

## Article

# Photoelectricity Theory-Based Concrete Crack Image Segmentation and Optimal Exposure Interval Research

Haijing Liu <sup>1</sup>, Renwei Luo <sup>2</sup>, Yan Liu <sup>2</sup>, Ji He <sup>2</sup>, Yongzhi Sang <sup>2</sup>, Jiayan Zheng <sup>2,\*</sup> and Zhixiang Zhou <sup>3</sup>

<sup>1</sup> China Merchant Chongqing Communication Research & Design Institute Co., Ltd., Chongqing 400067, China; liuhaijing@cmhk.com

<sup>2</sup> School of Civil Engineering, Chongqing Jiaotong University, Chongqing 400074, China; 622210970187@mails.cqjtu.edu.cn (R.L.); 622210030011@mails.cqjtu.edu.cn (Y.L.); 622220030009@mails.cqjtu.edu.cn (J.H.); 622230030003@mails.cqjtu.edu.cn (Y.S.)

<sup>3</sup> School of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China; zhixiangzhou@szu.edu.cn

\* Correspondence: jiayanzheng@cqjtu.edu.cn; Tel.: +86-18-1665-64128

**Abstract:** To solve the problem of low accuracy in automatic concrete crack image segmentation and the non-standardization of concrete crack image datasets, an exposure-based concrete crack image capture scene characterization method was proposed, and the optimal exposure interval for crack segmentation was presented by multiple scene image capture experiments. First, current public crack datasets were collected and analyzed, and it was shown that improper spatial resolution, mislabeling, overexposure, and defocus are frequent non-standardization problems in crack dataset production. Through the analysis of the photoelectric principle in concrete crack imaging, an equivalent exposure was set as a core indicator for scene characterization. Twenty-one indoor scenes were designed by varying the illumination intensity and exposure time, and the experimental results showed that an equivalent exposure can be a core control index for scene characterization. The grayscale distribution law of concrete crack images was analyzed with four specimens' images captured indoors in 50 exposure scenes, and the segmentation accuracy of an image from each scene was calculated through comparison with corresponding manually labeled binary files. The experiment's results revealed that 5~50 lx·s was the optimal equivalent exposure interval for concrete crack image segmentation, in which better segmentation accuracy was achieved with an F1 score of up to 96.3%.

**Keywords:** concrete crack image segmentation; dataset standardization; optimal exposure interval; photoelectricity theory



**Citation:** Liu, H.; Luo, R.; Liu, Y.; He, J.; Sang, Y.; Zheng, J.; Zhou, Z. Photoelectricity Theory-Based Concrete Crack Image Segmentation and Optimal Exposure Interval Research. *Appl. Sci.* **2024**, *14*, 1527. <https://doi.org/10.3390/app14041527>

Academic Editor: Sang-Hyo Kim

Received: 4 January 2024

Revised: 1 February 2024

Accepted: 6 February 2024

Published: 14 February 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

Cracks are the most frequent defects in concrete structures, and their formation and propagation significantly deteriorate the safety and durability of infrastructures [1]. The accurate detection of the location and severity of cracks is crucial for the operational safety and long-term performance of structures, which depend on the detection and quantification of crack parameters, and machine vision-based crack detection technology has been widely applied due to its high efficiency, convenience, and non-contact advantages [2]. Traditional image processing methods (white-box methods) and artificial intelligence techniques (black-box methods) are replacing manual inspection, measurement, and analysis [3] and have drawn considerable attention from researchers.

White-box methods have the advantages of a low computational cost, traceability, transparency, and interpretability; they do not require large volumes of crack image datasets, and they have been widely used in crack image detection for more than a decade. Wang introduced an algorithm for a multi-angle, multi-structure element morphological filter based on morphological filter techniques, and the method could preserve details in images and enhance the effectiveness of crack identification and extraction [4]. Xu et al. improved

detection accuracy using the OTSU method and by adaptively setting Canny threshold parameters [5]. Dow et al. proposed a skeleton-marker method to remove binary noise and segment concrete cracks and achieved more reliable crack detection results than previous methods [6]. The acquisition of concrete crack images often takes place in complex environments, including illumination variations, stains, and oil residues on the cracks' surfaces and, sometimes, outer object occlusion. Unfavorable environmental factors may influence the value of the gray threshold, the area threshold, and the connectivity in image processing and pose threats or challenges in crack detection and parameter calculation, all of which limit the generalization capability of image segmentation [7].

In AI-based crack segmentation methods, the deep learning method demonstrates excellent performance and has become popular as the mainstream method due to its high accuracy, robustness, and strong generalization capabilities [8–12]. Liu et al. first applied U-Net, a concrete crack segmentation method optimized with the Adam algorithm, and achieved more accurate crack image segmentation than prior to its integration, with effectiveness and robustness [8]. Xiang introduced a dual-coding network, DTrC-Net, and it outperformed other state-of-the-art segmentation networks and exhibited superior generalization in complex scenes [7]. Su et al. proposed the CBAM-Unet algorithm in bridge crack identification, which could effectively reduce detection costs and enhance efficiency [9]. Ren et al. presented an improved deep fully convolutional neural network named CrackSegNet with a higher accuracy and generalization capability, and it made tunnel detection and monitoring efficient and cost-effective [10]. In another study, a combination of white-box methods and black-box methods yielded superior pixel-level segmentation results [11]. Han et al. integrated deep learning with a digital image processing method in crack recognition [12], and the model could automatically locate and extract cracks by means of a deep convolutional neural network combined with local threshold image segmentation. Also, this approach could precisely locate the maximum crack width's position and calculate its width.

Deep learning-based concrete crack segmentation depends on deep convolutional neural network architecture, datasets, and evaluation metrics [13]. The dataset is the digital basis for the data-driven image segmentation method, also called the data-dependent image segmentation method, where dataset quality and quantity are equally important. High-quality concrete crack dataset fabrication is labor-intensive and expensive and causes the scarcity of large-scale datasets, which poses a great challenge to crack semantic segmentation algorithms and limits their robustness and generalizability. As the accuracy of the deep learning-based crack segmentation method is determined by dataset quality and labeling fineness, the standardization of concrete crack image capture scene and dataset fabrication is in urgent demand, yet there are still no standards or characterization parameters for concrete crack dataset production so far.

To standardize the concrete crack imaging scene and produce a high-quality dataset, the objective of this paper is to provide more insight into the difference in concrete crack image segmentability, with images from different exposure scenes, using current dataset standardization analysis, a photoelectric principal study, scene core characterization index selection, and indoor experimental validation. This paper is organized as follows: The current pavement and concrete crack datasets are collected in Section 2 to analyze their standardization levels. In Section 3, the photoelectric principle in concrete crack imaging is analyzed, and equivalent exposure, as a key control index for concrete crack scene characterization, is proposed and verified using images from 21 designed indoor scenes. In Section 4, the mean values and standard deviations of image histograms are analyzed to reveal the grayscale distribution law of concrete crack images from 50 equivalent exposure scenes, and the segmentation accuracy of images in each scene is calculated and compared to find out the optimal equivalent exposure interval for concrete crack image segmentation. Finally, this study is concluded in Section 5.

## 2. Dataset Scene Standardization Analysis

As the information carrier of infrastructure cracks, crack image datasets are prerequisite for white-box and black-box methods based on crack image segmentation and are of great importance to researchers [14–22]. Therefore, a series of crack image datasets of different materials, including concrete, pavement asphalt, and metals, were designed and produced. Although a number of public infrastructure crack image datasets were released, the image capture devices, infrastructure materials, image sizes, image numbers, and imaging environments of popular datasets are totally different, as list in Table 1, which indicates that there is no standard rule for the selection of above factors.

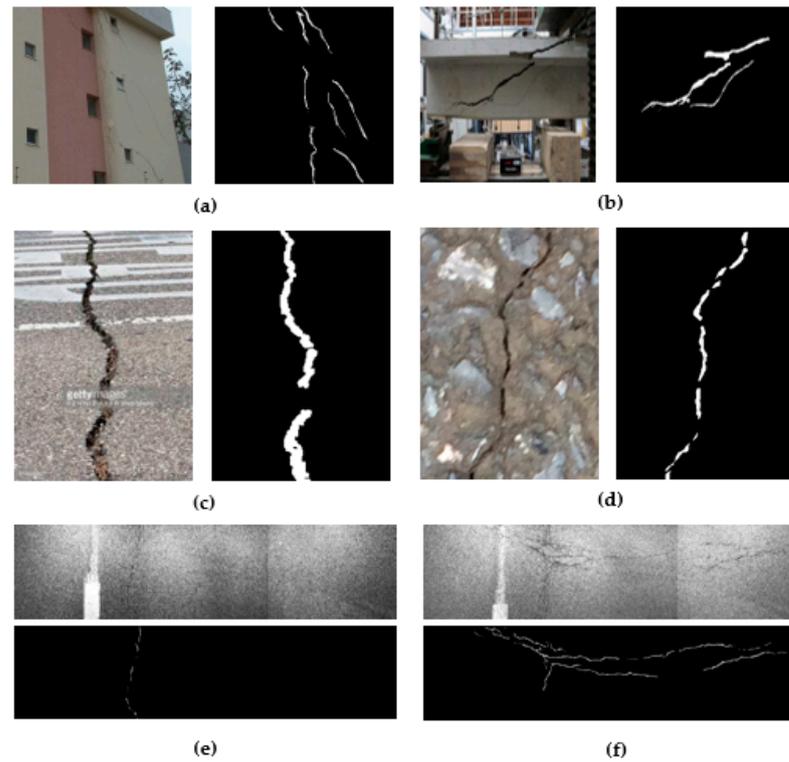
**Table 1.** Public crack segmentation datasets.

Dataset	Materials	Quantity	Resolution	Device
DeepCrack [17]	Hybrid	537	544 × 364	-
CFD [15]	Pavement	320	480 × 320	iPhone5
CrackTree [18]	Pavement	206	600 × 800	Area-array camera
TITS [19]	Pavement	269	Multi	2D laser
Crack500 [20]	Pavement	500	360 × 640	Mobile phone
GAPs384 [21]	Pavement	384	1920 × 1080	2D laser
Crack3238 [7]	Hybrid	3238	256 × 256	-
crack dataset [22]	Concrete	776	297 × 306	Mobile phone

Image capture devices could be mainly categorized into three classes including hand-held, vehicle-carried, and unmanned aerial vehicle-carried camera or laser radar [14], and the type and resolution of capture devices are also different from each other. For example, the capture devices of current public datasets in Table 1 are, respectively, an iPhone5 [15], a mobile phone [20,22], an area-array camera [18], and a 2D laser [19,21]. The classification hierarchy of these devices is very coarse and indistinct, which indicates that there is no standard for crack capture device selection. It can also be seen that the resolution and image number of each dataset are different, indicating non-standardization in the image size and number of datasets.

Moreover, there are non-standardization problems in the image capture environment, image sources, and image labeling in current public infrastructure crack image datasets. In the dataset list in Table 1, only images in Crack3238 [7], DeepCrack [17], and TITS [19] are high quality labeled, and there are mislabeling and incorrectness in other datasets. The images in Figure 1 are typical non-standardization image samples from the datasets in Table 1. In Figure 1a,b, the object building captured is too large, reasonable spatial resolution could not be used, and there are unnecessary background images in the visual field in the image, for example, the sky or unrelated buildings and objects. In Figure 1c, the watermarked pavement crack image is from a network with an unrecognized spatial resolution and capture device, and in Figure 1d, the out-of-focus pictures may have been captured due to operational issues, such as shaking of the acquisition equipment, improper shutter speed, and incorrect aperture. In Figure 1e,f, certain cracks are left unlabeled due to overexposure.

In the theory of photography, image quality is greatly influenced by camera type, image capture scene parameters including surrounding illumination intensity, wavelength and angle, and image capture parameters such as shooting angle, shutter speed, and aperture size. Insight into the photoelectric process of concrete crack image capture is necessary to explore crack imaging scene standardization.



**Figure 1.** Non-standard concrete crack datasets.

### 3. Photoelectric Scene Design and Verification

#### 3.1. The Photoelectric Principle and Concrete Crack Imaging

According to the principles of computational photography, the light energy that a camera sensor receives is determined by external illumination intensity, shutter speed, aperture size, shooting angle, and other camera capture parameters. The external light intensity is the result of environmental illuminance and the surface reflection coefficient of the object being photographed. If the energy loss through the lens is ignored, the radiant illuminance unit on the object surface  $E(\text{W}\cdot\text{m}^{-2})$  can be expressed as [23]:

$$E = L \frac{\pi}{4} \left( \frac{d}{f} \right)^2 \cos^4 \alpha \quad (1)$$

where  $L(\text{W}\cdot\text{m}^{-2}\cdot\text{sr}^{-1})$  is the emittance from the object face in the direction of the lens,  $d(\text{mm})$  is the aperture diameter,  $f(\text{mm})$  is the image distance, and  $\alpha(\text{sr})$  is the off-axis angle. The imaging process is physically a light, electrical, and numerical conversion process, and the energy is the guiding variable throughout. In the entire image capturing process, the energy captured per unit area on the image plane can be calculated as follows:

$$W = E \cdot s = L \frac{\pi}{4} \left( \frac{d}{f} \right)^2 \cos^4 \alpha \cdot s \quad (2)$$

where  $s(\text{ms})$  represents the exposure time. From Equation (2), it can be seen that the energy that camera sensors can capture is in proportion to exposure, as determined by shutter speed, aperture, and the camera's sensitivity, which are the three pillars of exposure triangle in photography. If the object surface follows the law of ideal diffuse reflection, the object face emittance  $L$  depends linearly on the environmental illuminance  $L_{in}(\text{lx})$  on the object surface. The diffusive reflection coefficient of the object surface  $\rho$ , object face emittance  $L$ , and environmental illuminance  $L_{in}$  approximately satisfy:

$$L = \rho \cdot L_{in} \quad (3)$$

when the camera works within the linear range of the photoelectric sensor, the pixel value  $I_p(px)$  approximately follows a linear relationship with  $L_{in}$ ,  $\rho$ , and  $d^2$ , and can be express as:

$$K = \frac{\pi}{4} \left(\frac{1}{f}\right)^2 \cos^4 \alpha \tag{4}$$

$$I_p = K \cdot K_1 \cdot \rho \cdot d^2 \cdot L_{in} \cdot s \tag{5}$$

where  $K_1$  is the linearity coefficient of the camera. If the object surface reflection coefficient  $\rho$ , aperture diameter  $d$ , image distance  $f$ , off-axis angle  $\alpha$ , the camera linearity coefficient  $K_1$ , and other parameters are fixed, the image pixel intensity value is approximately satisfied as follows:

$$K_c = \rho \cdot d^2 \cdot K \cdot K_1 \tag{6}$$

$$I_p = K_c \cdot L_{in} \cdot s \tag{7}$$

According to Equation (7), it can be seen that the gray value of each image pixel is linearly correlated with the external ambient illumination and exposure time. In Equations (2)–(7), the precision of imaging can be assured on condition that the light electric transfer of the photoelectric sensor works in the linear zone, and the error caused by noise, for example, dark current, can be ignored. Also, the precision of the above equation should be measured or validated using experiments only. Therefore, an equivalent exposure method that changes ambient illumination intensity or changes exposure time can be used to analyze the gray histogram distribution law of concrete in different illumination scenes.

### 3.2. Concrete Equivalent Exposure Scene Design and Validation

#### 3.2.1. Equivalent Exposure Scene Design

According to Equation (7), the product of  $L_{in}$  and  $s$  can serve as the equivalent exposure of the field environment, and equivalent exposure can be controlled by adjusting both ambient illuminance and camera exposure time. Twenty-one scenes of equivalent exposure were designed, as shown in Table 2. In Table 2, in the exposure time control scenes (EC), the illuminance remained constant (100 lx), and the camera exposure time was set from 100 ms to 2500 ms every 120 ms. In the illumination control scenes (IC), the exposure time remained constant (200 ms), and the ambient illumination varied from 50 lx to 1250 lx by increments of 60 lx.

**Table 2.** Equivalent exposure scenes.

Scene Number	Exposure Time Control		Illumination Control		Equivalent Exposure/lx·s
	Illuminance/lx	Exposure Time/ms	Exposure Time/ms	Illuminance Designed/lx	
1		100		50	10
2		220		110	22
3		340		170	34
4		460		230	46
5		580		290	58
6		700		350	70
7		820		410	82
8		940		470	94
9		1060		530	106
10		1180		590	118
11	100 ± 1	1300	200	650	130
12		1420		710	142
13		1540		770	154
14		1660		830	166
15		1780		890	178
16		1900		950	190
17		2020		1010	202
18		2140		1070	214
19		2260		1130	226
20		2380		1190	238
21		2500		1250	250

### 3.2.2. Experiment Devices and Setup

To validate photoelectric theory and equivalent exposure principles in concrete crack imaging, an indoor image capture experiment of a 400 mm × 300 mm × 50 mm-sized concrete specimen was conducted with acquisition resolution of 2400 pixels × 1600 pixels in the specified area, as shown in Figure 2. The specified area was 240 mm × 160 mm, and the image spatial resolution captured was 0.1 mm/pixel.

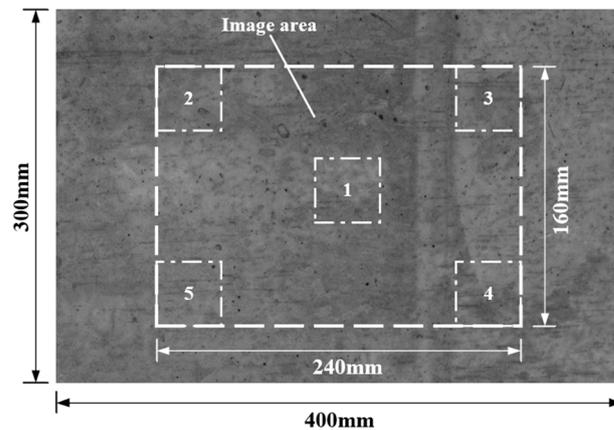


Figure 2. Concrete specimen image area.

To achieve uniformly distributed illumination on the specimen surface, two LED lights whose light intensity and location could be controlled were used, as shown in Figure 3. Furthermore, five verification zones were set to ensure illumination uniformity, as shown in Figure 2, of which the illumination was measured by a TES-1339R illuminometer. The error of the measured and designed illumination values in each verification zone should be less than 2 lx.



Figure 3. Arrangement of the experimental devices.

The camera, lens, illuminometer, LED light, and concrete member were set up as shown in Figure 3, and the device name, type, and parameters and the specification of the devices and shooting are listed in Table 3. According to Equations (6) and (7), aperture was considered constant and set to  $f/5.6$ , and the concrete specimen images were captured with a spatial resolution of 0.1 mm/pixel.

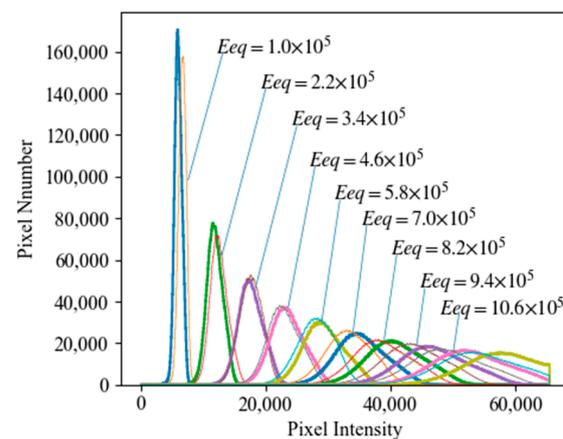
### 3.2.3. Concrete Image Photoelectric Behavior

In the 21 equivalent exposure scenes, images from the exposure time control scene and the corresponding illuminance control scene are theoretically identical, yet slight difference exists between equivalent exposure scenes due to the difficulty in illumination distribution uniformity on the specimen surface by LED illuminance control. Histograms of each equivalent exposure scene of two control style groups are shown in Figure 4,

and it can be seen that the concrete image histograms nearly follow the law of the normal distribution curve, and there is a slight difference in the histograms in each equivalent scene. Figure 4 indicates that the peak values and mean values of the image histograms showed no significant difference in exposure time-controlled scenes and illumination-controlled scenes at each equivalent exposure level. Therefore, the histograms for both control groups exhibit remarkable similarity, which serves as additional validation of the equivalent exposure scenario despite the control manner.

**Table 3.** Experimental devices and parameters.

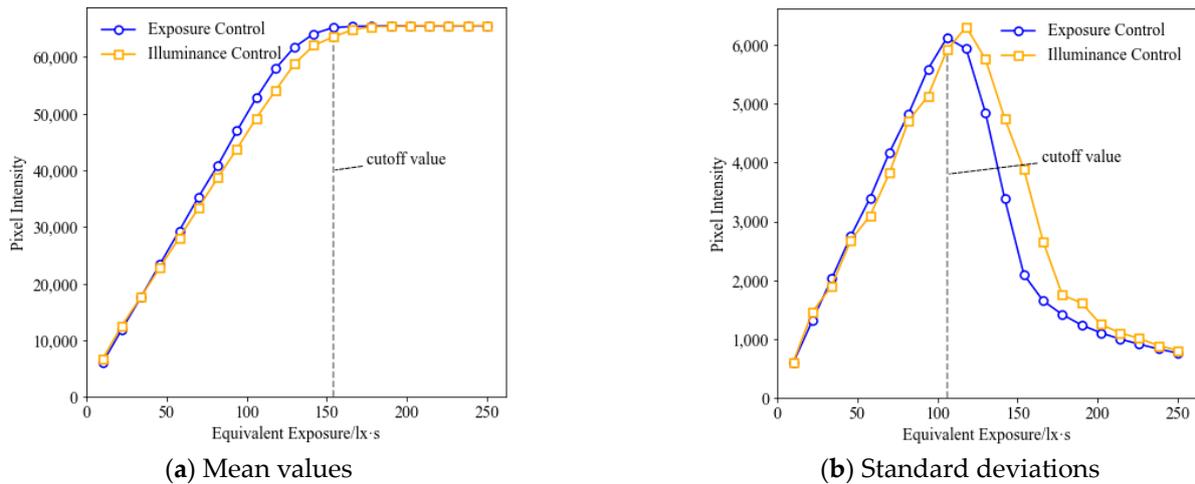
Device Name	Model	Parameter and Specification
Camera	QHY600	Shutter 40 us~3600 s Resolution 9600 × 3194 pixels Image format TIF Bit depth 8/10/12/16
Lens	Nikon AF Zoom-Nikkor	Aperture F2.8~22 Focal length 80~200 mm
Illuminometer	TES-1339R	Measurement range 0.1~999,900 lx Measurement speed five times per second Resolution 0.01 lx



**Figure 4.** Histogram of concrete images under different exposure scenes.

Moreover, to further understand concrete image photoelectric behavior with exposure changes, the relationships between the histogram mean value (HMV) and standard deviation (HSD) changes with exposure are plotted in Figure 5. From Figure 5a, it can be seen that in the low-exposure interval of 0–154 lx·s, the HMV is in linear proportion to equivalent exposure, while in a higher exposure interval, greater than 154 lx·s, the HMV remains at the maximum constant value of 65,535, which indicates extensive overexposure in image capture. In Figure 5b, the HSD also changes linearly with equivalent exposure in the low-exposure interval of 0–106 lx·s, but in a higher exposure interval, the HSD also decreases nearly quadratically with equivalent exposure. According to imaging experience, the HMV and HSD differences in the histograms were caused by the difficulty in uniform illumination control in each scene, which was achieved by the variation in light intensity and location shown in Figure 3.

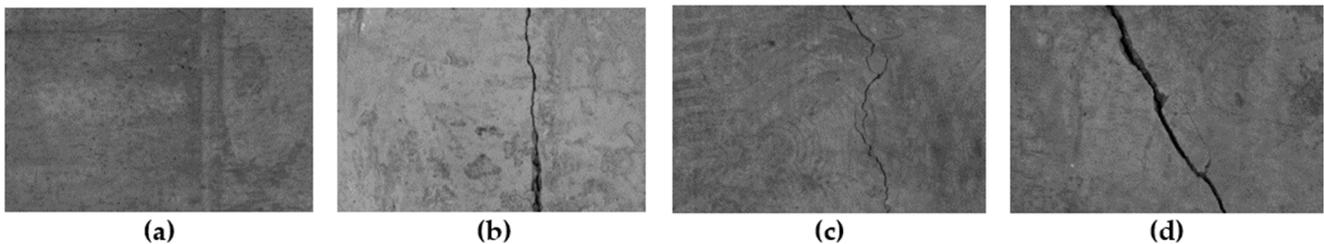
By comparing the mean values and standard deviations of concrete image capture from the 21 equivalent exposure scenes, it can be seen that images from the same exposure scene are nearly identical, and equivalent exposure can be considered a core control index for scene characterization.



**Figure 5.** Comparisons of the HSV and HSD from 21 equivalent exposure scenes in the EC and IC groups.

#### 4. Optimal Exposure Scene Analysis

All machine vision applications are influenced by image contrast, which indicates the difference in objects and is the basis for the image processing algorithm. Contrast is the direct consequence of illumination or exposure. To explore the influence of exposure on concrete crack image capture scenes and segmentation precision, images of four concrete specimens, shown in Figure 6a–d, were captured indoors under 50 equivalent exposures. The specimens were concrete boards with cracks except for specimen A, as shown in Figure 6a. Considering the difficulties in controlling illumination intensity and uniformity, the exposure time control method was applied in 50 different exposure scenes, shown in Table 4, and the exposure time was set from 50 to 2500 ms with an increment of 50 ms. For each specimen, one image was captured under each equivalent scene; in total, 200 images were captured. In the scenes controlled by exposure time, the illumination was set to a constant of 100 lx, and the image capture devices and verification zones were the same as described in Section 3.



**Figure 6.** Image acquisition specimens A/B/C/D.

**Table 4.** Equivalent exposure scenes.

Number	Illumination/lx	Exposure Time/ms	Equivalent Exposure/lx·s
1	100 ± 1	50	5
2		100	10
3		150	15
...		...	...
48		2400	240
49		2450	245
50		2500	250

#### 4.1. Concrete Imaging Mechanism Analysis

In fifty equivalent exposure scenes, the illuminance at the imaging sensor was different due to exposure time variation, and images of different gray levels were captured, as shown in Figure 7a–h, which includes images captured in the scene with exposures of 5lx·s, 40lx·s, 75lx·s, 110lx·s, 145lx·s, 180lx·s, 215lx·s, and 250lx·s, respectively. Figure 7a indicates obvious underexposure of the image, and Figure 7e–h indicates overexposure, so images from these exposure scenes did not have enough contrast for crack segmentation.

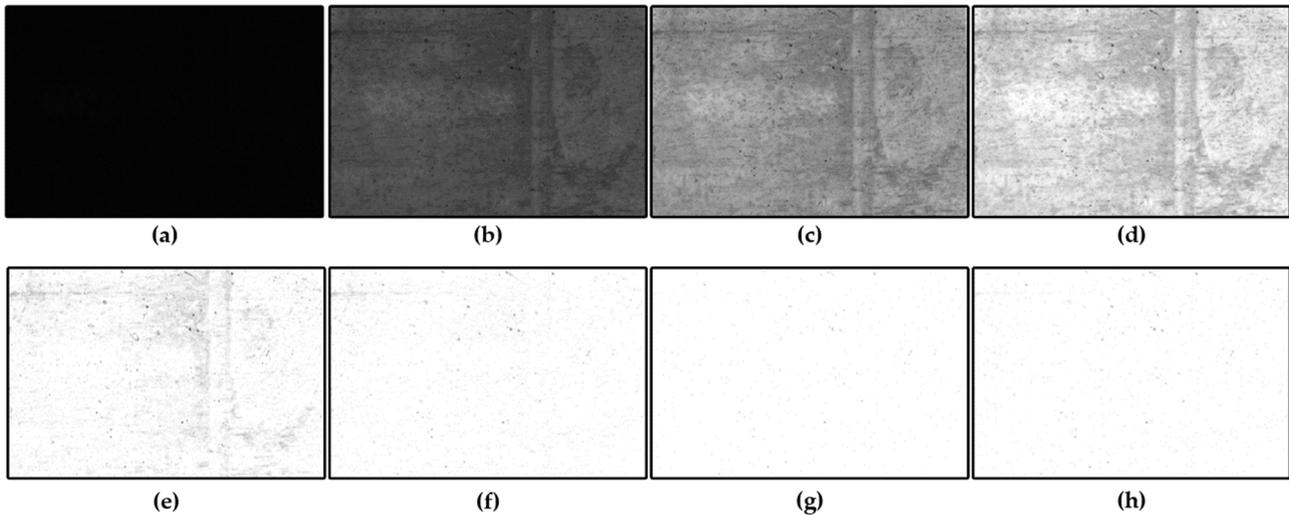


Figure 7. Images of concrete and crack changes with different equivalent exposures.

To clarify and quantify the gray level distribution law of the concrete images from different exposure scenes, image histograms from the first ten scenes are plotted in Figure 8, from which the mean value  $\mu_i$  and standard deviation  $\sigma_i$  of each image is calculated. The normal distribution curves are:

$$f(x) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}} \quad (i = 1, 2, \dots, 10) \tag{8}$$

where  $i$  is the scene number. The normal distribution curves of the first ten exposure scenes are plotted, in the case that overexposure existed in an image from larger equivalent exposure scenes, compared with the corresponding histograms shown in Figure 9.

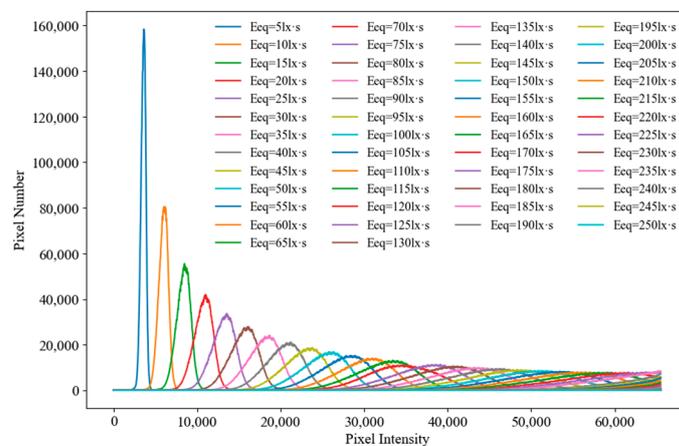


Figure 8. Histogram of images from 50 exposures scenes.

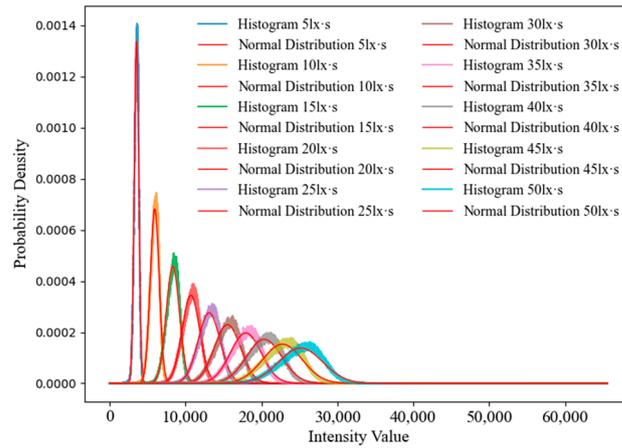


Figure 9. Concrete image histogram and normal distribution curve of ten scenes.

The HMV and HSD variation under the 21 scenes in Section 3 and the 50 equivalent exposure scenes in Table 4 are shown in Figure 10a,b. It can be seen that the two curves of the HMV and HSD in different exposure scenes are closely correlated. In Figure 10, the HMV curve goes up at a nearly constant gradient in the low-exposure interval until leveling off at the cutoff point in the higher exposure interval. After the cutoff exposure point, the HMV remains at the maximum constant value of 65,535. Figure 10 also shows more exposure scenes to help represent the HMV and HSD variation with exposure precisely. For the HMV exposure curve, the cut-off exposure point is 145 lx·s in 50 scenes instead of 154 lx·s, as in Figure 5, and for the HSD exposure curve, the cut-off exposure point is 115 lx·s in 50 scenes instead of 106 lx·s, as in Figure 5. Also, the experiment results show that for the HMV and exposure gradient, the cut-off exposure values are specimen-dependent, depending on the reflective coefficient of the concrete surface according to Equation (6), which is not discussed in this paper.

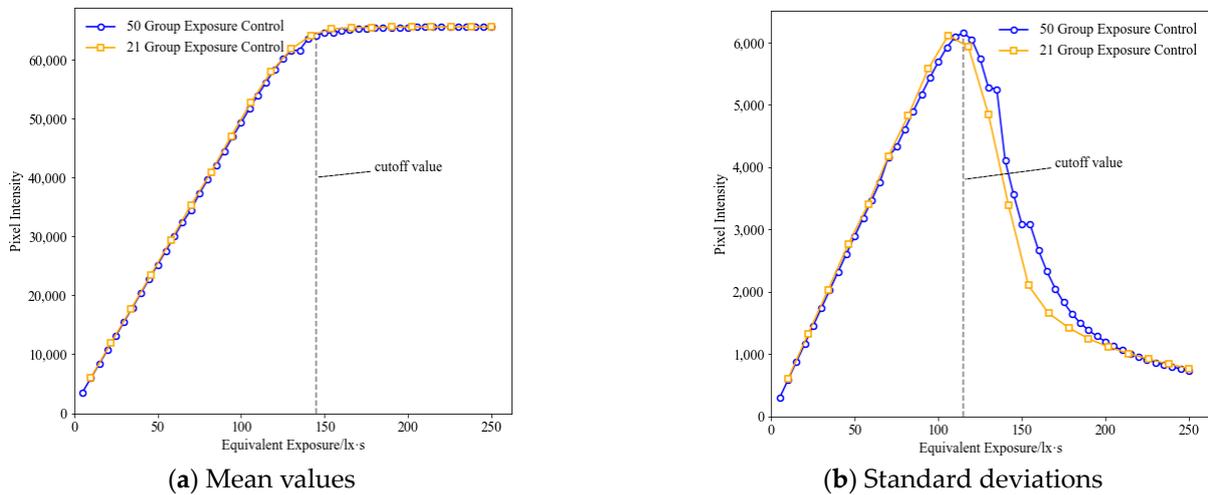
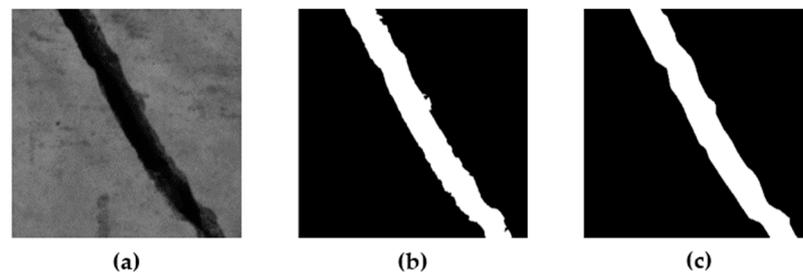


Figure 10. Comparison of the HSV and HSD of 21 and 50 equivalent exposures.

#### 4.2. Optimal Equivalent Exposure Analysis

In concrete crack image segmentation, there is a remarkable imbalance between concrete and the crack pixels in images, and the imbalance extent depends mainly on image size. According to dataset image size in previous research, the size of  $300 \times 300$  pixels for image segmentation was used, and the histogram of specified size images showed a good bimodal property, which is the basis for the Otsu method and iterative thresholding-based segmentation. Figure 11a is the original gray image, Figure 11b is the segmented binary

image produced by iterative thresholding, and Figure 11c is the ground-truth binary image produced by manual labeling with LabelMe4.5.6 software.



**Figure 11.** Crack image, segmentation file, and label file.

In segmentation accuracy metrics, precision, recall, and the F1 score are the most popular indices and are defined as follows:

$$Precision = \frac{TP}{TP + FP} \quad (9)$$

$$Recall = \frac{TP}{TP + FN} \quad (10)$$

In Equations (9) and (10), true positive ( $TP$ ), false positive ( $FP$ ), true negative ( $TN$ ), and false negative ( $FN$ ) are the elements in the confusion matrix of the segmented image and the ground truth image. The F1 score [24,25] is defined based on precision and recall as follows:

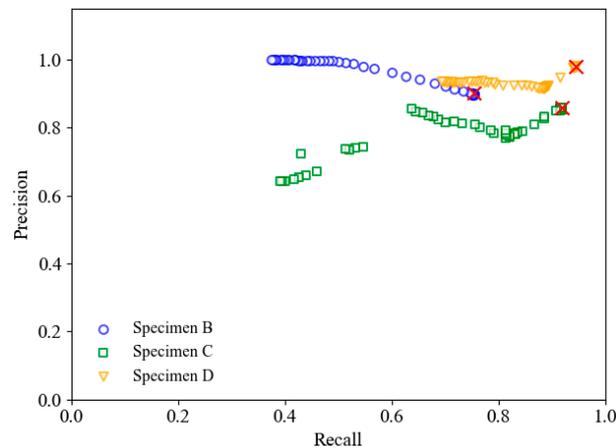
$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (11)$$

Figure 12 illustrates the *precision* and *recall* of the image segmentation for three samples under 50 exposure conditions. Referring to the precision and recall values in Table 5, it can be seen that the precision and recall are both large within an equivalent exposure range of 5–80 lx·s, suggesting effective crack segmentation using digital image processing within this range. Figure 13 reveals a similar law in the F1 scores of three specimens in relation to equivalent exposure. The F1 scores decrease from their peaks downward by 0.5 with the increase in exposure, where the optimal exposure is also 5~80 lx·s. In Figure 13, the largest F1 scores, indicating the most effective distinction between concrete and cracks, are achieved also within this exposure interval, where the scene is also the same as that marked by a red cross in Figure 12. Figure 13 indicates that a higher exposure level adversely affects crack detection accuracy. Moreover, the precision–recall curve and the F1 score of the three specimens show that the segmentation accuracy properties are specimen-dependent even in the same exposure scene, with the same image processing or segmentation method. The property that segmentation accuracy depends on specimen is called segmentability, which needs further exploration.

To be more specific, Table 5 presents the precision, recall, and F1 score values for the first 30 groups of images in the equivalent exposures scenes for specimen D. it can be seen that the F1 score values gradually decrease with the increase in equivalent exposure, and the highest value reaches 96.3% in the exposure range of 5–50 lx·s.

Metric comparisons between the iterative thresholding segmentation method and the ground truth of three specimens are plotted in Figure 14, and the  $TP$ ,  $FP$ ,  $FN$ , and  $TN$  areas are colored with numbers. Panel (a) displays the original images with real cracks and concrete pixel numbers, and panels (b) and (c) present the segmentation maps under the low-exposure scenes of 5 lx·s and 10 lx·s. It can be seen that low exposure reduced the contrast of the concrete crack images and resulted in false positives ( $FPs$ ) and lead to segmentation error. Panels (d) and (e) show the optimal segmentation maps of the scenes achieving the highest F1 score, and the equivalent exposure interval is in the range of 5~50 lx·s. Panels (f) and (g) show the segmentation result under the high-equivalent

exposure scenes of 245 lx·s and 250 lx·s. The accuracy of crack segmentation decreases in the high-exposure intervals, as higher exposure induces an increase in false negatives (FNs), midsegments part of the cracks as concrete, and causes cracks to appear thinner. The experimental results show that equivalent exposure is an important factor in characterizing the concrete image capture scene, and 5~50 lx·s is the optimal equivalent exposure interval for the best segmentation accuracy.



**Figure 12.** P-R curves of three specimens under different exposure scenes.

**Table 5.** Segmentation metrics of the crack images from specimen D.

Equivalent Exposure/lx·s	Precision	Recall	F1score
5	0.981	0.944	0.962
10	0.982	0.944	0.963
15	0.981	0.946	0.963
20	0.981	0.944	0.962
25	0.980	0.945	0.962
30	0.979	0.941	0.960
35	0.978	0.939	0.958
40	0.978	0.941	0.959
45	0.978	0.941	0.959
50	0.979	0.942	0.960
55	0.949	0.916	0.932
60	0.923	0.891	0.907
65	0.924	0.892	0.908
70	0.921	0.889	0.905
75	0.917	0.886	0.901
80	0.918	0.886	0.902
85	0.921	0.889	0.904
90	0.919	0.887	0.903
95	0.918	0.886	0.902
100	0.918	0.885	0.901
105	0.917	0.885	0.901
110	0.915	0.882	0.898
115	0.916	0.878	0.897
120	0.917	0.874	0.895
125	0.920	0.870	0.895
130	0.922	0.861	0.890
135	0.924	0.852	0.887
140	0.925	0.843	0.882
145	0.926	0.829	0.875
150	0.926	0.808	0.863

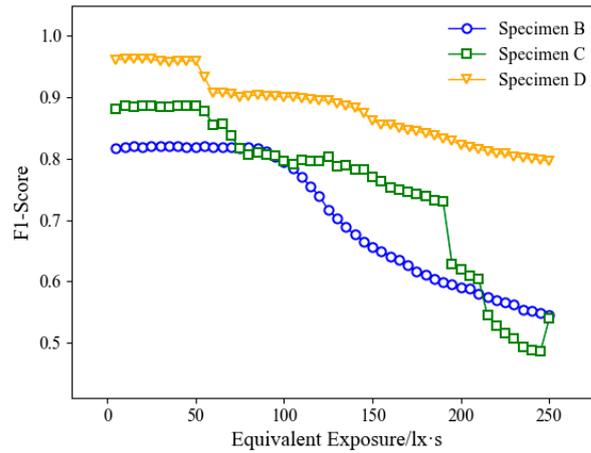


Figure 13. F1 score of three specimens under different exposure scenes.

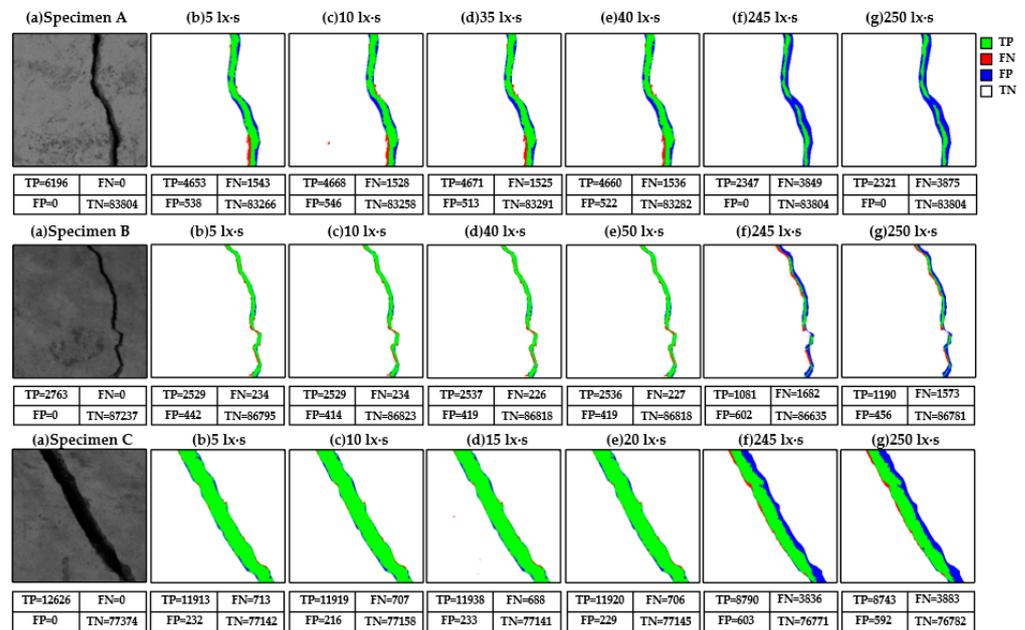


Figure 14. Segmentation of three specimens in different equivalent exposure scenes.

From Figure 14, it can be seen that, within low-exposure intervals, a minor number of false positives (FPs) frequently occur in crack image segmentation because the gray value of the cracks is proximate to that of concrete. Conversely, within high-exposure intervals, cracks are frequently misclassified as concrete and result in false negatives (FNs).

### 5. Conclusions

To standardize the concrete crack image capture scene and produce a high-quality concrete crack image dataset, this paper presents exposure as the scene characterizing parameter for concrete crack image capture based on a standardization investigation, the analysis of current public concrete crack segmentation dataset, and a photoelectric principal analysis. Through equivalent exposure scene design and validation, the optimal exposure interval is proposed in 50 scenes. The main conclusions of this paper are drawn as follows:

(1) The analysis of present publicly accessible datasets showed that non-uniform image capture devices, spatial resolution, image size and number, mislabeling, inappropriate spatial resolution or unnecessary backgrounds, sourced images, out-of-focus images, and motion blur are frequent problems. Thus, the standardization of concrete crack image acquisition scenes is a great challenge for high-precision concrete detection.

(2) Based on the photoelectric principle of the concrete crack imaging process, equivalent exposure was taken as the scene characterization parameter for machine-vision-based infrastructure crack detection. Twenty-one equivalent exposure scenes were designed, and the law of image histograms, mean values, and standard deviations were analyzed to validate the effectiveness of equivalent exposure or the equivalence of exposure time control and illumination control in crack detection.

(3) Concrete crack segmentation of images from 50 equivalent exposure scenes revealed that the highest segmentation precision happened within the 5–50 lx·s equivalent exposure interval, and the F1 score could reach 96.3%. In addition, high exposure was detrimental to concrete crack detection.

This paper found that the standardization of the concrete crack image scene was significant. Optimal equivalent exposure was the core characterization index, which could help to enhance crack segmentation accuracy. The factors that constitute the concrete crack image scene were multiple and complex. In this paper, the experimental specimen number was limited, the illumination was uniform, and the validation was an experiment conducted indoors. More outdoor or onsite experiments should be carried out on more real infrastructure members in complex illumination environments. In addition, the comprehensive understanding of concrete crack imaging scene still needs further exploration. At the same time, the dataset quality evaluation method and characteristic index are not totally clear and should be further explored, and there are factors that influence the quality of datasets, like imaging spatial resolution and image size, which still need more investigation and validation with deep learning algorithms.

**Author Contributions:** H.L.: conceptualization, methodology, and writing—review and editing. R.L.: data collection, software, and writing—original draft and editing. Y.L.: experiment, data collection, and algorithm. J.H.: data collection and image labeling. Y.S.: data collection, analysis, and image labeling. J.Z.: conceptualization, formal analysis, and writing—review. Z.Z.: conceptualization and formal analysis. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant No. 51778094), the Chongqing Municipal Education Commission (Grant No. JDLHPYJD2020019, Grant No. JDDSTD2022003), the Department of Transportation of Sichuan Province (Grant No. 2020-ZL-B1), and the Chongqing Municipal Education Commission (Grant No. CYS22398). We also acknowledge the State Key Laboratory of Mountain Bridge and Tunnel Engineering for providing experimental equipment.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** Haijing Liu was employed by the company of China Merchant Chongqing Communication Research & Design Institute Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Ji, A.; Xue, X.; Wang, Y.; Luo, X.; Xue, W. An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement. *Autom. Constr.* **2020**, *114*, 103176. [[CrossRef](#)]
2. Spencer, B.F.; Hoskere, V.; Narazaki, Y. Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering* **2019**, *5*, 199–222. [[CrossRef](#)]
3. Hsieh, Y.-A.; Tsai, Y.J. Machine Learning for Crack Detection: Review and Model Performance Comparison. *J. Comput. Civ. Eng.* **2020**, *34*, 04020038. [[CrossRef](#)]
4. Wang, Y. An Improved Filtering Algorithm for Bridge Crack Image. *J. Chongqing Jiao Tong Univ. (Nat. Sci.)* **2017**, *36*, 13–17.
5. Xu, H.; Li, Z.; Jiang, Y.; Huang, J.B. Pavement crack detection based on OpenCV and improved Canny operator. *Comput. Eng. Des.* **2014**, *35*, 4254–4258.

6. Dow, H.; Perry, M.; McAlorum, J.; Pennada, S.; Dobie, G. Skeleton-based noise removal algorithm for binary concrete crack image segmentation. *Autom. Constr.* **2023**, *151*, 104867. [[CrossRef](#)]
7. Xiang, C.; Guo, J.; Cao, R.; Deng, L. A crack-segmentation algorithm fusing transformers and convolutional neural networks for complex detection scenarios. *Autom. Constr.* **2023**, *152*, 104894. [[CrossRef](#)]
8. Liu, Z.; Cao, Y.; Wang, Y.; Wang, W. Computer vision-based concrete crack detection using U-net fully convolutional networks. *Autom. Constr.* **2019**, *104*, 129–139. [[CrossRef](#)]
9. Su, H.; Wang, X.; Han, T.; Wang, Z.; Zhao, Z.; Zhang, P. Research on a U-Net Bridge Crack Identification and Feature-Calculation Methods Based on a CBAM Attention Mechanism. *Buildings* **2022**, *12*, 1561. [[CrossRef](#)]
10. Ren, Y.; Huang, J.; Hong, Z.; Lu, W.; Yin, J.; Zou, L.; Shen, X. Image-based concrete crack detection in tunnels using deep fully convolutional networks. *Constr. Build. Mater.* **2020**, *234*, 117367. [[CrossRef](#)]
11. Dorafshan, S.; Thomas, R.J.; Maguire, M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Constr. Build. Mater.* **2018**, *186*, 1031–1045. [[CrossRef](#)]
12. Han, X.; Zhao, Z. Structural surface crack detection method based on computer vision technology. *J. Build. Struct.* **2018**, *39*, 418–427.
13. Lu, D.; Honghu, C.; Lizhi, L. Review of Deep Learning-based Crack Detection for Civil Infrastructures. *China J. Highw. Transp.* **2023**, *36*, 1–21.
14. Ai, D.; Jiang, G.; Lam, S.-K. Computer vision framework for crack detection of civil infrastructure—A review. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105478. [[CrossRef](#)]
15. Shi, Y.; Cui, L.; Qi, Z.; Meng, F.; Chen, Z. Automatic Road Crack Detection Using Random Structured Forests. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 3434–3445. [[CrossRef](#)]
16. Deng, J.; Singh, A.; Zhou, Y.; Lu, Y.; Lee, V.C.S. Review on computer vision-based crack detection and quantification methodologies for civil structures. *Constr. Build. Mater.* **2022**, *356*, 129238. [[CrossRef](#)]
17. 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](#)]
18. Zou, Q.; Cao, Y.; Li, Q.; Mao, Q.; Wang, S. CrackTree: Automatic crack detection from pavement images. *Pattern Recognit. Lett.* **2012**, *33*, 227–238. [[CrossRef](#)]
19. Amhaz, R.; Chambon, S.; Idier, J.; Baltazart, V. Automatic Crack Detection on Two-Dimensional Pavement Images: An Algorithm Based on Minimal Path Selection. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 2718–2729. [[CrossRef](#)]
20. Zhang, L.; Yang, F.; Zhang, Y.D.; 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; IEEE: Piscataway, NJ, USA, 2016; pp. 3708–3712.
21. Yang, F.; Zhang, L.; Yu, S.; Prokhorov, D.; Mei, X.; Ling, H. Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 1525–1535. [[CrossRef](#)]
22. 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](#)]
23. Svoboda, T.; Kybic, J.; Hlavac, V. *Image Processing, Analysis, and Machine Vision: A MATLAB Companion*; Nelson Engineering: Brevard County, FL, USA, 2007; pp. 89–95.
24. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A Review on Deep Learning Techniques Applied to Semantic Segmentation. *arXiv* **2017**, arXiv:1704.06857.
25. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]

**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.