Corrosion Assessment Using Computer Vision, Machine Learning, and Deep Learning on Imagery Data: Evaluation and Use for Bridge Inspections

Hana N. Herndon, S.M.ASCE[1]; and Iris Tien, Ph.D., M.ASCE[2]

Abstract: Corrosion causes irreversible damage to metal structures and is one of the most common causes of bridge failure. The advent of affordable unmanned aerial vehicles (UAVs) that can carry digital cameras close to a structure presents an opportunity to quickly inspect bridges for corrosion. However, imagery-based corrosion assessment techniques must be accurate, detailed, and comprehensive to be effectively used in bridge inspections. This paper reviews and evaluates the use of imagery data and machine learning for corrosion assessment on bridges, analyzing existing research on imagery-based corrosion assessment, examining the ongoing difficulties of proposed data analysis approaches, and determining the remaining gaps in the literature. This paper evaluates the existing work with an emphasis on image processing methods; the environment and conditions in which the methods have been tested; evaluation of metrics used for performance assessment; and distinctions between corrosion detection, localization, and segmentation, as are significant in supporting bridge inspection decisions. The analysis finds that image processing techniques are typically too prone to false positives to perform well on chaotic data, particularly with nonuniform lighting and misleading objects in the image, which would be collected during bridge inspections using UAVs. Deep learning algorithms have been reported to perform more accurately than the human benchmark on similar data, and multiple researchers have developed deep learning algorithms that work well on UAV-collected data. However, there remain gaps in performance between methods tested on laboratory data compared to field data and on human-collected compared to UAV-collected data. Algorithms need to be trained on larger, more variable data sets to demonstrate generalizability across bridge inspection applications. If addressed, these methods could improve the safety, time-, and cost-efficiency of bridge inspections; help inspectors understand the state of each asset; and support structural repair, reconstruction, and prioritization decisions before failures occur. **DOI: 10.1061/JBENF2.BEENG-7619**. *© 2025 American Society of Civil Engineers*.

Author keywords: Corrosion assessment; Computer vision; Machine learning; Deep learning; Structural health monitoring; Bridge inspection; Image segmentation; Image classification.

Introduction

Corrosion is a destructive phenomenon that alters metal properties and causes irreversible structural damage (Fumagalli 2021; Javaherdashti 2000). Without preventative measures, corrosion can lead to significant losses during a metal structure's service life (Hansson 2011). This is particularly concerning for civil infrastructure, as corrosion of steel and other alloys can result in major failures, including collapse (Aliasghar-Mamaghani et al. 2022; Nash et al. 2022). Bridges are especially vulnerable, with corrosion being one of the leading causes of failure, along with scour and construction errors (Lin and Yoda 2017). The 2021 ASCE Report Card states that 46,154, or 7.5% of bridges in the United States are structurally deficient and 42% are at least 50 years old (ASCE 2021), which face additional corrosion safety concerns due to structural design details (Zhang et al. 2019).

Understanding the state of each asset, including corrosion damage, is crucial for engineers to prioritize repairs and prevent failures. However, traditional bridge inspections can be subjective, inefficient, costly, and unsafe. Many emerging technologies, particularly computer vision, machine learning, and unmanned aerial vehicles (UAVs), are increasing the efficiency of bridge inspections and corrosion detection capabilities (Zhang et al. 2022). However, the performance of these approaches and the conditions under which they have been tested for use in bridge inspections vary. This paper reviews and evaluates the use of imagery data and machine learning for corrosion assessment in bridges, investigating the difficulties of corrosion assessment from imagery data that influence why these approaches have not yet been widely implemented in the field and determining the remaining gaps in the literature. This paper evaluates the existing work with an emphasis on data collection, cleaning, and processing methods; the environment and conditions in which the methods have been tested; and the assessment criteria of these approaches. Additionally, when possible, the authors test the methodologies on external data to see how they fare on UAV-collected bridge inspection images.

The rest of the paper is organized as follows: "Research Methodology" describes the methods for selecting the literature included in this paper. "Background" provides background on current bridge inspection techniques and advances in computer vision and machine learning. "Classic Computer Vision," "Classic Machine

[1]Graduate Research Assistant, School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30318 (corresponding author). ORCID: https://orcid.org/0009-0005-4722-9798. Email: hnherndon@gatech.edu

[2]Williams Family Associate Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30318. ORCID: https://orcid.org/0000-0002-1410-632X. Email: itien@ce.gatech.edu

Note. This manuscript was published online on October 8, 2025. Discussion period open until March 8, 2026; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Bridge Engineering***, © ASCE, ISSN 1084-0702.

Learning," and "Deep Learning" discuss approaches, studies, and experiments to assess corrosion from images using classic computer vision techniques, classic machine learning techniques, and deep learning techniques, respectively. Next, analyses are conducted across studies, with the advantages and limitations of the studies compared in "Comparisons across Studies," including assessment of performance evaluation methods; quantitative comparison of results across data sets; varying data preparation, cleaning, and preprocessing steps; and evaluation of the methods for use in decision-making in the bridge inspection process. Finally, "Discussion" concludes with an overview of imagery data analysis approaches for corrosion assessment and its outlook for bridge inspections.

Research Methodology

A keyword search-based approach was adopted to identify publications relevant to this review. The search attributes are presented in Table 1. Keywords such as "corrosion," "detection," "computer vision," "deep learning," "imaging," and "segmentation" were selected and combined using Boolean operators "AND" and "OR" to create the search string. The review included peer-reviewed journal articles and conference proceedings. While journal articles were prioritized, the limited availability of journal publications on this topic necessitated the inclusion of conference papers. An initial examination revealed that some conference papers contained unique contributions not represented in journal articles. To ensure quality, all conference papers were rigorously analyzed. Only publications written in English were included. After identifying possible relevant publications, the title and abstract of each article were analyzed based on well-defined inclusion and exclusion criteria, as given in Table 2.

An initial search identified many potentially relevant papers, but the titles of these articles were scanned to eliminate studies that did not meet the search criteria, resulting in 171 papers. After this initial filtering, the studies were reviewed to retain only those that used imagery data to assess structural defects. This step narrowed the pool to 127 articles. During this phase, many studies were excluded for focusing on damage types such as cracking or spalling rather than corrosion, or if they failed to specify the methods used for corrosion assessment, resulting in 46 studies. To ensure a comprehensive review, two additional manual searches were conducted using Google Scholar and the reference list of the screened articles. These efforts resulted in the additional inclusion of 14 journal articles in the first search, and 9 journal articles in the second.

The collection of papers analyzed in this study includes a selection of 68 publications, with the majority of these works (44 of the 68) published since 2020. These selected papers collectively investigate various methodologies for corrosion detection and assessment using imagery data. Specifically, they delve into the applications of image processing, machine learning (including deep learning), or a combination of both in addressing corrosion-related concerns. To see how the distribution of varying approaches, how these distributions have changed over time, and the recent growth of work in this area, Fig. 1 shows the number of papers across publication years and image assessment methods.

While some of the chosen studies focused on the examination of corrosion from images of bridges, it is essential to note that several of them did not explicitly target bridge applications. Some researchers collected images of pipes, while others examined images depicting corrosion in different contexts, including daily scenarios. However, the authors include these as essential to assess corrosion from imagery data. Additionally, researchers sought to evaluate corrosion to varying degrees. In this review, distinctions between corrosion detection (presence of corrosion or not), assessment (measure of degree of corrosion), localization (pinpointing specific locations of corrosion on a structure), and segmentation (outlining the corrosion in an image) are made as significant to supporting bridge inspections and repair or retrofit decisions. Covering all works, in this paper, we explore the potential applicability of these methodologies for bridge inspections, considering the diverse range of subjects investigated.

This paper offers a comprehensive and in-depth survey of relevant papers on imagery data and machine learning used in bridge corrosion damage detection, insights into the methods used, and in-depth explanations of achievements and areas that require more research. Articles were included in this review if they detect, localize, or segment corrosion in images. Articles were excluded if they assess corrosion using nonimagery data, assess damage that is not corrosion, or focus on predicting corrosion evolution with comparative analysis rather than assessing the current state of corrosion. Fig. 2 shows the search process to catalog the articles used in this work, and Table 2 presents the corresponding guidelines.

The use of imagery data and machine learning to assess corrosion has the potential to reduce bias, increase efficiency, and enhance the safety of bridge inspections. With recent advancements in deep learning techniques, the analyses conducted in this paper are particularly important to conduct now because there is a new range of models and architectures that can be used to improve image analysis. However, these methods also have drawbacks, such as the need for extensive training and testing on diverse image data sets before they can be used to support bridge inspections. Additionally, as shown later in this work, significant differences exist in the images collected using UAVs compared to those collected from human inspectors, which create additional challenges for these algorithms. These differences include increased background vegetation, misleading objects such as graffiti and insect hives, and nonuniform lighting. Fig. 3 shows examples of these differences between photographs captured by a human inspector and photographs captured by a UAV. Vegetation and nonuniform lighting conditions are clearly visible in Fig. 3(b). These differences would further vary by time of year and season, impacting the accuracy and generalizability of imagery data analysis

Table 1. Search attributes for articles in this study

Search attribute	Database	Keywords and Boolean operators	Search scope	Published year	Document type	Language
Value in this work	Google Scholar	("computer vision" OR "image-based" OR "image processing" OR "visual" OR "segmentation" OR "detection" OR "machine learning" OR "deep learning" OR "neural networks") AND ("inspection" OR "monitoring" OR "assessment" OR "health" or "damage" OR "defect") AND ("metal structure" OR "corrosion" OR "rust" OR "steel structure")	Article title, abstract, keywords	From all years to 2025	Journal article; conference proceeding	English

Table 2. Guidelines used for cataloging articles in this study

Guidelines for including an article	Guidelines for excluding an article
Article detects, localizes, or segments corrosion in images	Article detects or localizes corrosion using nonimagery data
Images could be compositions of structural steel	Article detects, localizes, or segments damage that is not corrosion
Images could be captured by UAVs or humans	Article predicts the future growth of corrosion using comparative analyses (including DIC)
Article uses computer vision, machine learning, deep learning, or a combination to assess corrosion	

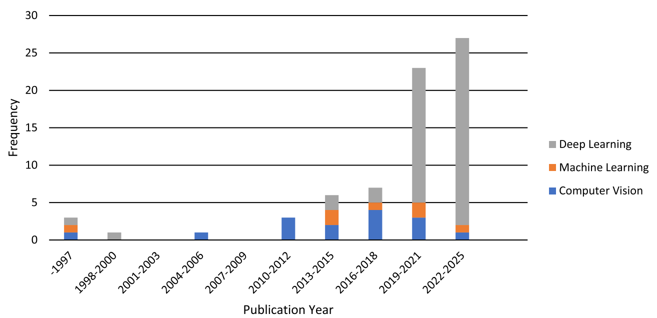


Fig. 1. Distribution of papers: publication years and image assessment method.

approaches. Therefore, this work separately analyzes corrosion assessment methods applied to laboratory data, field data collected by human inspectors, and field data collected by UAVs to evaluate their usefulness for bridge inspections.

Background

Nondestructive Evaluation Techniques for Corrosion Assessment in Bridges

Extensive and accurate bridge inspections support critical decisions in bridge asset management. However, current methods have several limitations. First, because human judgment–based evaluations are subjective and condition rating schemes can be vague, the information obtained from inspections is often inaccurate, subject to variability across individual inspectors and locations, and inadequate to aid in decision-making (Abdallah et al. 2022; Doshvarpassand et al. 2019; Lin et al. 2021). Second, current practices are time-consuming, expensive, and highly disruptive to traffic, requiring hours of manual labor and road closures (Abdallah et al. 2022; Jahanshahi et al. 2009).

Nondestructive evaluation (NDE) methods offer advantages over other inspection techniques because they do not permanently alter the asset under inspection, are less costly, reduce inspection subjectivity, and minimize traffic disruptions (Abdallah et al. 2022; Doshvarpassand et al. 2019; Gholizadeh 2016). Therefore, they can greatly aid corrosion detection, assessment, and monitoring on bridges (Liu et al. 2012). Common NDE methods include eddy-current, magnetic particle, liquid penetrant, radiographic, ultrasonic, and visual testing approaches (Dwivedi et al. 2018). Among these, visual testing methods, particularly optical imagery, hyperspectral imagery, and infrared thermography (IRT), are favored for corrosion detection due to their efficiency and minimal disruption (Dwivedi et al. 2018). These techniques are fully noncontact; however, optical image capture offers several advantages. It is generally more affordable and accessible than IRT and hyperspectral imagery, provides higher spatial resolution, and captures true-color images, which makes it well-suited for detecting early-stage corrosion indicators such as surface discoloration. Additionally, optical images are more compatible with existing computer vision and deep learning models, most of which are trained on standard RGB images (Nash et al. 2022). These qualities make optical imaging an attractive option for state agencies tasked with large-scale bridge inspections. However, a key limitation of optical systems is that they are limited to surface-level detection and cannot capture subsurface anomalies, which can be critical for identifying more advanced stages of corrosion. Table 3 further elaborates the strengths and limitations of optical imagery compared to infrared thermography and hyperspectral imagery.

Additionally, affordable UAVs equipped with digital cameras present an opportunity to revolutionize the bridge inspection process by conducting quick and safe bridge inspections from a distance (Lin et al. 2021; Tomiczek et al. 2019). Recent surveys show that an increasing number of state DOTs are interested in using UAVs for bridge inspections because they are noncontact, time-efficient, and cost-efficient (Jeong et al. 2020). UAVs may be equipped with various data-collection sensors, including a superpixel camera or infrared camera, and have inspection-enhancing abilities, such as rotating cameras that can take images of the underside of the bridge deck and obstacle avoidance in areas with insufficient GPS (Dorafshan et al. 2018; Tomiczek et al. 2019).

Improvements in computer vision and machine learning, combined with increased computing power, have enabled research into automated damage assessment from visual imagery (Nash et al. 2022). These algorithms have the potential to address many

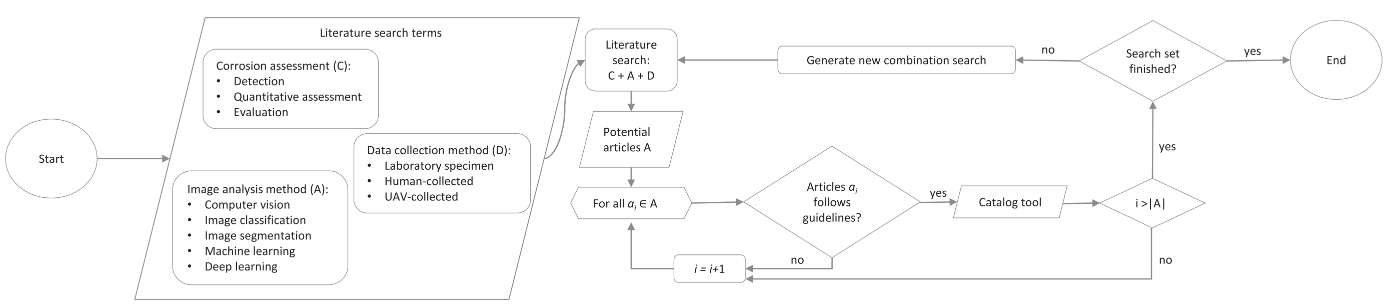


Fig. 2. Search process for cataloging articles in this study.

Fig. 3. (a) Human-collected (Bianchi and Hebdon 2021b) versus (b) UAV-collected images of corrosion on a bridge in Douglasville, Georgia.

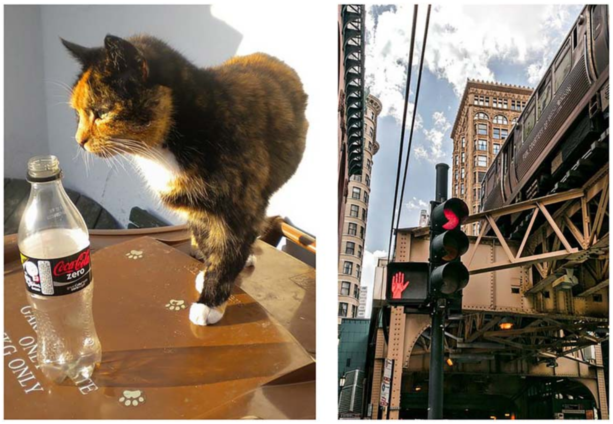
of the limitations of traditional bridge inspections noted previously. Previous studies have shown these technologies to be efficient and accurate in detecting and assessing cracking. For example, a deep learning framework using convolutional neural networks (CNNs) and naïve Bayes data fusion detected cracking in metallic surfaces with high accuracy (Chen and Jahanshahi 2018), while another study used a recurrent regional convolutional neural network (RRCNN) to segment cracking on bridges at the pixel level (Li et al. 2021). Additionally, a study by Kim et al. (2017) showed that these methodologies also work sufficiently on UAV imagery. These works achieve an accuracy of over 92%. To reliably and objectively train and evaluate these models, ground truth data must be rigorously defined. The most appropriate practice for establishing ground truth is to have multiple trained inspectors independently annotate the images, then aggregate—either through averaging, consensus, or by adopting the most conservative assessment—to reduce individual bias and increase label consistency, as done in Bianchi and Hebdon (2022). These works demonstrate the potential of optical imagery, combined with computer vision and machine learning, for automatic damage assessment during bridge inspections, although they only focus on cracking.

Compared to cracking, this study shows that current methods for corrosion assessment are less accurate. From a computer vision standpoint, this discrepancy can be attributed to key technical differences between crack and corrosion detection. Cracks typically present as well-defined linear features with high contrast against the background material, making them easier to detect using edge-based filters or segmentation algorithms. As demonstrated in studies by Ayele et al. (2020), Chen and Jahanshahi (2018), and Duque et al. (2018), deep learning models trained on annotated crack data sets achieve high accuracy due to the geometric simplicity of cracks.

Applied to the problem of corrosion assessment, it is postulated that uniform corrosion is a more difficult target using machine learning and computer vision techniques compared to other structural condition evaluations, such as crack evaluation, due to its lack of common shape-based features and its variance in color (Nash et al. 2022). However, pitting corrosion, which generally exhibits more uniform shapes and harsh edges, may be an easier problem than uniform corrosion. In computer vision terminology, uniform corrosion falls under the stuff category of subjects, along with the sky, the sea, and carpet, because it does not have discrete and distinct shape-based features compared to things such as cats and dogs. Fig. 4 shows a visual comparison of things versus stuff. In the figure, the cat can be recognized by features like its pointed ears, triangular nose, and round eyes. Meanwhile, the shape of the sky changes depending on the angle and location of the photo and cannot be defined by its shape. Convolution filters

Table 3. Comparison of noncontact, nondestructive evaluation methods for corrosion assessment

Method	Corrosion features leveraged	Strengths	Limitations
Optical image capture	Surface discoloration, surface texture changes, cracking	Low cost True-color imaging High spatial resolution Easy UAV integration Compatible with DL	Limited to surface information Affected by lighting
Infrared thermography (IRT)	Thermal conductivity differences, delamination, trapped moisture	Useful in low light or nighttime conditions Easy UAV integration Detects subsurface anomalies	Lower spatial resolution No color or texture data Requires calibration
Hyperspectral imagery	Spectral reflectance changes, surface discoloration, oxide layer formation	Captures rich spectral data Detects subtle material changes Easy UAV integration	High cost Large data volume Complex processing Limited adoption in field settings



(a) (b)

Fig. 4. (a) Things versus (b) Stuff to segment the cat compared to sky from the COCO data set (Lin et al. 2014).

used in deep learning models are well suited for detecting shape-based features, so these models detect things with better performance than "stuff." For example, a leading current model records F1 scores of 0.746 for things but just 0.488 for stuff (Nash et al. 2022).

For digital image techniques to be utilized in bridge inspections in the field, corrosion assessment techniques must show the same progress and be made just as accurate, detailed, and efficient (Lin et al. 2021). They can only be integrated into a system that streamlines data collection, analytics, and end-to-end reporting for bridge maintenance, repair, rehabilitation, and replacement decisions.

Computer Vision and Machine Learning for Corrosion Assessment

Recent advancements in computer vision and machine learning have increased capabilities for automated damage detection and evaluation research (Bianchi et al. 2021). Key tasks in computer vision include image classification, object detection, and image segmentation (IBM 2023a). Computer vision operates in tandem with machine learning, although it is essential to note that many computer vision techniques, such as edge detection, wavelet transformation, noise removal, and color analysis, do not rely on machine learning methods to assess images (Ahuja and Shukla 2018; Li et al. 2021; Nash et al. 2022). However, machine learning can be used for computer vision tasks like semantic segmentation or object detection. Damage assessment of bridges has seen the application of both supervised and unsupervised machine learning methods with varying results (Ahuja and Shukla 2018; Li et al. 2021). Additionally, deep learning, a subset of machine learning inspired by human brain behavior (IBM 2023b), can also be used for

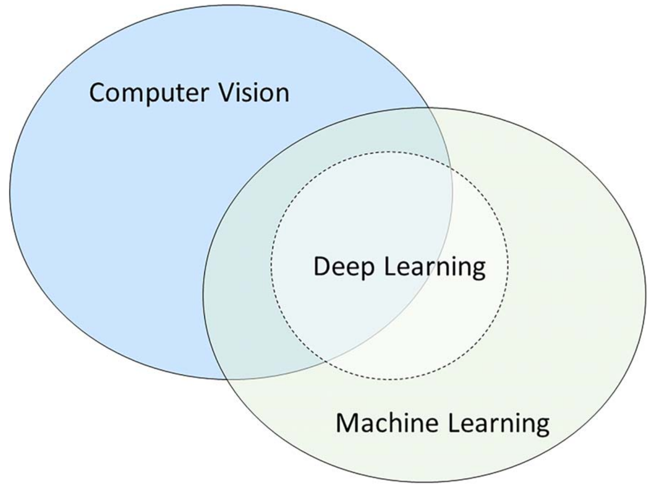


Fig. 5. Relationship between computer vision, machine learning, and deep learning.

computer vision tasks. Recent developments in deep learning architectures, such as DenseNet (Huang et al. 2017), FCN (Long et al. 2015), PSPNet (Liu et al. 2015), SegNet (Badrinarayanan et al. 2017), and DeepLab (Chen et al. 2018), have performed well in competitive computer vision exercises and have improved the accuracy of these tasks over the years (Nash et al. 2022). The relationship between computer vision, machine learning, and deep learning is shown in Fig. 5 and further elaborated in Table 4.

Corrosion Assessment Tasks

For bridge inspection applications, it is important to specify the particular data analysis task at hand, ranging from corrosion detection to localization to segmentation within an image. Detection algorithms are designed to determine whether corrosion exists in the image, but they do not determine its precise location or quantify its extent. Localization algorithms pinpoint the location of the surface defect. Segmentation algorithms determine both the location and size of the defect. Many previous studies begin with the task of detection corrosion in images (Gamarra Acosta et al. 2014; Itzhak et al. 1981; Jahanshahi and Masri 2013a; Lee et al. 2006). Detection is favored because it is the simplest type of algorithm, making it the natural starting point for corrosion assessment algorithms. This output can create a red flag for bridge inspectors, enabling them to identify potential corrosion. However, it does not provide comprehensive information about the location and extent of the defect, requiring inspectors to manually inspect the photos or the bridge to make informed decisions. Localization and segmentation are more complicated tasks, but they provide more information about the present defects. While detection is a useful first

Table 4. Overview of computer vision, machine learning, and deep learning

Domain	Description	Examples	In this study
Computer vision (CV)	Techniques enabling machines to interpret and process visual information	Edge detection (Canny, Sobel), SIFT, image thresholding	All included studies involve computer vision because all analyze image data; classic CV refers to rule-based, noniterative methods to distinguish corrosion from background features
Machine learning (ML)	Algorithms that learn patterns from data to make predictions or classifications	Decision trees, SVMs, *k*-nearest neighbors	Classic ML methods are iterative and data-driven but do not involve deep neural networks
Deep learning (DL)	A subfield of ML using multilayer neural networks to model complex relationships	CNNs, ViTs, GANs	DL methods use hidden layers and large numbers of parameters, which are learned from data during training

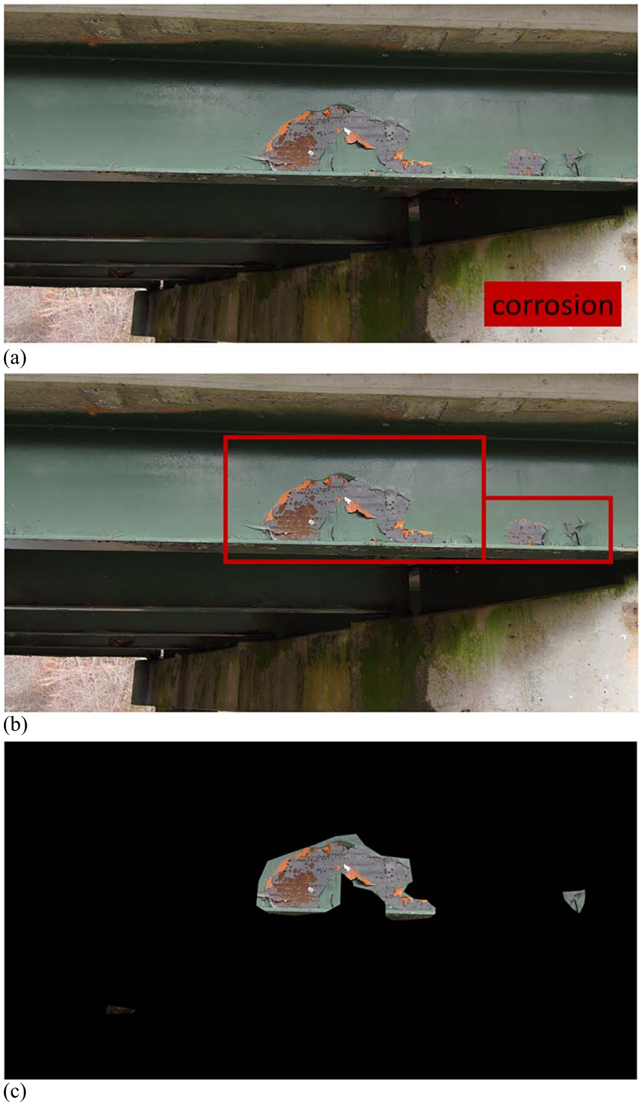


Fig. 6. (a) Corrosion detection; (b) localization; and (c) segmentation.

step, localization and segmentation algorithms are particularly valuable for bridge inspections because they provide inspectors with valuable information about the location and quantity of corrosion. Fig. 6 shows the differences between these corrosion assessment tasks. Fig. 6(a) detects corrosion as present in the image; Fig. 6(b) localizes the corrosion by providing a bounding box around the corrosion; and Fig. 6(c) segments the corrosion by outlining the corrosion, providing information about both location and size of the corrosion.

Another important task for bridge inspections is determining the severity of corrosion. Low levels of corrosion impact a structure's mechanical properties differently than high levels, so inspectors need to know the severity to make decisions about the structure's state and necessary protective actions. However, corrosion also affects the interior of the structure, yet imagery data only captures information about the surface, limiting the extent of evaluation. Additionally, there is no standard method to calculate the severity of corrosion in images. Some researchers have investigated the use of deep learning methods to classify corrosion according to the condition states provided by the American Association of State Highway and Transportation Officials (AASHTO) (Ameli et al. 2024; Bianchi and Hebdon 2022), while others have fused deep learning methods with classic computer vision to determine corrosion severity based on the darkness of the corroded pixels (Rahman et al. 2021; Zhou et al. 2022). Both of these methods to assess corrosion would provide more information for bridge inspectors than segmentation alone; however, it is important to note that these methods have not been widely explored, the results have not been compared to the results of in-person inspections, and their applicability in real-world scenarios has not been determined. The following sections provide detailed analyses of methods for corrosion assessment from imagery data with direct application to bridge inspections.

Novelty of the Presented Analysis

Given the rapid advancement of imagery data-based analysis approaches and the value of accurate and efficient corrosion assessment in bridges, the analysis presented in this paper is of timely and critical importance. In previous reviews of machine learning methods for corrosion, Coelho et al. (2022, Soomro et al. (2022, and Liu and Bao (2022) reviewed machine learning–based corrosion detection and assessment methods with multiple data sources, including imagery. However, they did not discuss recent advancements in imagery data collection in detail, especially regarding bridge inspections. Bridge imagery data are particularly complex because of the chaos in the background, nonuniform lighting, misleading objects, and because bridge elements overlap and can block other elements from view. Other papers focusing on corrosion detection using imagery data do not encompass recent deep learning advancements (Ahuja and Shukla 2018; Jahanshahi et al. 2009). Jahanshahi et al. (2009) discussed automatic crack and corrosion detection on bridges but predated recent advancements in deep learning. Ahuja and Shukla (2018) reviewed 29 studies but lacked an evaluation of their strengths, limitations, and applications to bridge inspections, in addition to not covering deep learning approaches. Additionally, they did not discuss deep learning approaches despite being published more recently.

Previous reviews related to artificial intelligence (AI) methods in structural engineering offer a broad perspective of vision-based structural inspection, focusing on multiple degradation mechanisms, without providing an in-depth review of corrosion assessment (Zhang et al. 2022), or multiple data collection mechanisms, without providing an in-depth analysis of computer vision methods (Abdallah et al. 2022). These reviews offering a broader perspective do not provide an in-depth understanding of corrosion assessment algorithms. For example, Zhang et al. (2022) reviewed 115 papers, 50 of which are related to all types of damage assessment, resulting in a small survey of papers for corrosion assessment and not providing a detailed analysis of the remaining struggles in this specific area. There remains a gap for a focused, up-to-date analysis of corrosion detection and assessment algorithms that centers on imagery data, evaluates approaches ranging from classic computer vision to deep learning, and critically assesses their performance and limitations specifically for bridge inspections. To address these gaps, this paper seeks to provide a current, comprehensive, in-depth, and detailed review of corrosion assessment using computer vision methods, including state-of-the-art deep learning techniques, tailored for bridge inspection purposes. The authors review the studies in detail, analyzing the evaluation metrics used, evaluating the data set characteristics, and strictly examining the studies' applicability for bridge inspections. This paper evaluates the performance of these approaches in different environmental conditions. For example, bridges can be in environments with heavy foliage, which can be misleading for computer vision algorithms. This paper also evaluates how the algorithms' performance changes when tested on UAV-collected images versus images collected from human inspectors, highlighting the current state regarding UAV-aided bridge inspections. Additionally, when possible, the authors test the methodologies on external data to see how they fare on UAV-collected bridge inspection images.

A state-of-the-art paper providing critical evaluation and analysis on this topic is needed at this time, with the recent advances in the field and the potential for the use of imagery data and machine learning to advance current processes in bridge inspections and bridge engineering. An in-depth analysis of corrosion assessment techniques is needed to assess current capabilities and determine the remaining gaps that could be overlooked in a broader review. If computer vision corrosion assessment methods are accurate, detailed, and comprehensive, they could improve the safety, time-, and cost-efficiency of bridge inspections; help inspectors better understand the state of each asset; and support structural repair, reconstruction, and prioritization decisions before failures occur.

The novelty of the review lies in its focused examination of image-based corrosion assessment techniques specifically for bridge inspections. This paper critically evaluates the performance of multiple methods, ranging from traditional computer vision to more recent deep learning methods, discussing the strengths and limitations of each when applied to real-world bridge imagery—particularly UAV-collected data—which remains underexplored. The study highlights how environmental factors, data sources, and algorithm design influence accuracy, generalizability, and practical usability. By benchmarking reported metrics across studies and identifying common failure points, this paper offers a clearer, more detailed understanding of where current methods succeed and where they fall short. This work equips researchers and practitioners with the knowledge needed to prioritize future research directions, advance algorithm development, and ultimately integrate vision-based corrosion assessment into existing inspection workflows. The insights provided support the transition of visual corrosion assessment methodologies from experiments to deployment, helping agencies adopt more efficient, objective, and safe bridge inspection strategies.

Classic Computer Vision

Classic computer vision refers to computer vision techniques that do not incorporate machine learning. These algorithms follow explicit rules from predetermined formulas to extract information from digital images instead of machine learning, which uses patterns in the data to iteratively find the optimal feature representation. Classic computer vision includes a wide range of algorithms, including color thresholding, color space optimization, image compression, entropy quantification, and texture thresholding. For corrosion assessment, researchers have used different techniques to capitalize on the visual properties of corrosion, particularly its texture and color. Corrosion produces a unique rough surface and reddish coloration in most materials, including steel (Medeiros et al. 2010). Classic computer vision methods to evaluate texture include gray-level covariance matrices (GLCMs) (Haralick et al. 1973), Laws texture-energy measures (Laws 1980), and Shannon entropy, but other methods exist and have been investigated as well (Yuan et al. 2015). To analyze the color, researchers typically evaluate the image in different color spaces (Jahanshahi and Masri 2013a). The following studies used one or multiple classic computer vision methods to assess corrosion in images.

Laboratory Experiments

One main distinction in previous studies using imagery data for corrosion assessment is whether the images were collected in the laboratory or the field. Laboratory environments offer quicker data preparation and collection and allow a greater understanding of the underlying processes. However, laboratory images lack the lighting irregularities, misleading objects, and scale differences that are prevalent in field images and therefore do not directly replicate real-world conditions. Therefore, a methodology that performs well in the laboratory may not be as effective in real applications (Forkan et al. 2022). This holds for bridge inspections, which occur in varied environmental conditions depending on lighting, seasons, and surrounding vegetation. While laboratory studies are suitable for preliminary investigations, these techniques need further testing before they can be reliably used in bridge inspections.

Itzhak et al. (1981) were the first to introduce digital image processing tools for corrosion assessment, manually selecting a brightness threshold to segment pitting corrosion in the image. Pakrashi et al. (2010) and Wang and Cheng (2016) continued this work using edge detection techniques to clarify the boundaries of the pits after segmentation using brightness descriptors. In general, edge detection methods identify regions of rapid intensity change in an image by approximating the gradient magnitude of pixel intensities using convolution filters, such as the Sobel filter, as shown in Eqs. (1) and (2). The efficacy of these approaches lies in the contrast of the damaged region compared to the surroundings and therefore is generally more applicable for cracking detection. Pakrashi et al. (2010) and Wang and Cheng (2016) showed that these methods can also be effective for pitting corrosion because pits tend to be dark and well-defined. In the controlled environments used in these studies, these approaches achieved a 0.1% false alarm rate and 94% accuracy in detecting pitting corrosion. However, generalized corrosion, which may have diffused and inconsistently defined boundaries, can present a greater challenge for edge detection methodologies. Instead of using brightness descriptors, Enikeev et al. (2017) turned to texture features, demonstrating that they can be useful in distinguishing corroded laboratory specimens. However, because they did not evaluate their methodology quantitatively, only qualitatively, it is difficult to compare their technique to others. Choi and Kim (2005) combined color and texture characteristics to detect corrosion using a probabilistic decision-making tool. This technique worked with 85% accuracy. An example of image thresholding for corrosion detection is shown in Fig. 7

$$G = \sqrt{G_x^2 + G_y^2} \tag{1}$$

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}; \quad G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{2}$$

Most studies presented in this section focus on corrosion detection rather than segmentation, do not assess the severity of corrosion, and do not investigate the techniques' ability to distinguish corrosion from misleading objects. Therefore, they do not provide enough information for bridge inspectors to make decisions and must be tested further before they can be used to improve the bridge inspection process.

J. Bridge Eng., 2025, 30(12): 03125001

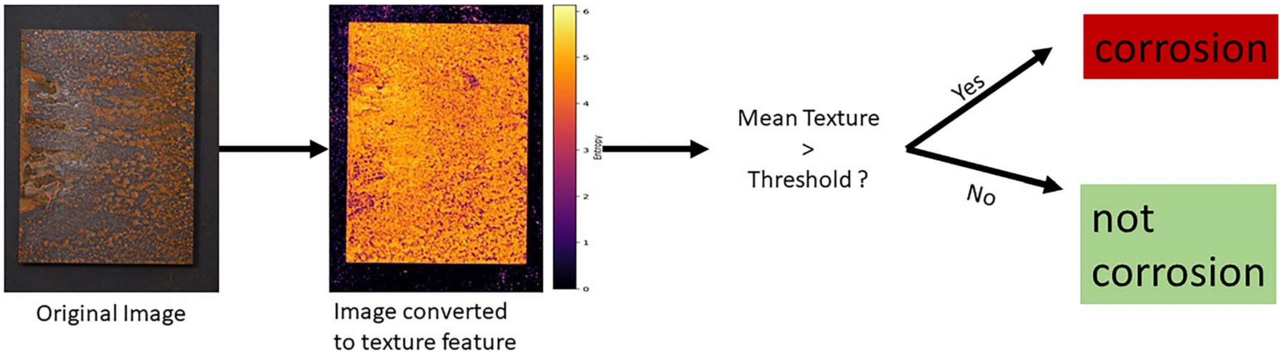


Fig. 7. Image thresholding for corrosion detection example.

Field Experiments

Bridges are often located in areas where misleading objects, including vegetation, graffiti, or insect hives, are present on and around the structure. Additionally, because they are located outdoors, the lighting and vegetation will change with the time of day and time of year, affecting the images collected at the site. Therefore, researchers must test their approaches on field data to demonstrate their performance in these environments, even if the approach has been shown to work well in the laboratory. The following studies use classic computer vision methods to assess corrosion on images captured in the field.

Ghanta and Tanja Karp (2011) and Petricca et al. (2016) used color thresholding to classify images as corroded or not corroded without providing information about the location or amount of corrosion in the image. Image thresholding is a simple yet widely used image processing technique in which pixel values are compared to a defined threshold $\boldsymbol{T}$. A pixel p in an image $\boldsymbol{I}(x, y)$ is classified if $\boldsymbol{I}(x, y) > \boldsymbol{T}$, and background otherwise. Thresholding can be applied to classify images based on the total number of pixels exceeding the threshold or to segment an image by evaluating each pixel individually. It may also be applied to various color channels, transformed color spaces, or texture-based features. However, a limitation of thresholding is its inability to distinguish between visually similar features. For example, when Petricca et al.'s (2016) algorithm was tested on misleading objects, they found that it could not distinguish between corrosion and other red objects. Therefore, this methodology would yield many false positives during a bridge inspection. Bondada et al. (2018) and Das et al. (2023) used color thresholding to segment corrosion in images, providing more information about the damage; however, it required significant manual preprocessing that limits its efficiency and applicability in the field. Fei et al. (2021) and Moranduzzo and Melgani (2014) also used automatic color thresholding to segment corrosion and conducted automatic preprocessing to make it less tedious for inspectors. Fei et al.'s (2021) technique achieved a precision of 77% on a data set of 10 images, and Moranduzzo and Melgani's (2014) technique performed with 1% error on a data set of unspecified size. While promising, these techniques must be tested on more varied data (these approaches were tested on images of pressure vessels) before their applicability to bridge inspections can be established. On average, classic computer vision algorithms tested in the field achieve 81.0% accuracy, while those tested in laboratory conditions reported up to 99%, showing that color-based image processing techniques are currently inadequate for corrosion assessment in bridge inspections due to their poor performance when presented with misleading objects.

Bonnin-Pascual and Ortiz (2014), Khayatazad et al. (2020), and Medeiros et al. (2010) used color and texture thresholding to evaluate corrosion in images. By requiring that both color and texture thresholds are met, these methods aim to reduce false positives. Medeiros et al. (2010) utilized GLCM texture features and hue–saturation–value (HSV) color features to detect corrosion in images. Bonnin-Pascual and Ortiz (2014) combined these features into the weak-classifier color-based corrosion detector (WCCD) to segment corrosion in images, augmenting the performance by integrating both techniques. Khayatazad et al. (2020) tested this algorithm on images with misleading objects and nonuniform lighting, discovering that the algorithm suffers in complicated environments. In certain lighting conditions, its ability to detect texture decreased, and it generally could not distinguish between corrosion and other red objects. Fig. 8 provides an example of classic computer vision methods to segment corrosion in images.

Marchewka et al. (2020) developed a comprehensive framework for UAV bridge inspections that includes UAV route planning, structural element segmentation, and corrosion detection. After automatically segmenting at-risk structural elements such as rivets, they detected corrosion using statistical measures of color features in the YCbCr color space, achieving 96% precision. The corrosion segmentation method performed well when structural elements were segmented properly. However, if misleading objects were present, the algorithm misclassified them as corrosion. This study suggests that image processing techniques could be valuable for bridge inspections if combined with other algorithms to filter out misleading objects. Otherwise, they are generally too prone to false positives to be useful for bridge inspections. Additionally, the methods discussed in this section were only tested on surface corrosion, not pitting or subsurface corrosion, and did not attempt to characterize the corrosion severity. Therefore, they are only marginally useful at this stage to support bridge inspection decisions. Tables 5 and 6 summarize the classic computer vision studies based on methodologies and performance, respectively. The last column of Table 6 provides a comparable accuracy metric to adequately compare studies. In cases where the accuracy was not reported, it was calculated from the given performance metrics. However, if the accuracy was reported, that value is repeated in the table.

Classic Machine Learning

In contrast to classic computer vision, which relies on predefined formulas and rule-based techniques to extract information from images, classic machine learning uses statistical algorithms to extract patterns from the data and make predictions. However, unlike deep learning, classic machine learning techniques cannot autonomously extract features from raw image data. Instead, researchers or bridge

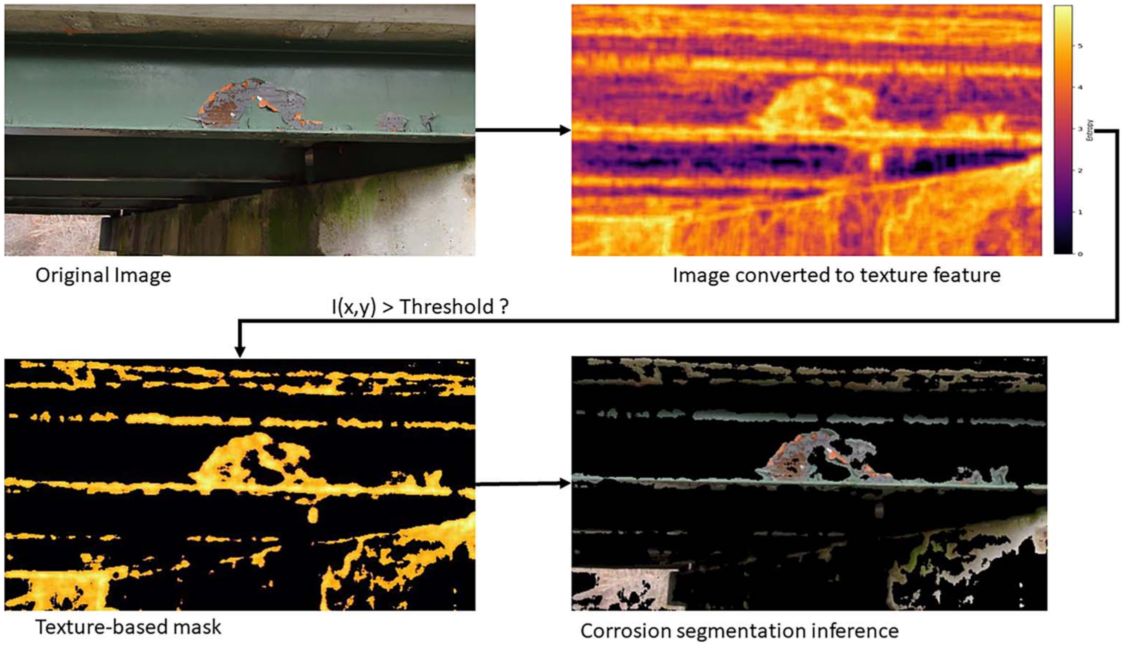


Fig. 8. Image thresholding for corrosion segmentation example.

inspectors must manually engineer these features before applying the algorithm. Several studies have used classic machine learning methods, or machine learning methods that do not involve deep neural networks, to assess corrosion in images.

Laboratory Experiments

The following studies investigate the use of machine learning approaches for corrosion assessment on laboratory specimens. Gunatilake et al. (1997) and Siegel and Gunatilake (1998) explored classic machine learning methods using texture features. Gunatilake et al. (1997) used a *k*-nearest neighbor (KNN) algorithm to classify the images as corroded or not corroded using texture features. KNN is a simple, nonparametric algorithm that classifies a data point based on the majority class of its *k*-nearest neighbors in a feature space. However, KNN is sensitive to the choice of *k*, the scale of features, and noise in the data—factors that can hinder performance in complex environments. Siegel and Gunatilake (1998) extended this to detect subsurface corrosion, which is characterized by surface pillowing rather than texture. They used KNN to differentiate between surface and subsurface corrosion, providing valuable information for bridge inspections. However, these methodologies lack quantitative evaluation, limiting their comparability and assessment for bridge inspections.

Table 5. Summary of classic computer vision studies: methodologies

No.	Study	Algorithm type	Methods	Application	Corrosion type	Corrosion features: Texture	Corrosion features: Color	Input features
1.	Itzhak et al. (1981)	Segmentation	Thresholding	Laboratory	Pitting	—	X	Brightness
2.	Pakrashi et al. (2010)	Segmentation	Thresholding	Laboratory	Pitting	X	X	Optical contrast, image edges
3.	Wang and Cheng (2016)	Detection	Hough transform	Laboratory	Pitting	X	—	—
4.	Enikeev et al. (2017)	Detection	Thresholding	Laboratory	Pitting	X	—	Brightness
5.	Choi and Kim (2005)	Detection	Probabilistic decision making	Laboratory	Uniform	X	X	GLCM and HSV
6.	Bondada et al. (2018)	Segmentation	Thresholding	Field	Uniform	—	X	HSV
7.	Ghanta and Tanja Karp (2011)	Detection	Thresholding	Field	Uniform	—	X	Luminance: energy of color layers
8.	Petricca et al. (2016)	Detection	Thresholding	Field	Uniform	—	X	HSV and red components
9.	Das et al. (2023)	Segmentation	Thresholding	Field	Uniform	X	X	YcbCr color
10.	Fei et al. (2021)	Segmentation	Thresholding	Field	Uniform	—	X	RGB range
11.	Moranduzzo and Melgani (2014)	Segmentation	Thresholding	Field	Uniform	—	X	Brightness
12.	Bonnin-Pascual and Ortiz (2014)	Segmentation	Thresholding	Field	Uniform	X	X	GLCM and HSV
13.	Khayatazad et al. (2020)	Segmentation	Thresholding	Field	Uniform	X	X	GLCM and HSV
14.	Medeiros et al. (2010)	Detection	Thresholding	Field	Uniform	X	X	GLCM and HSV
15.	Marchewka et al. (2020)	Segmentation	Thresholding	Field	Uniform	—	X	YCbCr color; bridge geometry

Table 6. Summary of classic computer vision studies: performance

No.	Study	Algorithm type	Methods	Application	Corrosion type	Eval. metric	Performance (%)	Accuracy (%)
1.	Pakrashi et al. (2010)	Segmentation	Thresholding	Laboratory	Pitting	Probability of false alarms	0.1	99.0
2.	Marchewka et al. (2020)	Segmentation	Thresholding	Field	Uniform	Precision	96.6	96.6
3.	Moranduzzo and Melgani (2014)	Segmentation	Thresholding	Field	Uniform	Error	1.0–3.4	96.6
4.	Wang and Cheng (2016)	Detection	Hough transform	Laboratory	Pitting	Accuracy	94.0	94.0
5.	Medeiros et al. (2010)	Detection	Thresholding	Field	Uniform	Accuracy F1 score	92.0 91.0	92.0
6.	Das et al. (2023)	Segmentation	Thresholding	Field	Uniform	Accuracy	90.0	90.0
7.	Choi and Kim (2005)	Detection	Probabilistic decision making	Laboratory	Uniform	Accuracy	85.0	85.0
8.	Bonnin-Pascual and Ortiz (2014)	Segmentation	Thresholding	Field	Uniform	False positives False negatives	9.8 5.9	84.3
9.	Fei et al. (2021)	Segmentation	Thresholding	Field	Uniform	Precision	77.4%	77.4
10.	Petricca et al. (2016)	Detection	Thresholding	Field	Uniform	Accuracy	69.0	69.0
11.	Khayatazad et al. (2020)	Segmentation	Thresholding	Field	Uniform	Accuracy	69–85	69.0
12.	Ghanta and Tanja Karp (2011)	Detection	Thresholding	Field	Uniform	Visual	X	X
13.	Bondada et al. (2018)	Segmentation	Thresholding	Field	Uniform	Visual	X	X
14.	Itzhak et al. (1981)	Segmentation	Thresholding	Laboratory	Pitting	Visual	X	X
15.	Enikeev et al. (2017)	Detection	Thresholding	Laboratory	Pitting	Visual	X	X

Field Experiments

Machine learning algorithms, which seek patterns in data rather than specific features, may reduce false positives compared to image processing methods. Laboratory studies show promise in corrosion assessment using machine learning; the following studies test these algorithms on field data.

Bonnin-Pascual and Ortiz (2014), Hoang (2020), Hoang and Tran (2019), and Son et al. (2014) used various machine learning algorithms to segment corrosion in images with good performance. Bonnin-Pascual and Ortiz (2014) used decision trees, while Hoang (2020) and Hoang and Tran (2019) employed support vector machines (SVMs) to localize corrosion. Son et al. (2014) compared multiple machine learning algorithms, including SVM, backpropagation neural network (BPNN), decision tree, naïve Bayes (NB), and KNN, across the same data set. They found that decision trees performed the best, with an accuracy of 97.5%. Decision trees are intuitive and interpretable models that use a hierarchical structure to recursively partition the data based on feature thresholds. They perform well on structured data sets with clear, separable patterns and have demonstrated success in crack identification (Frias and Hidalgo 2021; Sheng et al. 2018). Although corrosion presents more variability in shape, texture, and color than cracking, the hierarchical nature of decision trees can still manage multifeature data sets effectively. However, decision trees are known to overfit when trained on small or low-variance data sets, and their performance can degrade in noisy environments with high visual complexity. Because these studies used images collected close to the structure, minimizing lighting irregularities, misleading objects, and background chaos, they did not demonstrate whether decision trees can generalize to the more variable and uncontrolled imagery typically collected during UAV-based bridge inspections. A study by Herndon and Tien (2025) applied K-means clustering to segment corrosion in images collected by UAVs and the corrosion condition state data set developed by Bianchi and Hebdon (2021b), showing that this algorithm can perform with high recall across multiple data sets. K-means is an unsupervised machine learning algorithm that partitions data into clusters based on similarity, typically measured by Euclidean distance. Its unsupervised nature allows it to operate without training data, making it well-suited for unbalanced data sets and enhancing its generalizability. However, despite its strengths in recall and adaptability, the algorithm's mIoU and F1 scores remain suboptimal. As a result, while K-means may serve as a valuable tool for initial corrosion screening, further refinement or integration with other methods is needed before deployment in bridge inspections.

In conclusion, machine learning needs further testing and refinement on diverse and complex data to validate its efficacy for bridge inspections. In addition, machine learning methods have only been tested on their ability to detect surface corrosion, and other forms have not yet been addressed. Tables 7 and 8 summarize the machine learning studies based on methodologies and performance, respectively. Table 8 provides a comparable accuracy metric, determined from the reported metrics, to adequately compare studies.

Deep Learning

Deep learning, a subset of machine learning that requires less human intervention and more computational power, has seen significant advancements in recent years. Convolutional neural networks (CNNs) and vision transformers (ViTs), the most widely used deep learning algorithms for processing visual data, are specifically designed to process gridlike data such as images through convolution and patch operations, respectively (Sookpong et al. 2023; Yao et al. 2019). The main advantage of CNNs and ViTs is their ability to automatically extract relevant features from the data, which typically requires manual feature engineering in traditional machine learning algorithms. A known limitation of deep learning, however, is its lack of explainability; inspectors and engineers may not understand why a model identifies a region as corroded, which can lower confidence in its use for safety-critical decisions. Fortunately, results can often be verified through visual inspection in computer vision tasks, mitigating this concern. The elimination of the need for prior knowledge and human effort, coupled with the rise in available computing power, has spurred research into

Table 7. Summary of classic machine learning studies: methodologies

No.	Study	Algorithm type	Methods	Application	Corrosion type	Corrosion features		Input features
						Texture	Color	
1.	Gunatilake et al. (1997)	Segmentation	KNN	Laboratory	Pitting	X	—	Discrete wavelet transforms
2.	Siegel and Gunatilake (1998)	Segmentation	KNN	Laboratory	Subsurface	X	—	Discrete wavelet transforms
3.	Bonnin-Pascual and Ortiz (2014)	Segmentation	Decision tree	Field	Uniform	X	X	Laws' texture energy and HSV
4.	Hoang (2020)	Localization	SVM	Field	Uniform	X	X	GLCM, statistical measurements of RGB channels
5.	Hoang and Tran (2019)	Localization	SVM	Field	Uniform	X	X	GLCM, statistical measurements of RGB channels
6.	Son et al. (2014)	Segmentation	Decision tree	Field	Uniform	—	X	HSV
7.	Herndon and Tien (2025)	Segmentation	K-means	Field-UAV	Uniform	X	X	Shannon entropy, a- and b-chromaticities

Table 8. Summary of classic machine learning studies: performance

No.	Study	Algorithm type	Methods	Application	Corrosion type	Eval. metric	Performance (%)	Accuracy (%)
1.	Son et al. (2014)	Segmentation	Decision tree	Field	Uniform	Accuracy	97.5	97.5
2.	Hoang and Tran (2019)	Localization	SVM	Field	Uniform	Accuracy	92.8	92.8
3.	Hoang (2020)	Localization	SVM	Field	Uniform	Accuracy F1 score	91.8 92.0	92.0
4.	Bonnin-Pascual and Ortiz (2014)	Segmentation	Decision tree	Field	Uniform	False positives False negatives	17.2 3.4	79.4
5.	Herndon and Tien (2025)	Segmentation	K-means	Field-UAV	Uniform	mIoU	72.0	72.0
6.	Gunatilake et al. (1997)	Segmentation	KNN	Laboratory	Pitting	Visual	X	X
7.	Siegel and Gunatilake (1998)	Segmentation	KNN	Laboratory	Subsurface	Visual	X	X

deep learning (Atha and Jahanshahi 2018), making it a prominent choice in corrosion assessment studies.

Laboratory Experiments

Furuta et al. (1995) were one of the first to use multilayered neural networks for uniform corrosion detection nearly three decades ago. Petricca et al. (2016) compared classic computer vision algorithms to deep learning models for corrosion detection, finding that deep learning achieved 10% greater accuracy. Jahanshahi and Masri (2013a, b) studied different combinations of color channels, texture characteristics, and subimage window sizes for corrosion detection neural networks, finding a combination that achieves 90.5% accuracy. While these works show the promise of deep learning and its strengths over classic computer vision approaches, they fall short of realizing its full potential because they focus on detection rather than more complex tasks like segmentation.

Xu et al. (2020) and Yao et al. (2019) fine-tuned pre-existing CNNs, AlexNet, and Faster R-CNN to segment uniform corrosion in images with accuracies over 90%. Zhang et al. (2021) developed a CNN to classify and localize corroded regions of various degrees with an F1 score of 97.5%. Naik et al. (2020), Qian et al. (2025), and Son et al. (2024) used multilayer perceptron, a custom CNN, and Mask-RCNN, to segment pitting corrosion, also with accuracies over 90%. These studies show the promise of deep learning for more difficult corrosion assessment tasks across corrosion types, providing information about the location and extent of corrosion. However, before use in bridge inspections, these methods must be tested on images representative of bridge inspections that include variable lighting and complex backgrounds.

Field Experiments

Compared to classic computer vision and other machine learning algorithms, CNNs are typically better suited for chaotic data sets because they automatically learn relevant classification features that can be used to distinguish degradation from visually similar objects. Unlike traditional methods that rely on handcrafted features, CNNs extract low- to high-level features such as edges, textures, and complex patterns through layers of convolutional filters, pooling, and nonlinear activations. Mathematically, CNNs apply convolution operations between the input image and learned kernels, defined in the following equation:

$$S(i, j) = (I \times K)(i, j) = \sum_m \sum_n I(i - m, j - n)K(m, n) \quad (3)$$

where I = input image; K = convolution kernel; m and n = dimensions of the image; and S = resulting feature map. CNNs have outperformed many state-of-the-art algorithms in various computer vision tasks, including image classification (Bastian et al. 2019; Petricca et al. 2016), making them a strong candidate for detecting and segmenting corrosion in bridge inspection data.

However, effective application of CNNs requires training and validation on data that accurately represent the intended task. Therefore, it is important to distinguish between studies using human inspection photos and those using UAV-collected images because high performance on the latter will be needed to support the growing trend of UAV-aided bridge inspections. Another key limitation is overfitting: CNNs are prone to memorizing training data if the data set is small or lacks diversity. Therefore, to ensure generalizability across diverse bridge types, environments, and lighting conditions, deep learning algorithms must be trained and tested on large, diverse data sets and validated rigorously to confirm broad applicability. This section explores deep learning methods applied to field data, distinguishing approaches developed using human-collected versus UAV-collected imagery.

Human-Collected Photos

Alviz-Meza et al. (2024), Bastian et al. (2019), Holm et al. (2020), Malashin et al. (2024), and Nash et al. (2020) tested different CNN architectures for corrosion detection using human-collected images, all achieving accuracies above 90%, with Malashin et al. (2024) using the approach to detect pitting corrosion instead of uniform corrosion. Additionally, Nash et al. (2020) attempted to streamline the tedious data labeling process by developing a crowdsourced data collection website to create training data. These findings demonstrate that CNNs outperform classic computer vision methods in corrosion detection, significantly reducing false positives in data sets with nonuniform lighting and misleading objects, and that crowdsourcing can be a viable way to increase training data without sacrificing performance. However, these studies focus solely on detection, rather than on more complicated localization or segmentation tasks. Fig. 9 provides an example deep learning approach for corrosion detection; note that the shown architecture is an example and does not represent a specific algorithm.

Most recent deep learning studies focus on these more difficult tasks. For example, Khayatazad et al. (2023) used a five-layer neural network without convolution layers to segment corrosion at the pixel level, achieving an F1 score of 74%. Generally, CNN-based studies report better performance. Researchers have investigated the use of many CNN architectures for uniform corrosion segmentation, including HRNet (Nash et al. 2022), DeepLabV3 (Rahman et al. 2021), YOLO (Jia et al. 2023), U-Net (Chen et al. 2021), EfficientNet (Das et al. 2024), and custom architectures (Atha and Jahanshahi 2018), on data sets of various steel objects, such as storage tanks or street lamps. All of these studies surpass the accepted human-level performance of an F1 score of 81%, which is the pixel-level accuracy that can be expected from a human labeler, as determined from analysis of human labeling of the MS-COCO data set (Nash et al. 2022). Additionally, Das et al. (2023, 2024) showed that deep learning outperformed classic computer vision on their data set of traffic poles and street lamps.

Tan et al. (2024) showed that CNNs can effectively segment pitting corrosion, achieving an F1 score of 97% on a data set of tunnel images. Visually, pitting corrosion often has harsh edges and defined shapes, making it more similar to things compared to uniform corrosion, which behaves more like stuff. This distinction may make pitting corrosion easier for deep learning algorithms to detect. This is reflected in reported F1 scores across studies: while Tan et al. (2024) achieved 97% on pitting corrosion, Nash et al. (2022) and Rahman et al. (2021) reported F1 scores of 84% and 81%, respectively, for uniform corrosion. However, some studies report poorer performance using these methods. For example, Casas et al. (2024) achieved an F1 score of 59.4% using the YOLO architecture and Yuvaneswaren et al. (2024) reached only a 32% F1 score using YOLOv8. The differences in performance are not necessarily due to the network architecture because Jia et al. (2023) showed that YOLO can achieve an F1 score of 90.5%. Instead, these differences underscore the impact of factors beyond the model architecture. Hyperparameters such as batch size, learning rate, or loss functions play a crucial role in model performance. Additionally, the quality and diversity of training and validation data sets can significantly influence outcomes. The importance of hyperparameter tuning and data set preparation in segmentation performance has been well documented (Bianchi and Hebdon 2022). Therefore, when developing a model for bridge inspections, it is essential to consider these factors and ensure that the hyperparameters are tuned and the data set reflects the variability and complexity of real-world inspection conditions.

For bridge-specific photos, many studies have shown that CNNs can be trained to localize or segment corrosion on images of bridges with promising performance. For example, Ta et al. (2022) and Ta and Kim (2020) used region-based CNNs (RCNNs) to detect

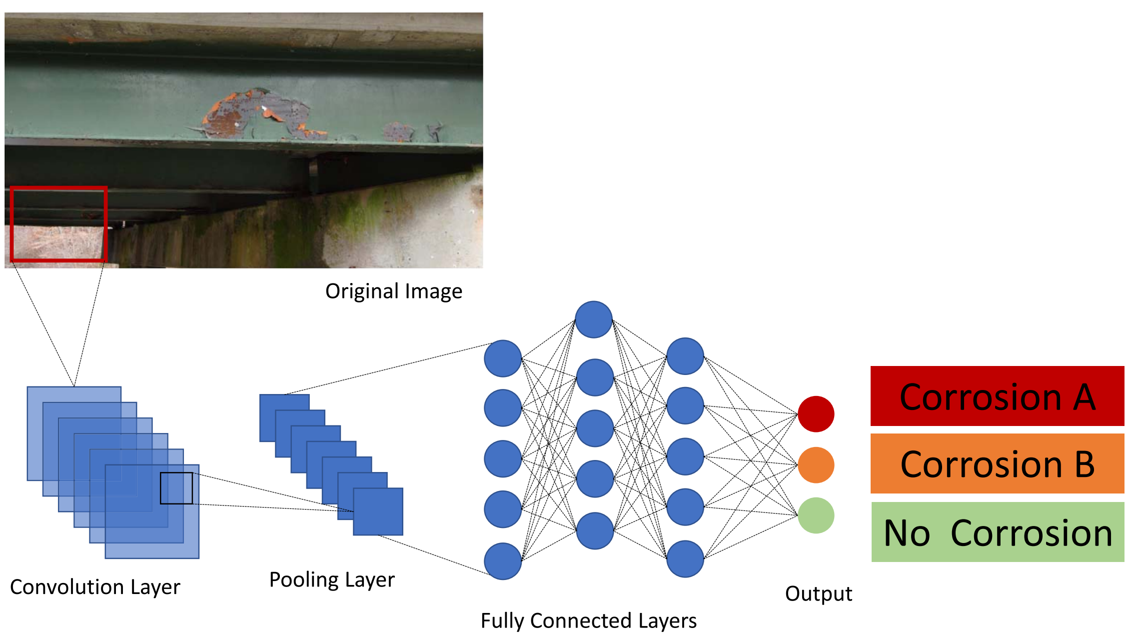


Fig. 9. Deep learning for corrosion detection example.

and localize corroded bolts and distinguish them from uncorroded ones. Nabizadeh and Parghi (2023) used YOLOv3 and Jin Lim et al. (2021) used Faster RCNN to localize corrosion in images, both achieving accuracies above 85%. Meanwhile, Fondevik et al. (2020) compared Mask RCNN and PSPNet for corrosion segmentation, achieving mean intersection-over-union (mIoU) scores of 73.2% and 77.5%, respectively. Miao et al. (2023) used YOLOv4 to localize corrosion on the bridge, then refined the corrosion boundaries using image processing on a reconstructed 3D model, achieving 79% average precision. Huang et al. (2022), Jiang et al. (2023), and Zhang et al. (2023) segmented corrosion on images of bridges using various CNN architectures, all achieving F1 scores above 89%. Bianchi and Hebdon (2021a, 2022) used DeepLabV3+ to segment corrosion condition states, determining the location and amount of fair, poor, and severe corrosion in the images with an F1 score of 86.6%. Safa et al. (2023) and Yan et al. (2025) tested SegFormer, a ViT architecture, on the same data set, achieving slightly lower performances. Fig. 10 shows an example of a deep learning approach for corrosion segmentation on images of bridges; note that the architecture is an example and not representative of any particular algorithm.

These studies show the potential of CNNs for segmenting corrosion in bridge images, providing crucial information about the location and extent of corrosion during inspections. With sufficient training data, CNNs can outperform traditional image processing and machine learning algorithms. However, these methods were all tested on human-collected data, which typically provides more detailed images of corrosion, larger portions of corrosion in the image, and is less variable. In addition, these deep learning methods require extensive training data, which can be difficult to obtain and labor-intensive to label. In practical bridge inspection applications, a CNN often encounters new data sets with environmental and lighting conditions different from the training data. For CNNs to perform well across diverse inspection data sets, they must be trained on large data sets encompassing various environmental and lighting conditions. Without this generalizability, CNNs may require retraining with site-specific data, increasing the time and effort required for the inspection process.

UAV-Collected Photos

Forkan et al. (2022) compared various CNN architectures for detecting corrosion, including those tested on human-collected photos in Atha and Jahanshahi (2018). When applied to a data set collected by UAVs, the accuracy of these algorithms dropped from 96% to 82%, due to the different compositions of the images. This highlights a key limitation of deep learning: while CNNs are powerful at learning patterns, they are also prone to overfitting if trained on narrow data sets. Transfer learning—where pretrained models are fine-tuned on new data—can partially address this issue, but only if the target data shares enough similarity with the source data.

Additionally, because UAVs collect images from various altitudes, angles, and distances, the images may have inconsistent viewpoints or scale distortion, which can lead to low-resolution views of structural elements. Lighting can also vary dramatically between images depending on the UAV's orientation relative to the sun or shaded regions of the bridge. These factors can cause damage to appear smaller, less distinct, or occluded, which can be confusing for corrosion detection algorithms that already cannot rely on clear shape-based features often needed for reliable segmentation. For instance, corrosion regions that occupy large areas in close-up human-collected photos will shrink to smaller regions in UAV-collected images, causing algorithms trained on human-captured regions to misinterpret the damage in UAV data. Collecting images with UAVs means that the algorithms can no longer rely on size information, but need to develop a contextual understanding of the structure and damage to determine the relative size of corrosion, which can be inconsistent.

To address some of these challenges, Forkan et al. (2022) developed a custom network that achieved 86.3% accuracy in segmenting corrosion on images of telecommunication structures collected by UAVs. This study shows that deep learning can be a powerful tool for corrosion segmentation, even on UAV images, but it also indicates that robust training must incorporate UAV-specific data to ensure reliable performance in the field. Generalizability across data sets remains a significant hurdle for deep learning algorithms and must be tested for and validated to support field use.

Ortiz et al. (2016) used neural networks (without the convolution layers) to segment corrosion at the pixel level but encountered

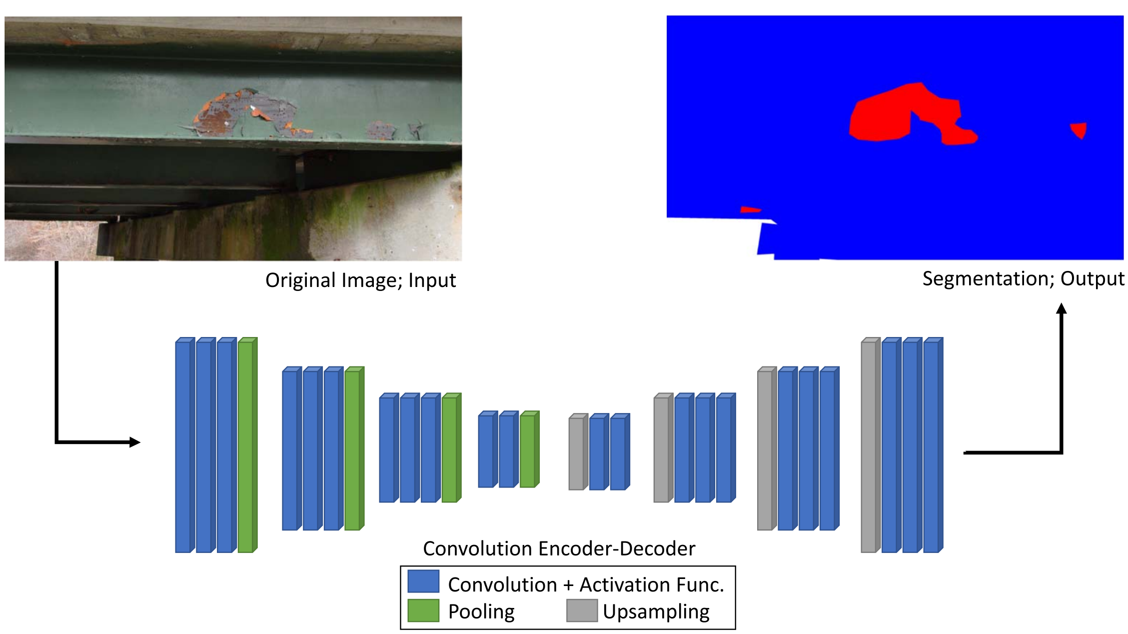


Fig. 10. Deep learning for corrosion segmentation example.

a high false positive rate, similar to classic computer vision approaches. This suggests that while neural networks are promising corrosion segmentation, deep learning methods with convolutional layers may be better suited for handling chaotic data sets.

Chen et al. (2019a), Lemos et al. (2023), Liu et al. (2019), and Sookpong et al. (2023) used various deep learning architectures to segment corrosion on images of various metal infrastructure, including cranes, roofs, and marine structures, all achieving performances above the accepted human-level F1 score of 81%. Zhou et al. (2022) extended this approach using YOLOv3 to localize corroded regions in images of cranes and then used image processing to define the precise contours. Additionally, Zhou et al. (2022) introduced methods to automatically assess corrosion severity using image processing techniques on the segments identified by deep learning, further enhancing the application of these technologies in infrastructure inspections.

Three studies have explored corrosion segmentation in UAV-collected images of bridges using deep learning, and one has explored corrosion localization. These studies represent the forefront of image-based corrosion assessment research for bridge inspection applications, highlighting both the challenges and potential of these methods. Han et al. (2021) showed that feature pyramid network (FPN) can localize corrosion in UAV-collected images of bridges with 97% accuracy. Ameli et al. (2024) showed the importance of robust testing, achieving a test F1 score of 74.5% but only 51.9% during validation when using Mask R-CNN to segment corrosion. Katsamenis et al. (2020) combined Mask R-CNN with image processing techniques to enhance corrosion contour accuracy, achieving an F1 score of 72.0%. Finally, Munawar et al. (2022) developed a deep learning framework using cycle-generative adversarial network (GAN) that achieved an F1 score of 83.0% and an accuracy of 98.9%, further demonstrating the potential of deep learning in this domain.

These studies show the potential of deep learning methods for corrosion segmentation during bridge inspections, including UAV-aided bridge inspections. Multiple CNN architectures have shown promising results on both bridge inspection data and UAV-collected data in specific cases. Additionally, some studies have extended these methods to assess corrosion severity, providing more information for bridge inspectors. However, the generalizability of these models remains a major limitation. The decreased performance when applying algorithms to data sets they were not trained on, as shown by Forkan et al. (2022), highlights the need for more diverse and comprehensive data sets. Without generalizability, these technologies cannot be integrated into bridge inspections without retraining, which limits their practical applicability and efficiency. Tables 9 and 10 summarize the deep learning studies based on methodologies and performance, respectively. Table 10 provides a comparable accuracy metric, determined from the reported metrics, to adequately compare studies.

Comparisons across Studies

Performance Evaluation Metrics and Methods

When comparing corrosion assessment methods, it is important to select performance metrics that reflect the specific task. The metrics for assessing performance vary across corrosion assessment tasks, making it difficult, for example, to compare detection algorithms to segmentation algorithms. Additionally, binary and multiclass problems require different evaluation metrics. Although corrosion is typically considered a binary problem, some studies classify different corrosion levels, creating a multiclass problem that requires separate metrics, further complicating comparisons. Direct comparison between different tasks is generally discouraged; however, in practice, inspectors may need to decide between using a detection algorithm or a segmentation algorithm, making these comparisons necessary.

Detection algorithms are typically evaluated using accuracy, precision, recall, and F1 score, on a per-image basis (Medeiros et al. 2010; Nash et al. 2020; Petricca et al. 2016). These metrics are effective for balanced data sets; however, since human inspectors can identify corrosion in images very well, detection algorithms must reach near-perfect performance to add value. Segmentation algorithms, which assign a corrosion label to each pixel, are evaluated at a finer scale. Common metrics include pixel-level accuracy, precision, recall, and F1 score. However, these metrics can be unforgiving—a few misclassified pixels can significantly lower the scores. Additionally, human performance in pixel-level segmentation is not perfect, with an F1 score of 81% (Nash et al. 2022). Segmentation algorithms must exceed this value to benefit bridge inspections, whereas detection algorithms must be close to perfect.

Many studies rely on these metrics to evaluate segmentation algorithms (Fondevik et al. 2020; Munawar et al. 2022; Nash et al. 2022; Rahman et al. 2021). However, segmentation data sets are often unbalanced, with noncorroded pixels vastly outnumbering corroded ones. In these scenarios, accuracy becomes misleading. For example, in a 10×10 image with only two corroded pixels, an algorithm that labels all pixels as background would achieve 98% accuracy despite failing completely. The F1 score, which would be 0, better captures its true performance. A study by Herndon and Tien (2025) demonstrated this issue, showing that some algorithms achieved high accuracy, yet the F1 score was 0, showing that other metrics are necessary to accurately reflect corrosion segmentation algorithm performance. A more representative metric for evaluating segmentation algorithms is the mean intersection over union (mIoU), which calculates the overlap between predicted and ground truth corroded regions, divided by their union. This metric is less affected by pixel-level noise and better reflects practical performance. Studies such as Bianchi and Hebdon (2022) and Fondevik et al. (2020) have employed mIoU in corrosion applications. Although mIoU is common in broader computer vision applications (Van Beers et al. 2019), it is underutilized in civil engineering contexts and is recommended here for structural inspection-type problems described in this paper.

Additionally, binary evaluation metrics and multiclass metrics differ. Binary metrics only consider the positive class and exclude true negatives, whereas multiclass metrics (macro, micro, or weighted averages) include all classes. For example, in the multiclass Common Objects in Context (COCO)-Bridge data set by Bianchi and Hebdon (2021a, 2022), over 70% of the pixels are background. The true positive ratio for the background is 97.1%, while for the corrosion classes, it is 55.1%, 44.9%, and 30.3%. This results in varying F1 score values, i.e., a weighted F1 score of 85.4%, a macro F1 score of 55.2%, and a binary F1 score of 67.0%—a 20-point spread depending on the evaluation method. It is important to note that classifying fair corrosion as poor corrosion yields a negative result in multiclass evaluations. These metrics are summarized in Table 11. These differences illustrate how metric choice can significantly affect the perceived success of a model.

Based on detailed evaluation across studies, this work recommends that researchers provide precision and recall, or combined metrics such as F1 score or mIoU, when evaluating segmentation algorithms. Accuracy alone is insufficient. Additionally, these metrics should be disaggregated by class so the performance

Table 9. Summary of deep learning studies: methodologies

No.	Study	Algorithm type	Methods	Application	Corrosion type	Number of images
1.	Furuta et al. (1995)	Segmentation	Custom NN	Laboratory	Uniform	400
2.	Petricca et al. (2016)	Detection	AlexNet	Laboratory	Uniform	3,500
3.	Jahanshahi and Masri (2013a, b)	Detection	Custom NN	Laboratory	Uniform	2,059
4.	Xu et al. (2020)	Segmentation	Faster R-CNN	Laboratory	Uniform	538
5.	Yao et al. (2019)	Detection	AlexNet	Laboratory	Uniform	330
6.	Zhang et al. (2021)	Localization	CAMCD	Laboratory	Uniform	1,719
7.	Naik et al. (2020)	Localization	Multilayer perceptron	Laboratory	Pitting	40
8.	Qian et al. (2025)	Segmentation	Custom CNN	Laboratory	Pitting	Unspecified
9.	Son et al. (2024)	Segmentation	Mask R-CNN	Laboratory	Pitting	35,753
10.	Bastian et al. (2019)	Detection	Custom CNN	Field	Uniform	142,000 (after augmentation)
11.	Holm et al. (2020)	Detection	VGG16	Field	Uniform	9,300
12.	Alviz-Meza et al. (2024)	Detection	EfficientNet	Field	Uniform	800
13.	Nash et al. (2020)	Detection	Custom CNN	Field	Uniform	2,539
14.	Khayatazad et al. (2023)	Segmentation	Custom NN	Field	Uniform	30
15.	Malashin et al. (2024)	Detection	Custom CNN	Field	Uniform	576,000
16.	Atha and Jahanshahi (2018)	Localization	Custom CNN	Field	Uniform	926
17.	Nash et al. (2022)	Segmentation	HRNet	Field	Uniform	225
18.	Rahman et al. (2021)	Segmentation	DeepLab V3+	Field	Uniform	10,000
19.	Jia et al. (2023)	Localization	YOLO	Field	Uniform	649
20.	Chen et al. (2021)	Segmentation	U-Net	Field	Uniform	213
21.	Das et al. (2024)	Segmentation	EfficientNetB7	Field	Uniform	300
22.	Tan et al. (2024)	Segmentation	Custom CNN	Field	Uniform	1,441
23.	Casas et al. (2024)	Segmentation	YOLOv5, v8	Field	Uniform	4,942
24.	Yuvaneswaren et al. (2024)	Segmentation	YOLOv8	Field	Uniform	4,942
25.	Ta and Kim (2020)	Localization	RCNN	Field	Pitting	419
26.	Ta et al. (2022)	Localization	Mask R-CNN	Field	Pitting	577
27.	Nabizadeh and Parghi (2023)	Localization	YOLOv3	Field	Pitting	129
28.	Jin Lim et al. (2021)	Localization	Faster RCNN	Field	Pitting	36
29.	Fondevik et al. (2020)	Segmentation	Mask R-CNN	Field	Uniform	603
—	—	—	PSPNet	—	—	—
30.	Miao et al. (2023)	Segmentation	YOLOv4	Field	Uniform	295
31.	Bianchi and Hebdon (2022)	Segmentation	DeepLabV3+	Field	Uniform (multiclass)	440
32.	Huang et al. (2022)	Segmentation	Custom CNN	Field	Uniform	440
33.	Zhang et al. (2023)	Segmentation	HRNet	Field	Uniform	1,719
34.	Jiang et al. (2023)	Segmentation	FAU-Net	Field	Uniform	77
35.	Safa et al. (2023)	Segmentation	SegFormer (ViT)	Field	Uniform	440
36.	Yan et al. (2025)	Segmentation	SegFormer (ViT)	Field	Uniform	2,703
37.	Forkan et al. (2022)	Localization	Custom CNN	Field-UAV	Uniform	573
38.	Ortiz et al. (2016)	Segmentation	Custom NN	Field-UAV	Uniform	25
39.	Chen et al. (2019b)	Localization	VGG 16	Field-UAV	Uniform	63
40.	Liu et al. (2019)	Localization	VGG16	Field-UAV	Uniform	1,900
41.	Lemos et al. (2023)	Segmentation	Mask R-CNN	Field -UAV	Uniform	8,400
42.	Sookpong et al. (2023)	Segmentation	ViT-Adapter	Field-UAV	Uniform	660
43.	Zhou et al. (2022)	Segmentation	YOLO V3	Field-UAV	Uniform	1,500
44.	Han et al. (2021)	Localization	FPN	Field-UAV	Pitting	341
45.	Katsamenis et al. (2020)	Segmentation	Mask R-CNN	Field-UAV	Uniform	116
46.	Ameli et al. (2024)	Segmentation	Mask R-CNN	Field-UAV	Uniform	514
47.	Munawar et al. (2022)	Segmentation	U-Net + CycleGAN	Field-UAV	Uniform	600

comparison between binary and multiclass problems is meaningful and not skewed by class imbalance.

Generalizability of Methods across Data Sets

In considering the generalizability of varying methods across data sets, comparing outcomes from classic computer vision to those from deep learning approaches reveals the advancements of new approaches. Classic computer vision algorithms can be prone to false positives, as revealed in Khayatazad et al. (2020) and Petricca et al. (2016). While these methods perform well in controlled environments, their performance tends to drop when applied to real-world environments where chaos and variability impact their effectiveness. Bridge inspections, in particular, involve high levels of noise and variation, so algorithms must be robust against these challenges. Classic computer vision algorithms can still be effective when misleading elements are filtered out, but manually doing so reduces the efficiency of the algorithm. Marchewka et al. (2020) addressed this by developing a system that automatically removes irrelevant elements before classic computer vision automatically segments corrosion in UAV-captured bridge elements, achieving a precision of more than 96%. However, a major limitation of the study is that it only detects corrosion in predefined areas, such as bolts or rivets, and excludes other regions from the analysis.

Deep learning methods are generally more robust against false positives, showing only minor performance drops compared to classic computer vision algorithms. They have demonstrated good performance across various environments, including in bridge inspections. For instance, Nash et al. (2022) used HRNet to segment corrosion in UAV-collected images of pipes, achieving an F1 score of 84%. Munawar et al. (2022) trained CycleGAN on

J. Bridge Eng., 2025, 30(12): 03125001

Table 10. Summary of deep learning studies: performance

No.	Study	Algorithm type	Methods	Appl.	Corrosion type	Evaluation metric	Performance (%)	Accuracy (%)
1.	Munawar et al. (2022)	Segmentation	U-Net + CycleGAN	Field-UAV	Uniform	Accuracy F1 score	98.9 83.3	98.9
2.	Yao et al. (2019)	Detection	AlexNet	Laboratory	Uniform	Accuracy	98.9	98.9
3.	Bastian et al. (2019)	Detection	Custom CNN	Field	Uniform	Accuracy	98.8	98.8
4.	Ta and Kim (2020)	Localization	RCNN	Field	Pitting	Accuracy	98.7	98.7
5.	Atha and Jahanshahi (2018)	Localization	Custom CNN	Field	Uniform	F1 score	98.5	98.5
6.	Malashin et al. (2024)	Detection	Custom CNN	Field	Uniform	Accuracy	98.0	98.0
7.	Das et al. (2024)	Segmentation	EfficientNetB7	Field	Uniform	F1 score	97.8	97.8
8.	Holm et al. (2020)	Detection	VGG16	Field	Uniform	Accuracy F1 score	97.7 95.5	97.7
9.	Zhang et al. (2021)	Localization	CAMCD	Laboratory	Uniform	F1 score	97.5	97.5
10.	Tan et al. (2024)	Segmentation	Custom CNN	Field	Uniform	F1 score	97.0	97.0
11.	Han et al. (2021)	Localization	FPN	Field-UAV	Pitting	Accuracy	97.0	97.0
12.	Ta et al. (2022)	Localization	Mask R-CNN	Field	Pitting	Accuracy	96.3	96.3
13.	Zhou et al. (2022)	Segmentation	YOLOv3	Field-UAV	Uniform	F1 score	96.1	96.1
14.	Sookpong et al. (2023)	Segmentation	ViT-Adapter	Field-UAV	Uniform	F1 score	94.5	94.5
15.	Chen et al. (2021)	Segmentation	U-Net	Field	Uniform	Accuracy	94.0	94.0
16.	Son et al. (2024)	Segmentation	Mask R-CNN	Laboratory	Pitting	Accuracy	94.0	94.0
17.	Alviz-Meza et al. (2024)	Detection	EfficientNet	Field	Uniform	F1 score	93.0	93.0
18.	Nash et al. (2020)	Detection	Custom NN	Field	Uniform	Accuracy	93.0	93.0
19.	Jiang et al. (2023)	Segmentation	FAU-Net	Field	Uniform	F1 score	91.9	91.9
20.	Naik et al. (2020)	Localization	Multilayer perceptron	Laboratory	Pitting	Accuracy	91.0	91.0
21.	Zhang et al. (2023)	Segmentation	HRNet	Field	Uniform	F1 score	90.5	90.5
22.	Jia et al. (2023)	Localization	YOLO	Field	Uniform	F1 score	90.5	90.5
23.	Jahanshahi and Masri (2013a, b)	Detection	Custom NN	Field	Uniform	Accuracy	90.5	90.5
24.	Xu et al. (2020)	Segmentation	Faster R-CNN	Laboratory	Uniform	Error	10	90.0
25.	Liu et al. (2019)	Localization	VGG16	Field-UAV	Uniform	Accuracy	89.5	89.5
26.	Huang et al. (2022)	Segmentation	Custom CNN	Field	Uniform	F1 score	89.1	89.1
27.	Jin Lim et al. (2021)	Localization	Faster RCNN	Field	Pitting	mAP	88.7	88.7
28.	Bianchi and Hebdon (2022)	Segmentation	DeepLabV3+	Field	Uniform (Multiclass)	F1 score	88.6	88.6
29.	Rahman et al. (2021)	Segmentation	DeepLab V3+	Field	Uniform	Accuracy F1 score	88.0 81.0	88.0
30.	Ortiz et al. (2016)	Segmentation	Custom NN	Field-UAV	Uniform	Success rate	87.0	87.0
31.	Forkan et al. (2022)	Localization	Custom CNN	Field-UAV	Uniform	Accuracy	86.3	86.3
32.	Nabizadeh and Parghi (2023)	Localization	YOLOv3	Field	Pitting	F1 score	85.0	85.0
33.	Nash et al. (2022)	Segmentation	HRNet	Field	Uniform	F1 score	84.0	84.0
34.	Lemos et al. (2023)	Segmentation	Mask R-CNN	Field-UAV	Uniform	F1 score	84.9	84.9
35.	Miao et al. (2023)	Segmentation	YOLOv4	Field	Uniform	mAP	79.0	79.0
36.	Petricca et al. (2016)	Detection	AlexNet	Laboratory	Uniform	Accuracy	78.0%	78.0
37.	Fondevik et al. (2020)	Segmentation	Mask R-CNN PSPNet	Field	Uniform	mIoU	73.2 77.5	77.5
38.	Khayatazad et al. (2023)	Segmentation	Custom NN	Field	Uniform	F1 score	74.0	74.0
39.	Katsamenis et al. (2020)	Segmentation	Mask R-CNN	Field-UAV	Uniform	mIoU	72.0	72.0
40.	Safa et al. (2023)	Segmentation	SegFormer (ViT)	Field	Uniform	mIoU	71.2	71.2
41.	Yan et al. (2025)	Segmentation	SegFormer (ViT)	Field	Uniform	mIoU	67.7	67.7
42.	Casas et al. (2024)	Segmentation	YOLOv5, v8	Field	Uniform	F1 score	59.4	59.4
43.	Ameli et al. (2024)	Segmentation	Mask R-CNN	Field-UAV	Uniform	F1 score	51.9	51.9
44.	Yuvaneswaren et al. (2024)	Segmentation	YOLOv8	Field	Uniform	F1 score	32.4	32.4
45.	Chen et al. (2019b)	Localization	VGG 16	Field-UAV	Uniform	Visual	X	X
46.	Qian et al. (2025)	Segmentation	Custom CNN	Laboratory	Pitting	Visual	X	X
47.	Furuta et al. (1995)	Detection	Multilayer NN	Laboratory	Uniform	Visual	X	X

UAV images of bridges, reaching an F1 score of 83%. Miao et al. (2023) used YOLO V4 to detect corrosion in bridge images, reaching 80% accuracy with 295 images. Sookpong et al. (2023) used ViT-Adapter to segment corrosion in UAV-collected images of offshore oil infrastructure, achieving an F1 score of 94.5%. Forkan et al. (2022) developed a custom CNN for corrosion segmentation on 573 UAV bridge images, achieving 86% accuracy. Finally, Rahman et al. (2021) created a semisupervised labeling mechanism to develop a 10,000-image data set, which helped train DeepLabV3 and resulted in an F1 score of 81.0% and an accuracy of 89%.

Table 11. Performance metrics of deep learning models on the COCO-Bridge data set

Model	Accuracy	Weighted F1 score	Macro F1 score	Positive F1 score
Bianchi and Hebdon (2021a)	0.860	0.854	0.552	0.670

While ViTs have shown strong performance in some domains, CNNs continue to outperform ViTs on corrosion segmentation in the COCO-Bridge data set (Bianchi and Hebdon 2021b), with CNNs achieving an mIoU of 83% but ViTs reaching 68% and 71% (Bianchi and Hebdon 2022; Safa et al. 2023; Yan et al. 2025). While these methods still need to be tested for their generalizability and replicability, these studies highlight the potential of deep learning in complex environments, showing that these methods can enhance corrosion detection beyond human inspection alone.

Fig. 11 illustrates these findings, comparing the segmentation accuracy of classic computer vision and deep learning methods, both tested with and without misleading objects, averaged across all studies discussed. Average accuracy is selected as the measure of comparison to facilitate evaluation across studies because it is the most commonly reported performance measure.

Fig. 11 shows that classic computer vision methods experience a more significant performance drop compared to deep learning methods when faced with misleading objects. This suggests that deep learning algorithms maintain stronger performance even in challenging environments, including UAV-based inspections. The similar performance of deep learning in studies with misleading objects and in UAV studies implies that these algorithms are more adaptable to use in UAV-aided bridge inspections. Note that classic computer vision appears to perform better in UAV studies because the only study that attempted this, (Marchewka et al. 2020), automatically identified regions of interest, specifically rivets and beams, and removed other elements. However, this algorithm is highly particular to riveted steel truss bridges and heavily relies on robust bridge element identification. If the region-of-interest identifier does not work correctly, it performs with many false positives. Meanwhile, over 10 studies have explored using deep learning to segment corrosion in UAV images. While the consistent performance of deep learning algorithms on data with misleading objects and UAV images suggests the suitability of these methods for UAV-based conditions, the overall decreased performance levels compared to data without misleading objects reveal an ongoing gap in performance, even for deep learning approaches.

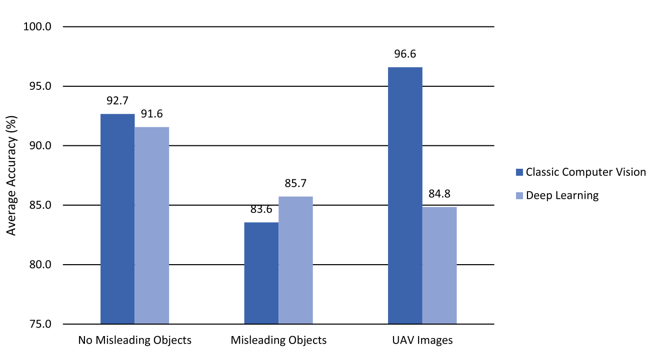


Fig. 11. Quantitative accuracy comparison of corrosion segmentation methods tested on data with and without misleading objects.

In addition, a key challenge for deep learning algorithms is their lack of generalizability across data sets. On average, these methods achieve accuracies above 85% and F1 scores above 82% (when reported) in scenarios that included misleading objects. However, they often only perform well on the data set they are trained on, struggling with external data sets even when measures to reduce overfitting, such as *k*-fold cross-validation and smaller batches, are implemented, as in Bianchi and Hebdon (2022). Forkan et al. (2022) demonstrated this issue, showing that the algorithm developed by Atha and Jahanshahi (2018) experienced a significant performance drop when applied to UAV images compared to the images it was originally trained on, and Herndon and Tien (2025) showed that classic machine learning algorithms produce more consistent results across data sets than deep learning algorithms. This highlights a broader limitation in the field: most studies do not test their algorithms on completely external data sets. While it is common to split a data set into training, validation, and testing subsets, all three generally originate from the same collection, which limits insight into how well the model generalizes to new environments. Testing on external data sets is rarely performed but crucial for assessing the generalizability and practical applicability of corrosion assessment algorithms.

For example, Nash et al.'s (2022) algorithm reached 84% accuracy on its training data set, but when applied to an external data set, the F1 score dropped to 40%, with a corresponding mIoU of 38%. Similarly, in the study by Bianchi and Hebdon (2022), the F1 score of the algorithm decreased from 88.6% to 58.8% when tested on an external data set. Fig. 12 illustrates this issue, showing how these algorithms perform on external UAV-collected bridge inspection images. In the image, the small regions highlighted in white represent true positives, the large regions highlighted in dark blue represent true negatives, and the other two regions represent false negatives and false positives as indicated. In this particular image, the F1 score when using the algorithm by Nash (2022) and Nash et al. (2022) was 18.5%, while the F1 score when using the algorithm by Bianchi and Hebdon (2021a) was 22.0%. Such performance reductions make applying these models in the field impractical without fine-tuning and retraining the algorithms, requiring additional labeled data and extensive time and effort to produce.

This lack of generalizability poses a significant hurdle for the practical use of deep learning in bridge inspections. Inspectors cannot expect the same high performance observed in the original study when applying the model to additional bridges without retraining, unless they fine-tune or retrain the algorithm on specific data from each new structure. These processes involve manually labeling new data, one of the most time-consuming tasks in deep learning. Including this step in the process significantly reduces the overall efficiency of these methods. Addressing the challenge of data labeling is crucial for improving the practical applicability of deep learning for corrosion segmentation. Semisupervised labeling techniques, such as those proposed by Rahman et al. (2021), could streamline the labeling process and enhance the generalizability of deep learning algorithms, making them more viable for widespread field use.

Data Preparation, Cleaning, and Preprocessing

In addition to the corrosion assessment task and algorithm performance, a key aspect that differentiates studies is the type and level of data cleaning and preprocessing. Jahanshahi and Masri (2013b) and Ortiz et al. (2016) investigated the effects of varying amounts of data preparation on neural network performance, finding that different combinations of color and texture descriptors influenced

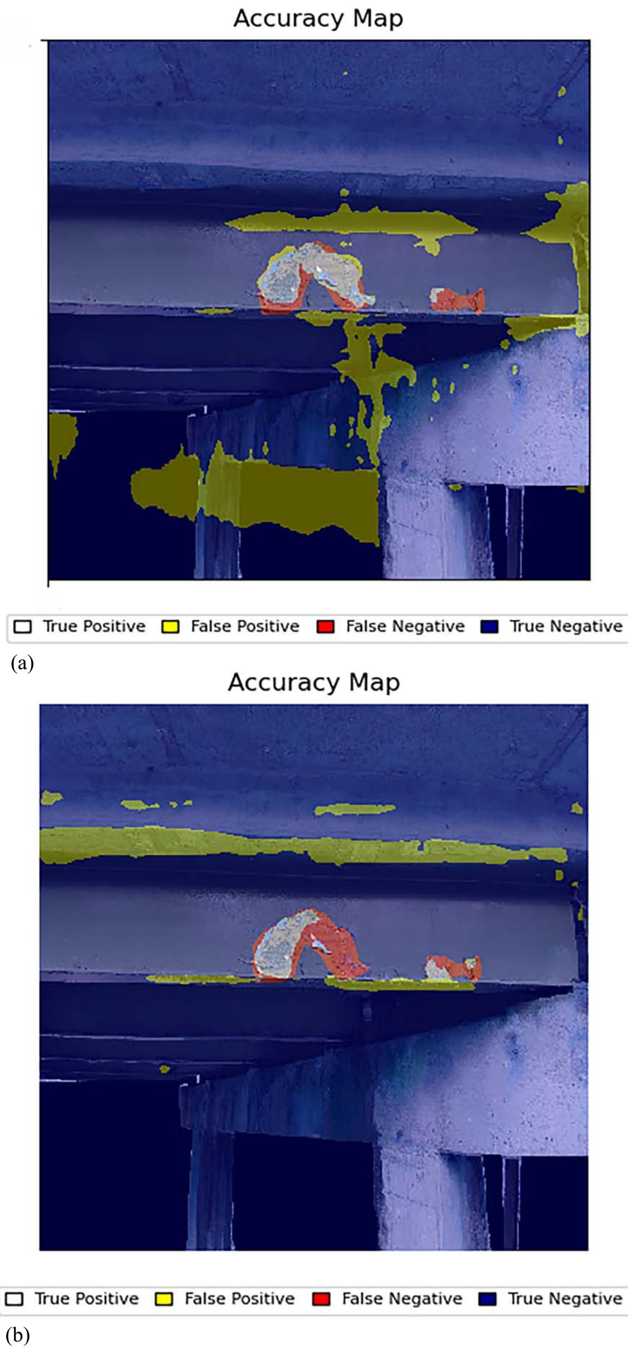


Fig. 12. Results of the corrosion segmentation models developed by (a) Nash (2022) and Nash et al. (2022); and (b) Bianchi and Hebdon (2021a, 2022) for corrosion segmentation applied to external data.

algorithm outcomes. For CNN frameworks, we would not expect image processing to enhance performance because feature extraction is included in the algorithm. However, Atha and Jahanshahi (2018) and Liu et al. (2019) investigated the impact of data preprocessing in CNN frameworks, yielding mixed results on its effectiveness. For example, Atha and Jahanshahi (2018) tested the effects of different color spaces on five CNN architectures, discovering that RGB (the original color space) and YCbCr perform best on most algorithms, although the improvements were not large. Liu et al. (2019) employed gamma correction, brightness adjustment, and histogram equalization to enhance the images before placing them through the CNN, finding that preprocessing improves the algorithm's performance on this data set. However, their methodology yields similar results to those not preparing their data this way. These results suggest that image processing does not offer significant advantages for deep learning methods.

In classic computer vision frameworks, however, data preparation plays a more significant role. Marchewka et al. (2020) demonstrated that well-executed preprocessing can make these algorithms viable for UAV-aided bridge inspections. In this study, the authors created a framework for UAV-aided bridge inspections that includes route planning and identification of bridge elements that are prone to corrosion. Using the scale invariant feature transform (SIFT) algorithm, they identified bolts and connections and then applied computer vision to assess corrosion on these regions. This approach achieved 96.6% precision on UAV-collected images, a stark improvement over other studies using classic computer vision methods, which faced too many false positives to be applicable. Their findings show that with informed, sufficient data preparation, computer vision methods can be effective in corrosion assessment during UAV-aided bridge inspections.

Some researchers have also developed hybrid frameworks that use deep learning to prepare images for classic computer vision algorithms. For example, Katsamenis et al. (2020) and Zhou et al. (2022) first localized corrosion using a trained CNN and then refined the contours using color thresholding. Zhou et al. (2022) achieved an impressive pixel-level F1 score of 96.1% on UAV-collected images. For comparison, the best F1 score for corrosion segmentation on UAV-collected images using purely deep learning was 84.9% (Lemos et al. 2023). This shows that pairing deep learning with classic computer vision can improve the individual performance of both. Additionally, the methodology from Zhou et al. (2022) was trained on 1,500 images, as opposed to 8,400 in the study by Lemos et al. (2023). This shows that when used correctly, combining CNNs and classic computer vision can not only outperform CNNs alone but also reduce the amount of labeling needed to achieve high performance.

Discussion

Despite the recent development of computer vision and deep learning algorithms and their applications to corrosion assessment, many challenges remain that prevent these technologies from full deployment in the field for bridge inspections. Fig. 13 outlines these challenges and presents a vision for automated corrosion assessment during bridge inspections.

While useful for developing and testing corrosion assessment technologies, corrosion detection algorithms have limitations in supporting decision-making during bridge inspections. For these algorithms to be effective, they must perform with near-perfect accuracy to surpass human inspectors. Even then, inspectors would still need to manually identify the precise location and extent of the damage. As a result, even the best-performing detection algorithms (Jahanshahi and Masri 2013b; Medeiros et al. 2010) would fail to improve the efficiency, accuracy, or objectivity of inspections. In contrast, corrosion localization and segmentation algorithms offer more practical value by providing inspectors with detailed information on the location and extent of corrosion (Bianchi et al. 2021; Marchewka et al. 2020; Munawar et al. 2022), aiding in more informed decision-making.

Some classic computer vision approaches have achieved valuable segmentation results. For example, Bonnin-Pascual and Ortiz (2014), Khayatazad et al. (2020), and Moranduzzo and Melgani (2014) used color and texture thresholding, achieving

Current State	Challenges	Proposed Vision
Classic computer vision algorithms for segmenting corrosion	• Prone to false positives • Bridge environments have irrelevant and misleading objects (foliage, graffiti, etc.) • Automatic removal methods might remove corroded areas	Image capture planning and multiple algorithms for reducing misleading objects
Deep learning algorithms for segmenting corrosion	• Require large amounts of labeled data to train • Surface corrosion lacks distinct shape-based features • Low generalizability across datasets • Require fine-tuning or retraining to be used on new bridges	Establish collaborative labeling efforts across DOTs and researchers to build diverse corrosion image datasets and promote training on larger datasets
UAV-aided bridge inspections	• UAV images capture large areas with small corroded regions, increasing the signal to noise ratio • Environmental clutter (debris, grime, foliage) visually resembles corrosion • Only 6/68 studies investigated corrosion segmentation on UAV-collected images of bridges • A robust corrosion segmentation methodology for bridge inspections has yet to be developed	UAV flight planning to mitigate image inconsistency; prioritize algorithm testing on UAV-collected imagery of corroded bridges
Corrosion severity assessment	• Corrosion degree estimation can be subjective • Some severity metrics are not correlated with bridge inspection specifications (BIRM, AASHTO) • Multiclass models often underperform compared to binary segmentation	Validate visual severity against corrosion penetration measurements aligned with bridge inspection specifications

Fig. 13. Challenges toward fully automated corrosion assessment during bridge inspections.

accuracies up to 96%, providing inspectors with useful data on the amount of corrosion present. However, classic computer vision algorithms are prone to false positives, which can lead to accuracy dropping below 70% in cases with misleading objects, particularly red objects, or nonuniform lighting (Khayatazad et al. 2023). This is a large limitation in bridge inspections, where misleading objects such as vegetation and graffiti are common. When paired with automatic object removal techniques, these algorithms can perform more reliably, as demonstrated in the study by Marchewka et al. (2020). However, there is a significant concern; these methods may remove corroded areas before they are detected, as the algorithms focus on components more prone to corrosion while eliminating other areas. A comprehensive strategy, involving image capture planning, multiple algorithms for reducing misleading objects, and then corrosion segmentation, may mitigate this. Incorporating varied removal techniques increases the likelihood of segmenting corrosion across more regions of the bridge.

Recent advancements in deep learning have shown promise for enhancing the usability of imagery-based algorithms in bridge inspections. Studies by Bianchi et al. (2021), Munawar et al. (2022), Nash et al. (2022), Rahman et al. (2021), Sookpong et al. (2023), and Zhou et al. (2022) have demonstrated the capacity of these methods to accurately segment corrosion in photos with misleading objects, surpassing the human benchmark with F1 scores of 88.6%, 83.0%, 84.0%, 81.0%, 94.5%, and 96.0%, respectively. These algorithms extract precise information about corrosion, such as its location, area, and severity, making them applicable for decision-making during bridge inspections. The results highlight the promise of deep learning approaches for bridge inspections because these algorithms are less prone to false positives and provide more comprehensive information. However, significant limitations remain, particularly the need for large labeled data sets and the low generalizability of the algorithms across different data sets. Currently, these models require fine-tuning or retraining when applied to bridges outside the original data set, which requires additional labeling. Increasing data sharing to train models on larger, more diverse data sets can improve their generalizability. Additionally, semisupervised labeling methods, like those used by Rahman et al. (2021), could speed up the labeling process required for fine-tuning or retraining, further increasing the practicality of these algorithms in real-world bridge inspections.

In addition, these algorithms must also be tested on UAV-collected images if they are to be used for UAV-aided bridge inspections. While it is possible for human-collected photos to have nonuniform lighting and misleading objects, UAV-collected images tend to be even more challenging for automated algorithms to decipher. They are often noisier, more chaotic, and vary widely in image quality. Forkan et al. (2022) showed that an algorithm that achieves a 96% accuracy on human-collected images drops to 82% on UAV-collected images. Similar drops in performance were observed when the high-performing algorithms developed by Nash et al. (2020) and Bianchi and Hebdon (2022) were tested on external UAV-collected data sets. This highlights the importance of testing algorithms on UAV data sets to properly evaluate their effectiveness for automated bridge inspections.

Of the 68 studies reviewed, only six investigated corrosion segmentation on UAV-collected bridge images. These studies reported varied results, with F1 scores ranging from 51.0% to 83.0%. Specifically, Ameli et al. (2024) achieved F1 scores of 51.0% and 74.5%, while Munawar et al. (2022) reported F1 scores of 83.0%. Herndon and Tien (2025) and Katsamenis et al. (2020) reported mIoUs of 72.0%. The other two studies, Han et al. (2021) and Marchewka et al. (2020) did not report F1 scores or mIoU, only accuracy or precision. The range of outcomes for these approaches and the presence of a single study with an F1 score above 80% suggests that a robust, consistently performing methodology for automatic corrosion segmentation during bridge inspections has yet to be developed. All other studies relied on human-collected photos or images collected of other objects such as pipes or vessels. These findings suggest that corrosion segmentation during UAV-aided bridge inspections is still underdeveloped compared to other corrosion segmentation applications. For comparison, methods perform with an average F1 score on UAV-collected bridge data of 69.8% compared to an average F1 score of 84.5% in other applications. Further testing and development are needed to bring UAV-based segmentation to the same level of readiness for deployment in bridge inspections.

The final step in applying automatic corrosion segmentation algorithms for bridge inspections is assessing corrosion severity. Studies by Rahman et al. (2021) and Zhou et al. (2022) quantify corrosion severity by analyzing the darkness of corroded pixels after segmentation, operating under the assumption that darker pixels indicate more severe corrosion. While straightforward, this approach may not align with bridge inspection standards, which define element condition ratings based on the presence of corrosion and section loss. Bianchi and Hebdon (2022) take a different approach, segmenting multiple corrosion classes based on the Bridge Inspector's Reference Manual (BIRM) (Ryan et al. 2023), identifying regions of fair, poor, and severe corrosion. To reduce subjectivity, they use the average of 20 human annotations, producing outputs that are more consistent with current guidelines. However, to clearly assess performance, both binary and multiclass segmentation metrics need to be assessed. In Bianchi and Hebdon's (2022) case, the weighted F1 score is 88.6%, but the binary F1 score for all corrosion is 67.0%, suggesting that this multiclass approach may not yet match the performance of binary methods. For multiclass methods, it is important to report both weighted and binary F1 scores to provide inspectors with a clearer assessment and understanding of expected algorithm performance. More research, data, and algorithm training are needed to bring corrosion severity assessment to the level of performance required for practical use in bridge inspections.

Despite recent advances, automated corrosion assessment remains far from full implementation in bridge inspections. This review underscores the need for targeted improvements in algorithm generalizability, data set diversity, and validation under real-world conditions. Future studies should prioritize testing on UAV-collected imagery because performance on these data sets is notably lower than on human-collected images, likely due to greater variability and noise. Research efforts should also focus on developing robust, hybrid methodologies that can allow models to adapt across environments. Expanding public data sets through data sharing and using semisupervised or self-supervised learning models will reduce the labeling burden while improving model robustness. Additionally, future algorithms must be benchmarked using both binary and multiclass metrics to better align with field conditions and inspection standards. Finally, severity estimation approaches must evolve to reflect structural guidelines, such as those from BIRM or AASHTO, and be validated across measured corrosion severity values. By addressing these gaps, future research can move the field closer to reliable, deployable solutions for UAV-aided bridge inspections.

Conclusion

The advent of affordable UAVs equipped with digital cameras presents an opportunity to enhance the safety and efficiency of bridge inspections. With this innovation, an increasing number of state DOTs are exploring the integration of UAVs into their bridge inspection procedures. Since corrosion is one of the most common causes of bridge failure, the automatic assessment of corrosion from UAV imagery holds immense value for bridge inspectors. Recent advancements in computer vision and deep learning have opened the door for research into automated corrosion assessment and its potential use for bridge inspections. This research provides a focused, systematic review of automatic corrosion assessment algorithms, analyzing 68 articles, discussing their applicability for automated or semiautomated bridge inspections, and informing the development and deployment of vision-based corrosion assessment tools that can ultimately support faster, safer, and more consistent bridge inspections.

Accurate and comprehensive corrosion assessment is crucial because subpar information can pose significant problems. Excessive false positives, for instance, do not narrow down the areas of potential damage and may give inspectors misleading information. On the other hand, false negatives may lead to unnoticed corrosion, with potentially dire consequences such as unresolved structural issues and, in the worst case, failure.

The analysis of proposed approaches found that classic computer vision techniques often fall short because they are typically too prone to false positives, especially on the inherently chaotic data that would be collected during bridge inspections via UAVs. Meanwhile, deep learning techniques have shown promising results, performing with pixel-level accuracies ranging from 80% to 98% on UAV-collected data and surpassing the human benchmark of 81%. However, these algorithms come with different limitations, including high computational cost, a need for substantial training data, and a lack of generalizability to new data sets. When inspectors prefer more computationally efficient methods, they could combine classic computer vision and machine learning techniques.

Multiple researchers have developed deep learning algorithms that perform well on UAV-collected data, which could be used to support decisions by human bridge inspectors. However, the generalizability of these algorithms must be improved before they can be deployed for bridge inspections across the country. When these algorithms are used to segment corrosion on an image external to the training data set, the performance drops considerably. One way for bridge inspectors to mitigate low generalizability is to fine-tune the deep learning framework to fit the application in question, which requires a smaller amount of labeled data than full training. Alternatively, researchers could improve generalizability by training and testing deep learning algorithms on larger, more diverse data sets. To achieve either of these goals, semisupervised labeling frameworks may reduce the manual effort and time required to label images.

Finally, although recent strides have been made in achieving accurate and efficient corrosion assessment from UAV-collected images, methods for determining corrosion severity, whether based on darkness-based assessments or multiclass segmentation, need further evaluation and validation before they can be integrated into bridge inspection methods. In some cases, these methods must determine their consistency with bridge inspection specifications, while others need improvements in performance before they can be adopted for bridge inspections. Future research should focus on improving the recall of localization and segmentation algorithms, followed by developing corrosion severity assessment methods that are consistent with bridge inspection specifications. Alternatively, if automatic corrosion severity continues to fall short, DOTs can use high-performing segmentation algorithms to identify regions of corrosion requiring further inspection and use other nondestructive evaluation tools to determine the corrosion severity.

Data Availability Statement

The collated accuracy values and segmentation results presented in this study are available from the corresponding author upon reasonable request.

Acknowledgments

This project was partially funded by the INSPIRE University Transportation Center (UTC). Financial support for INSPIRE UTC projects is provided by the US Department of Transportation, Office of the Assistant Secretary for Research and Technology (USDOT/OST-R) under Grant no. 69A3551747126 through the INSPIRE University Transportation Center at the Missouri University of Science and Technology. The views, opinions, findings, and conclusions reflected in this publication are solely those of the authors and do not represent the official policy or position of the USDOT/OST-R or any state or other entity. Additional funding support from the Georgia Department of Transportation through Grant no. RP22-01 is also acknowledged. Author Herndon acknowledges support by the National Science Foundation Graduate Research Fellowship under Grant no. DGE-2039655. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Support from the Williams Family Professorship for Author Tien is also acknowledged.

Author Contributions

Hana N. Herndon: Conceptualization, Formal analysis, Investigation, Software, Writing—original draft, Writing—review and editing; Iris Tien: Funding acquisition, Resources, Supervision, Writing—review and editing.

References

Abdallah, A. M., R. A. Atadero, and M. E. Ozbek. 2022. "A state-of-the-art review of bridge inspection planning: Current situation and future needs." *J. Bridge Eng.* 27 (2): 03121001. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001812.

Ahuja, S. K., and M. K. Shukla. 2018. "A survey of computer vision based corrosion detection approaches." *Information and communication technology for intelligent systems. Vol. 2 of smart innovation, systems and technologies*, edited by S. C. Satapathy and A. Joshi, 55–63. Cham, Switzerland: Springer.

Aliasghar-Mamaghani, M., I. Koutromanos, C. Roberts-Wollmann, and M. Hebdon. 2022. "Analytical evaluation of corrosion-induced strength degradation in prestressed bridge girders." *J. Struct. Eng.* 148 (11): 04022188. https://doi.org/10.1061/(ASCE)ST.1943-541X.0003503.

Alviz-Meza, A., L. L. Hadechini-Meza, and D. Y. Peña-Ballesteros. 2024. "Deep neural networks for external corrosion classification in industrial above-ground storage tanks." *Heliyon* 10 (15): e34882. https://doi.org/10.1016/j.heliyon.2024.e34882.

Ameli, Z., S. J. Nesheli, and E. N. Landis. 2024. "Deep learning-based steel bridge corrosion segmentation and condition rating using mask RCNN and YOLOv8." *Infrastructures* 9 (1): 3. https://doi.org/10.3390/infrastructures9010003.

ASCE. 2021. *ASCE infrastructure report card—Bridges*. Reston, VA: ASCE.

Atha, D. J., and M. R. Jahanshahi. 2018. "Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection." *Struct. Health Monit.* 17 (5): 1110–1128. https://doi.org/10.1177/1475921717737051.

Ayele, Y. Z., M. Aliyari, D. Griffiths, and E. L. Droguett. 2020. "Automatic crack segmentation for UAV-assisted bridge inspection." *Energies* 13 (23): 6250. https://doi.org/10.3390/en13236250.

Badrinarayanan, V., A. Kendall, and R. Cipolla. 2017. "SegNet: A deep convolutional encoder-decoder architecture for image segmentation." *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (12): 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615.

Bastian, B. T., N. Jaspreeth, S. K. Ranjith, and C. V. Jiji. 2019. "Visual inspection and characterization of external corrosion in pipelines using deep neural network." *NDT & E Int.* 107: 102134. https://doi.org/10.1016/j.ndteint.2019.102134.

Bianchi, E., A. L. Abbott, P. Tokekar, and M. Hebdon. 2021. "COCO-Bridge: Structural detail data set for bridge inspections." *J. Comput. Civil Eng.* 35 (3): 04021003. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000949.

Bianchi, E., and M. Hebdon. 2021a. *Trained model for the semantic segmentation of corrosion condition states*. Blacksburg, VA: Univ. Libraries, Virginia Tech.

Bianchi, E., and M. Hebdon. 2021b. *Corrosion condition state semantic segmentation dataset*. Blacksburg, VA: Univ. Libraries, Virginia Tech.

Bianchi, E., and M. Hebdon. 2022. "Development of extendable open-source structural inspection datasets." *J. Comput. Civil Eng.* 36 (6): 04022039. https://doi.org/10.1061/(ASCE)CP.1943-5487.0001045.

Bondada, V., D. K. Pratihar, and C. S. Kumar. 2018. "Detection and quantitative assessment of corrosion on pipelines through image analysis." *Procedia Comput. Sci.* 133: 804–811. https://doi.org/10.1016/j.procs.2018.07.115.

Bonnin-Pascual, F., and A. Ortiz. 2014. "Corrosion detection for automated visual inspection." In *Developments in corrosion protection*, edited by M. Aliofkhazraei, 619–632. London, UK: InTech.

Casas, E., L. Ramos, C. Romero, and F. Rivas-Echeverría. 2024. "A comparative study of YOLOv5 and YOLOv8 for corrosion segmentation tasks in metal surfaces." *Array* 22: 100351. https://doi.org/10.1016/j.array.2024.100351.

Chen, F.-C., and M. R. Jahanshahi. 2018. "NB-CNN: Deep learning-based crack detection using convolutional neural network and naïve Bayes data fusion." *IEEE Trans. Ind. Electron.* 65 (5): 4392–4400. https://doi.org/10.1109/TIE.2017.2764844.

Chen, L.-C., G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. 2018. "Deeplab: Semantic image segmentation with deep convolutional nets, Atrous convolution, and fully connected CRFs." *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4): 834–848. https://doi.org/10.1109/TPAMI.2017.2699184.

Chen, Q., X. Wen, S. Lu, and D. Sun. 2019a. "Corrosion detection for large steel structure base on UAV integrated with image processing system." *IOP Conf. Ser.: Mater. Sci. Eng.* 608 (1): 012020. https://doi.org/10.1088/1757-899X/608/1/012020.

Chen, S., D. F. Laefer, E. Mangina, S. M. I. Zolanvari, and J. Byrne. 2019b. "UAV bridge inspection through evaluated 3D reconstructions." *J. Bridge Eng.* 24 (4): 05019001. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001343.

Chen, S.-K., I.-F. Huang, and P.-H. Chen. 2021. "Applying fully convolutional neural networks for corrosion semantic segmentation for steel bridges: The use of U-Net." In *Bridge maintenance, safety, management, life-cycle sustainability and innovations*, edited by H. Yokota and D. M. Frangopol, 341–346. Boca Raton, FL: CRC Press.

Choi, K. Y., and S. S. Kim. 2005. "Morphological analysis and classification of types of surface corrosion damage by digital image processing." *Corros. Sci.* 47 (1): 1–15. https://doi.org/10.1016/j.corsci.2004.05.007.

Coelho, L. B., D. Zhang, Y. Van Ingelgem, D. Steckelmacher, A. Nowé, and H. Terryn. 2022. "Reviewing machine learning of corrosion prediction in a data-oriented perspective." *NPJ Mater. Degrad.* 6 (1): 8. https://doi.org/10.1038/s41529-022-00218-4.

Das, A., S. Dorafshan, and N. Kaabouch. 2024. "Autonomous image-based corrosion detection in steel structures using deep learning." *Sensors* 24 (11): 3630. https://doi.org/10.3390/s24113630.

Das, A., E. Ichi, and S. Dorafshan. 2023. "Image-based corrosion detection in ancillary structures." *Infrastructures* 8 (4): 66. https://doi.org/10.3390/infrastructures8040066.

Dorafshan, S., R. J. Thomas, and M. Maguire. 2018. "Fatigue crack detection using unmanned aerial systems in fracture critical inspection of steel bridges." *J. Bridge Eng.* 23 (10): 04018078. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001291.

Doshvarpassand, S., C. Wu, and X. Wang. 2019. "An overview of corrosion defect characterization using active infrared thermography." *Infrared Phys. Technol.* 96: 366–389. https://doi.org/10.1016/j.infrared.2018.12.006.

Duque, L., J. Seo, and J. Wacker. 2018. "Bridge deterioration quantification protocol using UAV." *J. Bridge Eng.* 23 (10): 04018080. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001289.

Dwivedi, S. K., M. Vishwakarma, and A. Soni. 2018. "Advances and researches on non destructive testing: A review." *Mater. Today:. Proc.* 5 (2): 3690–3698. https://doi.org/10.1016/j.matpr.2017.11.620.

Enikeev, M., I. Gubaydullin, and M. Maleeva. 2017. "Analysis of corrosion process development on metals by means of computer vision." *Eng. J.* 21 (4): 183–192. https://doi.org/10.4186/ej.2017.21.4.183.

Fei, Z., E. Yang, B. Yang, and L. Yu. 2021. "Image enhancement and corrosion detection for UAV visual inspection of pressure vessels." In *Intelligent life system modelling, image processing and analysis, communications in computer and information science*, edited by M. Fei, L. Chen, S. Ma, and X. Li, 145–154. Singapore: Springer.

Fondevik, S. K., A. Stahl, A. A. Transeth, and O. O. Knudsen. 2020. "Image segmentation of corrosion damages in industrial inspections." In *Proc., 2020 IEEE 32nd Int. Conf. on Tools with Artificial Intelligence*, 787–792. New York: IEEE.

Forkan, A. R. M., Y.-B. Kang, P. P. Jayaraman, K. Liao, R. Kaul, G. Morgan, R. Ranjan, and S. Sinha. 2022. "CorrDetector: A framework for structural corrosion detection from drone images using ensemble deep learning." *Expert Syst. Appl.* 193: 116461. https://doi.org/10.1016/j.eswa.2021.116461.

Frias, D., and J. Hidalgo. "A high accuracy image hashing and random forest classifier for crack detection in concrete surface images." Preprint, submitted June 10, 2021. https://doi.org/10.48550/arXiv.2106.05755.

Fumagalli, M. 2021. "The corrosion of the Morandi Bridge: The story of a predictable collapse?" In *International Paint and Coating Magazine*, November 30, 2021.

Furuta, H., T. Deguchi, and M. Kushida. 1995. "Neural network analysis of structural damage due to corrosion." In *Proc., 3rd Int. Symp. on Uncertainty Modeling and Analysis and Annual Conf. of the North American Fuzzy Information Processing Society*. New York: IEEE.

Gamarra Acosta, M. R., J. C. Vélez Díaz, and N. Schettini Castro. 2014. "An innovative image-processing model for rust detection using Perlin Noise to simulate oxide textures." *Corros. Sci.* 88: 141–151. https://doi.org/10.1016/j.corsci.2014.07.027.

Ghanta, S., and S. L. Tanja Karp. 2011. *Wavelet domain detection of rust in steel bridge images*. Piscataway, NJ: IEEE.

Gholizadeh, S. 2016. "A review of non-destructive testing methods of composite materials." *Procedia Struct. Integrity* 1: 50–57. https://doi.org/10.1016/j.prostr.2016.02.008.

Gunatilake, P., M. Siegel, A. G. Jordan, and G. W. Podnar. 1997. "Image understanding algorithms for remote visual inspection of aircraft surfaces." In *Proc., SPIE 3029, Machine Vision Applications in Industrial Inspection*, edited by A. R. Rao and N. S. Chang, 2–13. San Jose, CA: SPIE Digital Library.

Han, Q., N. Zhao, and J. Xu. 2021. "Recognition and location of steel structure surface corrosion based on unmanned aerial vehicle images." *J. Civ. Struct. Health Monit.* 11 (5): 1375–1392. https://doi.org/10.1007/s13349-021-00515-7.

Hansson, C. M. 2011. "The impact of corrosion on society." *Metall. Mater. Trans. A* 42 (10): 2952–2962. https://doi.org/10.1007/s11661-011-0703-2.

Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. "Textural features for image classification." *IEEE Trans. Syst. Man Cybern.* SMC-3 (6): 610–621. https://doi.org/10.1109/TSMC.1973.4309314.

Herndon, H., and I. Tien. 2025. "Creating a first-pass algorithm for corrosion assessment in bridge inspections using machine learning and UAV-collected imagery data." *Struct. Infrastruct. Eng.* 1–15. https://doi.org/10.1080/15732479.2025.2470857.

Hoang, N.-D. 2020. "Image processing-based pitting corrosion detection using metaheuristic optimized multilevel image thresholding and machine-learning approaches." *Math. Probl. Eng.* 2020: 1–19. https://doi.org/10.1155/2020/6765274.

Hoang, N.-D., and V.-D. Tran. 2019. "Image processing-based detection of pipe corrosion using texture analysis and metaheuristic-optimized machine learning approach." *Comput. Intell. Neurosci.* 2019: 1–13. https://doi.org/10.1155/2019/8097213.

Holm, E., A. A. Transeth, O. O. Knudsen, and A. Stahl. 2020. "Classification of corrosion and coating damages on bridge constructions from images using convolutional neural networks." In *Proc., 12th Int. Conf. on Machine Vision*, edited by W. Osten and D. P. Nikolaev. Amsterdam, The Netherlands: SPIE Digital Library.

Huang, G., Z. Liu, L. van der Maaten, and K. Q. Weinberger. 2017. "Densely connected convolutional networks." In *Proc., IEEE Conf. on Computer Vision and Pattern Recognition*. New York: IEEE.

Huang, J., Q. Liu, L. Xiang, G. Li, Y. Zhang, and W. Chen. 2022. "A lightweight residual model for corrosion segmentation with local contextual information." *Appl. Sci.* 12 (18): 9095. https://doi.org/10.3390/app12189095.

IBM. 2023a. "What is computer vision?" *IBM Topics*. Accessed September 21, 2023. https://www.ibm.com/topics/computer-vision.

IBM. 2023b. "What is deep learning?" *IBM Topics*. Accessed September 21, 2023. https://www.ibm.com/topics/deep-learning.

Itzhak, D., I. Dinstein, and T. Zilberberg. 1981. "Pitting corrosion evaluation by computer image processing." *Corros. Sci.* 21: 17–22. https://doi.org/0010-938X/81/0101-0017.

Jahanshahi, M. R., J. S. Kelly, S. F. Masri, and G. S. Sukhatme. 2009. "A survey and evaluation of promising approaches for automatic image-based defect detection of bridge structures." *Struct. Infrastruct. Eng.* 5 (6): 455–486. https://doi.org/10.1080/15732470801945930.

Jahanshahi, M. R., and S. F. Masri. 2013a. "Effect of color space, color channels, and sub-image block size on the performance of wavelet-based texture analysis algorithms: An application to corrosion detection on steel structures." *Comput. Civ. Eng.* 685–692. https://doi.org/10.1061/9780784413029.086.

Jahanshahi, M. R., and S. F. Masri. 2013b. "Parametric performance evaluation of wavelet-based corrosion detection algorithms for condition assessment of civil infrastructure systems." *J. Comput. Civil Eng.* 27 (4): 345–357. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000225.

Javaherdashti, R. 2000. "How corrosion affects industry and life." *Anti-Corros. Methods Mater.* 47 (1): 30–34. https://doi.org/10.1108/00035590010310003.

Jeong, E., J. Seo, and J. Wacker. 2020. "Literature review and technical survey on bridge inspection using unmanned aerial vehicles." *J. Perform. Constr. Facil.* 34 (6): 04020113. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001519.

Jia, Z., M. Fu, X. Zhao, and Z. Cui. 2023. "Intelligent identification of metal corrosion based on corrosion-YOLOv5s." *Displays* 76: 102367. https://doi.org/10.1016/j.displa.2022.102367.

Jiang, F., Y. Ding, Y. Song, F. Geng, and Z. Wang. 2023. "Automatic pixel-level detection and measurement of corrosion-related damages in dim steel box girders using Fusion-Attention-U-net." *J. Civ. Struct. Health Monit.* 13 (1): 199–217. https://doi.org/10.1007/s13349-022-00631-y.

Jin Lim, H., S. Hwang, H. Kim, and H. Sohn. 2021. "Steel bridge corrosion inspection with combined vision and thermographic images." *Struct. Health Monit.* 20 (6): 3424–3435. https://doi.org/10.1177/1475921721989407.

Katsamenis, I., E. Protopapadakis, A. Doulamis, N. Doulamis, and A. Voulodimos. 2020. "Pixel-level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation." In *Advances in visual computing*, edited by G. Bebis, Z. Yin, E. Kim, J.

Bender, K. Subr, B. C. Kwon, J. Zhao, D. Kalkofen, and G. Baciu, 160–169. Cham, Switzerland: Springer. Lecture Notes in Computer Science.
Khayatazad, M., L. De Pue, and W. De Waele. 2020. "Detection of corrosion on steel structures using automated image processing." *Dev. Built Environ.* 3: 100022. https://doi.org/10.1016/j.dibe.2020.100022.
Khayatazad, M., M. Honhon, and W. De Waele. 2023. "Detection of corrosion on steel structures using an artificial neural network." *Struct. Infrastruct. Eng.* 19: 1860–1871. https://doi.org/10.1080/15732479.2022.2069272.
Kim, H., J. Lee, E. Ahn, S. Cho, M. Shin, and S.-H. Sim. 2017. "Concrete crack identification using a UAV incorporating hybrid image processing." *Sensors* 17 (9): 2052. https://doi.org/10.3390/s17092052.
Laws, K. I. 1980. "Rapid texture identification." *Image Processing for Missile Guidance* 238: 376–381. https://doi.org/10.1117/12.959169.
Lee, S., L.-M. Chang, and M. Skibniewski. 2006. "Automated recognition of surface defects using digital color image processing." *Autom. Constr.* 15 (4): 540–549. https://doi.org/10.1016/j.autcon.2005.08.001.
Lemos, R., R. Cabral, D. Ribeiro, R. Santos, V. Alves, and A. Dias. 2023. "Automatic detection of corrosion in large-scale industrial buildings based on artificial intelligence and unmanned aerial vehicles." *Appl. Sci.* 13 (3): 1386. https://doi.org/10.3390/app13031386.
Li, G., X. Li, J. Zhou, D. Liu, and W. Ren. 2021. "Pixel-level bridge crack detection using a deep fusion about recurrent residual convolution and context encoder network." *Measurement* 176: 109171. https://doi.org/10.1016/j.measurement.2021.109171.
Lin, J. J., A. Ibrahim, S. Sarwade, and M. Golparvar-Fard. 2021. "Bridge inspection with aerial robots: Automating the entire pipeline of visual data capture, 3D mapping, defect detection, analysis, and reporting." *J. Comput. Civil Eng.* 35 (2): 04020064. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000954.
Lin, T.-Y., M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. 2014. "Microsoft COCO: Common objects in context." In *Computer vision—ECCV 2014*, edited by D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, 740–755. Cham, Switzerland: Springer. Lecture Notes in Computer Science.
Lin, W., and T. Yoda. 2017. *Bridge engineering*. Oxford, UK: Elsevier.
Liu, L., E. Tan, X. J. Yin, Y. Zhen, and Z. Q. Cai. 2019. "Deep learning for coating condition assessment with active perception." In *Proc., 3rd High Performance Computing and Cluster Technologies Conf.*, 75–80. New York: Association for Computing Machinery (ACM).
Liu, W., A. Rabinovich, and A. C. Berg. "ParseNet: Looking wider to see better." Preprint, submitted June 15, 2015. https://doi.org/10.48550/arXiv.1506.04579.
Liu, Y., and Y. Bao. 2022. "Review on automated condition assessment of pipelines with machine learning." *Adv. Eng. Inf.* 53: 101687. https://doi.org/10.1016/j.aei.2022.101687.
Liu, Z., M. Genest, and D. Krys. 2012. "Processing thermography images for pitting corrosion quantification on small diameter ductile iron pipe." *NDT & E Int.* 47: 105–115. https://doi.org/10.1016/j.ndteint.2012.01.003.
Long, J., E. Shelhamer, and T. Darrell. 2015. "Fully convolutional networks for semantic segmentation." In *Proc., IEEE Conf. on Computer Vision and Pattern Recognition*. New York: IEEE.
Malashin, I., et al. 2024. "Deep learning approach for pitting corrosion detection in gas pipelines." *Sensors* 24 (11): 3563. https://doi.org/10.3390/s24113563.
Marchewka, A., P. Ziółkowski, and V. Aguilar-Vidal. 2020. "Framework for structural health monitoring of steel bridges by computer vision." *Sensors* 20 (3): 700. https://doi.org/10.3390/s20030700.
Medeiros, F. N. S., G. L. B. Ramalho, M. P. Bento, and L. C. L. Medeiros. 2010. "On the evaluation of texture and color features for nondestructive corrosion detection." *EURASIP J. Adv. Signal Process.* 2010 (1): 817473. https://doi.org/10.1155/2010/817473.
Miao, P., G. Xing, S. Ma, and T. Srimahachota. 2023. "Deep learning–based inspection data mining and derived information fusion for enhanced bridge deterioration assessment." *J. Bridge Eng.* 28 (8): 04023048. https://doi.org/10.1061/JBENF2.BEENG-6053.
Moranduzzo, T., and F. Melgani. 2014. "Monitoring structural damages in big industrial plants with UAV images." In *Proc., 2014 IEEE Int. Geoscience and Remote Sensing Symp*. New York: IEEE.
Munawar, H., F. Ullah, D. Shahzad, A. Heravi, S. Qayyum, and J. Akram. 2022. "Civil infrastructure damage and corrosion detection: An application of machine learning." *Buildings* 12 (2): 156. https://doi.org/10.3390/buildings12020156.
Nabizadeh, E., and A. Parghi. 2023. "Automated corrosion detection using deep learning and computer vision." *Asian J. Civ. Eng.* 24 (8): 2911–2923. https://doi.org/10.1007/s42107-023-00684-4.
Naik, D. L., H. U. Sajid, R. Kiran, and G. Chen. 2020. "Detection of corrosion-indicating oxidation product colors in steel bridges under varying illuminations, shadows, and wetting conditions." *Metals* 10 (11): 1439. https://doi.org/10.3390/met10111439.
Nash, W. 2022. "SpotRust model." *Github*. Accessed December 12, 2022. https://github.com/StuvX/SpotRust.
Nash, W., L. Zheng, and N. Birbilis. 2022. "Deep learning corrosion detection with confidence." *NPJ Mater. Degrad.* 6 (1): 26. https://doi.org/10.1038/s41529-022-00232-6.
Nash, W. T., C. J. Powell, T. Drummond, and N. Birbilis. 2020. "Automated corrosion detection using crowdsourced training for deep learning." *Corrosion* 76 (2): 135–141. https://doi.org/10.5006/3397.
Ortiz, A., F. Bonnin-Pascual, E. Garcia-Fidalgo, and J. P. Company. 2016. "Visual inspection of vessels by means of a micro-aerial vehicle: An artificial neural network approach for corrosion detection." In *Proc., Robot 2015: Second Iberian Robotics Conf., Advances in Intelligent Systems and Computing*, edited by L. P. Reis, A. P. Moreira, P. U. Lima, L. Montano, and V. Muñoz-Martinez, 223–234. Cham, Switzerland: Springer.
Pakrashi, V., F. Schoefs, J. B. Memet, and A. O'Connor. 2010. "ROC dependent event isolation method for image processing based assessment of corroded harbour structures." *Struct. Infrastruct. Eng.* 6 (3): 365–378. https://doi.org/10.1080/15732470701718072.
Petricca, L., T. Moss, G. Figueroa, and S. Broen. 2016. "Corrosion detection using A.I : A comparison of standard computer vision techniques and deep learning model." In *Computer science & information technology*, edited by Z. Jan and N. Dhinaharan, 91–99. Vienna, Austria: Academy & Industry Research Collaboration Center.
Qian, G., Z. Wu, Z. Hu, and M. D. Todd. 2025. "Pitting corrosion diagnostics and prognostics for miter gates using multiscale simulation and image inspection data." *Struct. Health Monit.* 24: 2008–2030. https://doi.org/10.1177/14759217241264291.
Rahman, A., Z. Y. Wu, and R. Kalfarisi. 2021. "Semantic deep learning integrated with RGB feature-based rule optimization for facility surface corrosion detection and evaluation." *J. Comput. Civil Eng.* 35 (6): 04021018. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000982.
Ryan, T., C. Lloyd, M. Pichura, D. Tarasovich, and S. Fitzgerald. 2023. *FHWA bridge inspectors reference manual*. Washington, DC: Federal Highway Administration (FHWA).
Safa, A., A. Mohamed, B. Issam, and H. Mohamed-Yassine. 2023. "SegFormer: Semantic segmentation based transformers for corrosion detection." In *Proc., 2023 Int. Conf. on Networking and Advanced Systems*. New York: IEEE.
Sheng, P., L. Chen, and J. Tian. 2018. "Learning-based road crack detection using gradient boost decision tree." In *Proc., 2018 13th IEEE Conf. on Industrial Electronics and Applications*, 1228–1232. New York: IEEE.
Siegel, M., and P. Gunatilake. 1998. "Remote enhanced visual inspection of aircraft by a Mobiloe robot." In *Proc., IEEE Workshop on Emerging Technologies, Intelligent Measurement, and Virtual Systems for Instrumentation and Measurement*. New York: IEEE.
Son, E.-Y., D. Jeong, and M.-J. Oh. 2024. "Corrosion area detection and depth prediction using machine learning." *Int. J. Nav. Archit. Ocean Eng.* 16: 100617. https://doi.org/10.1016/j.ijnaoe.2024.100617.
Son, H., N. Hwang, C. Kim, and C. Kim. 2014. "Rapid and automated determination of rusted surface areas of a steel bridge for robotic maintenance systems." *Autom. Constr.* 42: 13–24. https://doi.org/10.1016/j.autcon.2014.02.016.
Sookpong, S., S. Phimsiri, T. Tosawadi, P. Choppradit, V. Suttichaya, C. Utintu, and E. Thamwiwatthana. 2023. "Comparison of corrosion segmentation techniques on oil and gas offshore critical assets." In *Proc., 20th Int. Conf. on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. New York: IEEE.

Soomro, A. A., A. A. Mokhtar, J. C. Kurnia, N. Lashari, H. Lu, and C. Sambo. 2022. “Integrity assessment of corroded oil and gas pipelines using machine learning: A systematic review.” *Eng. Fail. Anal.* 131: 105810. https://doi.org/10.1016/j.engfailanal.2021.105810.

Ta, Q.-B., T.-C. Huynh, Q.-Q. Pham, and J.-T. Kim. 2022. “Corroded bolt identification using mask region-based deep learning trained on synthesized data.” *Sensors* 22 (9): 3340. https://doi.org/10.3390/s22093340.

Ta, Q.-B., and J.-T. Kim. 2020. “Monitoring of corroded and loosened bolts in steel structures via deep learning and hough transforms.” *Sensors* 20 (23): 6888. https://doi.org/10.3390/s20236888.

Tan, L., X. Chen, D. Yuan, and T. Tang. 2024. “DSNet: A computer vision-based detection and corrosion segmentation network for corroded bolt detection in tunnel.” *Struct. Control Health Monit.* 2024: 1898088. https://doi.org/10.1155/2024/1898088.

Tomiczek, A. P., T. J. Whitley, J. A. Bridge, and P. G. Ifju. 2019. “Bridge inspections with small unmanned aircraft systems: Case studies.” *J. Bridge Eng.* 24 (4): 05019003. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001376.

Van Beers, F., A. Lindström, E. Okafor, and M. Wiering. 2019. “Deep neural networks with intersection over union loss for binary image segmentation.” In *Proc., 8th Int. Conf. on Pattern Recognition Applications and Methods*, 438–445. Prague, Czech Republic: Science and Technology Publications (SCITEPRESS).

Wang, Y., and G. Cheng. 2016. “Application of gradient-based Hough transform to the detection of corrosion pits in optical images.” *Appl. Surf. Sci.* 366: 9–18. https://doi.org/10.1016/j.apsusc.2015.12.207.

Xu, X., Y. Wang, and J. Wu. 2020. *Intelligent corrosion detection and rating based on faster region-based convolutional neural network*. New York: IEEE.

Yan, B., C. Wang, and X. Hao. 2025. “Research on the performance of the SegFormer model with fusion of edge feature extraction for metal corrosion detection.” *Sci. Rep.* 15 (1): 8134. https://doi.org/10.1038/s41598-025-92531-6.

Yao, Y., Y. Yang, Y. Wang, and X. Zhao. 2019. “Artificial intelligence-based hull structural plate corrosion damage detection and recognition using convolutional neural network.” *Appl. Ocean Res.* 90: 101823. https://doi.org/10.1016/j.apor.2019.05.008.

Yuan, J., D. Wang, and A. M. Cheriyadat. 2015. “Factorization-based texture segmentation.” *IEEE Trans. Image Process.* 24 (11): 3488–3497. https://doi.org/10.1109/TIP.2015.2446948.

Yuvaneswaren, R. S., S. T. Prakash, and K. Sudha. 2024. “A YOLOv8-based model for precise corrosion segmentation in industrial imagery.” In *Proc., 3rd Int. Conf. on Artificial Intelligence for Internet of Things*. New York: IEEE.

Zhang, C., M. M. Karim, and R. Qin. 2023. “A multitask deep learning model for parsing bridge elements and segmenting defect in bridge inspection images.” *Transp. Res. Rec.* 2677 (7): 693–704. https://doi.org/10.1177/03611981231155418.

Zhang, C., Y. Zou, F. Wang, E. del Rey Castillo, J. Dimyadi, and L. Chen. 2022. “Towards fully automated unmanned aerial vehicle-enabled bridge inspection: Where are we at?” *Constr. Build. Mater.* 347: 128543. https://doi.org/10.1016/j.conbuildmat.2022.128543.

Zhang, S., X. Deng, Y. Lu, S. Hong, Z. Kong, Y. Peng, and Y. Luo. 2021. “A channel attention based deep neural network for automatic metallic corrosion detection.” *J. Build. Eng.* 42: 103046. https://doi.org/10.1016/j.jobe.2021.103046.

Zhang, Y., R. DesRoches, and I. Tien. 2019. “Impact of corrosion on risk assessment of shear-critical and short lap-spliced bridges.” *Eng. Struct.* 189: 260–271. https://doi.org/10.1016/j.engstruct.2019.03.050.

Zhou, Q., S. Ding, Y. Feng, G. Qing, and J. Hu. 2022. “Corrosion inspection and evaluation of crane metal structure based on UAV vision.” *Signal, Image Video Process.* 16 (6): 1701–1709. https://doi.org/10.1007/s11760-021-02126-7.