	Composite sample	Image	Not mentioned
[33]	VGG16 CNN	Calcium fluoride	Image	Not mentioned
[38]	MLP KNNDT RF ET	Steel plate	Signal	The change in the highest-temperature point on the defect.
[64]	CNN	Laser welding	Image	20 features, describing the specimen's response to laser heat welding excitation, both in time and space.
[66]	DEEPLAB-V3+ CNN	CFRP	Image	3 channels and 30 channel images.
[67]	BI-LSTM	Steel flat-bottom hole	Image and Signal	Time series data.
[57]	SVM	Metal	Image	Pixels of the image.
[65]	DNN	Five different metal alloys	Image and Signal	Not mentioned

Table A2. *Cont.*

Ref.	Machine Learning Model	Experimental Object	Data Type	Features
[85]	CNN	Mild steel specimen	Signal	Linear fitted temporal thermal profile
[27]	SVM ANN	FEA plant root	Signal	Temperature data from the cooling process
[34]	ANN	Thermal barrier coating	Signal	Temperature evolution for each instant of the acquisition time
[62]	SVM KNN LDA DT	Electronic	Signal	Noisy data of anomaly models
[18]	SVM	Nylon and PLA	Signal	Regions of interest (ROI) of 4 × 4 pixels, one for the defect area and the other for a non-defect area
[44]	SVM	CFRP	Signal	Not mentioned
[70]	CNN (ResNet50)	CFRP	Image	Not mentioned
[93]	KNN	Composite sample	Image	Damage contour in image
[61]	RNN	Plexiglass	Signal	Temporal evolutions occurring at specified points on the investigated surface
[71]	SVM ANN	Wheat leaf	Signal	Wheat leaf nitrogen content
[60]	Based on regressive neural network	Plexiglass	Signal	Temporal evolution of temperature rise recorded at sampling instants
[12]	CNN	Plexiglass, CFRP and steel	Image	Defect localized and segmented by the mask
[69]	RCNN	Flat-bottom hole defects	Image	Region of interest from the raw image
[72]	ANN	CFRP	Image	Not mentioned
[73]	MLP	Switchboards of different buildings	Image	Regions of interest (ROI) for both reference and hot components
[49]	ANN	CFRP	Image	Thermal pattern obtained from pseudo-static sequence is divided into small sections
[95]	Hybrid neural algorithm	Surface of a coated sample	Signal	Not mentioned
[48]	RCNN	Fiberboards	Image	Not mentioned
[74]	TBSS and STSDL	SS-316L plates and one INC718 plate	Image	Condensed 2D data from image pixels
[69]	GAN	CFRR	Image	ROI containing 308 × 212 pixels
[56]	LSTM-RNN	GFRP	Signal	Not mentioned
[37]	GNB	Stator Imbalance	Signal	Selection of the necessary feature vector by PCA
[76]	ERF EDT ELR	Breast cancer image	Image	Extracted by combining ResNet101 transfer learning and VGG16 transfer learning with their combined functionalities
[79]	GAN	CFRP	Image	Original image resized to 256 × 256 pixels
[46]	GAN	CFRP	Image	An 308 × 212 pixel thermogram was selected as the region of interest
[78]	GAN	CFRP	Image	Images of 64 × 64, 128 × 128, and 256 × 256 pixels compressed from the original image

Table A2. *Cont.*

Ref.	Machine Learning Model	Experimental Object	Data Type	Features
[77]	CNN	Eight wood species within the Fabaceae family	Signal	Data matrix extracted by matrix
[81]	SK-means	CFRP	Image	Image after TSR processing
[80]	U-Net	Composite film material plate	Image	Grayscale image
[82]	CNN	Acrylonitrile butadiene styrene (ABS) polymer	Image	Defect area and sound area images cropped from the original image

References

1. Liu, Y.; Wu, J.Y.; Liu, K.; Wen, H.L.; Yao, Y.; Sfarra, S.; Zhao, C. Independent component thermography for non-destructive testing of defects in polymer composites. *Meas. Sci. Technol.* **2019**, *30*, 044006. [CrossRef]
2. Huang, J.; Wei, Q.; Zhuo, L.; Zhu, J.; Li, C.; Wang, Z. Detection and quantification of artificial delaminations in CFRP composites using ultrasonic thermography. *Infrared Phys. Technol.* **2023**, *130*, 104579. [CrossRef]
3. Ali, Z.; Addepalli, S.; Zhao, Y. Thermal Material Property Evaluation Using through Transmission Thermography: A Systematic Review of the Current State-of-the-Art. *Appl. Sci.* **2024**, *14*, 6791. [CrossRef]
4. Wu, D.; Busse, G. Lock-in thermography for nondestructive evaluation of materials. *Rev. Générale Therm.* **1998**, *37*, 693–703. [CrossRef]
5. Jones, R.; Lo, M.; Dorman, M.; Bowler, A.; Roles, D.; Wade, S. Lock-in Thermography for the Wide Area Detection of Paint Degradation and Incipient Corrosion. In *Non-Destructive Evaluation of Corrosion and Corrosion-Assisted Cracking*; John Wiley & Sons, Ltd.: New York, NY, USA, 2019; Chapter 5, pp. 122–159. [CrossRef]
6. *EN 17119*; Non-Destructive Testing—Thermographic Testing—Active Thermography. European Committee for Standardization: Brussels, Belgium, 2018; p. 15.
7. Schmid, S.; Reinhardt, J.; Grosse, C.U. Spatial and temporal deep learning for defect detection with lock-in thermography. *NDT E Int.* **2024**, *143*, 103063. [CrossRef]
8. Oliveira, B.C.F.; Seibert, A.A.; Borges, V.K.; Albertazzi, A.J.; Schmitt, R.H. Employing a U-net convolutional neural network for segmenting impact damages in optical lock-in thermography images of CFRP plates. *Nondestruct. Test. Eval.* **2021**, *36*, 440–458. [CrossRef]
9. Cheng, L.; Tian, G.Y. Comparison of Nondestructive Testing Methods on Detection of Delaminations in Composites. *J. Sens.* **2012**, *2012*, 408437. [CrossRef]
10. Chung, Y.; Lee, S.; Kim, W. Latest Advances in Common Signal Processing of Pulsed Thermography for Enhanced Detectability: A Review. *Appl. Sci.* **2021**, *11*, 12168. [CrossRef]
11. Garcia Rosa, R.; Barella, B.P.; Garcia Vargas, I.; Tarpani, J.R.; Herrmann, H.; Fernandes, H. Advanced Thermal Imaging Processing and Deep Learning Integration for Enhanced Defect Detection in Carbon Fiber-Reinforced Polymer Laminates. *Materials* **2025**, *18*, 1448. [CrossRef]
12. Fang, Q.; Ibarra-Castanedo, C.; Garrido, I.; Duan, Y.; Maldague, X. Automatic Detection and Identification of Defects by Deep Learning Algorithms from Pulsed Thermography Data. *Sensors* **2023**, *23*, 4444. [CrossRef]
13. Mulaveesala, R.; Tuli, S. Theory of frequency modulated thermal wave imaging for nondestructive subsurface defect detection. *Appl. Phys. Lett.* **2006**, *89*, 191913. [CrossRef]
14. Vesala, G.T.; Ghali, V.S.; Subhani, S.; Vijaya Lakshmi, A.; Naik, R.B. Convolution Neural Networks Based Automatic Subsurface Anomaly Detection and Characterization in Quadratic Frequency Modulated Thermal Wave Imaging. *SN Comput. Sci.* **2022**, *3*, 219. [CrossRef]
15. Maldague, X.; Marinetti, S. Pulse phase infrared thermography. *J. Appl. Phys.* **1996**, *79*, 2694–2698. [CrossRef]
16. Kuzy, J.; Li, C. A Pulsed Thermographic Imaging System for Detection and Identification of Cotton Foreign Matter. *Sensors* **2017**, *17*, 518. [CrossRef] [PubMed]

17. Kamińska, P.; Ziemkiewcz, J.; Synaszko, P.; Dragan, K. Comparison of Pulse Thermography (PT) and Step Heating (SH) Thermography in Non-Destructive Testing of Unidirectional GFRP Composites. *Fatigue Aircr. Struct.* **2019**, *2019*, 87–102. [CrossRef]
18. Rodríguez-Martín, M.; Fueyo, J.; Pisonero, J.; López-Rebollo, J.; Gonzalez-Aguilera, D.; García-Martín, R.; Madruga, F. Step heating thermography supported by machine learning and simulation for internal defect size measurement in additive manufacturing. *Measurement* **2022**, *205*, 112140. [CrossRef]
19. Almond, D.P.; Angioni, S.L.; Pickering, S.G. Long pulse excitation thermographic non-destructive evaluation. *NDT E Int.* **2017**, *87*, 7–14. [CrossRef]
20. Marani, R.; Campos-Delgado, D.U. Depth classification of defects in composite materials by long-pulsed thermography and blind linear unmixing. *Compos. Part B Eng.* **2023**, *248*, 110359. [CrossRef]
21. Muzika, L.; Svantner, M.; Honner, M.; Houdkova, S. Laser Scanning Thermography for Coating Thickness Inspection. *Eng. Proc.* **2021**, *8*, 17. [CrossRef]
22. Mund, M.; Blass, D.; Dilger, K. Laser excited Thermography—Simulation based Determination of detection thresholds in aluminum welds depending on geometrical and excitation Properties. In Proceedings of the 12th European Conference on Non-Destructive Testing (ECNDT 2018), Gothenburg, Sweden, 11–15 June 2018.
23. Li, T.; Almond, D.P.; Rees, D.S. Crack imaging by scanning pulsed laser spot thermography. *NDT E Int.* **2011**, *44*, 216–225. [CrossRef]
24. Aujeszky, T.; Korres, G.; Eid, M. Material classification with laser thermography and machine learning. *Quant. InfraRed Thermogr. J.* **2019**, *16*, 181–202. [CrossRef]
25. Roemer, J.; Pieczonka, L.; Uhl, T. Laser spot thermography of welded joints. *Diagnostyka* **2014**, *15*, 43–49.
26. Roemer, J.; Uhl, T.; Pieczonka, L. Laser Spot Thermography for Crack Detection in Aluminum Structures. In Proceedings of the 7th International Symposium on NDT in Aerospace, Bremen, Germany, 16–18 November 2015.
27. Shi, X.; Hsieh, S.J. *Infrared Imaging and Machine Learning Techniques for Plant Root Location and Depth Prediction*; SPIE: Philadelphia, PA, USA, 2021; Volume 11743, p. 1174303.
28. Qiu, Q. Effect of internal defects on the thermal conductivity of fiber-reinforced polymer (FRP): A numerical study based on micro-CT based computational modeling. *Mater. Today Commun.* **2023**, *36*, 106446. [CrossRef]
29. Aldrin, J.; Coughlin, C.; Forsyth, D.; Welter, J. Progress on the Development of Automated Data Analysis Algorithms and Software for Ultrasonic Inspection of Composites. *AIP Conf. Proc.* **2014**, *1581*, 1920–1927. [CrossRef]
30. Bossi, R.H. *ASNT Industry Handbook: Aerospace NDT*; American Society for Nondestructive Testing: Columbus, OH, USA, 2014.
31. Booth, A.; Martyn-St James, M.; Clowes, M.; Sutton, A. *Systematic Approaches to a Successful Literature Review*; SAGE Publications Ltd.: London, UK, 2021.
32. Marani, R.; Palumbo, D.; Galietti, U.; Stella, E.; D'Orazio, T. Automatic detection of subsurface defects in composite materials using thermography and unsupervised machine learning. In Proceedings of the 2016 IEEE 8th International Conference on Intelligent Systems (IS), Sofia, Bulgaria, 4–6 September 2016; pp. 516–521. [CrossRef]
33. Klamert, V.; Schmid-Kietreiber, M.; Bublin, M. A deep learning approach for real time process monitoring and curling defect detection in Selective Laser Sintering by infrared thermography and convolutional neural networks. In Proceedings of the 12th CIRP Conference on Photonic Technologies [LANE 2022], Furth, Germany, 4–8 September 2022; Volume 111, pp. 317–320. [CrossRef]
34. Halloua, H.; Elhassnaoui, A.; Saifi, A.; Elamiri, A.; Obbadi, A.; Errami, Y.; Sahnoun, S. *Neural Networks and Genetic Algorithms for the Evaluation of Coatings Thicknesses in Thermal Barriers by Infrared Thermography Data*; Elsevier: Amesterdam, The Netherlands, 2017; Volume 5, pp. 997–1004. [CrossRef]
35. Zhan, Y.; Hu, D.; Wang, Y.; Yu, X. Semisupervised Hyperspectral Image Classification Based on Generative Adversarial Networks. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 212–216. [CrossRef]
36. Shakeel, P.M.; Tobely, T.E.E.; Al-Feel, H.; Manogaran, G.; Baskar, S. Neural Network Based Brain Tumor Detection Using Wireless Infrared Imaging Sensor. *IEEE Access* **2019**, *7*, 5577–5588. [CrossRef]
37. Idrissi, A.E.; Derouich, A.; Mahfoud, S.; Ouanjli, N.E.; Byou, A.; Banakhr, F.A.; Mosaad, M.I. Stator Imbalance Defects Diagnosis of Induction Machine Using Thermography and Machine Learning Algorithms. *IEEE Access* **2024**, *12*, 51606–51618. [CrossRef]
38. Chen, H.; Zhang, Z.; Zhao, C.; Liu, J.; Yin, W.; Li, Y.; Wang, F.; Li, C.; Lin, Z. Depth Classification of Defects Based on Neural Architecture Search. *IEEE Access* **2021**, *9*, 73424–73432. [CrossRef]
39. Zhang, X.; Saniie, J.; Bakhtiari, S.; Heifetz, A. Compression of Pulsed Infrared Thermography Data with Unsupervised Learning for Nondestructive Evaluation of Additively Manufactured Metals. *IEEE Access* **2022**, *10*, 9094–9107. [CrossRef]
40. Daghigh, V.; Naraghi, M. Machine learning-based defect characterization in anisotropic materials with IR-thermography synthetic data. *Compos. Sci. Technol.* **2023**, *233*, 109882. [CrossRef]
41. Atwya, M.; Panoutsos, G. Transient Thermography for Flaw Detection in Friction Stir Welding: A Machine Learning Approach. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4423–4435. [CrossRef]

42. Gupta, K.K.; Vijay, R.; Pahadiya, P.; Saxena, S.; Gupta, M. Novel Feature Selection Using Machine Learning Algorithm for Breast Cancer Screening of Thermography Images. *Wirel. Pers. Commun.* **2023**, *131*, 1929–1956. [CrossRef]
43. Mahami, A.; Rahmoune, C.; Zair, M.; Bettahar, T.; Benazzouz, D. Automated Transformer fault diagnosis using infrared thermography imaging, GIST and machine learning technique. *Proc. Inst. Mech. Eng. Part E J. Process Mech. Eng.* **2022**, *236*, 1747–1757. [CrossRef]
44. Parvez M, M.; Mishra, S.K.; Nandini, K.; Ahammad, S.H.; Inthiyaz, S.; Altahan, B.R.; Smirani, L.K.; Hossain, M.A.; Rashed, A.N.Z. Machine learning based models for defect detection in composites inspected by Barker coded thermography: A qualitative analysis. *Adv. Eng. Softw.* **2023**, *178*, 103425. [CrossRef]
45. Liu, K.; Ma, Z.; Liu, Y.; Yang, J.; Yao, Y. Enhanced Defect Detection in Carbon Fiber Reinforced Polymer Composites via Generative Kernel Principal Component Thermography. *Polymers* **2021**, *13*, 825. [CrossRef]
46. Liu, K.; Wang, F.; He, Y.; Liu, Y.; Yang, J.; Yao, Y. Data-Augmented Manifold Learning Thermography for Defect Detection and Evaluation of Polymer Composites. *Polymers* **2023**, *15*, 173. [CrossRef]
47. Usamentiaga, R.; Ibarra-Castanedo, C.; Klein, M.; Maldague, X.; Peeters, J.; Sanchez-Beato, A. Nondestructive Evaluation of Carbon Fiber Bicycle Frames Using Infrared Thermography. *Sensors* **2017**, *17*, 2679. [CrossRef]
48. Garrido, I.; Erazo-Aux, J.; Lagüela, S.; Sfarra, S.; Ibarra-Castanedo, C.; Pivarčiová, E.; Gargiulo, G.; Maldague, X.; Arias, P. Introduction of Deep Learning in Thermographic Monitoring of Cultural Heritage and Improvement by Automatic Thermogram Pre-Processing Algorithms. *Sensors* **2021**, *21*, 750. [CrossRef]
49. Fernandes, H.; Zhang, H.; Figueiredo, A.; Malheiros, F.; Ignacio, L.H.; Sfarra, S.; Ibarra-Castanedo, C.; Guimaraes, G.; Maldague, X. Machine Learning and Infrared Thermography for Fiber Orientation Assessment on Randomly-Oriented Strands Parts. *Sensors* **2018**, *18*, 288. [CrossRef]
50. Vesala, G.T.; Ghali, V.S.; Lakshmi, A.V.; Naik, R.B. Deep and Handcrafted Feature Fusion for Automatic Defect Detection in Quadratic Frequency Modulated Thermal Wave Imaging. *Russ. J. Nondestruct. Test.* **2021**, *57*, 476–485. [CrossRef]
51. Ullah, I.; Yang, F.; Khan, R.; Liu, L.; Yang, H.; Gao, B.; Sun, K. Predictive Maintenance of Power Substation Equipment by Infrared Thermography Using a Machine-Learning Approach. *Energies* **2017**, *10*, 1987. [CrossRef]
52. Ullah, I.; Khan, R.U.; Yang, F.; Wuttisittikulkij, L. Deep Learning Image-Based Defect Detection in High Voltage Electrical Equipment. *Energies* **2020**, *13*, 392. [CrossRef]
53. Choudhary, A.; Goyal, D.; Letha, S.S. Infrared Thermography-Based Fault Diagnosis of Induction Motor Bearings Using Machine Learning. *IEEE Sensors J.* **2021**, *21*, 1727–1734. [CrossRef]
54. Séguin-Charbonneau, L.; Walter, J.; Théroux, L.D.; Scheed, L.; Beausoleil, A.; Masson, B. Automated defect detection for ultrasonic inspection of CFRP aircraft components. *NDT E Int.* **2021**, *122*, 102478. [CrossRef]
55. Benítez, H.D.; Loaiza, H.; Caicedo, E.; Ibarra-Castanedo, C.; Bendada, A.; Maldague, X. Defect characterization in infrared non-destructive testing with learning machines. *NDT E Int.* **2009**, *42*, 630–643. [CrossRef]
56. Hu, C.; Duan, Y.; Liu, S.; Yan, Y.; Tao, N.; Osman, A.; Ibarra-Castanedo, C.; Sfarra, S.; Chen, D.; Zhang, C. LSTM-RNN-based defect classification in honeycomb structures using infrared thermography. *Infrared Phys. Technol.* **2019**, *102*, 103032. [CrossRef]
57. Wei, Z.; Osman, A.; Gross, D.; Netzelmann, U. Artificial intelligence for defect detection in infrared images of solid oxide fuel cells. *Infrared Phys. Technol.* **2021**, *119*, 103815. [CrossRef]
58. Schmid, J.; Bicat, D.; Elfner, M.; Bauer, H.J. Improved in-situ calibration applied to infrared thermography under high angles of view. *Infrared Phys. Technol.* **2021**, *119*, 103952. [CrossRef]
59. Zhang, S.; Rao, P.; Chen, X.; Li, Y. An FCN-based transfer-learning method for spatial infrared moving-target recognition. *Infrared Phys. Technol.* **2024**, *137*, 105145. [CrossRef]
60. Dudzik, S. Analysis of the accuracy of a neural algorithm for defect depth estimation using PCA processing from active thermography data. *Infrared Phys. Technol.* **2013**, *56*, 1–7. [CrossRef]
61. Dudzik, S. Two-stage neural algorithm for defect detection and characterization uses an active thermography. *Infrared Phys. Technol.* **2015**, *71*, 187–197. [CrossRef]
62. Liu, H.; Tinsley, L.; Addepalli, S.; Liu, X.; Starr, A.; Zhao, Y. Detectability evaluation of attributes anomaly for electronic components using pulsed thermography. *Infrared Phys. Technol.* **2020**, *111*, 103513. [CrossRef]
63. Siljama, O.; Koskinen, T.; Jessen-Juhler, O.; Virkkunen, I. Automated Flaw Detection in Multi-channel Phased Array Ultrasonic Data Using Machine Learning. *J. Nondestruct. Eval.* **2021**, *40*, 67. [CrossRef]
64. Buongiorno, D.; Prunella, M.; Grossi, S.; Hussain, S.M.; Rennola, A.; Longo, N.; Stefano, G.D.; Bevilacqua, V.; Brunetti, A. Inline Defective Laser Weld Identification by Processing Thermal Image Sequences with Machine and Deep Learning Techniques. *Appl. Sci.* **2022**, *12*, 6455. [CrossRef]
65. Adamczyk, W.; Pawlak, S.; Durejko, T.; Łazińska, M.; Białecki, R.; Orlande, H.R.B.; Widuch, A.; Polański, M. A methodology to determine the thermal diffusivity of additively manufactured jet engine blades based on laser-induced temperature field. *Measurement* **2022**, *203*, 111955. [CrossRef]

66. Pedrayes, O.D.; Lema, D.G.; Usamentiaga, R.; Venegas, P.; García, D.F. Semantic segmentation for non-destructive testing with step-heating thermography for composite laminates. *Measurement* **2022**, *200*, 111653. [CrossRef]
67. Wu, Z.; Chen, S.; Feng, F.; Qi, J.; Feng, L.; Tao, N.; Zhang, C. Automatic defect detection and three-dimensional reconstruction from pulsed thermography images based on a bidirectional long-short term memory network. *Eng. Appl. Artif. Intell.* **2023**, *124*, 106574. [CrossRef]
68. Li, J.; Gopalakrishnan, K.; Piao, G.; Pacha, R.; Walia, P.; Deng, Y.; Chakrapani, S.K. Classification of adhesive bonding between thermoplastic composites using ultrasonic testing aided by machine learning. *Int. J. Adhes. Adhes.* **2023**, *125*, 103427. [CrossRef]
69. Liu, H.; Li, W.; Yang, L.; Deng, K.; Zhao, Y. Automatic reconstruction of irregular shape defects in pulsed thermography using deep learning neural network. *Neural Comput. Appl.* **2022**, *34*, 21701–21714. [CrossRef]
70. Deng, K.; Liu, H.; Yang, L.; Addepalli, S.; Zhao, Y. Classification of barely visible impact damage in composite laminates using deep learning and pulsed thermographic inspection. *Neural Comput. Appl.* **2023**, *35*, 11207–11221. [CrossRef]
71. Yao, X.; Huang, Y.; Shang, G.; Zhou, C.; Cheng, T.; Tian, Y.; Cao, W.; Zhu, Y. Evaluation of Six Algorithms to Monitor Wheat Leaf Nitrogen Concentration. *Remote Sens.* **2015**, *7*, 14939–14966. [CrossRef]
72. Balasubramaniam, K.; Sikdar, S.; Ziaja, D.; Jurek, M.; Soman, R.; Malinowski, P. A global-local damage localization and quantification approach in composite structures using ultrasonic guided waves and active infrared thermography. *Smart Mater. Struct.* **2023**, *32*, 3. [CrossRef]
73. Huda, A.S.N.; Taib, S.; Ghazali, K.H.; Jadin, M.S. A new thermographic NDT for condition monitoring of electrical components using ANN with confidence level analysis. *ISA Trans.* **2014**, *53*, 717–724. [CrossRef]
74. Zhang, X.; Saniie, J.; Cleary, W.; Heifetz, A. Quality Control of Additively Manufactured Metallic Structures with Machine Learning of Thermography Images. *JOM* **2020**, *72*, 4682–4694. [CrossRef]
75. Li, X.; Ying, X.; Zhu, W.; Liu, W.; Hou, B.; Zhou, L. Nondestructive detection and analysis based on data enhanced thermography. *Meas. Sci. Technol.* **2022**, *33*, 064006. [CrossRef]
76. Priyadarshni, V.; Sharma, S.K.; Rahmani, M.K.I.; Kaushik, B.; Almajalid, R. Machine Learning Techniques Using Deep Instinctive Encoder-Based Feature Extraction for Optimized Breast Cancer Detection. *Comput. Mater. Contin.* **2024**, *78*, 2441–2468. [CrossRef]
77. Pan, X.; Yu, Z.; Yang, Z. A Multi-Scale Convolutional Neural Network Combined with a Portable Near-Infrared Spectrometer for the Rapid, Non-Destructive Identification of Wood Species. *Forests* **2024**, *15*, 556. [CrossRef]
78. Cheng, L.; Kersemans, M. Dual-IRT-GAN: A defect-aware deep adversarial network to perform super-resolution tasks in infrared thermographic inspection. *Compos. Part B Eng.* **2022**, *247*, 110309. [CrossRef]
79. Cheng, L.; Tong, Z.; Xie, S.; Kersemans, M. IRT-GAN: A generative adversarial network with a multi-headed fusion strategy for automated defect detection in composites using infrared thermography. *Compos. Struct.* **2022**, *290*, 115543. [CrossRef]
80. Zhang, Z.; Hu, J.; Wang, R.; Chen, X.; Yang, D.; Vavilov, V.P.; Duan, Y.; Zhang, H. Automatic segmentation of microporous defects in composite film materials based on the improved attention U-Net module. *Quant. InfraRed Thermogr. J.* **2025**, *22*, 313–328. [CrossRef]
81. Liu, M.; Liu, T.; Liu, B.; Tang, H.; Zhang, H. Defect detection in CFRP combining SK-means clustering and probability of detection analysis. *Quant. InfraRed Thermogr. J.* **2025**, 1–13. [CrossRef]
82. Szymanik, B.; Psuj, G.; Hashemi, M.; Lopato, P. Detection and Identification of Defects in 3D-Printed Dielectric Structures via Thermographic Inspection and Deep Neural Networks. *Materials* **2021**, *14*, 4168. [CrossRef]
83. Dong, Y.; Zhao, B.; Yang, J.; Cao, Y.; Cao, Y. Two-stage convolutional neural network for joint removal of sensor noise and background interference in lock-in thermography. *NDT E Int.* **2023**, *137*, 102816. [CrossRef]
84. Matarrese, T.; Marani, R.; Palumbo, D.; D'Orazio, T.; Galietti, U. Influence of lock-in thermography set-up parameters on the capability of a temporal convolutional neural network to characterize defects in a CFRP. *Opt. Lasers Eng.* **2024**, *182*, 108455. [CrossRef]
85. Vesala, G.T.; Ghali, V.S.; Naik, R.B.; Lakshmi, A.V.; Suresh, B. *Deep Learning in Quadratic Frequency Modulated Thermal Wave Imaging for Automatic Defect Detection*; Springer: Singapore, 2021; pp. 433–442.
86. Helwing, R.; Mrzljak, S.; Hülsbusch, D.; Walther, F. Cycle-consistent generative adversarial networks for damage evolution analysis in fiber-reinforced polymers based on synthetic damage states. *Compos. Sci. Technol.* **2024**, *254*, 110695. [CrossRef]
87. Shepard, S.M. Advances in pulsed thermography. In Proceedings of the Thermosense XXIII, Orlando, FL, USA, 16–19 April 2001; Rozlosnik, A.E., Dinwiddie, R.B., Eds.; International Society for Optics and Photonics: San Diego, CA, USA; SPIE: Philadelphia, PA, USA, 2001; Volume 4360, pp. 511–515. [CrossRef]
88. Katunin, A.; Dragan, K.; Dziendzikowski, M. Damage identification in aircraft composite structures: A case study using various non-destructive testing techniques. *Compos. Struct.* **2015**, *127*, 1–9. [CrossRef]
89. Wang, Y.; Li, Q.; Chu, M.; Kang, X.; Liu, G. Application of infrared thermography and machine learning techniques in cattle health assessments: A review. *Biosyst. Eng.* **2023**, *230*, 361–387. [CrossRef]
90. Jalloul, R.; Krishnappa, C.H.; Agughasi, V.I.; Alkhatib, R. Enhancing Early Breast Cancer Detection with Infrared Thermography: A Comparative Evaluation of Deep Learning and Machine Learning Models. *Technologies* **2025**, *13*, 7. [CrossRef]

91. Keprate, A.; Ratnayake, R.M. Probability of Detection as a Metric for Quantifying NDE Reliability: The State of The Art. *J. Pipeline Eng.* **2015**, *14*, 199–209.
92. Martinho, M.T.; Booth, N.; Attard, N.; Dillenburger, K. A systematic review of the impact of precision teaching and fluency-building on teaching children diagnosed with autism. *Int. J. Educ. Res.* **2022**, *116*, 102076. [CrossRef]
93. Zhou, J.; Du, W.; Yang, L.; Deng, K.; Addepalli, S.; Zhao, Y. Pattern Recognition of Barely Visible Impact Damage in Carbon Composites Using Pulsed Thermography. *IEEE Trans. Ind. Inform.* **2022**, *18*, 7252–7261. [CrossRef]
94. Virkkunen, I.; Koskinen, T.; Jessen-Juhler, O.; Rinta-aho, J. Augmented Ultrasonic Data for Machine Learning. *J. Nondestruct. Eval.* **2021**, *40*, 4. [CrossRef]
95. Halloua, H.; Saifi, A.; El-amiri, A.; Obbadi, A.; Errami, Y.; Sahnoun, S.; Elhassnaoui, A. The Processing of Active Infrared Thermography Data by a Hybrid Neural Algorithm for the Evaluation of Thermal Barrier Coating Thicknesses. In Proceedings of the 2018 6th International Renewable and Sustainable Energy Conference (IRSEC), Rabat, Morocco, 5–8 December 2018; pp. 1–6. [CrossRef]

Machine learning in thermography non-destructive testing: a systematic review

Peng, Shaoyang

2025-09-01

Peng S, Addepalli S, Farsi M. (2025) Machine learning in thermography non-destructive testing: a systematic review. Applied Sciences, Volume 15, Issue 17, September 2025, Article number 9624
https://doi.org/10.3390/app15179624