

Article

Deep Learning-Based Superpixel Texture Analysis for Crack Detection in Multi-Modal Infrastructure Images

Sara Shahsavarani ^{1,*} , Clemente Ibarra-Castanedo ¹ , Fernando Lopez ²  and Xavier P. V. Maldague ¹ 

¹ Computer Vision and Systems Laboratory (CVSL), Department of Electrical and Computer Engineering, Faculty of Science and Engineering, Laval University, Quebec City, QC G1V 0A6, Canada; clemente.ibarra-castanedo@gel.ulaval.ca (C.I.-C.); xavier.maldague@gel.ulaval.ca (X.P.V.M.)

² Tornegats Services Techniques, 200 Boul. du Parc-Technologique, Quebec City, QC G1P 4S3, Canada

* Correspondence: sara.shahsavarani.1@ulaval.ca

Abstract: Infrared and visible imaging play crucial roles in non-destructive testing, where accurate defect segmentation and detection are paramount. However, the scarcity of annotated training data or the limited number of data availability often poses a challenge. To address this, we propose an innovative framework tailored to the domain of infrared and visible imaging, integrating segmentation and detection tasks. The proposed approach eliminates the dependency on annotated defect data during training, enabling models to adapt to real-world scenarios with limited annotations. By utilizing super-pixel segmentation and texture analysis, the proposed method enhances the accuracy of defect detection. Concrete structures, globally subjected to aging and degradation, demand constant monitoring for structural health. Traditional manual crack detection methods are labor-intensive, necessitating automated systems. The proposed approach combines deep learning-based super-pixel segmentation with texture analysis, offering a solution for limited-defect-data situations. Utilizing convolutional neural networks (CNNs) for super-pixel segmentation and texture features for defect analysis, the proposed methodology improves the efficiency and accuracy of crack detection, especially in scenarios with limited labeled data or a limited number of data available. Evaluation on public benchmark datasets have validated the effectiveness of the proposed approach in detecting cracks in concrete structures.

Keywords: non-destructive testing; defect detection; convolutional neural networks; super-pixel segmentation; image segmentation; texture analysis; limited defect data



Citation: Shahsavarani, S.; Ibarra-Castanedo, C.; Lopez, F.; Maldague, X.P.V. Deep Learning-Based Superpixel Texture Analysis for Crack Detection in Multi-Modal Infrastructure Images. *NDT* **2024**, *2*, 128–142. <https://doi.org/10.3390/ndt2020008>

Academic Editor: Fabio Tosti

Received: 1 May 2024

Revised: 31 May 2024

Accepted: 3 June 2024

Published: 14 June 2024



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

1. Introduction

Numerous critical concrete structures worldwide, including bridges, roads, and other infrastructure components, have endured decades of service, and are confronting the inevitable challenges of aging and degradation [1]. Over time, these structures have become increasingly vulnerable to failure, posing significant risks to public safety and property. Thus, routine maintenance and structural health monitoring (SHM) are crucial to ensure their ongoing functionality and safety. SHM plays a pivotal role not only during the operational lifespan of these structures but also during their construction phase, where early detection of potential issues can prevent future failures. The need for automated crack detection has grown apparent over the years, given the reliance on effective management practices and the specialized expertise required for structural maintenance, particularly in remote or inaccessible areas where manual inspection is impractical.

Cracks are a significant concern for the safety, durability, and serviceability of concrete structures, necessitating their detection and monitoring. In addition to structural considerations, cracks raise aesthetic and financial concerns for building owners and structural engineers. The presence of cracks can allow harmful chemicals to penetrate buildings, compromising their integrity and visual appeal. Surface cracks serve as critical indicators

of structural damage and durability, requiring thorough inspection of building elements. However, manual crack detection remains prevalent in many developing countries, resulting in time-consuming and subjective assessments lacking cost-effectiveness and accuracy. Consequently, there is a critical need for automated crack detection systems to enhance efficiency and reliability in structural monitoring and maintenance practices.

In addressing these challenges, researchers have highlighted the significance of employing advanced crack detection techniques, particularly in the context of structural health assessment in concrete structures. Despite such recommendations, manual visual inspection remains prevalent in many developing regions, despite its inherent limitations, in terms of time, effort, and accuracy. Automated crack detection systems have emerged as a promising remedy to these issues, offering a more efficient and objective means of structural monitoring. Leveraging automation technologies, such as computer vision and machine learning, enables these systems to enhance the precision and dependability of crack detection. Moreover, with the integration of infrared and visible images (fusion images) [2,3], these automated systems demonstrate even greater potential in detecting and characterizing cracks within concrete structures, thereby advancing the field of structural health monitoring.

This paper proposes an effective and robust crack detection methodology. We concentrate on both deep learning-based super-pixel segmentation and texture analysis for accurate crack detection. The main contributions in this work are as follows:

1. We aim to assess the performance of the proposed method in accurately detecting cracks under conditions of limited data availability. Through this investigation, we endeavor to contribute to the advancement of non-destructive testing methodologies for structural integrity assessment and defect identification.
2. The proposed approach involves a multi-step process, beginning with the segmentation of images using a deep learning-based super-pixel method. Subsequently, we apply texture analysis techniques using the Mahotas Python library to identify cracks present in the images.
3. Additionally, we aim to investigate the effectiveness of accurate segmentation on crack detection performance. By evaluating the influence of precise segmentation on the effectiveness of our detection method, we seek to understand the importance of segmentation quality in defect identification and localization.
4. Furthermore, we explore the feasibility of utilizing thermal and visible image fusion as part of our detection strategy. This investigation aims to determine whether fusion images offer advantages over individual modalities in terms of crack-detection accuracy and reliability. By integrating thermal and visible images, we seek to enhance the robustness and versatility of our detection method, particularly in scenarios characterized by limited training data.

This paper is organized as follows. Section 2 presents the related works. Section 3 introduces the proposed defect methodology. Then, Section 4 outlines the conducted comparative experiments on the algorithms of each part. Finally, we summarize the main outcomes in Section 5.

2. Literature Review

Detection of surface cracks in concrete holds paramount importance for preserving concrete structures. Traditional visual inspection techniques are inherently intricate, cumbersome, and labor-intensive [4]. The manual inspection of cracks poses limitations, in terms of accuracy and efficiency, necessitating the adoption of automated approaches through digital image processing [4]. Detecting and monitoring cracks is crucial for ensuring the structural integrity and longevity of concrete structures [4]. The presence of cracks not only affects the aesthetics of the structure but also accelerates corrosion of reinforcement, leading to premature aging and a reduced lifespan [5]. Moreover, the carrying capacity of the structure is compromised by the dimensions of the cracks [5].

Various techniques have been employed for crack detection on concrete surfaces, including manual inspection, photogrammetry, fiber optics, and ultrasonic methods. However, manual inspection is laborious and time-consuming, prompting the development of alternative approaches, such as the stereovision-based crack width detection method [6]. Pavements are integral components of transportation infrastructure, and the detection of pavement distress, including cracks, is essential for maintenance and rehabilitation purposes [7]. Cracks in welds are considered severe defects as they can lead to weld failure under stress. Despite being typically surface-initiated, small cracks can be challenging to visually identify. Traditional non-destructive testing (NDT) techniques, such as radiography and ultrasound, may encounter difficulties in detecting surface cracks. Thermography, a novel NDT technique, offers advantages, such as speed, non-contact operation, and full-field data provision [8].

Various techniques utilizing deep learning have been developed for defect segmentation. The authors in [9] proposed an auto-encoder with conditional random fields and guided filtering methods. Escalona et al. [10] investigated the effect of the depth of the U-Net architecture on crack segmentation performance. Additionally, the authors in [11] proposed an auto-encoder with conditional random fields and guided filtering methods. Yang et al. [12] were the first to apply the U-Net architecture in the field of crack detection to address several limitations associated with using CNNs for this task. Fan et al. [13] proposed a modified encoder-decoder architecture based on U-Net, incorporating multi-dilation and hierarchical feature learning to enhance crack-segmentation performance. König et al. [14] introduced an encoder-decoder structure based on the U-Net architecture, combined with attention gating and residual connections to improve performance. Zhang et al. [15] applied deep learning to the task of crack segmentation, using ConvNet for feature extraction on raw images.

The challenge arises when there is insufficient data, as deep learning typically requires a substantial amount of data to be effective. In the following, a method is proposed to address this problem.

3. Materials and Methods

The proposed methodology introduces an innovative approach to address the challenge of defect detection for limited training data. By combining deep learning-based super-pixel segmentation [16,17] followed by texture analysis [18,19] automated defect detection is empowered in scenarios where annotated training data are limited. This novel method not only overcomes the limitations of traditional approaches but also opens up new possibilities for advancements in various industrial and scientific applications.

Super-pixel segmentation, a critical component within image processing and computer vision, is essential for grouping pixels with similar attributes into coherent and perceptually meaningful regions [20]. In this paper, we confront the inherent challenge of limited training data for defect detection by employing a sophisticated deep learning-based super-pixel segmentation method [21]. This innovative approach, which employs convolutional neural networks (CNNs) to effectively partition images into cohesive regions, lays the foundation for subsequent texture analysis. By employing the capabilities of CNNs, our objective is to accurately delineate image components, thereby facilitating defect detection through texture analysis. This practical application of the proposed methodology can significantly enhance the efficiency and effectiveness of defect detection in various industries.

Texture analysis, on the other hand, is a powerful technique for characterizing the spatial arrangement of pixel intensities within an image [18,19]. By extracting statistical features from the local neighborhood, texture analysis reveals crucial insights into the underlying structures and patterns present in the image [18].

Figure 1 indicates that the proposed methodology uses deep learning super-pixel segmentation followed by texture analysis to tackle the challenge of defect detection with limited training data. Specifically, texture features are renowned for their ability to capture essential textural properties, such as contrast, entropy, and homogeneity. In the proposed

method, these features are the foundation for detecting anomalies, such as cracks, within multi-modal images. By employing deep learning super-pixel segmentation, we aim to create a structured representation of the image that preserves spatial context and crack localization while reducing the complexity of subsequent analysis [22]. This strategy enables us to partition the image into coherent regions [23]. Using the inherent spatial coherence provided by super-pixel segmentation, the efficiency and effectiveness of texture analysis are enhanced for defect detection. Furthermore, by separating the segmentation task from the requirement of annotated data, it is not necessary to spend as much time manually labeling defect images. This makes it easier for the proposed method to be used for detecting defects in various situations where labeled data might not be available as well as for scenarios with limited numbers of data.

This section first presents the deep learning-based super-pixel segmentation approach. Subsequently, it presents the texture analysis method. Moreover, the related mathematical formulations and algorithms are explained.

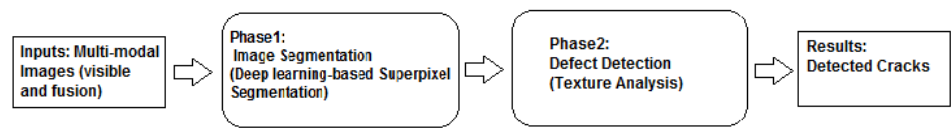


Figure 1. The proposed crack detection methodology pipeline.

3.1. Deep Learning-Based Super-Pixel Segmentation Phase

The hypothesis behind employing a CNN-based super-pixel segmentation method for defect detection, particularly crack detection, is that it can effectively delineate regions of interest within images, facilitating accurate identification and localization of defects [21]. By employing convolutional neural networks, which are adept at capturing intricate patterns and features in images, the method aims to enhance the precision and efficiency of defect detection compared to traditional segmentation approaches. The primary objective in employing deep learning-based super-pixel segmentation is to enhance the accuracy of crack defect segmentation and optimizing the over-segmentation, particularly in distinguishing between very narrow cracks and thicker cracks.

This section introduces the CNN-based super-pixel segmentation method, tailored specifically for detecting cracks in concrete structures. We start by presenting the approach to directly predicting pixel–super-pixel associations on a regular grid, followed by detailing the architecture of the super-pixel network.

3.1.1. Learning Super-Pixels on a Regular Grid

Super-pixel segmentation is traditionally performed by dividing images into a grid of cells, where each cell is treated as an initial super-pixel “seed”, and then assigning each pixel to one of these seeds. However, this assignment for all the pixels of a super-pixel pair is computationally intensive. To mitigate this, the search is restricted to neighboring grid cells for each pixel, improving computational efficiency. This is achieved by training a deep neural network to learn the mapping, replacing hard assignment with a soft-association map. This map represents the probability of each pixel being assigned to a specific grid cell, allowing for the determination of super-pixels based on the highest probability assignment.

Mathematical Formulation: The hard assignment tensor, denoted as G , represents the pixel-to-superpixel mapping. For a given pixel p , the search is confined to the set of neighboring grid cells N_p . The soft-association map Q is computed, where $q_s(p)$ signifies the probability of pixel p being assigned to each superpixel s in N_p . Finally, super-pixels are derived by assigning each pixel to the grid cell with the highest probability: $s^* = \arg \max_s q_s(p)$.

3.1.2. Deep Learning-Based Super-Pixel Architecture

Figure 2 indicates encoder–decoder architecture with skip connections employed to predict the super-pixel association map. The encoder processes input color images

and generates high-level feature maps through convolutional layers, while the decoder progressively upsamples these features to make the final prediction, incorporating features from corresponding encoder layers. Leaky rectified linear units (ReLU) are utilized for all layers except the prediction layer, where softmax is applied. The end-to-end trainable super-pixel network offers flexibility in defining loss functions, enabling customization based on specific application requirements.

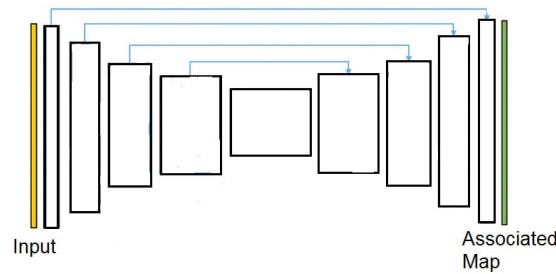


Figure 2. Deep learning-based super-pixel architecture.

3.2. Deep Learning-Based Super-Pixel Texture Analysis Phase

The core of our proposed methodology lies in deep learning-based super-pixel texture-based analysis for crack detection. Algorithm 1 outlines the procedure for conducting texture analysis based on deep learning-based super-pixels. In order to enhance the visibility of texture information within intensity variations, pre-processing steps are applied to the original image. Subsequently, individual super-pixel regions are isolated, and texture features are extracted from the respective grayscale regions masked by each super-pixel.

Algorithm 1 Deep learning-based Super-pixel Texture Analysis for Crack Detection

Require: Deep learning-based Super-pixel segmentation result, Original Defect Image

Ensure: An Image with the highlighted crack area

```

1: function CREATEMASK(superpixel_label)
2:   mask ← superpixel_label
3:   return mask
4: end function
5: function APPLYMASK(mask, grayscale image)
6:   masked_image ← Apply mask to grayscale image
7:   return masked_image
8: end function
9: function EXTRACTFEATURES(masked_image)
10:  features ← Extract features from masked_image
11:  return features
12: end function
13: function HIGHLIGHTCRACK(superpixel_label)
14:  Highlight crack area corresponding to superpixel_label
15: end function
16: function RESULTIMAGES(superpixel boundaries, original image)
17:  Result_image ← Combine superpixel boundaries with original image
18:  return Result_image
19: end function
20: function SAVEIMAGE(Result_image)
21:  Save Result_image
22: end function
23: Initialize Threshold1, Threshold2
24: for each superpixel label superpixel_label in unique(superpixel result) do
25:   mask ← CREATEMASK(superpixel_label)
26:   masked_image ← APPLYMASK(mask, grayscale image)
27:   features ← EXTRACTFEATURES(masked_image)
28:   if features.mean() > Threshold1 and features.var() > Threshold2 then
29:     HIGHLIGHTCRACK(superpixel_label)
30:   end if
31: end for
32: Result_image ← RESULTIMAGES(superpixel boundaries, original image)
33: SAVEIMAGE(Result_image)

```

Mathematical Analysis

The thresholds for `features.mean()` and `features.var()` can be determined in a professional and mathematically sound manner by employing statistical methods to analyze the distribution of texture features extracted from the dataset. The process can be formulated as follows.

3.3. Mathematical Analysis for Mean Threshold

Let μ_{crack} denote the population mean of the texture feature values extracted from super-pixels corresponding to crack areas in an image. Similarly, let $\mu_{\text{non-crack}}$ represent the population mean of the texture feature values extracted from super-pixels corresponding to non-crack areas.

The threshold, `Threshold1`, can be determined by computing a confidence interval for the population mean of crack texture features. A common approach is to construct a confidence interval using the sample mean $\bar{x}$ and sample standard deviation s of the crack texture feature values, typically at a confidence level of 95% or higher.

Mathematically, `Threshold1` can be formulated as

$$\text{Threshold1} = \mu_{\text{crack}} + z \times \frac{s_{\text{crack}}}{\sqrt{n}}$$

where z is the z-score corresponding to the desired confidence level, s_{crack} is the sample standard deviation of crack texture feature values, and n is the number of samples.

3.4. Mathematical Analysis for Variance Threshold

Similarly, the threshold, `Threshold2` for `features.var()` can be determined by analyzing the distribution of texture feature variances within the dataset.

Let σ_{crack}^2 denote the population variance of texture feature values extracted from super-pixels corresponding to crack areas, and $\sigma_{\text{non-crack}}^2$ denote the population variance of texture feature values extracted from non-crack areas.

Like `Threshold1`, `Threshold2` can be determined using a confidence interval approach, ensuring that it captures the variability in texture features observed within crack areas with a certain level of confidence [24,25].

Mathematically, the threshold `Threshold2` can be formulated as

$$\text{Threshold2} = \sigma_{\text{crack}}^2 + z \times \frac{s_{\text{crack}}^2}{\sqrt{2(n-1)}}$$

where z is the z-score corresponding to the desired confidence level, s_{crack}^2 is the sample variance of crack texture feature values, and n is the number of samples.

By formulating the thresholds for `features.mean()` and `features.var()` using statistical methods such as confidence intervals, robust criteria for classifying super-pixels as crack areas can be established, ensuring that the algorithm's decisions are based on sound statistical principles and providing confidence in its performance across different datasets.

4. Results and Discussions

This section first presents the results of the proposed crack detection approach, using visible image data within the constraints of limited available data. Subsequently, it assesses the efficacy of infrared and visible fusion datasets while examining their advantages and disadvantages.

We begin by presenting the segmentation process using deep learning-based super-pixel segmentation and comparing its accuracy with the SLIC method [26]. Next, we investigate the detection phase, providing qualitative and quantitative crack detection assessments using visible images. Finally, we explore the potential benefits of fusion images for enhancing crack detection accuracy.

The datasets utilized in this paper are introduced in the following subsection. Subsequently, the details of the segmentation phase and its implementation are outlined in this

section. The crack detection results are presented in the following subsection. Furthermore, the results are analyzed, compared, and discussed.

4.1. Dataset

This paper utilizes two public datasets. The first dataset comprises visible images including cracked and non-cracked samples. The second dataset includes visible, infrared, and fusion images. Detailed information about both datasets is provided below.

The first dataset [27] serves the dual purpose of crack detection and segmentation, making it a valuable resource for further research. It encompasses many infrastructural crack types, such as pavements, bridges, and buildings. Figure 3 indicates some samples of this dataset.

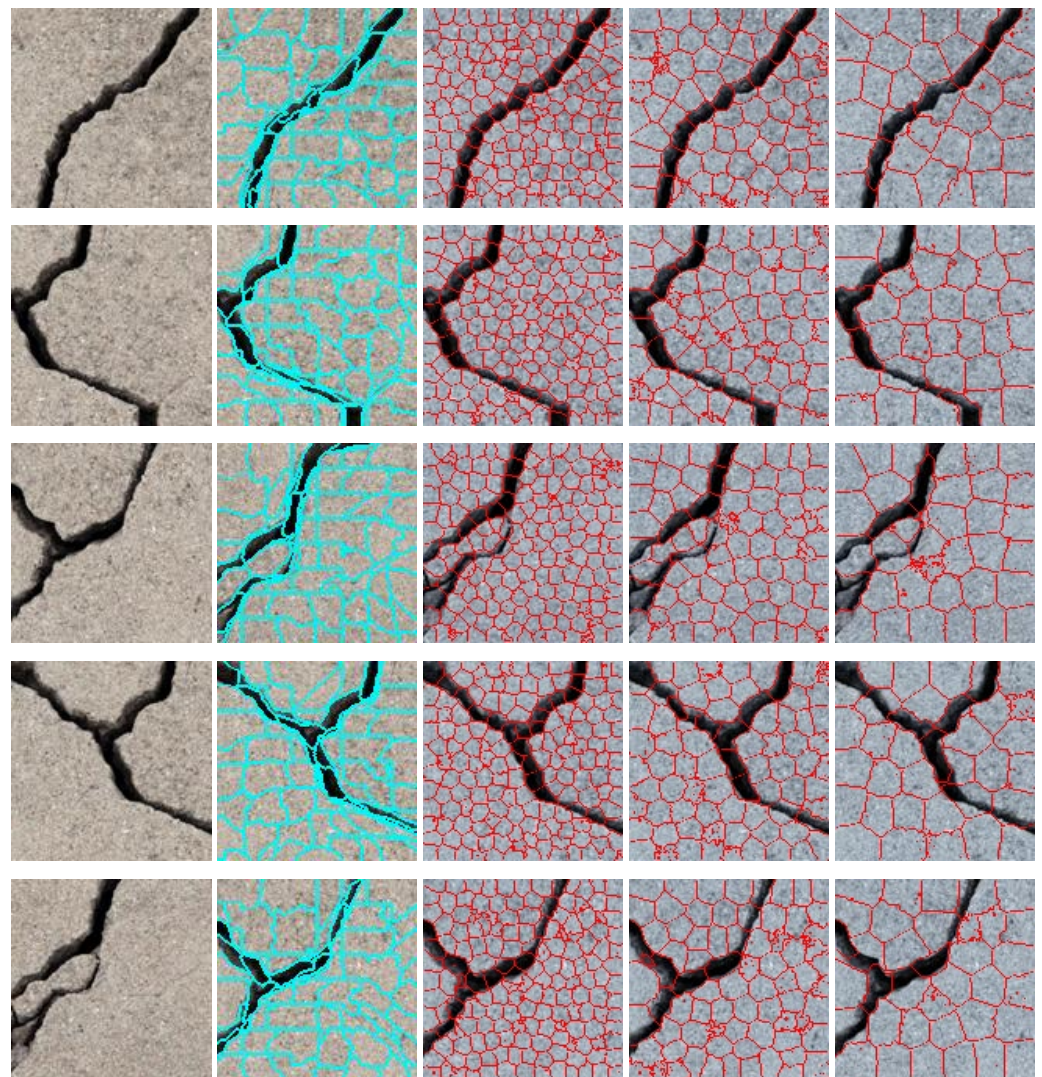


Figure 3. The first column shows original images. The second column shows the deep learning-based superpixel segmentation. The third to fifth columns indicate SLIC superpixel segmentation with 10, 15, and 20 pixels in each superpixel (cluster), respectively.

The second dataset [28] consists of crack images collected from Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, USA, specifically from asphalt pavements near the Thomas M. Murray Structures and Materials Laboratory. To capture visible and infrared images simultaneously, the FLUKE TiX580 [28] was utilized to capture infrared camera, which features both visible light and infrared camera lenses. Additionally, IR-Fusion™ technology facilitated the creation of fusion images by blending visible and

infrared images at various ratios to produce composite images. This fusion capability could also be achieved using the SmartView software [28]. Consequently, three images—visible, infrared, and fusion—were simultaneously captured for the same object. More detailed information about the images used in this study can be found in the literature [28,29]. Figure 4 shows several samples of visible and fusion images of the dataset.

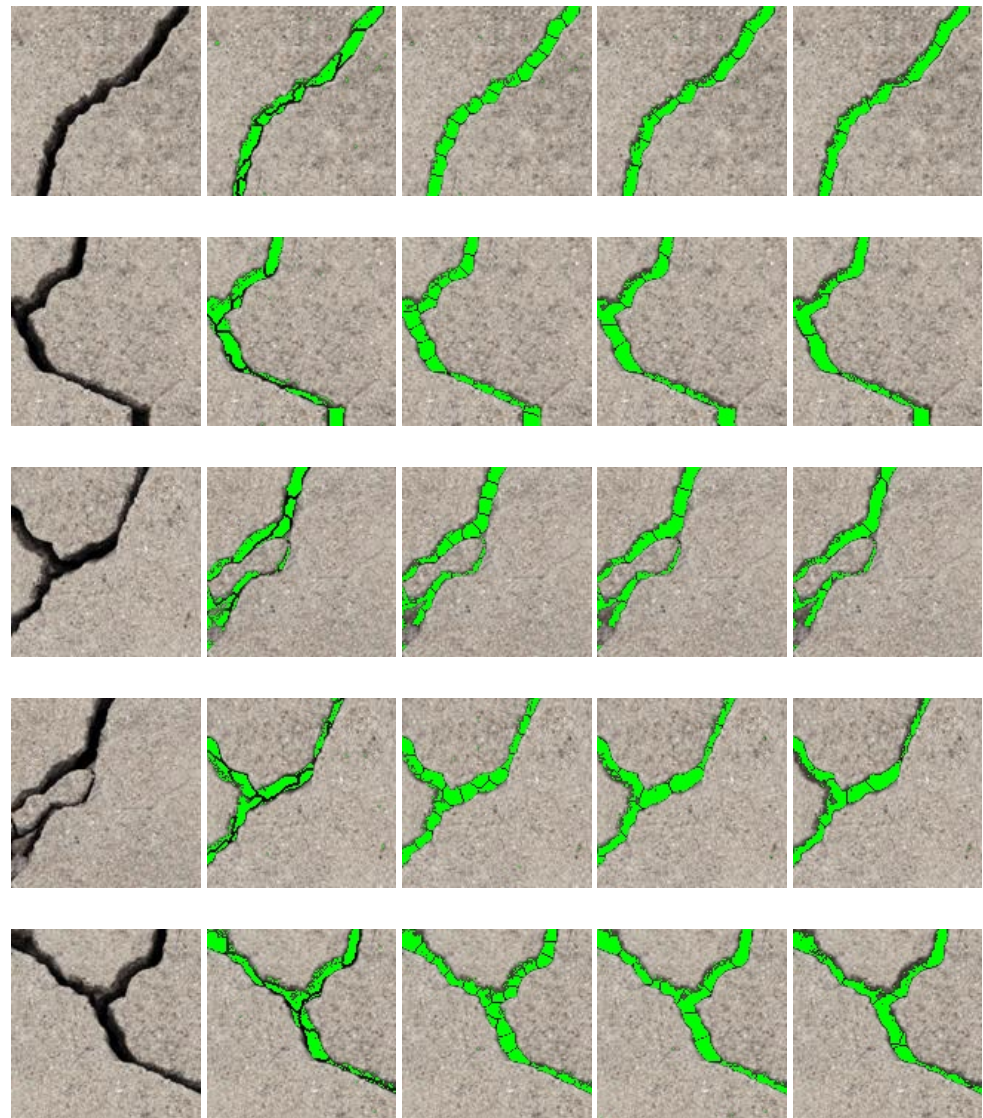


Figure 4. The first column shows the original images. The second column illustrates the results of crack detection using the proposed method. The third to fifth is the results of using SLIC+texture analysis with 10, 15, and 20 pixels in each superpixel (cluster), respectively.

In this paper, a subset of data is randomly selected from each of the datasets introduced earlier. The first dataset is utilized to validate the functionality of the proposed method. Subsequently, the visible and fusion data selected from the second dataset are employed to assess the effectiveness of the proposed methodology on fusion images.

4.2. Results of Image Segmentation Phase

This subsection first presents implementation details of the deep learning-based super-pixel method. After showcasing the segmentation phase results, we compare and analyze the results with the SLIC super-pixel method with different parameters.

4.2.1. Implementation Details

Employing the PyTorch framework facilitated the implementation of this method, providing extensive support for developing and training deep neural networks. GPU acceleration, particularly using the NVIDIA GeForce RTX 2080Ti GPU, addressed the computational demands, enhancing the efficiency of the segmentation model.

4.2.2. Results of Deep Learning-Based Super-Pixel Segmentation Phase

A deep learning model for super-pixel segmentation was employed to precisely segment multi-modal image data [21], driven by the hypothesis that accurate segmentation is crucial for subsequent crack detection. In the segmentation phase, the multi-modal images were partitioned into homogeneous regions conducive to crack detection. The second column in Figure 3 presents the results of segmentation obtained using the proposed approach, showcasing the delineation of crack regions within the visible images. Moreover, Figure 3 indicates deep learning-based segmentation facilitates more accurate boundary delineation than traditional methods, such as SLIC.

Following the training and validation phases using general images, not necessarily defect images, the model underwent testing using crack defect images. This testing phase aimed to evaluate the effectiveness of the segmentation approach in accurately delineating crack regions within the multi-modal images. The method demonstrated superior segmentation by developing deep learning-based super-pixel segmentation, particularly in capturing intricate crack patterns and minimizing over-segmentation. These results underscore the importance of accurate segmentation as a foundational step for effective crack detection methodologies.

4.2.3. Comparative Analysis of Proposed Method with SLIC Method

To further evaluate the efficacy of our segmentation approach, we compared it with the SLIC method using varying parameters. Figure 3 illustrates the segmentation results obtained using both techniques across different parameter settings. Visual inspection reveals that the deep learning-based approach achieved superior segmentation quality, particularly in capturing intricate crack patterns and minimizing over-segmentation.

4.3. Results of Crack Detection Phase

Following successful segmentation, the project progressed to the crack detection phase, where the objective was to identify and localize cracks within the segmented images. The crack detection methodology involved several steps, which are presented in the following subsections.

4.3.1. Implementation Details

The Python programming language, specifically version 3.9, was utilized for implementing the crack detection algorithm. Texture analysis, a key component of crack detection, was performed using the Mahotas library, which provided essential functionality for examining texture characteristics within individual super-pixels.

The Mahotas library facilitated texture analysis to identify patterns indicative of cracks within individual super-pixels. This analysis enabled the differentiation between super-pixels which contained true crack textures and those that did not, thus facilitating accurate crack detection.

4.3.2. Crack Detection Using Visible Images

We evaluated the performance of the proposed crack detection method, employing visible image data, using the segmented regions obtained through deep learning-based super-pixel segmentation.

4.3.3. Performance Metrics

Table 1 presents the quantitative performance metrics obtained from crack detection using visible images, including IOU, precision, and recall. These metrics provide insights into the effectiveness of the proposed approach in accurately identifying and delineating cracks in visible imagery.

Table 1. Evaluation metric results for SLIC-superpixel+ texture analysis and the proposed method.

Sample	Method	Intersection of Union (IOU)	Precision	Recall
Sample 1	Proposed Method	93.72	93.81	99.90
	SLIC10	92.13	92.20	99.91
	SLIC15	92.98	92.44	99.92
	SLIC20	92.98	93.05	99.92
Sample 2	Proposed Method	94.02	94.08	99.93
	SLIC10	93.47	93.58	99.88
	SLIC15	92.44	92.48	99.95
	SLIC20	92.73	92.74	96.22
Sample 3	Proposed Method	94.02	94.08	99.93
	SLIC10	93.47	93.58	99.83
	SLIC15	93.28	93.32	99.95
	SLIC20	93.88	93.93	99.95
Sample 4	Proposed Method	94.08	94.08	99.93
	SLIC10	92.16	92.21	99.93
	SLIC15	92.34	92.37	99.95
	SLIC20	93.40	93.44	99.94
Sample 5	Proposed Method	92.79	92.83	99.96
	SLIC10	90.81	90.87	99.92
	SLIC15	90.42	90.49	99.92
	SLIC20	91.80	91.85	99.94

4.3.4. Qualitative Analysis

Figure 4 depicts qualitative examples of crack detection results obtained from the visible image dataset. Visual inspection of these examples illustrates the ability of the proposed method to detect cracks of varying sizes, orientations, and intensities within visible images.

4.4. Multi-Modal Images for Crack Detection

To further investigate the potential benefits of multi-modal fusion, we utilized datasets containing both visible and fusion images. We analyzed the effectiveness of fusion images on the proposed crack detection performance and explored if fusion images can deal with the situations of blurriness in images [30].

4.5. Effectiveness of Fusion Images

Figures 4 and 5 present qualitative examples illustrating the effectiveness of fusion images on the proposed crack detection method. These examples highlight instances where fusion images enhance crack visibility and aid in the accurate delineation of crack boundaries compared to visible images alone.

4.6. Comparative Analysis

Tables 2 and 3 provide a comparative analysis of crack detection metrics obtained from visible images and fusion images. By comparing the performance of our method across these datasets, we aim to determine the extent to which fusion improves crack detection accuracy under different conditions.

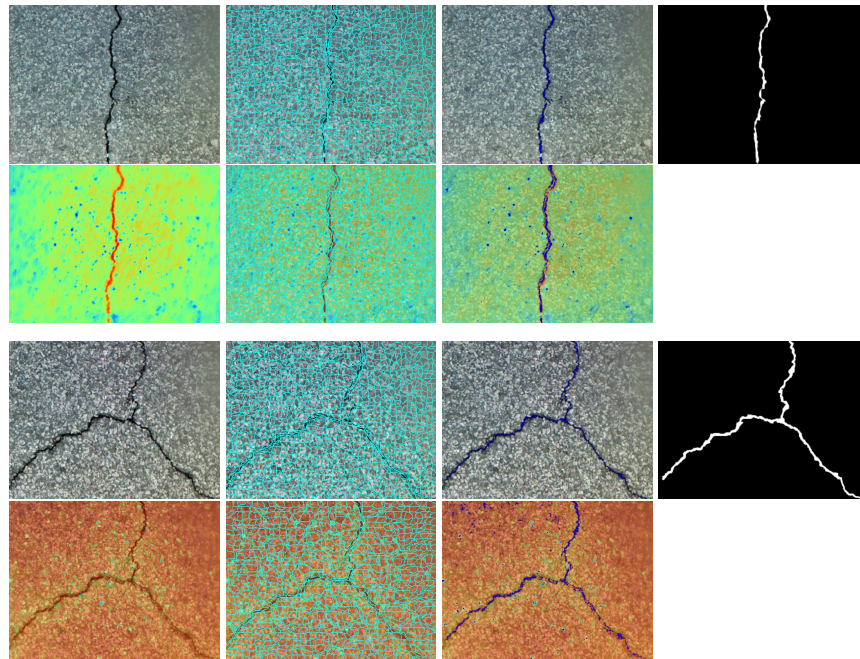


Figure 5. Crack detection results for Sample 1 (top) and Sample 2 (bottom): visible image, deep learning-based super-pixel segmentation for visible image, detected crack on visible image, ground-truth, fusion image, deep learning-based super-pixel segmentation for fusion image, detected crack on fusion image.

Table 2. Evaluation metrics results on visible and fusion images using proposed method.

Sample	Image Spectrum	Intersection of Union (IOU)	Precision	Recall
Sample 2	Visible	97.15	99.12	79.99
	Fusion	99.60	99.99	99.60
Sample 3	Visible	99.67	99.82	99.84
	Fusion	98.93	99.56	99.36

Table 3. Evaluation metrics results on visible and fusion images using proposed method for worst-case scenario.

Sample	Image Spectrum	Intersection of Union (IOU)	Precision	Recall
Sample 1	Visible	73.56	98.87	98.66
	Fusion	95.51	99.99	95.55
Sample 2	Visible	76.15	99.12	79.99
	Fusion	99.60	99.99	99.60
Sample 3	Visible	75.67	99.82	76.84
	Fusion	98.93	99.56	99.36

4.7. Discussion

In this section, we interpret the results obtained from our crack detection experiments and discuss their implications for non-destructive testing methodologies. We analyze the effectiveness of our method using visible images, evaluate the added value of multi-modal fusion for improving crack detection accuracy, and discuss the conditions under which fusion images provide the most significant benefits.

Figure 3 investigated the trade-off between computational efficiency and segmentation accuracy inherent in the simple linear iterative clustering (SLIC) method. We found that while reducing the seed size in SLIC can lead to more detailed segmentation results, it also increases computational time significantly. Conversely, larger seeds can expedite the

segmentation process but may sacrifice accuracy. In contrast, our experiments utilized fixed parameters for deep learning-based super-pixel segmentation, ensuring consistency in results across different samples. This approach allowed us to explore the balance between computational efficiency and segmentation quality, providing valuable insights for practical applications in image processing and computer vision.

This study extended the investigation, to encompass the fusion of infrared and visible images, aiming to evaluate the segmentation of crack regions and assess its impact on crack detection efficacy. Figures 4–6 and Tables 1–3 reveal the method’s remarkable accuracy in detecting crack regions within both thin and thick crack formations in visible images, as depicted in the accompanying figure. Furthermore, we explored the fusion counterparts of these images, considering the potential benefits of integrating infrared data with visible imagery. Given the inherent presence of cracks in visible images, this study sought to ascertain whether the fusion of infrared and visible images could enhance crack detection capabilities. This comprehensive analysis sheds light on the potential synergistic effects of fusion images in bolstering crack detection methodologies, thereby advancing non-destructive testing and defect identification.

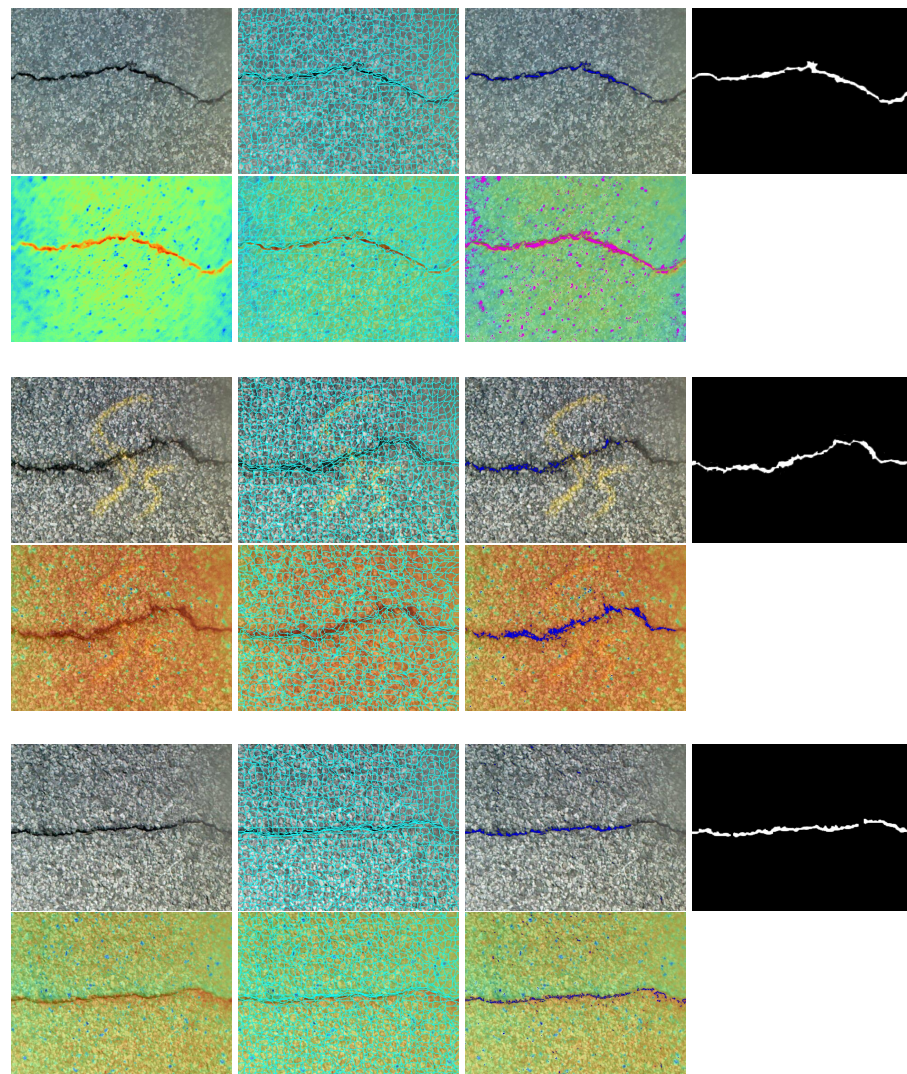


Figure 6. Crack detection results for Sample 3 (top), Sample 4 (middle), and Sample 5 (bottom): visible image, deep learning-based super-pixel segmentation of visible image, detected crack on visible image, ground-truth, fusion image, deep learning-based super-pixel segmentation of fusion image, detected crack on fusion image.

Our proposed crack detection methodology has exhibited remarkable performance across various scenarios, particularly excelling in challenging cases where traditional methods struggle. Figure 6 and Table 3 show that in the worst-case scenario, where one side of the image suffered from blur, the proposed methodology showcased exceptional resilience. By using fusion images and the proposed approach, we effectively mitigated the adverse effects of blurriness, demonstrating the robustness and adaptability of the proposed method to real-world imperfections commonly encountered in infrastructure inspection scenarios.

Furthermore, the segmentation results in Figure 3, columns 3, 4, and 5, indicate segments containing both crack and non-crack regions using the SLIC method. Conversely, Column 1, which employs a deep learning-based super-pixel approach, notably enhances the segmentation precision, distinctly illustrating crack and non-crack areas. Moreover, in columns 3, 4, and 5, SLIC exhibited tendencies towards over-segmentation, resulting in fragmented regions and potentially compromising the accuracy of subsequent crack detection. Conversely, the proposed methodology, appropriate for optimal performance even in limited-data scenarios, demonstrated superior segmentation quality, providing a coherent and precise representation of image regions conducive to accurate crack detection. Table 3 illustrates that the IOU scores for Sample 1 through Sample 5 were 94.02%, 94.02%, 94.8%, and 92.79%, respectively. Conversely, the IOU results for the SLIC methods indicate that, at best, they performed 1.29% worse on average compared to these scores. Also, the recall scores for Sample 1 through Sample 5 were 93.81%, 94.08%, 94.8%, and 94.08%, 92.83%, respectively. Conversely, the recall results for the SLIC methods indicate that, at best, they performed approximately 1% worse. There were no significant differences in recall.

It is essential to emphasize that the proposed approach is specifically designed to address the challenges posed by limited data availability, a common constraint in real-world applications of crack detection. By decoupling the segmentation task from reliance on annotated data, we have effectively alleviated the burden of manual annotation, thereby facilitating the widespread applicability of our methodology. This significant advancement not only enhances the efficiency of crack detection processes but also contributes to the scalability and accessibility of infrastructure inspection technologies, particularly in contexts where resources for data annotation are limited.

In short, our crack detection methodology, augmented by fusion image utilization and deep learning-based super-pixel segmentation, represents a significant step forward in the field of non-destructive testing for infrastructure assessment. Through rigorous evaluation and validation, our approach has demonstrated exceptional performance, particularly in mitigating the challenges posed by adverse imaging conditions and limited data availability.

5. Conclusions

In conclusion, this paper presents a comprehensive methodology for automated crack detection in concrete structures. By integrating deep learning-based super-pixel segmentation with texture analysis, we address the challenges associated with manual inspection methods, offering a more efficient and objective approach to crack detection. Our methodology leverages state-of-the-art techniques in image processing and computer vision, enabling accurate identification and characterization of cracks, even in scenarios with limited training data. Through systematic evaluation and comparison with existing methods, using public benchmark datasets, we demonstrate the effectiveness and superiority of our approach. As a future work, focusing on the infrared and visible fusion of stitched images is recommended so that we can simultaneously detect surface and subsurface defects and have sufficient understanding of the defect location and continuity.

Author Contributions: Conceptualization, S.S., F.L., C.I.-C. and X.P.V.M.; data curation, S.S., F.L. and X.P.V.M.; formal analysis, S.S., F.L., C.I.-C. and X.P.V.M.; methodology, S.S., F.L., C.I.-C. and X.P.V.M.; project administration, F.L., C.I.-C. and X.P.V.M.; resources, C.I.-C., F.L. and X.P.V.M.; software, S.S.; supervision, X.P.V.M.; validation, S.S., F.L., C.I.-C. and X.P.V.M.; visualization, S.S.; writing—original draft, S.S., F.L., C.I.-C. and X.P.V.M.; writing—review and editing, S.S., F.L., C.I.-C. and X.P.V.M. All authors have read and agreed to the published version of the manuscript.

Funding: The Natural Sciences and Engineering Council of Canada (NSERC), CREATE-oN DuTy Program (funding reference number 496439-2017), the Canada Research Chair in Multipolar Infrared Vision (MIVIM), and the Canada Foundation for Innovation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data for this research are available at the following links: <https://github.com/KangchengLiu/Crack-Detection-and-Segmentation-Dataset-for-UAV-Inspection> (accessed on 1 May 2024), <https://github.com/lfangyu09/IR-Crack-detection> (accessed on 1 May 2024).

Acknowledgments: We acknowledge the support of Torngats Services Techniques for providing the required equipment and support for performing the experiments.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

NDT	non-destructive testing
CNN	convolutional neural network
SLIC	simple linear iterative clustering
ReLU	leaky rectified linear units
SHM	structural health monitoring

References

1. Sfarra, S.; Cicone, A.; Yousefi, B.; Ibarra-Castanedo, C.; Perilli, S.; Maldague, X. Improving the detection of thermal bridges in buildings via on-site infrared thermography: The potentialities of innovative mathematical tools. *Energy Build.* **2019**, *182*, 159–171. [CrossRef]
2. Maldague, X. Applications of infrared thermography in nondestructive evaluation. *Trends Opt. Nondestruct. Test.* **2000**, 591–609. [CrossRef]
3. Ghiass, R.S.; Arandjelović, O.; Bendada, H.; Maldague, X. A unified framework for thermal face recognition. In *Neural Information Processing, Proceedings of the 21st International Conference, ICONIP 2014, Kuching, Malaysia, 3–6 November 2014, Proceedings, Part II*; Springer International Publishing: Berlin/Heidelberg, Germany, 2014; pp. 335–343.
4. Qu, Z.; Ju, F.R.; Guo, Y.; Bai, L.; Chen, K. Concrete surface crack detection with the improved pre-extraction and the second percolation processing methods. *PloS ONE* **2018**, *13*, e0201109. [CrossRef] [PubMed]
5. Lins, R.G.; Sidney, N.G. Automatic crack detection and measurement based on image analysis. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 583–590. [CrossRef]
6. Shan, B.; Zheng, S.; Ou, J. A stereovision-based crack width detection approach for concrete surface assessment. *KSCE J. Civ. Eng.* **2016**, *20*, 803–812. [CrossRef]
7. Salman, M.; Mathavan, S.; Kamal, K.; Rahman, M. Pavement crack detection using the Gabor filter. In *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, The Hague, The Netherlands, 6–9 October 2013; pp. 2039–2044.
8. Broberg, P. Surface crack detection in welds using thermography. *NDT E Int.* **2013**, *57*, 69–73. [CrossRef]
9. Liu, Y.; Yao, J.; Lu, X.; Xie, R.; Li, L. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing* **2019**, *338*, 139–153. [CrossRef]
10. Escalona, U.; Arce, F.; Zamora, E.; Sossa, A. Fully convolutional networks for automatic pavement crack segmentation. *Comput. Syst.* **2019**, *23*, 451–460. [CrossRef]
11. Jenkins, M.D.; Carr, T.A.; Iglesias, M.I.; Buggy, T.; Morison, G. A Deep Convolutional Neural Network for Semantic Pixel-Wise Segmentation of Road and Pavement Surface Cracks. In *Proceedings of the 2018 26th European Signal Processing Conference (EUSIPCO)*, Rome, Italy, 3–7 September 2018; pp. 2120–2124.

12. Yang, X.; Li, H.; Yu, Y.; Luo, X.; Huang, T.; Yang, X. Automatic Pixel-Level Crack Detection and Measurement Using Fully Convolutional Network. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 1090–1109. [[CrossRef](#)]
13. Fan, Z.; Li, C.; Chen, Y.; Wei, J.; Loprencipe, G.; Chen, X.; Di Mascio, P. Automatic Crack Detection on Road Pavements Using Encoder-Decoder Architecture. *Materials* **2020**, *13*, 29–60. [[CrossRef](#)] [[PubMed](#)]
14. König, J.; Jenkins, M.D.; Barrie, P.; Mannion, M.; Morison, G. A Convolutional Neural Network for Pavement Surface Crack Segmentation Using Residual Connections and Attention Gating. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019.
15. Zhang, L.; Yang, F.; Zhang, D.Y.; Zhu, Y.J. Road crack detection using deep convolutional neural network. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 3708–3712.
16. Xue, H.; Chen, X.; Zhang, R.; Wu, P.; Li, X.; Liu, Y. Deep learning-based maritime environment segmentation for unmanned surface vehicles using superpixel algorithms. *J. Mar. Sci. Eng.* **2021**, *9*, 1329. [[CrossRef](#)]
17. Yuan, X.; Shi, J.; Gu, L. A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Syst. Appl.* **2021**, *169*, 114417. [[CrossRef](#)]
18. Tuceryan, M.; Jain, A. Texture analysis. In *Handbook of Pattern Recognition and Computer Vision*; World Scientific Publishing: Singapore, 1993; pp. 235–276.
19. Chen, Y.; Ding, Y.; Zhao, F.; Zhang, E.; Wu, Z.; Shao, L. Surface defect detection methods for industrial products: A review. *Appl. Sci.* **2021**, *11*, 7657. [[CrossRef](#)]
20. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Susstrun, S. Slic superpixels. *Technical Report 149300 EPFL*, June 2010.
21. Yang, F.; Sun, Q.; Jin, H.; Zhou, Z. Superpixel segmentation with fully convolutional networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 13964–13973.
22. Yu, Y.; Wang, C.; Fu, Q.; Kou, R.; Huang, F.; Yang, B.; Yang, T.; Gao, M. Techniques and challenges of image segmentation: A review. *Electronics* **2023**, *12*, 1199. [[CrossRef](#)]
23. Neupane, B.; Horanont, T.; Aryal, J. Deep learning-based semantic segmentation of urban features in satellite images: A review and meta-analysis. *Remote Sens.* **2021**, *13*, 808. [[CrossRef](#)]
24. Khoriphan, W.; Niwitpong, S.A.; Niwitpong, S. Confidence Intervals for the Ratio of Variances of Delta-Gamma Distributions with Applications. *Axioms* **2022**, *11*, 689. [[CrossRef](#)]
25. Mehrtash, A.; Wells, W.M.; Tempny, C.M.; Abolmaesumi, P.; Kapur, T. Confidence calibration and predictive uncertainty estimation for deep medical image segmentation. *IEEE Trans. Med. Imaging* **2020**, *39*, 3868–3878. [[CrossRef](#)] [[PubMed](#)]
26. Shahsavarani, S.; Lopez, F.; Ibarra-Castaneda, C.; Maldague, X. Semantic Segmentation of Defects in Infrastructures through Multi-modal Images. In *Thermosense: Thermal Infrared Applications XLV*; SPIE: Bellingham, WA, USA, 2024.
27. Liu, K.; Han, X.; Chen, B.M. Deep learning based automatic crack detection and segmentation for unmanned aerial vehicle inspections. In Proceedings of the 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), Dali, China, 6–8 December 2019; IEEE: New York, NY, USA, 2019.
28. Liu, F.; Liu, J.; Wang, L. Asphalt pavement crack detection based on convolutional neural network and infrared thermography. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 22145–22155. [[CrossRef](#)]
29. Liu, F.; Liu, J.; Wang, L. Deep learning and infrared thermography for asphalt pavement crack severity classification. *Autom. Constr.* **2022**, *140*, 104383. [[CrossRef](#)]
30. Shahsavarani, S.; Lopez, F.; Maldague, X. Multi-modal image processing pipeline for NDE of structures and industrial assets. In *Thermosense: Thermal Infrared Applications XLV*; SPIE: Bellingham, WA, USA, 2023; Volume 12536, pp. 255–264.

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