

A Bibliometric Study on the Applications of Neural Networks in Metal Surface Defect Inspection

Um Estudo Bibliométrico sobre as Aplicações de Redes Neurais na Inspeção de Defeitos em Superfícies Metálicas

Sandro Bessa de Lima^{1*}, Marcos dos Santos¹, Júlio César Duarte¹

Abstract: The transformation industries are tasked with the process of converting vast amounts of chemical compounds into metal sheets. By applying artificial intelligence to detect anomalies and enhance the scope of visual inspection on metal surfaces, these industries have significantly advanced their digital transformation. In this research, we explore and highlight extensively used neural network architectures, since, despite the existing literature, recent bibliometric studies on this important topic remain scarce. Keeping this in mind, this work presents an analysis of a quantitative approach based on bibliometric metrics to evaluate the application of neural network architectures aimed at inspecting defects on metal alloy surfaces, using data from the Scopus and Web of Science databases as references. It primarily focuses on identifying aspects related to the research in this field, such as the distribution of scientific publications, the most relevant sources, the interaction dynamics among these sources, the most prominent subjects in recent research (trend topics), and the application of machine learning techniques to the detection of metal surfaces' failures. Finally, our results show a lack of relevant public datasets of adequate size, as well as structured methods to apply the combined neural network architectures and techniques of machine learning in real-world applications within the scope of this research.

Keywords: Machine Learning — Convolutional Neural Networks — Segmentation — Transfer Learning — Data Augmentation — Bibliometrix — Computer Vision — Metal Surfaces

Resumo: As indústrias de transformação são encarregadas do processo de conversão de grandes quantidades de compostos químicos em chapas metálicas. Ao aplicar inteligência artificial para detectar anomalias e ampliar o escopo da inspeção visual em superfícies metálicas, essas indústrias avançaram significativamente em sua transformação digital. Nesta pesquisa, exploramos e destacamos arquiteturas de redes neurais amplamente utilizadas, uma vez que, apesar da literatura existente, estudos bibliométricos recentes sobre esse tema relevante ainda são escassos. Com isso em mente, este trabalho apresenta uma análise de uma abordagem quantitativa baseada em métricas bibliométricas para avaliar a aplicação de arquiteturas de redes neurais voltadas para a inspeção de defeitos em superfícies de ligas metálicas, utilizando dados das bases Scopus e Web of Science como referência. O foco principal do trabalho é identificar aspectos relacionados à pesquisa nessa área, como a distribuição de publicações científicas, as fontes mais relevantes, a dinâmica de interação entre essas fontes, os temas mais proeminentes na pesquisa recente (temas em tendência) e a aplicação de técnicas de aprendizado de máquina na detecção de falhas em superfícies metálicas. Finalmente, nossos resultados mostram uma escassez de conjuntos de dados públicos relevantes e de tamanho adequado, bem como de métodos estruturados para aplicar arquiteturas de redes neurais combinadas e técnicas de aprendizado de máquina em aplicações do mundo real no escopo desta pesquisa.

Palavras-Chave: Aprendizado de Máquina — Redes Neurais Convolucionais — Segmentação — Aprendizado por Transferência — *Data Augmentation* — Bibliometrix — Visão Computacional — Superfícies Metálicas

¹ Post Graduation Program in Systems and Computing, Military Institute of Engineering (IME), Brazil

*Corresponding author: lima.sandro@ime.eb.br

DOI: <http://dx.doi.org/10.22456/2175-2745.142111> • Received: 26/08/2024 • Accepted: 23/02/2025

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

The production process of metallic alloys is important for various industries, such as aviation and defense [1]. For example, a small flaw in the metal surface can compromise an entire project or cause catastrophic impacts, such as an air accident. However, inspecting these flaws is still an expensive and manual process that, in most cases, is susceptible to errors.

Additionally, the application of computer vision for detecting flaws in metallic alloys has several applications in the military sector [2], such as equipment inspection, predictive maintenance, materials and welds analysis, quality control, security of critical infrastructures, disaster response, training and simulations, and perimeter security [3]. This makes the application of such techniques in the process of flaw detection in metal surfaces important for improving the safety, efficiency, and operational reliability of petroleum platforms, for instance, playing an essential role in the maintenance of equipment and the safety of critical infrastructure [4].

This inspection process includes steps such as identifying flaws like scratches, holes, dents, corrosion, and shear [5]. However, many problems are identified throughout the production chain [6], related to factors such as reliability, productivity, maintainability, the risk to the lives of inspectors, and fraud during the process of manually pointing out non-conformities identified during the inspection. These aspects can be negatively affected by several factors, including insufficient time for adequate analysis, technical and managerial problems, lack of resources, environmental and meteorological conditions, and even procedural fraud [7].

The main objective of this study is to build a bibliometric study of related research, providing the academic community with possibilities for new approaches and scenarios, not limited to the scope of this research. Although several applications for detecting flaws on metal surfaces are available in scientific articles, no studies have been found that perform a bibliometric analysis of these applications. Therefore, to address this gap in the specialized literature, this work presents a bibliometric study on applications to detect flaws in metal alloys using neural networks and other machine learning techniques. In this sense, this research aims to answer the following questions:

- How are publications distributed (over time and by publication type)?
- What are the most relevant sources (publication vehicles, articles, references, authors, institutions, and countries)?
- What are the interaction dynamics among these sources?
- What are the most prominent subjects in recent research, or trending topics?
- What is the most commonly used method for detecting failure on a metal surface?

With this analysis, this paper provides an overview of this topic with state-of-the-art scientific research that indicates the trend toward using machine learning for this problem, helping managers and researchers in the transformation industry. Based on the proposed research questions, this manuscript is structured as follows: [Section 2](#) provides a theoretical framework for the basic concepts involved in the research, [Section 3](#) details the methodology used, [Section 4](#) presents and analyzes the results obtained in the bibliometric study, [Section 5](#) concludes with our final considerations and some future works.

2. Theoretical Framework

This section covers the key concepts and methods used in this research. It outlines Digital Image Processing and Analysis, which focuses on techniques for enhancing and analyzing images; Segmentation, which involves dividing images into meaningful parts; Deep Learning, which applies neural networks to detect patterns; and Datasets, which highlights the importance of quality data for training machine learning models.

2.1 Digital Image Processing and Analysis

Digital Image Processing (DIP) uses the numerical representation of the image by applying mathematical operations to the values of its pixels to improve its manipulation and analysis. On the other hand, Digital Image Analysis (DIA) is characterized by the automatic analysis of digital images using methods that enhance visual information and facilitate human interpretation. They can be used in several fields such as medicine, geography, physics, and biology. Currently, DIP and DIA are interdependent processes, identified as Digital Image Processing and Analysis (DIPA), as outlined in [Figure 1](#).

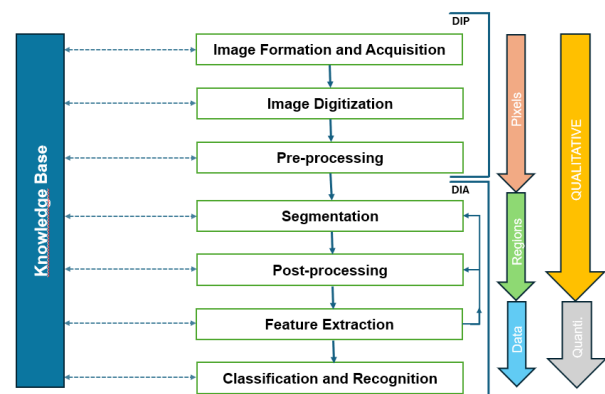


Figure 1. DIPA Sequence. Adapted from [8]

DIPA can be defined as the application of mathematical algorithms to the mathematical representations of images, thus being a specific part of Digital Signal Processing (DSP), which deals with the representation, transformation, and manipulation of signals and the information contained within them [8].

2.2 Segmentation

One of the most common tasks performed by deep learning and machine learning methods is scene classification. This process is built on the premise that an algorithm should be capable of accurately predicting the content or the underlying meaning of an image [9].

Semantic segmentation is a type of scene classification where the input is not an entire image but each pixel. In this way, the algorithms need to infer the content of each pixel in the image, which can be used to determine which objects are present in an image, their spatial localization, shapes, and other detailed information [10]. For instance, in Figure 2, semantic segmentation was applied to perform pixel-level scene analysis.



Figure 2. Segmentation Example. Source: Hazirbas[11]

2.3 Deep Learning

Deep Learning is a field of machine learning that is inspired by how living beings represent and interpret visual information [12]. In deep learning, abstraction is performed through multiple learning layers, with each layer progressively extracting simpler and more abstract representations of the data or features. The most commonly used deep learning model for computer vision is the Convolutional Neural Network (CNN), especially for tasks such as image recognition and analysis.

2.3.1 Convolutional Neural Networks

The human brain consists of over 10 billion neurons, which constantly communicate with one another through electrical impulses transmitted via synapses, enabling complex functions and cognitive abilities [13]. Inspired by this concept, neural Networks are machine learning computational models with the ability to recognize different types of patterns, such as images, words, sounds, and videos [13].

2.3.2 Data Augmentation

CNNs are excellent for the context of image classification; however, these networks require a large amount of data to provide reliable results and avoid overfitting, which occurs when the model has learned excessively from the training data but does not accurately predict unseen data. Considering that the availability of labeled data is one of the main constraints in industrial failure problems, data augmentation is an effective and common technique that enables the expansion of the training dataset, mitigating overfitting. Thus, using data augmentation techniques becomes necessary for building deep learning pipelines.

There are many data augmentation techniques, such as geometric transformations, color spaces, and other methods

across different taxonomies [14]. For basic image manipulations, several categories encompass specific data augmentation techniques, as depicted in Figure 3:

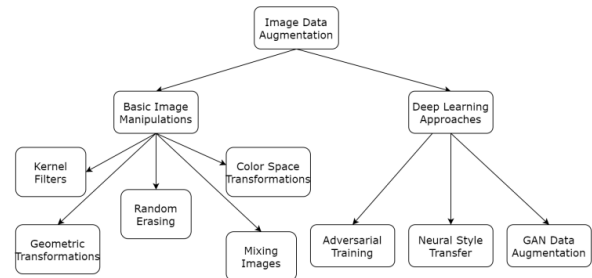


Figure 3. Data Augmentation. Source: [14]

For instance, Farady et al.[15] employed data augmentation to increase the image dataset for corrosive pattern detection by adjusting brightness, rotation, and image movement, applying the following effects: dark spots, scratches, and dents, as well as performing image expansion [16]. The effectiveness of data augmentation in image classification can be demonstrated in the learning model, generally showing better results when manipulating the brightness of the image [17].

2.3.3 Evaluation Metrics

Fundamentally, performance evaluation for classifiers seeks to verify if the model is underfitting or overfitting. Underfitting occurs when the classifier has not sufficiently trained, reflecting an inadequate level of learning that needs improvement, as it fails to generalize effectively in practice. A recent definition of overfitting in machine learning describes it as a situation where the model excessively fits the training data, capturing both the signal and noise. This leads to a specific adaptation that undermines the model's ability to generalize to new, unseen data. Overfitting is particularly problematic in complex models, as they can adapt to noisy data without performing well in broader test scenarios [18].

In the context of evaluating the performance of machine learning models, several key metrics are used to quantify their effectiveness. These include accuracy (1), precision (2), recall (3), and F1-score (4), which provide insights into different aspects of a model's performance, especially in binary classification tasks. As follows, the formulas for each of these metrics are presented:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (3)$$

$$f1\text{-score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

It is often convenient to combine precision and recall into a single metric called the F1-score, especially when it is necessary to compare different classifiers.

In studies involving corrosion detection in metals, the literature commonly addresses some evaluation metrics associated with convolutional models for detecting corrosive patterns, such as those presented.

3. Methodology

To evaluate the application of neural network architectures aimed at inspecting defects on metal alloy surfaces and to answer the research questions, a systematic literature review was carried out as proposed by Thomé e Scavarda[19], adopting a quantitative approach based on bibliometric analysis [20]. Therefore, the following six steps were applied:

1. **Planning:** Including the conception of the scope of bibliometric research based on the definition and choice of the Databases. The Scopus and Web of Science (WoS) databases were used, as suggested by Mongeon e Paul-Hus[21].
2. **Search Strategy:** The definition of the database search parameters:

$$((\text{TITLE-ABS-KEY} ("corrosion\ detection")) \text{ OR } \text{TITLE-ABS-KEY} ("metallic\ surfaces") \text{ OR } \text{TITLE-ABS-KEY} ("corrosion\ rates")) \text{ AND } (\text{TITLE-ABS-KEY} ("transfer\ learning") \text{ OR } \text{TITLE-ABS-KEY} ("deep\ learning")) \text{ AND } \text{PUBYEAR} > 2018$$
3. **Search:** After defining the strategy, the search returned 135 documents from the Scopus database and 48 from the WoS database, totaling 183 documents.
4. **Refinement:** The initial search result was filtered to exclude documents that were not directly related to the application of detecting failures on metal surfaces. This resulted in the removal of 46 items from the initial search, narrowing down the scope of the study to 137 documents.
5. **Synthesis:** After the refinement, data from the 137 documents, obtained from the Scopus and WoS databases, were consolidated using the Bibliometrix software, an R tool for comprehensive scientific mapping analysis [22]. The synthesis process involved the elimination of 29 redundancies, resulting in a final set of 108 unique documents.
6. **Results and Analysis:** After data synthesis, the results were generated with the assistance of the Biblioshiny software [23]. The results and their analysis are presented in Section 4.

Figure 4 illustrates the co-occurrence network of terms extracted from the articles analyzed in this research, representing the connections between terms that frequently appear together in the selected publications. Each node in the network represents a key term, while the lines connecting the nodes (edges) indicate the co-occurrence of these terms within the same documents. The thickness of the edges reflects the frequency with which pairs of terms appear together, providing a visualization of the intensity of each relationship.

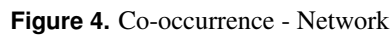
In Figure 4, the most relevant items are identified by the size of the nodes and the density of their connections to other terms. These elements represent the central and recurring themes in the co-occurrence network. The main themes are as follows:

1. **Deep Learning:** This is the most prominent term in the image, indicating that it is the central focus of the represented research. It connects various areas, including neural networks, segmentation, object detection, and corrosion analysis.
2. **Surface Defects:** Another highly relevant term, connected to "deep learning" and many other themes, indicating that surface defect detection is an important goal in the context of analysis with deep learning.
3. **Convolutional Neural Networks:** CNNs appear prominently, indicating their central role as a primary method for addressing surface analysis and detection challenges.
4. **Semantic Segmentation:** A relevant term in the image, suggesting that semantic segmentation is a widely used technique to identify and classify surface defects and corrosion.
5. **Corrosion Detection:** This specific node indicates that corrosion detection is an important theme and frequently co-occurs with other terms, such as "corrosion rate" and "corrosion evaluation."
6. **Object Detection:** With several connections, this item indicates that object detection techniques are applied in surface analysis, possibly to identify specific anomalies or defects.

These items form the research core, demonstrating the relevance of deep learning, convolutional networks, semantic segmentation, and corrosion detection as the main methods and objectives in surface defect analysis.

Figure 5 presents the co-occurrence network of terms extracted from the articles analyzed in this research, highlighting the co-occurrence relationships and the density of connections between terms frequently found together in the selected publications.

The areas of higher density reflect thematic concentration cores, suggesting subfields and recurring central topics in the field's literature. These cores highlight strong and relevant relationships between concepts, providing a structural view



analysis aids in understanding the key references underpinning the development of the field, highlighting central authors, studies, and their interrelationships.

The six steps of the employed methodology are depicted in Figure 7.

Table 1 presents the research questions, which guided the selection of articles that address them.

Another relevant point is the databases used, as in the con-

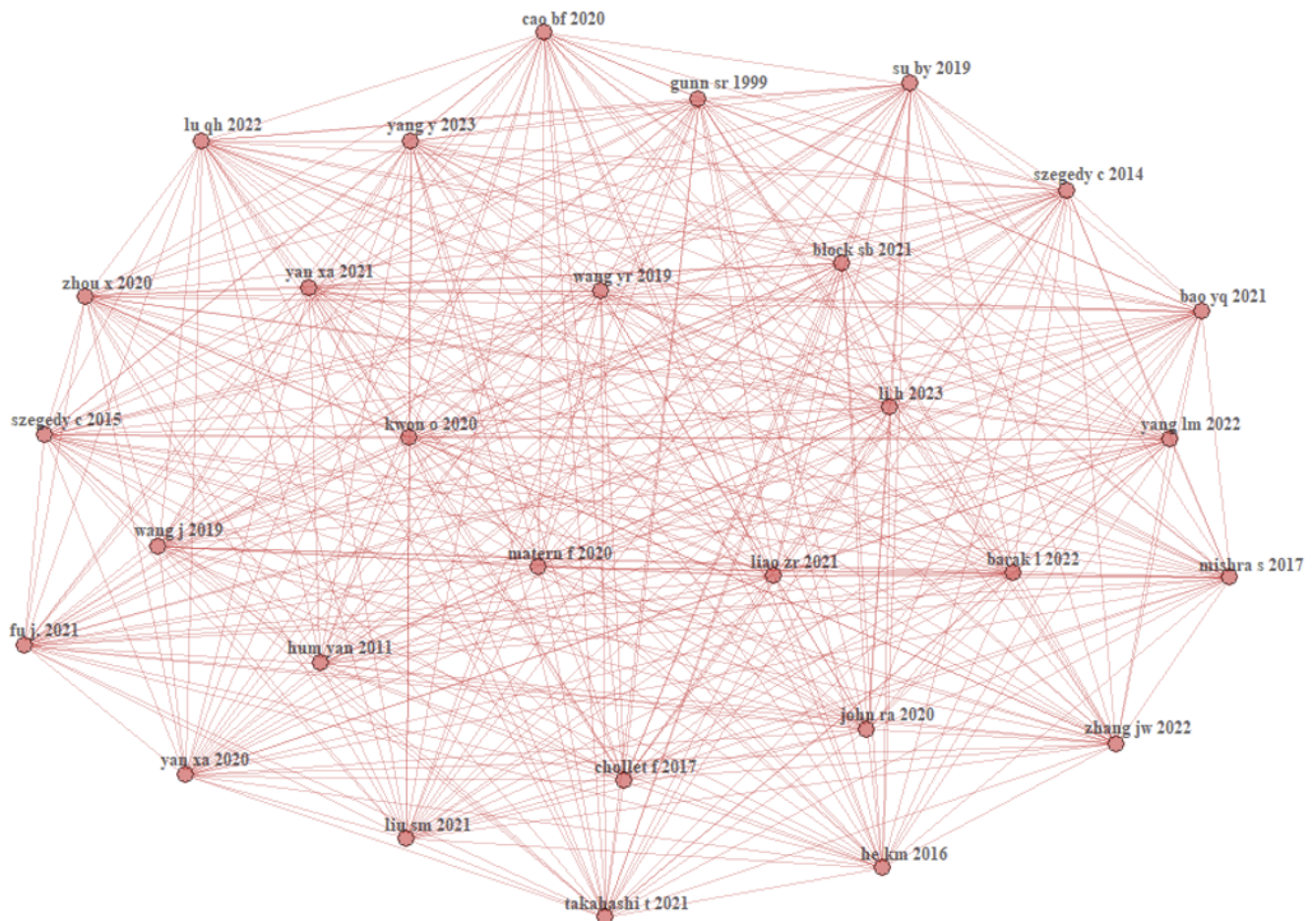


Figure 6. Co-citation - Network

Table 1. Research questions

Questions
1 Is there an efficient neural network architecture for fault detection in metal alloys on IOT devices?
2 Can transfer learning be applied to fault detection in metal alloys?
3 What computer vision techniques can be applied to detect faults in metal alloys?

text of this research, the Scopus and Web of Science databases provided the most substantial results. However, additional databases, such as IEEE Xplore and ACM Digital Library, could have also been consulted for further investigation.

3.1 Inclusion and Exclusion Criteria

To ensure the relevance and quality of the sources analyzed in this study, specific inclusion criteria were established for selecting scientific articles. These criteria guarantee that only publications aligned with the research objectives and with methodological rigor are considered. We present the main inclusion and exclusion criteria applied in selecting references, covering aspects such as the thematic scope of the research:

• Inclusion Criteria:

1. Works that address how corrosion detection is performed on metal surfaces;
2. Works that address how knowledge transfer is carried out in digital image processing;
3. Works that use various machine learning models, preferably CNNs applied to metal surfaces;
4. Works that perform image analysis based on images captured by three-dimensional photogrammetry.

• Exclusion Criteria:

1. Works whose focus does not align with the research question;
2. Works that only describe methods of capturing images;
3. Works that only describe digital image processing outside the research scope;
4. Works describing flaw detection on metal surfaces without characterizing corrosive patterns.

Regarding gray literature, sources such as technical reports, case studies, guides, and manuals were also excluded from the results.

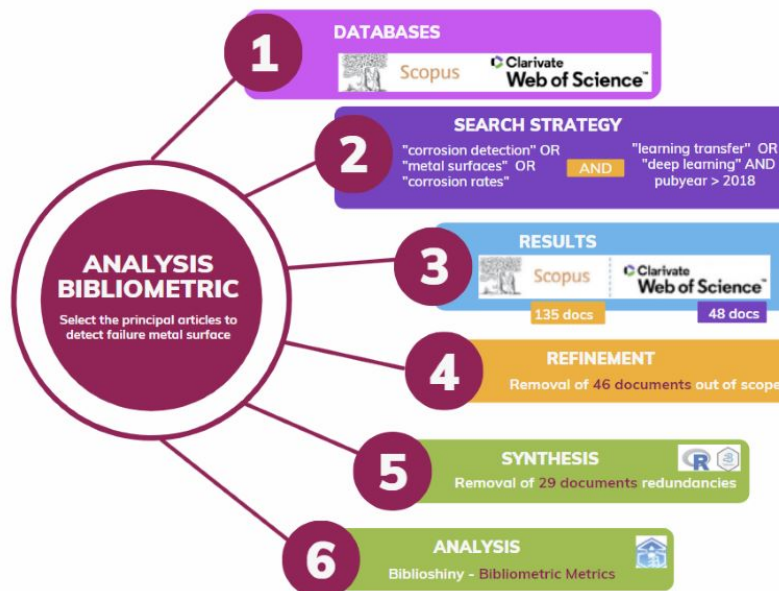


Figure 7. Summary of the methodology used in the research

4. Results and Analysis

In the analysis phase, a table was generated in Bibliometrix with data filtered by Biblioshiny, enabling the consolidation of the most relevant bibliometric information in the scope of the research on detecting flaws in metal surfaces. Using the Bibliometrix library, it is possible to import the results from Scopus and WoS into the R environment while converting them into data frames. After joining the two tables with the search data from the databases, the result is recorded in an Excel table for the final cleaning of redundant data. Once this process is carried out, data is analyzed, selected, and discussed throughout the research.

Based on the research questions, the results and their corresponding analysis were segmented into five parts: Subsection 4.2, which shows the distribution of scientific publications; Subsection 4.3, which shows the most relevant sources; Subsection 4.4, which shows the interaction Dynamics among sources; Subsection 4.5, which shows the most prominent subjects in recent research (trend topics); and Subsection 4.6, which shows the application to detect failure on metal surfaces.

One of the main challenges in applying machine learning to metal surface inspection is the lack of quality and unavailability of data. There is often a lack of public and standardized datasets, which makes it difficult to compare and validate different models. In addition, the wide variety of conditions and types of corrosion that metal surfaces can experience makes it difficult to create a representative dataset.

Another significant issue is data annotation. Manually annotating large volumes of data is both costly and time-consuming, and ensuring consistency in data annotation, especially in complex images, can also be challenging. Technically, deep learning models, such as CNNs, are complex and

require significant computational resources for training and inference. Interpreting the results of machine learning models can be difficult, particularly for personnel without a technical background.

Variable environments pose additional challenges, since changes in lighting conditions can significantly affect the accuracy of the models, and industrial environments often introduce significant noise and adverse conditions that complicate the capture of high-quality images. Moreover, integrating new machine learning systems with existing inspection systems can be technically challenging.

Despite these challenges, there are numerous opportunities for developing new applications in this field. One promising area is the development of hybrid models that combine CNNs with other machine learning techniques to improve accuracy and robustness in fault detection. Additionally, transfer learning can be applied to use pre-trained models on large datasets and adapt them to specific contexts of metal surface inspection.

Creating public and standardized datasets through academic and industrial collaboration could greatly facilitate the comparison of different methods and promote research replicability. Synthetic datasets generated using simulation techniques can complement real-world data, addressing the issue of limited data availability.

Automated annotation tools that utilize machine learning can reduce the costs and time associated with data annotation. Crowdsourcing platforms can also be employed to distribute annotation tasks among a large number of collaborators, ensuring consistency and efficiency.

Also, integrating machine learning models with IoT devices and drones offers the potential for automated and real-time inspections. This would enable continuous monitoring of metal surfaces, allowing for detecting issues before they

become critical. Such systems could be particularly beneficial in sectors like civil construction, pipelines, tanks, aeronautics, and naval industries, where early corrosion detection is essential for safety.

There is also an opportunity to develop new metrics tailored to the specific needs of the transformation sector and industrial requirements. Real-time evaluation metrics could also be developed to continuously assess the condition of metal surfaces.

CNNs usually achieve high precision and recall in defect detection on metal surfaces because they can capture detailed visual features. When trained with sufficient datasets and regularization techniques like dropout and data augmentation, CNNs tend to generalize well. They are particularly effective because they can capture detailed visual features while maintaining computational efficiency. The choice of the appropriate technique depends on the specific application context, available resources, and performance requirements.

Detecting defects on metal surfaces is a critical task in various industries, including automotive, aerospace, and manufacturing. The practical applications of defect detection techniques are vast. In the automotive industry, for instance, it can be used to inspect body panels and engine components for manufacturing defects. In aerospace, the system can ensure that critical structural elements are free from flaws that could compromise safety. In manufacturing, it can be integrated into production lines to continuously monitor product quality, reducing the reliance on manual inspection and increasing overall efficiency.

These promising research areas not only have the potential to improve the effectiveness of metal surface inspections but also promote innovation and safety across various industrial sectors.

4.1 Comparison with Other Bibliometric Studies and Literature Reviews

Comparing the results of this study on flaw detection in metal surfaces with other bibliometric studies or literature reviews in the area can provide valuable information about the similarities, differences, and possible discrepancies. The following are the main aspects discussed, compared, and highlighted, providing references where applicable.

Predominance of CNNs: Similar to other studies, this research identifies CNNs as the most commonly used architecture for defect detection on metal surfaces. For instance, a review by Bai et al.[24] found that CNNs were extensively applied in various image classification and object detection tasks, including corrosion detection on metallic surfaces [24].

Both this study and other literature reviews emphasize the role of high-quality and extensive datasets in training effective machine learning models. A study by Akbas, Arin e Onal[25] also highlighted data availability as a significant challenge in the application of machine learning for industrial inspection [25].

Focus on Hybrid Models: While this study suggests the development of hybrid models that combine CNNs with other machine-learning techniques, other reviews have predominantly focused on using standalone CNNs or DNNs. This difference highlights an area for further exploration where hybrid approaches may offer improved performance and robustness.

Integration with IoT and Drones: This study uniquely emphasizes the integration of machine learning models with IoT devices and drones for real-time and automated inspections. In contrast, other literature reviews, such as the one by Latif et al.[26], have primarily focused on the application of machine learning in static environments without considering the dynamic and real-time capabilities offered by IoT and drones [26].

4.2 Distribution of scientific publications

In terms of the timeline distribution, the research focused on works from the last 5 years. Figure 8 identifies an increase in publications over the last few years. The geographic distribution reveals that the majority of the research included in this systematic review is concentrated in China. In this particular graph, it is worth noting a growing trend in applying neural network architectures for detecting corrosion on metal surfaces.

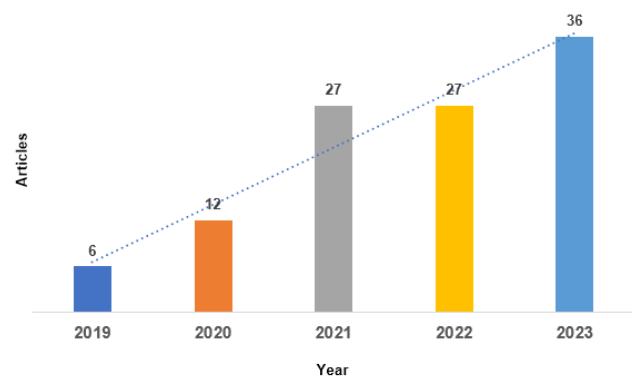


Figure 8. Distribution of scientific publications over time

As for the publication types, Table 2 presents the distribution of the documents obtained in the research by type of scientific publications.

Table 2. Distribution of documents by publication type.

Publication Type	# Documents	Percentage
Article	83	61.48%
Conference Paper	40	29.62%
Conference Review	6	4.44%
Book chapter	3	2.22%
Review	2	1.48%
Retracted	1	0.74%

Analyzing the types of documents used in research is important because it reveals the most common sources and

methods in the literature. For example, the majority of documents are articles (61.48%), indicating that this is the most frequent form of publication. Other types, such as book chapters and reviews, are less common. This suggests a need for more comprehensive reviews and the exploration of different publication formats to better consolidate existing knowledge.

The predominance of articles suggests that there is a greater focus on original studies and less on reviews that synthesize knowledge. Conversely, the low presence of reviews (1.48%) and book chapters (2.22%) indicates the need for systematic reviews to consolidate current knowledge and suggests that the field is still in a developmental phase.

4.3 Most relevant sources

The following aspects were analyzed to identify the most relevant sources: primary publication vehicles; most influential works; and high-impact authors, institutions, and countries applying neural network architectures and other techniques in machine learning. Regarding publication vehicles, Table 3 highlights the ten most relevant journals in this field of research based on the h-index. This bibliometric metric combines productivity (number of articles) with impact (number of citations) [27]. The relationship between the number of citations (NC), number of articles (NA), and number of years (NY) was used as a tiebreaker. This relation between citations, publications, and time is widely used in computing the impact factor of scientific publication vehicles [28].

Concerning the most relevant studies, Table 4 brings together ten articles within the research scope with the highest rates of global citations per year. A global citation means the total number of documents indexed in a bibliographic database that cited a paper, in the context of this research, the Scopus and WoS databases [29].

Table 5 lists the ten most prominent authors, obtained by following a similar approach to that used with publication vehicles. This analysis was based on the h-index, and the relationship between the number of citations, articles, and years of activity was used as a tiebreaker.

Table 6, on the other hand, presents the ten most influential research centers. Following a similar approach to that used with publication vehicles and authors, the analysis was based on the h-index, and the relationship between the number of citations, the number of articles, and the number of years of activity (NC/NA/NY) was used as a tiebreaker.

Finally, regarding the most influential countries, Figure 9 highlights the roles of China, India, and Australia as the most productive nations in this research area.

The result reflects the trend pointed out by the analysis of the relevance of authors and institutions, presented in Table 5 and Table 6, respectively, where it is possible to observe the prominence of these first three countries on the list.

4.4 Interaction Dynamics among sources

Considering the authors as the primary agents in the interaction dynamics among the sources, Figure 10 presents the

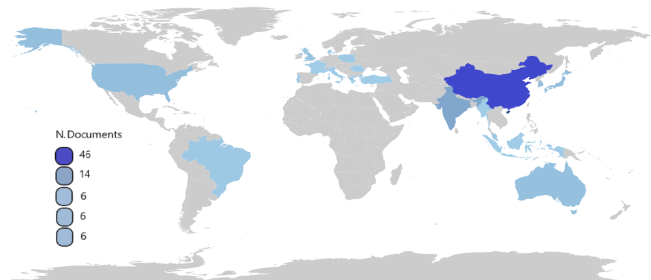


Figure 9. Relevance of Countries in terms of publications

main research clusters related to the application of detecting failures on metal surfaces.

Figure 10 highlights twelve research clusters: six in China, three in India, one in Australia, one in Greece, and one in Russia. It is noteworthy that this result reflects the trend pointed out by the analysis of the relevance of the sources presented in Figure 9, considering that the prominent clusters bring together authors, institutions, or countries of high impact in research related to the considered research area.

4.5 Most prominent subjects in recent research (trend topics)

After a global analysis of the frequency of the words presented in the title, abstract, and keywords of the papers within the scope of the study, Figure 11 shows the most prominent terms in recent research (trend topics). A growth trend is observed related to subjects associated with emphasizing failure on metal surfaces, corrosion, and segmentation.

4.6 Application to detect failure on metal surfaces

Presented here is a list of the 20 articles evaluated in Table 7, offering an overview of the type of research, neural network architectures, dataset type, machine learning strategies, evaluation metrics, and IoT approaches in the most recent and pertinent research conducted in the field of detecting flaws in metal surfaces.

Upon analyzing the technological applications aimed at detecting failures on metal surfaces, several scenarios were found in the collection of articles. Seventy-five percent of the articles are focused on corrosion. Few studies address “data segmentation” (30%), and the vast majority of research uses private datasets (75%). There is potential for further exploration of transfer learning (15%), data augmentation (25%), and IoT (25%). The most commonly used performance metrics are accuracy (35%), precision (50%), recall (25%), and F1-score (25%). Figure 12 presents the main scenarios found in the research, providing an overview of the theme.

This list encompasses approaches that include the application of various CNN models, data segmentation techniques, data augmentation methods, and industry-specific evaluation metrics. The following sections provide detailed descriptions of the neural network architectures, datasets used, and summaries of the enumerated research.

Table 3. Most relevant publication vehicles

Publication vehicle	H-Index	NC	NA	NY	NC/NA/NY
IEEE transactions	3	44	5	4	2.20
Lecture Notes in Computer Science	2	37	5	5	1.48
Journal of ambient intelligence	2	32	2	3	5.33
IEEE access	2	19	4	3	1.58
Automation and Construction	2	18	2	3	3.00
IEEE transactions on industrial	2	18	2	2	4.50
Applied sciences	2	11	3	2	1.83
Journal of building engineering	1	13	2	3	2.16
Corrosion	1	11	2	4	1.37
Lecture Notes in Electric	1	3	2	2	0.75

NC: Number Citations NA: Number Articles NY: Number Years
 NC/NA/NY: The ratio NC/NA/NY essentially measures the average number of citations per article per year

Table 4. Most relevant studies (global citations)

Reference	Title	NC	NY	NC/NY
Yun et al.[30]	Automated defect inspection system for metal surfaces based on deep learning and data augmentation	122	4	30.50
Bastian et al.[31]	Visual inspection and characterization of external corrosion in pipelines using deep neural network	70	5	14.00
Block et al.[32]	Inspection of imprint defects in stamped metal surfaces using deep learning and tracking	37	4	12.33
Katsamenis et al.[33]	Pixel level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation	31	4	7.75
Mao, Rajan e Chia[34]	Deep residual pooling network for texture recognition	31	2	10.33
Aslam, Santhi e Ramasamy[35]	Localization and segmentation of metal cracks using deep learning	30	3	10.00
Forkan et al.[36]	Corrdetector: A framework for structural corrosion detection from drone images using ensemble deep learning	28	2	14.00
Liu et al.[37]	Visual defect inspection of the metal part surface via deformable convolution and concatenate feature pyramid neural networks	25	4	6.25
Nash Liang Zheng[38]	Deep learning corrosion detection with confidence	18	2	9.00
Munawar et al.[39]	Civil infrastructure damage and corrosion detection: An application of machine learning	18	2	9.00

NC: Number Citations NA: Number Articles NY: Number Years
 NC/NY: The ratio NC/NY essentially measures the average number of citations per year

Deep Learning AI Corrosion Detection [12]:

- **Technologies:** U- Net, with 5 layers;
- **Dataset:** Private, personalized, available at corrosiondetector.com;
- **Description:** The model was generated and trained based on labeling images taken by users of corrosiondetector.com.
- **Metric:** Accuracy.

Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection [40]:

- **Technologies:** ZF-Net, Vgg16, Corrosion7, Corrosion5, Vgg15;

- **Dataset:** The dataset is copyrighted and not made available. 926 images were collected with different digital cameras. These images were cropped into regions with corrosion and without corrosion;
- **Description:** This article presents different convolutional neural network-based approaches for corrosion assessment on metallic surfaces.
- **Metric:** Recall, Precision, F1 Score.

Evaluation of Data Augmentation on Surface Defect Detection [41]

- **Technologies:** Faster R-CNN, Cascade R-CNN, ResNet-50);

Table 5. Most relevant authors

Author	Country	H-Index	NC	NA	Publication period	NY	NC/NA/NY
Liu Z	China	3	40	3	2020-2022	3	4.44
Wang Y	China	3	25	5	2020-2023	4	1.25
Birbilis N	Australia	3	39	3	2019-2022	4	3.25
Zhang Z	China	2	14	7	2021-2023	3	0.66
Wang C	China	2	8	4	2020-2023	4	0.50
Liu C	China	2	7	3	2020-2023	4	0.58
Cao Y	China	2	4	4	2021-2023	3	0.33
Chen Z N	China	2	12	3	2021-2023	3	1.33
Li C	China	1	8	4	2023-2023	1	2.00
Liu Y	China	1	1	3	2022-2023	2	0.16

NC: Number Citations

NA: Number Articles

NY: Number Years

NC/NA/NY: The ratio NC/NA/NY essentially measures the average number of citations per article per year

Table 6. Most relevant institutions

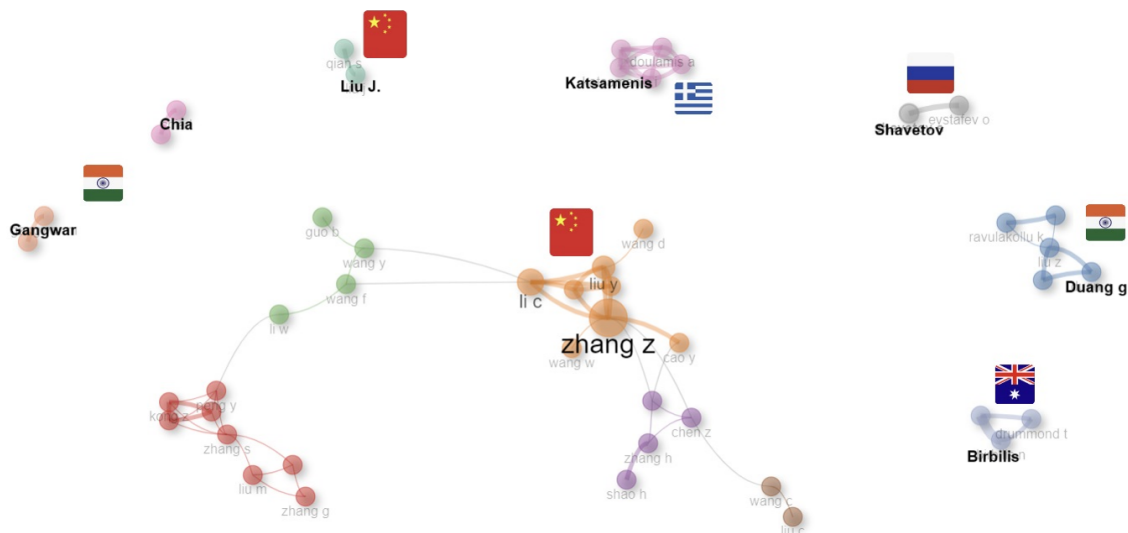
Institution	Country	H-Index	NC	NA	Publication Period	NY	NC/NA/NY
Monash Univ.	Australia	3	39	3	2020-2023	4	3.25
Zhejiang Univ.	China	2	41	5	2021-2023	3	2.73
Nanyang Univ.	Singapura	2	39	4	2019-2023	5	1.95
Australian Univ.	Australia	2	20	3	2020-2022	2	3.33
Tongji Univ.	China	2	19	3	2022-2023	2	3.16
Amity Univ.	India	2	9	4	2020-2023	4	0.56
Nanjing Univ.	China	1	3	4	2021-2021	4	0.56
Bhabha Univ.	India	1	2	3	2023-2023	1	0.66
Harbin Institute	China	1	1	3	2023-2023	1	0.33
Huazhong Univ.	China	1	5	3	2022-2023	2	0.83

NC: Number Citations

NA: Number Articles

NY: Number Years

NC/NA/NY: The ratio NC/NA/NY essentially measures the average number of citations per article per year

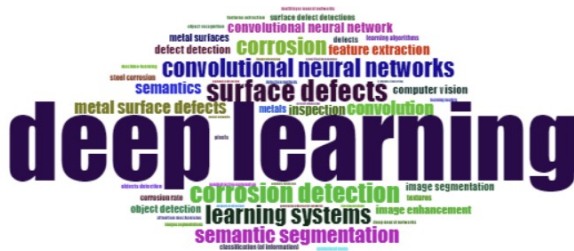
**Figure 10.** Main research clusters

- **Dataset:** Severstal steel surface public dataset from Kaggle DAGM [52] and NEU [53];
- **Description:** This article presents different convolutional neural network–based approaches for corrosion assessment on metallic surfaces.

Table 7. Overview of the most recent and pertinent research

Research	Type	Machine Learning Techniques	Data Augmentation	Dataset	Metrics	Data Segmentation	IOT	Transfer Learning
Nash, Drummond e Birbilis[12]	CO	UN		PR	AC			
Atha e Jahanshahi[40]	CO	CN VG ZN		PR	RC PN F1			
Farady et al.[41]	FL	RTRN	✓	PB	MP			
Chen et al.[42]	CO	CN		PR	PN			
Jian et al.[43]	FL	DE AN DN		PR	RC PN F1			
Xu et al.[44]	CO	RT RN VG		PR	MP			
Yun et al.[30]	CO	DN	✓	PB	PN RCF1			
Duy et al.[10]	FL	FN RN EN UN	✓	PB	IU	✓	✓	
Katsamenis et al.[33]	CO	FN RN UN		PR	PN F1	✓		
Tian et al.[45]	CO	FN RN UN		PR	PN F1			
Bastian et al.[4]	CO	CN VN ZN	✓	PR	PN RCF1		✓	
S Shukla M[46]	CO	MN RT		PR	PN			✓
Zuchniak et al.[1]	CO	CN RT		PR	PN RCF1		✓	
Cao et al.[17]	CO	YO		PR				
Dong, Gang e Zhengkun[47]	CO	SN SFL YO MN		PB	PN RC MP		✓	
Ahuja, Shukla e Ravulakollu[48]	CO	UN		PR	PN SY SE F1	✓		
Chouhad et al.[49]	FL	P R DT CN MN		PR		✓	✓	✓
Rahman, Wu e Kalfarisi[50]	CO	CN		PR		✓		
Ramani, Zhang e Kuang[2]	CO	FP DL MN RT UN		PB	PN RCMU	✓		
Li et al.[51]	FL	YO	✓	PR	MP			

CO: Corrosion FL: Failure CN: Custom CNN P: Perceptron DT: Decision Tree R: Random Forest
 FP: FPNCNN DL: Deep Lab SN: Squeeze Net SFN: Shuffle Net YO: Yolov MN: Mobile Net
 VN: VGG Net DN: DCNN RT: Resnet FN: FCN RN: RCNN PR: Private PB: Public
 AC: Accuracy PN: Precision SY: Specificity SE: Sensitivity RC: Recall F1: F1 Score
 MP: MAP MU: mLoU IU: IoU DE: Dense Net AN: Alex Net DN: Dark Net
 VG: VGG EN: E-Net UN: U-Net ZN: ZF-Net

**Figure 11.** Trending Topics

- **Metric:** MAP.

Metal Surface Corrosion Grade Estimation from Single image [42]

- **Technologies:** CNN, computer vision techniques using OpenCV);
- **Dataset:** The dataset is copyrighted and not made available. A metal surface correction dataset was created;
- **Description:** Estimation of the degree of corrosion of the

metal surface comparing the performance of computer vision techniques using OpenCV and CNN.

- **Metric:** Accuracy.

Deep Learning Model for Determining Defects of Vision Inspection Machine using only a few samples [43]

- **Technologies:** DarkNet-53, AlexNet, DenseNet-201;
- **Dataset:** The dataset is copyrighted and not made available. An image sensor was used to take photos of a specific area of the component;
- **Description:** The method was able to increase the defect recognition rate. An entropy calculation was proposed to choose the most appropriate Kernel size for the convolution layer to increase the component recognition rate and reduce the model training time.
- **Metric:** Precision, Recall, F1 Score.

Intelligent Corrosion detection and rating based on faster region-based convolutional neural network [44]

- **Technologies:** Faster-RCNN, Resnet50, VGG16;

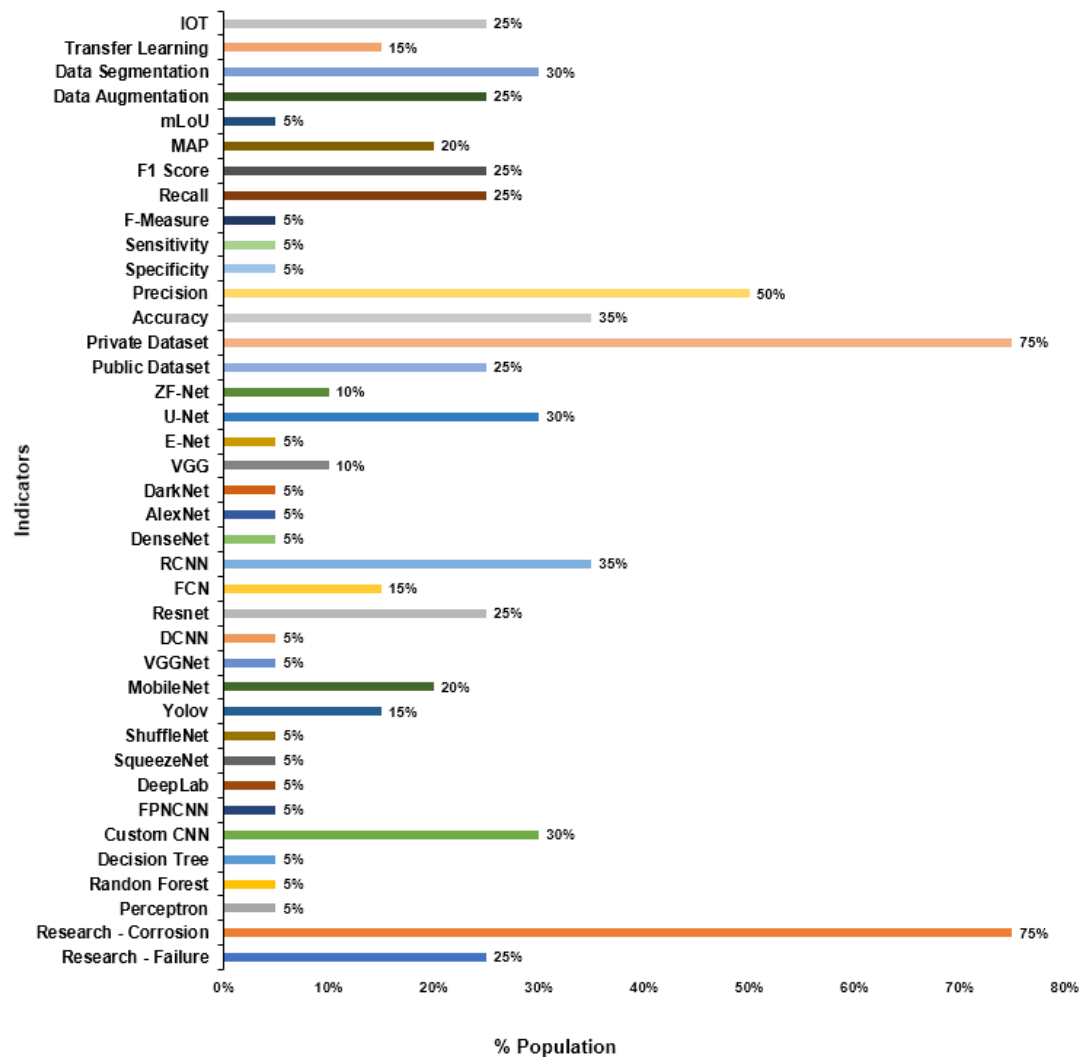


Figure 12. Overview of the most recent and pertinent research

- **Dataset:** The dataset is copyrighted and not made available. A dataset was developed from photos of metal plates with coatings exposed to the corrosive environment and regularly monitored and photographed. 538 images were collected to construct the dataset;
- **Description:** The Faster-RCNN technique was used, which accurately finds the corrosion of the metal plate covered by coatings. The experiment shows that mAP value > 94 and error rate $\leq 10\%$.
- **Metric:** MAP.

Automated Defect Inspection System for Metal Surfaces based on deep learning and Data augmentation [30]

- **Technologies:** DCNN;
- **Dataset:** Dataset was generated from surface images obtained from real metal production lines. 657 samples were used for training;

- **Description:** A novel algorithm named discriminant manifold regularized local descriptor was proposed to conduct the defect classification for steel surfaces.

- **Metric:** Precision, Recall, F1 Score.

Deep Learning in Semantic Segmentation of Rust in Images [10]

- **Technologies:** Densenet, U-Net, E-Net, FCN, RCNN;
- **Dataset:** The Dataset consists of 270 training images with random augmentation at each stage and another new set of 180 test images;
- **Description:** Three different types of corrosion detection models (DenseNet, U-Net, E-Net) were compared. The approach was focused on IOT (Drones). Semantic segmentation was applied (PSPNet) Pyramid Scene Parsing Network, performing scene analysis at the pixel level (decoder).
- **Metric:** IoU.

Pixel Level Corrosion Detection on Metal Constructions by Fusion of Deep Learning Semantic and Contour Segmentation [33]

- **Technologies:** FCN, U-Net (designed from scratch), Mask R-CNN (model trained under the COCO ensemble);
- **Dataset:** Used a real data dataset annotated by engineering experts. The dataset was captured from the European H2020 PANOPTIS project;
- **Description:** Performs data projection based on corrosion pixel regions to identify the semantics of a defect.
- **Metric:** Accuracy, F1 Score.

Corrosion Identification of Fittings Based on Computer Vision [45]

- **Technologies:** R-CNN, Faster R-CNN, FCN, U-Net;
- **Dataset:** Dataset is composed of a set of labeled data. Rust categories based on ISO 8501-1. The Dataset is obtained from heterogeneous data sources (DSLR, cameras, UAVs, cell phones), containing 116 images;
- **Description:** The HSI color feature is studied. The color space conversion method is applied to convert RGB to the HSI color model. The bottom is removed, leaving only the corrosion area. A manual labeling is applied using the LabellImg tool.
- **Metric:** Accuracy, F1 Score.

Visual Inspection and Characterization of External Corrosion in Pipelines Using Deep Neural Network [4]

- **Technologies:** Custom-CNN, ZFNet, VGGNet;
- **Dataset:** An image bank with images of pipelines in different degrees of corrosion is coated and labeled. After applying data augmentation, the dataset was left with approximately 142,400 images. The research classifies images into four classes: corrosion, low-level corrosion, medium and high-level corrosion;
- **Description:** The proposed classifier outperformed ZFNet and VGGNet for different evaluation metrics, with Custom-CNN being better than both. The method may apply to the IOT (aerial robots) scope.
- **Metric:** Recall, Precision, F1 Score.

Surface Corrosion Grade Classification using Convolution Neural Network [46]

- **Technologies:** R-CNN. The Mask R-CNN framework is based on Faster R-CNN with layer enhancements. MobileNets, ResNet;
- **Dataset:** The MS COCO dataset available for training was limited, which limits the model's performance;

- **Description:** Processing-based methodology for segmenting corrosion into different degrees using unsupervised learning was proposed. The R-CNN mask with transfer learning was used for the study.

- **Metric:** Precision, Accuracy.

Corrosion Detection on Aircraft Fuselage with Multi-Teacher Knowledge Distillation [1]

- **Technologies:** Resnet50, CustomCNN, Single, Ensemble, Averaging Student, Geometric Student Distillation, Multi-output student;
- **Dataset:** Dataset is made up of images from the DAIS system. The samples were labeled as light corrosion and moderate corrosion; The entire dataset was labeled by experts.
- **Description:** A method of aggregating and updating knowledge derived from various machine learning models was developed.
- **Metric:** Recall, Precision, Accuracy, F1 Score.

Corrosion Detection in Transformers Based on Hierarchical Annotation [17]

- **Technologies:** Faster R-CNN, YOLOV5;
- **Dataset:** The dataset is copyrighted and not made available. The dataset available for training was small, which limited the model's performance;
- **Description:** The research resulted in detection performance with the hierarchical annotation approach being better than the traditional annotation approach for the Faster R-CNN and YOLOV5 models.

- **Metric:** Recall, Precision, and mAP

Surface Corrosion Detection of Quayside Crane based on Improved MobileNetV2SSDLite [47]

- **Technologies:** MobileNetV2 SSDV2 Lite, ShuffleNetSSD, SqueezeNetSSD, MobileNetV2SSDLiteV1, MobileNetV2, YOLOV3-Tiny;
- **Dataset:** Data extracted from the LmDB dataset;
- **Description:** Comparison of neural network models suitable for IOT.

- **Metric:** Recall, Precision, and mAP

Optimized Deep Learning Framework for Detecting Pitting Corrosion based on Image Segmentation [48]

- **Technologies:** U-Net;
- **Dataset:** The dataset is copyrighted and not made available. A metal surface correction dataset was created;

- **Description:** The text addresses pitting corrosion detection by applying the residual U-Net method and then applying a Bidirectional Conv-LSTM technique to provide better classification results.

- **Metric:** Accuracy, precision, specificity, sensitivity, and F1-Score.

Smart Data-Driven Defect Detection Method for Surface Quality Control in Manufacturing [49];

- **Technologies:** Decision tree, random forest, multilayer perceptron, CNN, Mobilenet-v1;
- **Dataset:** The dataset is copyrighted and not made available. A metal surface correction dataset was created;
- **Description:** Segmentation carried out with the Weka plugin. Transfer learning was carried out with retraining of the Mobilenet-v1 model. Learning process that can be applied to IOT devices.
- **Metric:** Accuracy, Precision, Recall, F1-Score

Semantic Deep Learning Integrated with RGB Feature Based Rule Optimization for Facility Surface Corrosion Detection and Evaluation [50]

- **Technologies:** CNN
- **Dataset:** The dataset is made up of images collected by professional inspection engineers.
- **Description:** The text addresses the application of a deep learning method with semantic segmentation, in addition to applying image labeling through an unsupervised texture-based image segmentation method, integrated with a feature-based classifier (RGB).
- **Metric:** Accuracy, Precision, Recall, F1-Score, AUC-ROC

Probabilistic Assessment of Time to Cracking of Concrete Cover Due to Corrosion Using Semantic Segmentation of Imaging Probe Sensor Data [2]

- **Technologies:** DeepLabV3, resnet34, U-Net, FPNCNN, mobilenet-v1;
- **Dataset:** Dataset formed by images with different levels of corrosion being labeled using the LabelMe annotation tool. The dataset consists of 237 images, of which 102 are from corroded regions;
- **Description:** The text addresses the application of a deep learning method (DeepLabV3) with semantic segmentation.
- **Metric:** Accuracy, recall, precision, mIoU.

Deep learning based online metallic surface defect detection method for wire and arc additive manufacturing Data [51]

- **Technologies:** Yolov v4;
- **Dataset:** The dataset produced contains 760 images of welds taken by a CMOS camera, with a total of 8583 objects. These images contain data from industrial production and harsh conditions created by manual experiments;
- **Description:** You only look once (YOLO)-attention, based on YOLOv4, which achieves both fast and accurate defect detection for WAAM. The evaluation of the WAAM defect dataset shows that the model obtains a 94.5 mean average precision (mAP) with at least 42 frames per second.
- **Metric:** mAP

4.7 Possible Discrepancies

Variation in Performance Metrics: Discrepancies in performance metrics used to evaluate models can lead to different conclusions about their effectiveness. For example, while this study emphasizes metrics like accuracy, precision, recall, and F1 score, other studies might use additional metrics such as robustness, computational efficiency, and real-time applicability. These variations can cause differences in the reported performance of similar models.

Divergence in Application Contexts: Different studies may focus on varying industrial contexts and types of defects. For instance, some reviews may concentrate on corrosion detection in pipelines, while others might focus on aircraft fuselages or marine structures. These contextual differences can impact the applicability and generalization of findings across studies.

5. Conclusion

The transformation industries have accelerated their business digital transformation by using artificial intelligence to identify anomalies and gain insights into the range of visual inspection of metal surfaces. This study investigates and highlights the prevalent neural network topologies in academic literature.

Despite their significance, there is a lack of contemporary bibliometric studies on this topic. Using the Scopus and Web of Science databases as references, this research presents an examination of a quantitative technique based on bibliometric measures to evaluate the deployment of neural network designs aimed at detecting flaws on metal alloy surfaces.

For instance, studies from different areas applied data augmentation to overcome the limitations of restricted or private datasets. In the search for more accurate classifications, data segmentation techniques were also used.

However, it was observed that the vast majority of the research relied solely on using CNNs. Therefore, it is believed that there is a significant opportunity for investigations using new machine learning techniques. Additionally, there is

potential for the creation of new metrics to evaluate assets in the transformation industry.

The choice of “transfer learning” and “deep learning” in corrosion detection is based on several technical and practical reasons. Transfer learning allows the use of pre-trained models on large datasets, which can be adapted for specific corrosion detection tasks with less data and less training time. This approach is particularly advantageous when dealing with small and specific corrosion datasets, as it enables effective training by employing knowledge acquired from similar tasks.

Furthermore, pre-trained models already possess robust features that can be fine-tuned for corrosion detection, resulting in superior performance and flexibility in adjusting models for different types of corrosion and environments. Deep learning, on the other hand, offers complex modeling capabilities essential for capturing and recognizing intricate patterns in images, which is beneficial for corrosion detection where variations can be subtle. With multilayer structures, deep neural networks can learn hierarchical data representations, allowing for more precise and detailed detection.

The versatility of deep learning is evident in various architectures, such as CNN, U-Net, and Faster R-CNN, which can be applied to different aspects of corrosion detection, from semantic segmentation to the classification of corrosion grades. Additionally, deep learning can be easily integrated with IoT devices and drones for automated, real-time inspections in hard-to-reach locations.

Deep learning models have demonstrated superior performance metrics, such as high precision, recall, and F1 score, compared to traditional methods. Deep learning architectures can also be continuously enhanced with new techniques and data, ensuring the system becomes increasingly robust. Therefore, the combination of transfer learning and deep learning is strategic for corrosion detection due to its efficiency, complex modeling capabilities, and flexibility.

After the presentation and analysis of the results, the bibliometric study was able to satisfactorily answer the proposed research questions. As a limitation of the proposed method, it can be pointed out that the study is restricted to academic literature, specifically the Scopus and Web of Science databases.

It is important to note that several of these studies were exhaustively analyzed during this review [42, 46, 50]. Given that CNNs stood out in the scope of the research, a bibliometric study focused on this method would be beneficial.

Additionally, for future studies, it is recommended to carry out a broader study that includes other academic databases, such as the Institute of Electrical and Electronics Engineers (IEEE) Xplore and the Association for Computing Machinery (ACM) Digital Library.

Some possible future works include the systematization of reviews, which would consolidate findings and comparisons between different approaches and technologies, with publications also exploring book chapters and detailed reviews to strengthen existing knowledge. Another important avenue is the improvement of models for various applications, such as

developing and testing new deep learning models that combine existing techniques to enhance the detection and classification of failures on metal surfaces. This could also involve creating a large, standardized public dataset to facilitate comparison between methods and promote research applicability. Furthermore, there is potential in exploring the integration of failure detection techniques on metal surfaces with IoT and drones for automated, real-time inspections, as well as focusing on improving semantic segmentation techniques for more precise detection of failures on different types of metal surfaces.

Additionally, future research could focus on cost-benefit analyses of implementing various failure detection technologies in industrial environments, particularly studying the impact of techniques like transfer learning and data augmentation on the effectiveness of deep learning models. It would also be valuable to develop robust methods that perform well even with small datasets, which are common in industrial settings. Finally, extending the application of failure detection models beyond metallurgy to other sectors, such as construction, pipelines, aeronautics, and naval industries, could open up significant opportunities for broader impact.

References

- 1 ZUCHNIAK, K. et al. Corrosion detection on aircraft fuselage with multi-teacher knowledge distillation. In: PASZYNSKI, M. et al. (Ed.). *Computational Science – ICCS 2021*. Cham: Springer International Publishing, 2021. p. 318–332. ISBN 978-3-030-77980-1. Disponível em: https://link.springer.com/content/pdf/10.1007/978-3-030-77980-1_25.pdf?pdf=inline%20link.
- 2 RAMANI, V.; ZHANG, L.; KUANG, K. S. C. Probabilistic assessment of time to cracking of concrete cover due to corrosion using semantic segmentation of imaging probe sensor data. *Automation in Construction*, v. 132, p. 103963, 2021. ISSN 0926-5805. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0926580521004143>.
- 3 ASARI, V. K. et al. Recent progress in wide-area surveillance: protecting our pipeline infrastructure. In: LIN, Q.; ALLEBACH, J. P.; FAN, Z. (Ed.). *Imaging and Multimedia Analytics in a Web and Mobile World 2015*. 2015. v. 9408, p. 940802. Disponível em: <https://doi.org/10.1117/12.2086905>.
- 4 BASTIAN, B. T. et al. Visual inspection and characterization of external corrosion in pipelines using deep neural network. *NDT & E International*, v. 107, p. 102134, 2019. ISSN 0963-8695. Disponível em: <https://www.sciencedirect.com/science/article/pii/S096386951930060X>.
- 5 FARADY, I. et al. Evaluation of data augmentation on surface defect detection. In: *2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*. [s.n.], 2021. p. 1–2. Disponível em: <https://ieeexplore.ieee.org/document/9603212/keywords#keywords>.

- 6 YUN, J. P. et al. Automated defect inspection system for metal surfaces based on deep learning and data augmentation. *Journal of Manufacturing Systems*, v. 55, p. 317–324, 2020. ISSN 0278-6125. Disponível em: <https://www.sciencedirect.com/science/article/pii/S027861252030042X>.
- 7 BOŽIČ, J.; TABERNIK, D.; SKOČAJ, D. Mixed supervision for surface-defect detection: From weakly to fully supervised learning. *Computers in Industry*, Elsevier BV, v. 129, p. 103459, ago. 2021. ISSN 0166-3615. Disponível em: <http://dx.doi.org/10.1016/j.compind.2021.103459>.
- 8 IGNACIO; JULIANO. *Processamento e análise digital de imagens em estudos da cinética de recristalização de ligas Al-Mg-X*. 2013. 1-106 p. Dissertação (Mestrado) — Instituto de Pesquisas Energéticas e Nucleares, Universidade de São Paulo, São Paulo, 2013.
- 9 CAIO, J. *Segmentação semântica de imagens de sensoriamento remoto em cenário de conjunto aberto*. 1–98 p. Dissertação (Mestrado) — Universidade Federal de Minas Gerais (UFMG), 2019. Dissertação de Mestrado.
- 10 DUY, L. D. et al. Deep learning in semantic segmentation of rust in images. In: *Proceedings of the 2020 9th International Conference on Software and Computer Applications*. New York, NY, USA: Association for Computing Machinery, 2020. (ICSCA '20), p. 129–132. ISBN 9781450376655. Disponível em: <https://doi.org/10.1145/3384544.3384606>.
- 11 HAZIRBAS, C. *Feature Selection and Learning for Semantic Segmentation*. 1-80 p. Dissertação (Mestrado) — Technical University Munich, Germany, June 2014.
- 12 NASH, W.; DRUMMOND, T.; BIRBILIS, N. Deep Learning AI for Corrosion Detection. In: ONEPETRO (Ed.). *Corrosion 2019*. [s.n.], 2019. (NACE CORROSION, All Days), p. NACE-2019–13267. Disponível em: <https://onepetro.org/NACECORR/proceedings-abstract/CORR19/All-CORR19/NACE-2019-13267/127373>.
- 13 ABRAHAM, A. Adaptation of fuzzy inference system using neural learning. In: _____. *Fuzzy Systems Engineering: Theory and Practice*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. p. 53–83. ISBN 978-3-540-32397-6. Disponível em: https://doi.org/10.1007/11339366_3.
- 14 SHORTEN, T. M. K. C. A survey on image data augmentation for deep learning”, journal of big data. *Journal of Big Data*, v. 6, p. 1–48, 2019.
- 15 FARADY, I. et al. Evaluation of data augmentation on surface defect detection. In: . [S.l.: s.n.], 2021. p. 1–2.
- 16 KATSAMENIS, I. et al. Pixel-level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation. In: BEBIS, G. et al. (Ed.). *Advances in Visual Computing*. Cham: Springer International Publishing, 2020. p. 160–169.
- 17 CAO, Y. et al. Corrosion detection in transformers based on hierarchical annotation. In: SUN, X. et al. (Ed.). *Artificial Intelligence and Security*. Cham: Springer International Publishing, 2021. p. 575–584. ISBN 978-3-030-78609-0. Disponível em: https://link.springer.com/content/pdf/10.1007/978-3-030-78609-0_49.pdf?pdf=inline%20link.
- 18 BERRADA, L.; BENGIO, Y.; LAJOIE, G. A meta-analysis of overfitting in machine learning. In: *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*. [s.n.], 2022. Disponível em: <https://proceedings.neurips.cc/paper/2022>.
- 19 THOMÉ, L. F. S. A. M. T.; SCAVARDA, A. J. Conducting systematic literature review in operations management. *Production Planning & Control*, Taylor & Francis, v. 27, n. 5, p. 408–420, 2016. Disponível em: <https://doi.org/10.1080/09537287.2015.1129464>.
- 20 MIGUEL, P. A. C. et al. *Metodologia de pesquisa em engenharia de produção e gestão de operações*. [S.l.], 2018. v. 3rd ed.
- 21 MONGEON, P.; PAUL-HUS, A. The journal coverage of web of science and scopus: a comparative analysis. *Scientometrics*, Springer Science and Business Media LLC, v. 106, n. 1, p. 213–228, out. 2015. ISSN 1588-2861. Disponível em: <http://dx.doi.org/10.1007/s11192-015-1765-5>.
- 22 ARIA, M.; CUCCURULLO, C. bibliometrix: An r-tool for comprehensive science mapping analysis. *Journal of Informetrics*, v. 11, n. 4, p. 959–975, 2017. ISSN 1751-1577. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157717300500>.
- 23 BIBLIOMETRIX.ORG. *Biblioshiny: the shiny app for bibliometrix*. 2023. <https://www.bibliometrix.org/home/index.php/layout/biblioshiny>.
- 24 BAI, J. et al. *A Comprehensive Survey on Machine Learning Driven Material Defect Detection: Challenges, Solutions, and Future Prospects*. 2024. Disponível em: <https://arxiv.org/abs/2406.07880>.
- 25 AKBAS, C.; ARIN, I. S.; ONAL, S. *A PRISMA Driven Systematic Review of Publicly Available Datasets for Benchmark and Model Developments for Industrial Defect Detection*. 2024. Disponível em: <https://arxiv.org/abs/2406.07694>.
- 26 LATIF, S. et al. Deep learning for the industrial internet of things (iiot): A comprehensive survey of techniques, implementation frameworks, potential applications, and future directions. *Sensors*, v. 21, n. 22, 2021. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/21/22/7518>.
- 27 ALONSO, S. et al. h-index: A review focused in its variants, computation and standardization for different scientific fields. *Journal of Informetrics*, v. 3, n. 4, p. 273–289, 2009. ISSN 1751-1577. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157709000339>.

- 28 GUO, X.; LI, X.; YU, Y. Publication delay adjusted impact factor: The effect of publication delay of articles on journal impact factor. *Journal of Informetrics*, v. 15, n. 1, p. 101100, 2021. ISSN 1751-1577. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157720301474>.
- 29 YANG, J.; LIU, Z. The effect of citation behaviour on knowledge diffusion and intellectual structure. *Journal of Informetrics*, v. 16, n. 1, p. 101225, 2022. ISSN 1751-1577. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157721000961>.
- 30 YUN, J. P. et al. Automated defect inspection system for metal surfaces based on deep learning and data augmentation. *Journal of Manufacturing Systems*, v. 55, p. 317–324, 2020. ISSN 0278-6125. Disponível em: <https://www.sciencedirect.com/science/article/pii/S027861252030042X>.
- 31 BASTIAN, B. T. et al. Visual inspection and characterization of external corrosion in pipelines using deep neural network. *NDT & E International*, v. 107, p. 102134, 2019. ISSN 0963-8695. Disponível em: <https://www.sciencedirect.com/science/article/pii/S096386951930060X>.
- 32 BLOCK, S. B. et al. Inspection of imprint defects in stamped metal surfaces using deep learning and tracking. *IEEE Transactions on Industrial Electronics*, v. 68, n. 5, p. 4498–4507, 2021.
- 33 KATSAMENIS, I. et al. Pixel level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation. In: BEBIS, G. et al. (Ed.). *Advances in Visual Computing*. Cham: Springer International Publishing, 2020. p. 160–169. ISBN 978-3-030-64556-4. Disponível em: https://link.springer.com/content/pdf/10.1007/978-3-030-64556-4_13.pdf?pdf=inline%20link.
- 34 MAO, S.; RAJAN, D.; CHIA, L. T. Deep residual pooling network for texture recognition. *Pattern Recognition*, v. 112, p. 107817, 2021. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320321000042>.
- 35 ASLAM, Y.; SANTHI, N.; RAMASAMY, N. Localization and segmentation of metal cracks using deep learning. *Journal of Ambient Intelligence and Humanized Computing*, v. 12, p. 4205–4213, 2021. ISSN 1868-5145. Disponível em: <https://link.springer.com/content/pdf/10.1007/s12652-020-01803-8.pdf>.
- 36 FORKAN, A. R. M. et al. Corrdetector: A framework for structural corrosion detection from drone images using ensemble deep learning. *Expert Systems with Applications*, v. 193, p. 116461, 2022. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417421017437>.
- 37 LIU, Z. et al. Visual defect inspection of metal part surface via deformable convolution and concatenate feature pyramid neural networks. *IEEE Transactions on Instrumentation and Measurement*, v. 69, n. 12, p. 9681–9694, 2020. Disponível em: <https://ieeexplore.ieee.org/document/9115064>.
- 38 NASH LIANG ZHENG, N. B. W. Deep learning corrosion detection with confidence. *npj Materials Degradation*, v. 26, n. 6, p. 2397–2106, 2022. Disponível em: <https://doi.org/10.1038/s41529-022-00232-6>.
- 39 MUNAWAR, H. S. et al. Civil infrastructure damage and corrosion detection: An application of machine learning. *Buildings*, v. 12, n. 2, 2022. ISSN 2075-5309. Disponível em: <https://www.mdpi.com/2075-5309/12/2/156>.
- 40 ATHA, D. J.; JAHANSHAH, M. R. Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. *Structural Health Monitoring*, v. 17, n. 5, p. 1110–1128, 2018. Disponível em: <https://doi.org/10.1177/1475921717737051>.
- 41 FARADY, I. et al. Evaluation of data augmentation on surface defect detection. In: *2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*. [S.l.: s.n.], 2021. p. 1–2.
- 42 CHEN, Y. et al. Metal surface corrosion grade estimation from single image. In: YU, H.; DONG, J. (Ed.). *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)*. SPIE, 2018. v. 10615, p. 106154Z. Disponível em: <https://doi.org/10.1117/12.2302776>.
- 43 JIAN, B.-L. et al. Deep learning model for determining defects of vision inspection machine using only a few samples. *Sensors and Materials*, v. 32, p. 4217, 12 2020.
- 44 XU, X. et al. Intelligent corrosion detection and rating based on faster region-based convolutional neural network. In: *2020 Global Reliability and Prognostics and Health Management (PHM-Shanghai)*. [S.l.: s.n.], 2020. p. 1–5.
- 45 TIAN, Z. et al. Corrosion identification of fittings based on computer vision. In: *2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM)*. [s.n.], 2019. p. 592–597. Disponível em: <https://ieeexplore.ieee.org/document/8950822>.
- 46 S SHUKLA M, R. K. A. Surface corrosion grade classification using convolution neural network. *International Journal of Recent Technology and Engineering (IJRTE)*, Volume 8, Issue 3., p. Pages 7645 – 7649, September 2019. ISSN 2277-3878. Disponível em: <https://www.ijrte.org/wp-content/uploads/papers/v8i3/C6196098319.pdf>.
- 47 DONG, H.; GANG, T.; ZHENGKUN, Z. Surface corrosion detection of quayside crane based on improved mobilenetv2ssdlite. *Laser & Optoelectronics Progress*, v. 58, p. 1615006, 2021. Disponível em: <https://www.researching.cn/articles/OJ13de7d125dd5d517>.
- 48 AHUJA, S. K.; SHUKLA, M. K.; RAVULAKOLLU, K. K. Optimized deep learning framework for detecting pitting corrosion based on image segmentation. *International Journal of Performability Engineering*, Int J Performability Eng, v. 17, n. 7, p. 627, 2021. Disponível em: http://www.ijpe-online.com/EN/abstract/article_4605.shtml.

- 49 CHOUHAD, H. et al. Smart data driven defect detection method for surface quality control in manufacturing. *Measurement Science and Technology*, IOP Publishing, v. 32, n. 10, p. 105403, jul 2021. Disponível em: <https://dx.doi.org/10.1088/1361-6501/ac0b6c>.
- 50 RAHMAN, A.; WU, Z.; KALFARISI, R. Semantic deep learning integrated with rgb feature-based rule optimization for facility surface corrosion detection and evaluation. *Journal of Computing in Civil Engineering*, v. 35, n. 6, p. 04021018, 08 2021.
- 51 LI, W. et al. Deep learning based online metallic surface defect detection method for wire and arc additive manufacturing. *Robotics and Computer-Integrated Manufacturing*, v. 80, p. 102470, 2023. ISSN 0736-5845. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0736584522001521>.
- 52 JAGER, M.; KNOLL, C.; HAMPRECHT, F. A. Weakly supervised learning of a classifier for unusual event detection. *Trans. Img. Proc.*, IEEE Press, v. 17, n. 9, p. 1700–1708, sep 2008. ISSN 1057-7149. Disponível em: <https://doi.org/10.1109/TIP.2008.2001043>.
- 53 HE, Y. et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. *IEEE Transactions on Instrumentation and Measurement*, v. 69, n. 4, p. 1493–1504, 2020.