Article

Crack Width Recognition of Tunnel Tube Sheet Based on YOLOv8 Algorithm and 3D Imaging

Xunqian Xu [1], Qi Li [1], Shue Li [2], Fengyi Kang [2], Guozhi Wan [1,*], Tao Wu [1] and Siwen Wang [1]

[1] School of Transportation and Civil Engineering, Nantong University, Nantong 226019, China;
    xunqian_xu@ntu.edu.cn (X.X.); 2118310006@stmail.ntu.edu.cn (Q.L.); 2333320029@stmail.ntu.edu.cn (T.W.);
    2333310005@stmail.ntu.edu.cn (S.W.)
[2] Nantong Highway Development Center, Nantong 226007, China; seli_nt@jsgl.cn (S.L.);
    fykang_nt@jsgl.cn (F.K.)
* Correspondence: 2233310007@stmail.ntu.edu.cn

Abstract: Based on the tunnel crack width identification, there are operating time constraints, limited operating space, high equipment testing costs, and other issues. In this paper, a large subway tunnel is a research object, and the tunnel rail inspection car is an operating platform equipped with industrial cameras in order to meet the requirements of the tunnel tube sheet crack width recognition of more than 0.2 mm, with the measuring instrument to verify that the tunnel rail inspection car in the state of uniform motion camera imaging quality has the reliability through the addition of laser rangefinders, the accurate measurement of the object distance and the calculation of the imaging plane and the angle of the plane to be measured, to amend the three-dimensional cracks. The pixel resolution of the image is corrected, the images imaged by the industrial camera are preprocessed, the YOLOv8 algorithm is used for the intelligent extraction of crack morphology, and finally, the actual width is calculated from the spacing between two points of the crack. The crack detection width obtained by image processing using the YOLOv8 algorithm is basically the same as the value of crack width obtained by manual detection, and the error rate of crack width detection ranges from 0% to 11%, with the average error rate remaining below 4%. Compared with the crack detection error rate of the Support Vector Machine (SVM), the crack extraction model is reduced by 1%, so using the tunnel inspection vehicle as a platform equipped with an industrial camera, YOLOv8 is used to realize the recognition of the shape and width of the cracks on the surface of the tunnel tube sheet to meet the requirements of a higher degree of accuracy. The number of pixels and the detection error rate are inversely proportional to each other. The angle between the imaging plane and the plane under test is directly proportional to the detection error rate. The angle between the vertical axis where the lens midpoint is located and the line connecting the shooting target and the lens center point is α_i and the angle θ_i between the measured plane and the imaging plane is reciprocal, i.e., $\alpha_i + \theta_i = 90°$. Therefore, using the inspection vehicle as a mobile platform equipped with an industrial camera and based on the YOLOv8 algorithm, the crack recognition of the tunnel tube sheet has the feasibility and the prospect of wide application, which provides a reference method for the detection of cracks in the tunnel tube sheet.

Keywords: tunnel tube sheet; laser range finder; image processing; YOLOv8 algorithm; crack width

1. Introduction

Up to now, more than 545 cities in 78 countries and regions around the world have opened urban rail transit systems, with a total mileage of more than 41,386.12 km. In China, for example, a total of 53 cities have opened and operated 290 urban rail transit lines, with an operating mileage of 9584 km [1]. Among them, the metro systems in cities with larger metro networks, such as Shanghai, Beijing, and Guangzhou, have entered a period of both new construction and maintenance [2]. At present, some subway tunnels

have more diseases on the surface of their tube sheets due to physical damage and chemical corrosion, among which cracks are one of the common apparent diseases in subway tunnels. At this stage, there are two main methods for crack width testing and morphology description [3]: the first type is manual testing, and the second type is large-scale equipment testing. Both programs are unable to meet the needs of regular inspection of tunnels. Therefore, image crack recognition technology based on fixed or hand-held imaging devices is rapidly developing, which is more efficient, convenient, and faster than traditional manual detection and large equipment detection.

Numerous organizations at home and abroad have developed tunnel inspection equipment in conjunction with cameras or LiDAR to detect and identify tunnel cracks at various start-up speeds. Japan Keizai Inspection Co Ltd. [4]. developed the MIMMR-type tunnel comprehensive inspection vehicle. The equipment fully integrated camera, light source, scanner, and geological radar while completing the tunnel three-dimensional section laser point cloud and tunnel lining image information acquisition. The maximum running speed is 30 km/h, and the maximum crack recognition accuracy can reach 0.2 mm. Huang Hongwei's team from Tongji University and Shanghai Tong Rieske Company [5,6] independently developed the MT200a tunnel crack image inspection vehicle based on a CCD camera. This device can achieve 0.2 mm crack detection accuracy at 5 km/h running speed under the actual engineering environment.

Meanwhile, crack detection technology based on digital image processing is also developing vigorously. Li et al. [7] used a CCD camera to collect pavement crack images and used the NeighShrink algorithm to detect the crack width. The maximum deviation is only 0.31 mm. H.N.D [8] and Lee [9] et al. combined the CNN image preprocessing model to develop a model for automatic pavement disease detection, the image cracked/non-cracked states for feature extraction and classification. Wu [10] and Li [11] et al. used an improved U-Net network to identify cracks in highway pavements and to improve the detection accuracy of cracks. Li [12] and Cha [13] et al. trained the Faster R-CNN model to achieve the recognition of multiple types of damages in building structures. Kim [14] et al. used an improved Mask R-CNN model to accurately recognize structural surface cracks. J.M et al. [15] developed convolutional neural networks for crack recognition and conducted several sets of experiments to validate the proposed model for crack recognition in images in the presence of interference. G.M et al. [16] developed a CNN model to identify cracks and used image processing methods to determine crack quantification. A.W.A et al. [17] proposed a new technique for the automatic identification of corrosion cracks in pipelines by incorporating deep learning algorithms and three-dimensional shading modeling (3D-SM), which improves the accuracy and reliability of crack identification. Kim et al. [18] proposed a crack identification strategy using RGB-D and a combination of high-resolution digital cameras to accurately measure cracks. Kim et al. [19] utilized a crack identification method for aging concrete bridges using a commercial drone equipped with a high-resolution vision sensor. Deep learning algorithms were used to detect cracks on the surface of the structure and to calculate their thicknesses and lengths. J.T et al. [20] conducted a comprehensive analysis of YOLO's evolution, examining the evolution from the original YOLO to the YOLOv8, YOLO-NAS, and YOLO with transformer with innovations and contributions in each iteration. Yang [21], Yu [22], and Ling [23] et al. showed the feasibility of traditional YOLOv5 in crack detection by using the traditional YOLOv5 model and the improved YOLOv5 network, respectively, in crack detection. Improved YOLOv5 has fewer parameters during training, faster speed, and more accurate results. Wang [24], Wu [25], and S.R [26] et al. used improved YOLOv8 for real-time pixel crack detection based on UAV imaging, showing that the YOLOv8 model outperforms the other base algorithms in different scenarios.

In this paper, the tunnel inspection vehicle is a mobile platform equipped with an industrial camera to test the imaging quality of the camera. The camera is equipped with a laser rangefinder and the camera shutter synchronization range; through the calculation of the measured plane and the angle of the imaging plane and correcting the pixel resolution,

the imaging characteristics of the industrial camera for the development of the corresponding image processing methods, and based on the YOLOv8 algorithm to identify the cracks in the tunnel tube sheet, output the bounding box coordinates of the crack location to calculate the minimum normal width of the crack two-point spacing. The detected crack width obtained in this study is compared with the traditional manual detection of crack width, indicating that the detection vehicle equipped with industrial camera imaging crack recognition has good accuracy; this paper based on the YOLOv8 algorithm for crack width detection is compared with the error rate of crack detection based on the SVM model, and indicates that the YOLOv8 algorithm has higher crack detection accuracy. The specific system architecture is shown in Figure 1.
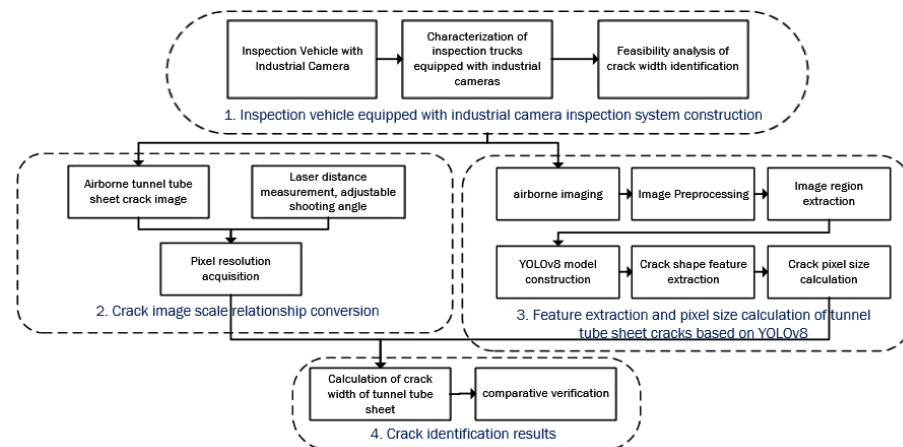

Figure 1. System architecture flowchart.

2. Characterization Test of Industrial Cameras on Inspection Vehicles

Shield Tunnel Engineering Design Standards (GB/T 51438-2021) [27] stipulate that the surface cracks of the tube sheet in the process of maintenance and use should not be greater than 0.2 mm. Tunnel inspection using industrial cameras with low resolution cannot meet the accuracy requirements. The inspection vehicle carrying a camera in the process of moving will cause relative movement between the target to be photographed and the photographic element, resulting in reduced clarity, fuzzy imaging, and affecting the imaging quality. Therefore, it is necessary to test the imaging quality of the industrial camera carried on the inspection vehicle.

2.1. Characterization Test Preparation

Figure 2 shows the relative position of the tunnel internal inspection vehicle carrying a camera to detect the tunnel cracks, shooting the lens midpoint where the vertical axis and the shooting target and the lens center point of the line between the angle of α, the lens is in the plane, and the shooting target to form a tilt angle is fixed to β, R is the radius of the tunnel, so that the lens distance from the shooting position of the object distance can be expressed as $R\cos\alpha\sec\beta$. Figure 3 shows a picture of the tunnel inspection vehicle site survey. The inspection vehicle is traveling at a fixed speed on the track and carrying a camera along the way to shoot pictures of the interior of the tunnel tube sheet.

Three marking points are set at the front of the camera lens, as shown in Figure 4. When the camera and light source are mounted on the inspection vehicle at a constant speed to photograph the cracks inside the tunnel, a measuring instrument is used to measure the line displacement at different time points in the direction of each axis of the lens marking points during the photographing process.
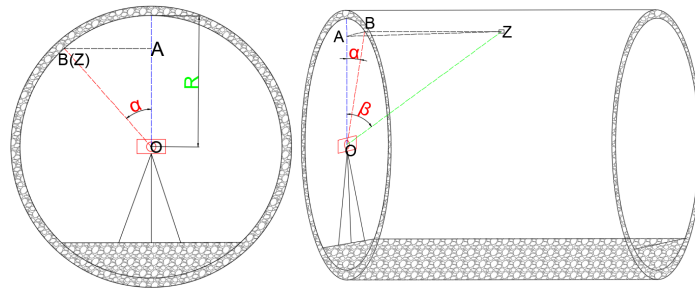
Figure 2. Schematic diagram of image acquisition by the camera. In the figure, point O indicates the center point of the camera, point Z is the point of the target position in the tunnel tube sheet, point B is the projection point of the target position on the cross-section of the tunnel tube sheet where the camera lens is located, and point A is the projection point of point B on the vertical axis of the camera lens.

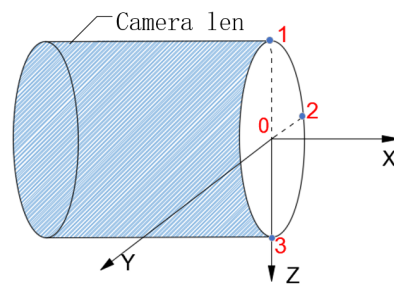**Figure 3.** Tunnel inspection vehicle site survey map.

Figure 4. Map of points marked by the camera. In the figure, point 0 is the midpoint of the camera lens, 1 is the upper vertex of the camera lens, 2 is the horizontal side point of the midpoint of the camera lens, 3 is the lower vertex of the camera lens.

2.2. Feasibility Analysis of Imaging Crack Identification

Imaging quality refers to the consistent level of intensity and chromaticity spatial distribution between image and object without considering the magnification. The optical Fourier modulation transfer function (*MTF*) in the scanning method is one of the most commonly used methods to analyze the imaging quality, and its modulation transfer function (*MTF*) expression [28,29] is as follows:

$$MTF(N) = \sin(\pi\delta N)/\pi\delta N \tag{1}$$

where δ is the image shift during the shutter time and N is the pixel point density.

In practice, different optical systems may have different requirements, so the exact range of *MTF* values may vary. In general, *MTF* values between 0.8 and 1 are considered to be high-quality imaging, providing images with sharp detail and high contrast.

In order to analyze the quality of the image, choose to shoot a fixed target, so the shooting $\alpha = 19.7°$, $\beta = 43°$. The camera uses the Teledyne Dalsa industrial camera Genie Nano. ML-5540-62M35 lens is selected, camera resolution: 4112×3008, assembled with a

36 mm $\times$ 24 mm full-frame CMOS sensor, pixel dot density: $N = 130\,\text{p}/\text{mm}$. focal length: $f = 55\,\text{mm}$.

(1) Effect of line displacement along OY and OZ axes on imaging quality.

The camera adopts a co-axial spherical optical system, the camera shooting direction is the OX coordinate axis direction, the optical system is symmetrical in the OY and OZ directions, and the laws of motion in the OY and OZ directions, and the calculation method are regarded as the same [30]. As shown in Figure 5, there is the following relation equation:

$$d = D\frac{f}{R\cos\alpha\sec\beta} \tag{2}$$

$$d' = D'\frac{f}{R\cos\alpha\sec\beta} \tag{3}$$

$$\delta = d' - d = \Delta D\frac{f}{R\cos\alpha\sec\beta} \tag{4}$$

where d is the position of the image without displacement; d' is the position of the image after camera displacement; f is the focal length of the optical system; R is the radius of the shield tunnel; D is the position of the target before displacement; D' is the relative position of the target after displacement; ΔD is the line shift of the industrial lens; and δ is the image shift produced on the photographic sensor. R, the radius of the tunnel, is 2.7 m. For profiling the imaging quality to choose a fixed shooting target, shooting $\alpha = 19.7°$, $\beta = 43°$ so the object distance is $R\cos\alpha\sec\beta = 3.4757$ m. The sampling frequency of the measuring instrument during the test is 50 Hz, the shutter speed is 1/500 s, and the focal length $f = 55$ mm. Extract 0.1 times the maximum, minimum, and average values of the line displacement data in the OY and OZ directions of the measuring instrument at the 3 and 2 marking points shown in Figure 4. The ΔD is obtained and substituted into Equation (4) to obtain the corresponding image shift δ. The relevant camera parameters and image shift are substituted into Equation (1) to obtain the modulation transfer function (*MTF*) value. Specific data are shown in Table 1 below.
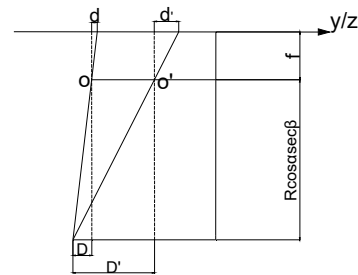

Figure 5. Vibration along the axis produces image shift OY.

Table 1. Analysis and calculation of the influence of OY, OZ axis displacement on image quality.

Items	Maximum Value	Minimum	Average
Y-line displacement (mm)	0.1215	0.0165	0.0674
Z-line displacement (mm)	0.1252	0.0152	0.0756
Resultant displacement of Y and Z direction vectors (mm)	0.1745	0.0224	0.1013
δ (microns)	2.7607	0.3545	1.6029
MTF	0.8012	0.9965	0.9301

(2) Analysis of the impact of line displacement along the OX axis on imaging quality, according to Figure 6:

$$r_1 = D\frac{f}{R\cos\alpha\sec\beta} \tag{5}$$

$$r_2 = D \frac{f}{R\cos\alpha\sec\beta - S_x} \tag{6}$$

$$\delta = \Delta r = r_2 - r_1 = \frac{r_1 S_x}{R\cos\alpha\sec\beta - S_x} \tag{7}$$

where r_1 and r_2 are the distance of the image point from the center of the CMOS full-frame sensor without displacement and after displacement, respectively; S_x is the line displacement of the camera lens along the OX axis; the maximum value of r_1 can be taken as the 1/2 diagonal physical size of the CMOS full-frame sensor, i.e., $r_1 = 21.63$ mm; the object distance is 3.4757 m; and the image shift δ. Similarly, obtain 0.1 times the maximum value, minimum value, and average value of the line displacement data in the OX direction of the 0-marker point gage shown in Figure 3 to obtain S_x. Substitute the relevant data with the camera parameters Equations (5)–(7) and (1) to obtain the modulation transfer function (*MTF*) value. Specific data are shown in Table 2 below.
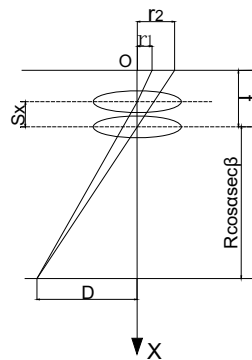

Figure 6. Image OX displacement generated by vibration along the axis.

Table 2. Analysis and calculation of the influence of OX axis displacement on image quality.

Items	Maximum Value	Minimum	Average
S_x (mm)	0.1208	0.0538	0.0762
δ (microns)	0.7518	0.3348	0.4742
MTF	0.9844	0.9969	0.9938

(3) Analysis of the effect of angular displacement of rotation around OY and OZ axes on imaging quality.

As shown in Figure 7, the rotation of the optical system around the OY and OZ axes can be regarded as the rotation of the system around the target at a small angle. Take the rotation around the OZ axis as an example. In the XOY plane, the OY axis rotates clockwise σ_Z. As shown in Figure 6, two points are taken on the OY axis, σ_Z as follows:

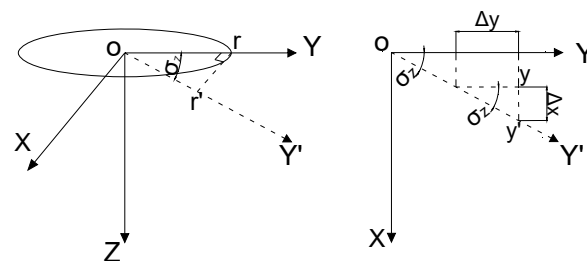$$\tan\sigma_Z = \frac{\Delta x}{\Delta y} \tag{8}$$

Figure 7. Image OZ displacement generated by rotation along the axis.

As shown in Figure 6, the angular displacement and its image shift are as follows:

$$\delta = r' - r = r(\sec \sigma_Z - 1) \tag{9}$$

where δ is the image shift; r can be taken as the physical size of the 1/2 long side of the CMOS full-frame sensor, $r = 18$ mm; Δy is the physical radius of the lens, $\Delta y = 23$ mm; and the 0, 2 on the OY axis in Figure 4 marker points in the OX direction of the line displacement difference is Δx. Similarly, extract the 4, 2 marker points shown in Figure 4 gauge in the OX direction of the line displacement difference of the maximum value, the minimum value, and 0.1 times the average value to get Δx. In the XOZ plane, as shown in Figure 4, extract the 0, 1 marker points in the OX direction of the line displacement difference of the maximum value, the minimum value, and 0.1 times the average value to get Δx. Substituting the relevant data with the camera parameters Equations (1), (8) and (9), the modulation transfer function (*MTF*) value is obtained. Specific data are shown in Table 3 below.

Table 3. Analysis and calculation of the influence of angular displacement of rotation around the OY, OZ axis on imaging quality.

Items	Maximum Value	Minimum	Average
Δx (XOY plane/mm)	0.1108	0.0596	0.0874
σ_Z (rad)	4.813×10^{-3}	2.591×10^{-3}	3.799×10^{-3}
δ_z (microns)	0.2089	0.0604	0.1299
Δx (XOZ plane/mm)	0.1095	0.0686	0.0727
σ_y (rad)	8.478×10^{-3}	2.982×10^{-3}	3.161×10^{-3}
δ_y (microns)	0.204	0.0801	0.09
After synthesis δ (µm)	0.292	0.101	0.158
MTF	0.9976	0.9997	0.9993

(4) Analysis of the effect of angular displacement of rotation around the OX axis on imaging quality.

As shown in Figure 8, in the YOZ plane, two points with a known separation distance of Δy are taken on the OY axis, and the difference of the line displacements of the two points along the OZ direction is Δz, there is

$$\tan \sigma_x = \frac{\Delta z}{\Delta y} \tag{10}$$

where δ is the image shift; r can take the 1/2 diagonal physical size of the CMOS full-frame sensor, $r = 21.63$ mm; and y is the physical radius of the lens, $\Delta y = 23$ mm. Similarly extracted as shown in Figure 4, 0, 2 marking point gauge in the OX direction of the line displacement difference between the maximum value, the minimum value, and 0.1 times the average value to get Δx. In the XOZ plane, the 1, 2 marking point shown in Figure 4 in the OZ direction of the difference between the maximum value, the minimum value, and 0.1 times the average value to get Δz. Substitute the relevant data with the camera parameters of Equation (10), Equation (11), and Equation (1) to obtain the modulation transfer function (*MTF*) value. The specific data are shown in Table 4 below. The angular displacement has its image shift as follows:

$$\delta = r\sigma_x \tag{11}$$

Tables 1–4 show that the *MTF* value of the modulation transfer function is between 0.8012 and 0.9997. According to the literature [29], in tunnel inspection vehicles equipped with an industrial camera, the rotation and translational image shift of the industrial camera is much smaller than the image element size, which will not cause the imaging quality to become fuzzy. The camera imaging quality of the tunnel inspection vehicle equipped

with an industrial camera has the reliability to satisfy the requirements of the 0.2 mm crack recognition accuracy.
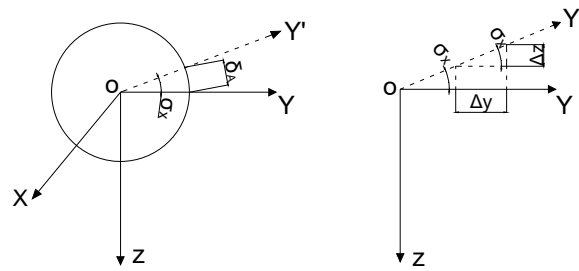

Figure 8. Image shift produced by rotation around the OX axis.

Table 4. Analysis and calculation of the influence of angular displacement around OX axis on imaging quality.

Items	Maximum Value	Minimum	Average
Δz (mm)	0.0029	0.0015	0.0024
σ_x (rad)	0.1261×10^{-3}	0.0653×10^{-3}	0.1043×10^{-3}
δ (microns)	2.7273	1.4124	2.256
MTF	0.8058	0.945	0.8644

3. Pixel Resolution Correction

As shown in Figure 9, let the XY plane be the camera imaging plane, and the Z-axis coincide with the camera centerline. The laser is AC, AD, C, and D are the reflection color spots of the laser on the measured plane, and AB is the measured object distance in the tunnel using $R\cos\alpha\sec\beta$. Point B represents the measured target corner point. Not considering the influence of laser reflection spot size. The laser AD is in the YZ plane, and the angles between the laser AC and AB and AD are, respectively, ξ and λ. The angle between AB and AD is γ. Make AC = l_2, AD l_1. Through the signal control to realize the laser rangefinder ranging and the airborne camera shutter imaging time synchronization, based on the time synchronization to determine the camera imaging and ranging information correspondence. The actual measured plane *BCD* and imaging plane cannot be completely parallel, B' is the point where point B is located on the XY imaging plane. Obtain the angle between the $B'CD$ plane and the measured plane *BCD* based on the geometric relationship between laser ranging and laser angle, θ indicates the angle between the actual measured plane and the imaging plane and corrects the pixel resolution [30].

1-Simple tunnel detection rail car; 2-Laser rangefinder; 3-Industrial camera; 4-Laser beam axis; 5-The measured plane; 6-xy imaging plane
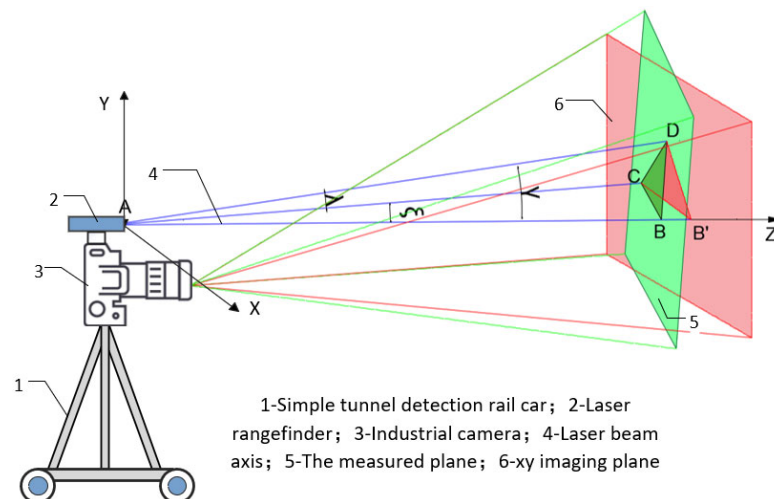
Figure 9. Schematic diagram of detecting the surface of the segment structure by the camera.

Since the plane BCD where the measured plane is an oblique plane, so we should correct it to the plane $B'CD$ perpendicular to the camera optical axis, and find out the angle between the plane BCD and the plane $B'CD$. According to the angular correction schematic shown in Figure 10, from the plane $B'CD$ and the camera optical axis plane ACD perpendicular to $\angle ACB'$ for the right angle, from the right triangle ACB' can be AB':

$$AB' = AC/\cos BAC = l_2/\cos \xi \tag{12}$$

$$B'C = \frac{AC}{\tan BAC} = \frac{l_2}{\tan \xi} \tag{13}$$
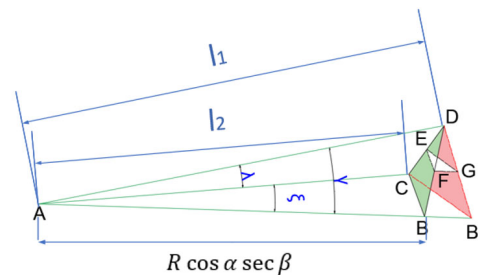
Figure 10. Schematic diagram of angle correction.

Within the triangle $AB'D$, $B'D$ can be calculated to obtain the following:

$$B'D^2 = (l_2/\cos \xi)^2 + l_1{}^2 - 2 \times l_1 \times l_2/\cos \xi \times \cos \gamma \tag{14}$$

Since it $\angle ACB'$ is a right triangle, $B'C$ can be obtained within $AB'C$:

$$B'B = AB' - AB = \frac{l_2}{\cos \xi} - R\cos \alpha \sec \beta \tag{15}$$

As shown in Figure 11, take the midpoint E on the CD and draw the perpendicular line EF and BD intersecting with F in the triangle BCD. Draw the perpendicular line EG and $B'D$ intersecting with G in the triangle $B'CD$ through E, connecting with FG. As shown in the figure below, $\angle FEG$ is the angle between the measured slant plane and the projection plane perpendicular to the camera's optical axis. Note $\angle BDC$ as $\angle a$, $\angle BDB'$ as $\angle b$, $\angle CDB'$ as $\angle c$, and $\angle FEG$ as $\angle \theta$.
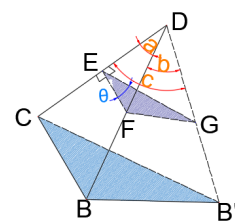

Figure 11. Schematic diagram of angle correction calculation.

cos a, EF, FD can be obtained from cosine theorem within triangle BCD:

$$\cos a = \frac{BD^2 + CD^2 - BC^2}{2 \times BD \times CD} \tag{16}$$

$$EF = \frac{1}{2} \times CD \times \tan a \tag{17}$$

$$DF = \frac{1}{2} \times \frac{CD}{\cos a} = \frac{BD \times CD^2}{BD^2 + CD^2 - BC^2} \tag{18}$$

Similarly, the cosine theorem can be used to obtain $\cos c$, EG, and GD within the triangle $B'CD$:

$$\cos c = \frac{B'D^2 + CD^2 - B'C^2}{2 \times B'D \times CD} \tag{19}$$

$$EG = \frac{1}{2} \times CD \times \tan c \tag{20}$$

$$DG = \frac{1}{2} \times \frac{CD}{\cos c} = \frac{B'D \times CD^2}{B'D^2 + CD^2 - B'C^2} \tag{21}$$

In the same way, in triangle $B'CD$, $\cos b$ and FG can be obtained from the cosine theorem:

$$\cos b = \frac{B'D^2 + BD^2 - B'B^2}{2 \times B'D \times BD} \tag{22}$$

$$FG^2 = DF^2 + DG^2 - 2 \times FD \times GD \times \cos b \tag{23}$$

Finally, the cosine theorem can be used to obtain $\cos \theta$ within the triangle FEG, is the angle between two planes

$$\cos \theta = \frac{EF^2 + EG^2 - FG^2}{2 \times EF \times EG} = \frac{\cos b - \cos a \times \cos c}{\sin a \times \sin c} \tag{24}$$

Sorting can be obtained with the following:

$$\cos \theta == \frac{d\sqrt{g}}{\sqrt{[l_2 df + efl_2]^2 + e^2 g + gd^2}} \tag{25}$$

Among them: $e = R\cos \alpha \sec \beta \cos \gamma - l_1$; $f = \frac{\cos \lambda - \cos \xi \cdot \cos \gamma}{\sin \gamma}$; $g = l_2^2(\sin^2 \xi - f^2)$; $d = R\cos \alpha \sec \beta \sin \gamma$.

In the camera imaging axial direction with a laser rangefinder and camera shutter synchronization, accurate measurement of the object distance H, that is, the distance between the rangefinder and the target can be measured according to the lens imaging principle:

$$\frac{1}{H'} + \frac{1}{H} = \frac{1}{f} \rightarrow H' = \frac{Hf}{H - f} \tag{26}$$

where H' is the image distance and f is the focal length of the lens. Let A be the actual size of the target that is, the actual physical width of the crack; Magnification ε,

$$\varepsilon = \frac{A'}{A} = \frac{H'}{H} \rightarrow A = \frac{H}{H'}A' \tag{27}$$

A' is the imaging size. Substitute Equation (12) into Equation (13) to get the following:

$$A = \frac{H - f}{f}A' \tag{28}$$

The imaging size A' is as follows:

$$A' = \frac{d}{D}A'' \tag{29}$$

where A'' is the number of imaging pixels; d is the physical size of the long side of the image sensor; and D is the number of pixels on the long side of the image sensor. The pixel resolution is as follows:

$$J = \frac{(H - f)d}{fD}\cos \theta = \frac{A}{A''}\cos \theta \tag{30}$$

J represents the actual physical size represented by the unit pixel and is the conversion factor between the actual physical size and the number of pixels. The digital image is processed to obtain the number of pixels of the measured target in the whole image, and the actual physical size of the measured target (crack) can be calculated by Equation (30).

4. Experimental Analysis of Tunnel Tube Sheet Crack Imaging

4.1. Imaging Test of Shield Segment Crack

The test is based on a large subway project, and the experiment collected image data of a typical environment with surface crack diseases in the subway tunnel section. The imaging equipment is an industrial camera mounted on the tunnel detection vehicle, the test vehicle is traveling at 10 km/h, and the lens center of the industrial camera coincides with the cross-section center of the tunnel. The camera adopts the Teledyne Dalsa industrial camera Genie Nano, with the ML-5540-62M35 lens selected. The focal length of the lens is 55 mm, the shutter speed is 1/500 s, and the object distance is about 3.5 m. The experiment is equipped with a laser rangefinder and shutter synchronous measurement, and a theodolite is used to measure the angle between the shooting target and the lens axis.

In the test process, the use of tunnel inspection trucks on the camera elevation is fixed, the light has a certain impact on the camera imaging, in the pipe sheet surface cracks to adjust the relevant parameters of the camera, if necessary, the use of flash or hand-held light to ensure imaging quality [31].

4.2. Pipe Sheet Crack Image Processing

The crack image processing program of the shield segment was compiled based on MATLAB. During the test, the crack region [32] was captured by man–machine interaction in Figure 12a–c, as shown in Figure 13, and image processing such as equalization, gray-scale transformation, filtering, and denoising was carried out in Figure 13 according to the steps shown in Figure 14. The gray linear transformation is used for the image in Figure 13, combined with the actual effect of image preprocessing, the weight of the three primary colors in the image is determined, the weighted value is taken as the final gray value, and the result of gray-scale processing is compared and enhanced to improve the image quality, enhance the gray difference between the crack characteristics and the background, and highlight the crack characteristics. Then, high–low cap transform combined filtering was used to enhance the filtering. The crack image after gray-scale transformation and filtering enhancement is shown in Figure 15. In view of the obvious gray-scale changes between the crack region and the background region of the crack image, the binarization of Figure 15 was carried out to ensure recognition accuracy. The results are shown in Figure 16.
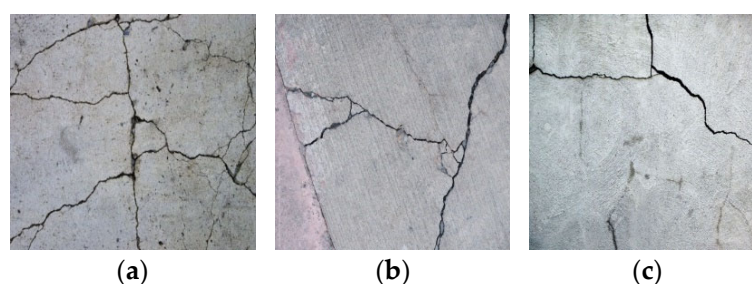

(a) (b) (c)

Figure 12. Fracture diagram of shield segment. (**a**) To take the first picture; (**b**) For the second picture taken; (**c**) For the third picture taken.
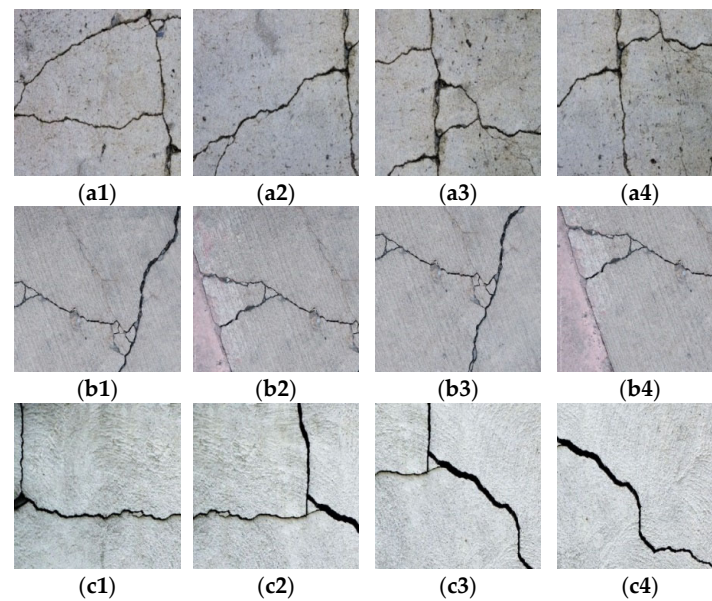
Figure 13. Crack area. The (**a1**–**a4**) group of images is the crack zoning map taken in the first photo; The (**b1**–**b4**) group of images is the crack zoning map taken in the second photo; The (**c1**–**c4**) group of images is the crack zoning map taken in the third photo.
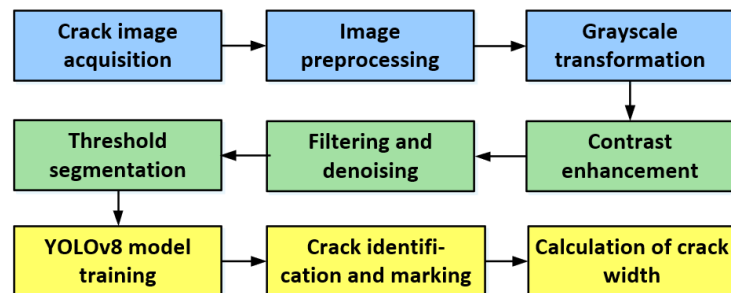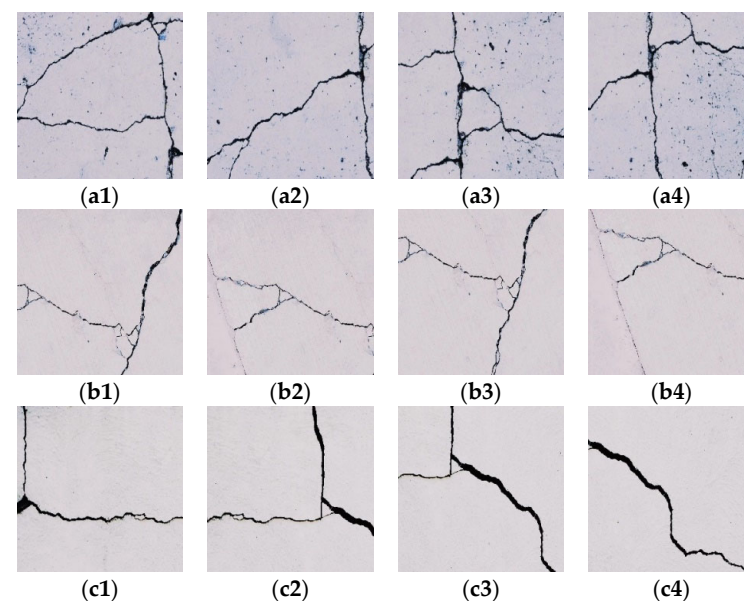

Figure 14. Image processing program.

Figure 15. Filter enhancement after gray level transformation. The (**a1**–**a4**) group of images is the crack zoning map taken in the first photo; The (**b1**–**b4**) group of images is the crack zoning map taken in the second photo; The (**c1**–**c4**) group of images is the crack zoning map taken in the third photo.
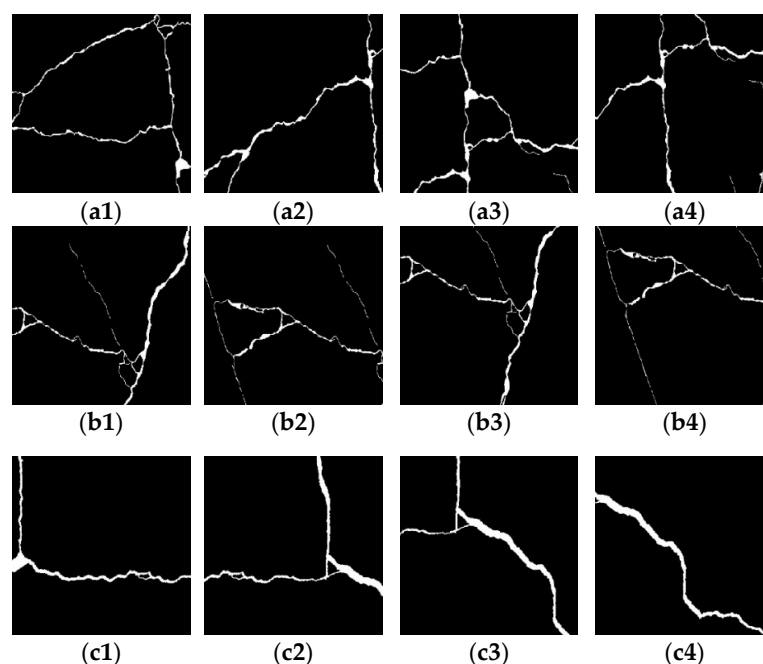
Figure 16. Binarization processing diagram. The (**a1**–**a4**) group of images is the crack zoning map taken in the first photo; The (**b1**–**b4**) group of images is the crack zoning map taken in the second photo; The (**c1**–**c4**) group of images is the crack zoning map taken in the third photo.

4.3. Pipe Sheet Crack Shape Extraction Based on YOLOv8 Modeling

The binarized crack image will be white with the crack due to the stains, holes, and spots in the tunnel tube sheet itself, so denoising of these whites, which are not related to the crack, is required. In order to identify the accurate crack information, the binarized image needs to be region-marked. The region labeling is done by marking the crack region as 1, which indicates that the region contains the crack target. The non-cracked region is labeled as 0, indicating that the region does not contain the crack target.

Compared with the C3 module of YOLOv5, the C2f module in YOLOv8 stacks more Bottleneck structures [33], which allows better utilization of features at different scales to give better performance. YOLOv8 has a high detection speed, which can be achieved in real time on GPUs, and is suitable for the detection of image targets in motion mode. Its network structure is shown in Figure 17, and the YOLOv8 model includes input, backbone, neck, and head 4 parts. Among them, the mosaic data enhancement method is chosen for input, and some hyperparameters will be modified for models of different sizes, such as large models will turn on mix up and copy paste data enhancement, which can enrich the dataset and improve the model's generalization ability and robustness. The backbone is mainly used for extracting information from the image, which is provided to the neck and head to use, which consists of multiple Conv, C2f modules, and SPPF at the tail. Conv module consists of a single Conv2d, BatchNorm2d, and activation function, which is used for extracting features and organizing the feature maps; YOLOv8 has designed the C2f structure by referring to the residual structure of the C3 module as well as the ELAN idea of YOLOv7, which can ensure a lightweight while obtaining richer gradient flow information and adjusting the number of channels according to the model scale, which greatly improves the model performance; SPPF is spatial pyramid pooling, which is able to merge features at different scales. Th neck part mainly plays the role of feature fusion, which makes full use of features extracted from the backbone network, and adopts the structure of FPN+PAN, which is able to enhance semantic expression and localization at multiple scales. The output of the head acquires the category and location information of the detected target according to the features obtained from the first two parts of processing and makes the recognition, which is replaced by the current mainstream decoupled head

structure, separating the classification and detection heads, solving the problem of different focuses of classification and localization, and also adopting the anchor-free target detection (Anchor-Free), which can improve the speed of detection. For loss calculation, in terms of the positive and negative samples, a dynamic allocation strategy is adopted, using VFL Loss as the classification loss and DFL Loss + CIOU Loss as the regression loss [34].
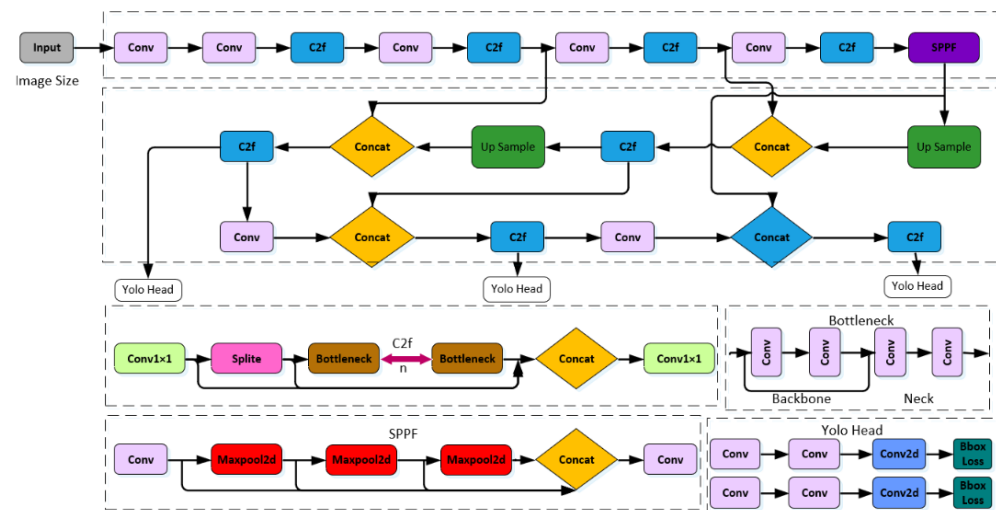

Figure 17. YOLOv8 network architecture.

In this paper, 80% of the 3630 sample set is taken as the training set and 20% as the validation set. Using Python 3.6 to write the program to complete the sample training set training, in the model training, the parameter batch size is 96, the number of iterations is 300, momentum is 0.937, weight decay value is 0.0001, the initial value of the learning rate is 0.01, and the optimizer selects the stochastic gradient descent (Stochastic Gradient Descent, SGD). The learning model is cross-validated against the validation set, and the crack recognition accuracy rate is about 97.8% in the test results, indicating that the model detection accuracy is high.

In this paper, the image shown in Figure 16 is input to the trained YOLOv8 model for crack multi-region labeling, and for the identified crack regions, different colors or markers are used to mark them in multiple regions. According to the recognized multiple regions, the regions can be merged according to the location and size of the cracks, and the obtained crack widths of each region are shown in Figure 18.
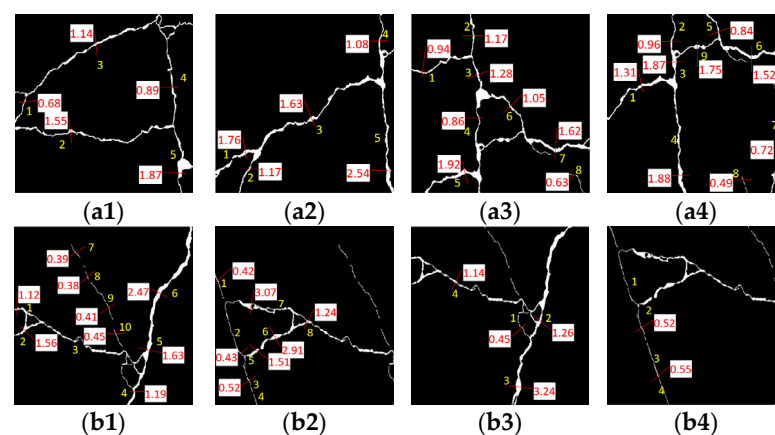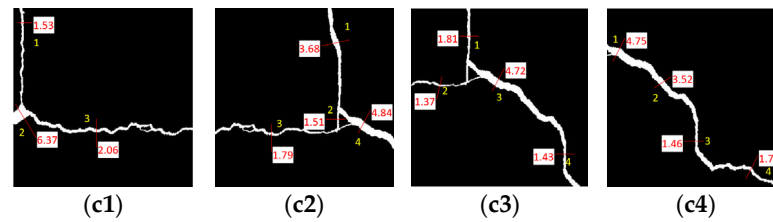

Figure 18. *Cont.*

| (**c1**) | (**c2**) | (**c3**) | (**c4**) |

Figure 18. Results of crack identification. The (**a1**–**a4**) group of images is the crack zoning map taken in the first photo; The (**b1**–**b4**) group of images is the crack zoning map taken in the second photo; The (**c1**–**c4**) group of images is the crack zoning map taken in the third photo. The numbers in the figure indicate the width of the detected cracks.

4.4. Calculation of Crack Width

Crack identification is performed using the YOLOv8 model, which will output the bounding box coordinates of the crack location. Obtaining information about the coordinates of the crack bounding box, as shown in Figure 19, is a schematic diagram of the edge of the crack shape region at the pixel level. After edge detection, the pixel value of the entire region is only 0 or 1, the resulting edge of the crack is a square pixel with a pixel value of 1 connected to the pixel. The pixel coordinates are essentially pixel centers of the pixel so that the width of the crack is calculated as the minimum distance between the lower edge of the upper edge pixel and the upper edge of the lower edge pixel [35].
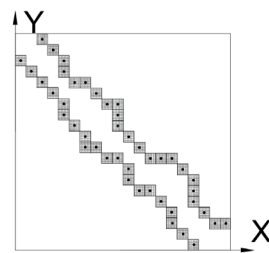

Figure 19. Schematic diagram of crack edge enlargement.

(1) Scan the entire image and mark the points with a pixel value of 1 as the upper edge coordinate points (x_n, y_n); (2) Scan the left and right 5 columns of the nth column and respectively record their lower edge points (x_{n+i}, y_{n+i}), $i = -5 \sim 5$. Using the formula between two points, calculate the distance from each of the five adjacent lower edge points on the left and right to the upper edge of the nth column as follows:

$$w = \sqrt{(x_{n+i} - x_n)^2 + (y_{n+i} - y_n)^2} \tag{31}$$

The smallest of them is taken as the actual crack width of the point, and the calculation diagram of the minimum crack width is shown in Figure 20.
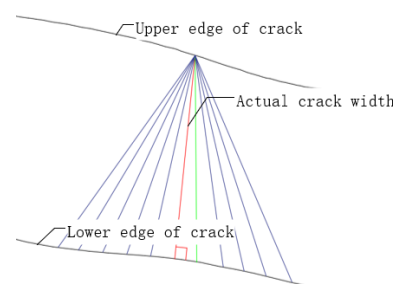

Figure 20. Schematic diagram of crack width.

5. Test and Verify the Analysis

Based on the camera imaging obtained by a large subway tunnel shield tube sheet cracks after the above image analysis method to obtain the crack morphology extraction results shown in Figure 18 map. For the convenience of calculation, the angle $\alpha_a = 55°$ and $\beta_a = 0°$ is set for (a1)–(a4) group images, $\alpha_b = 50°$ and $\beta_b = 0°$ is set for (b1)–(b4) group images, $\alpha_c = 45°$ and $\beta_c = 0°$ is set for (c1)–(c4) group images, and the distance between imaging axes and the captured images is equal to the radius of the tunnel, which is 2.7 m, i.e., $AB = 2700$ mm in Figure 8. The synchronized laser ranging of the imaging crack (a1)–(a4) group images shown is $l_1^{a} = 2742$ mm, $l_2^{a} = 2735$ mm. (b1)–(b4) group image, $l_1^{b} = 2733$ mm, $l_2^{b} = 2738$ mm. (c1)–(c4) group image, $l_1^{c} = 2745$ mm, $l_2^{c} = 2739$ mm. Imaging cracks (a1)–(a4) group image pre-set angle between lasers $\xi_a = 3°$, $\lambda_a = 3°$, $\gamma_a = 4°$ (b1)–(b4) group images settings: angle between lasers $\xi_b = 2°$, $\lambda_b = 2°$, $\gamma_b = 3°$. (c1)–(c4) group images settings: angle between lasers $\xi_c = 4°$, $\lambda_c = 3°$, $\gamma_c = 5°$. The Equation (17) can be calculated by $\cos\theta_a = 0.8148$, $\theta_a = 35.43°$; $\cos\theta_b = 0.7513$, $\theta_b = 41.29°$; $\cos\theta_c = 0.6912$, $\theta_c = 46.28°$.

Teledyne Dalsa industrial camera Genie Nano has a 36 mm × 24 mm full-frame CMOS sensor, camera resolution: 4112 × 3008, and the actual object distance is 2.7 m shield tunnel radius. When calculating, take the sensor's long side dimension $d = 36$ mm, the number of long side pixels $D = 4112$, and the focal $f = 55$ mm, $\theta_a = 35.43°$, $\theta_b = 41.29°$, $\theta_c = 46.28°$. Pixel resolution obtained by substituting into Equation (30): $J_1 = 0.343$ mm, $J_2 = 0.316$ mm, $J_3 = 0.291$ mm.

According to the corrected J, the full-domain calculation is performed in Figure 18 according to Equation (31) to extract the full-width information of the cracks, and the normal width of the cracks is the minimum value of w calculated by Equation (31). The error rate between the manual detection of the width information of the cracks and the machine detection is shown in Figure 21; the relationship between the number of pixels of the cracks and the error rate with respect to the number of pixels of the cracks is shown in Figure 22.
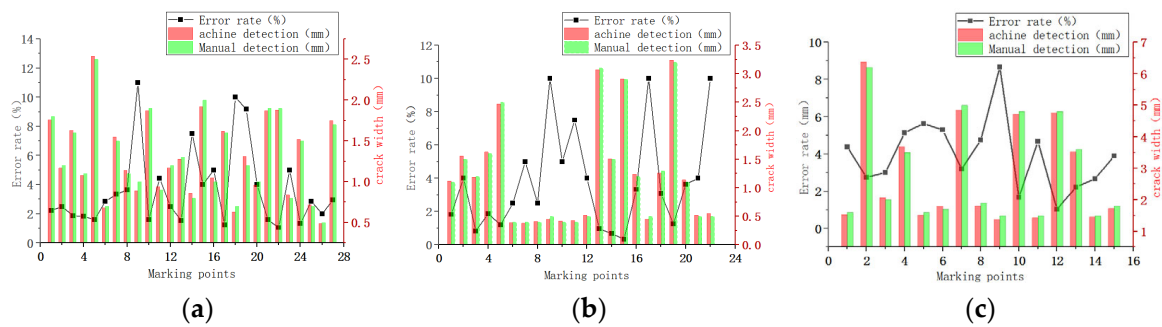

(a)   **(b)**   **(c)**

Figure 21. Comparison of crack width measurements. (**a**) The image shows a comparison of the width measurement of the first crack image; (**b**) The image shows a comparison of the width measurement of the second crack image; (**c**) The image shows a comparison of the width measurement of the third crack image.

Crack detection reports based on traditional manual inspection often fail to label all the width information of the cracks, which demonstrates the inadequacy of traditional crack width measurement. Figure 21 shows that the shield tube sheet crack detection system based on camera imaging, compared with the manual inspection data, the error of the results obtained ranges from 0% to 11%, with the average error rate remaining below 4%, and it can represent the specific crack widths in different areas of the entire through crack, indicating the feasibility of shield tube sheet crack width recognition based on the camera imaging of the tunnel inspection vehicle. According to Figure 22, the higher the number of pixels, the lower the detection error rate, and the smaller the pixel resolution correction coefficient $cos\theta$, the average value of the error rate increases, indicating that

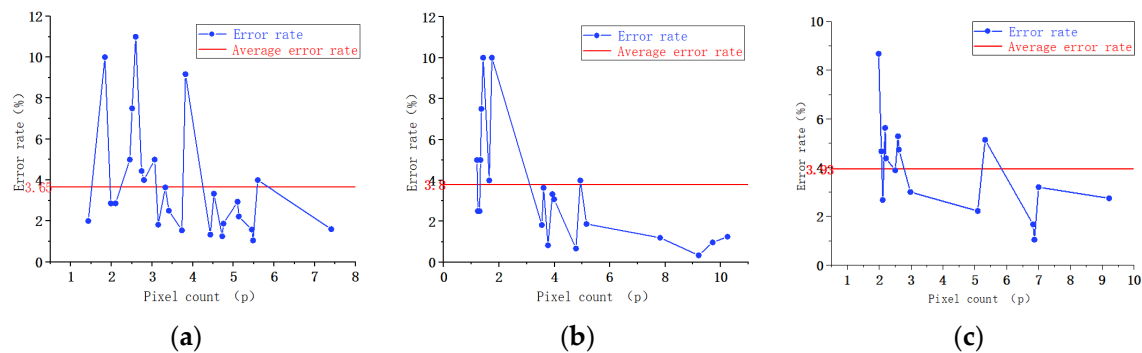the smaller the angle between the imaging plane and the measured plane, the lower the detection error rate.

(a)   **(b)**   **(c)**

Figure 22. The relative relationship between the number of crack width identification pixels and the error rate. (**a**) The figure shows the relationship between the number of pixels in the width of the first crack image and the error rate; (**b**) The figure shows the relationship between the number of pixels in the width of the second crack image and the error rate; (**c**) The figure shows the relationship between the number of pixels in the width of the third crack image and the error rate.

6. Conclusions

(1) Take the tunnel inspection vehicle as a mobile working platform equipped with an industrial camera, and the imaging quality of the industrial camera is reliable under the state of uniform motion. By adding a laser to the industrial camera and synchronizing the camera shutter to accurately measure the object distance, the three-dimensional crack image of the tunnel tube sheet is obtained, and the three-dimensional crack image is subjected to angular correction, preprocessing, and the training model for intelligent extraction of crack morphology is constructed, so as to obtain the crack width that meets the requirements of the crack width recognition accuracy of the tunnel tube sheet.

(2) By analyzing Figures 21 and 22, the error rate of crack detection based on the YOLOv8 algorithm is 0–11% compared with traditional manual detection, and the average error rate is kept below 4%. The error rate of crack detection based on the YOLOv8 algorithm is reduced by 1% compared with crack detection of the support vector machine (SVM) crack extraction model [30]. Therefore, using the tunnel inspection vehicle as a platform equipped with an industrial camera, the use of YOLOv8 to realize the recognition of the shape and width of cracks on the surface of the tunnel tube sheet has higher precision, feasibility, and a wide range of prospects.

(3) From Figures 21 and 22, it can be seen that the number of pixels is inversely related to the detection error rate, i.e., the higher the number of pixels, the lower the error rate. Therefore, when the camera parameters are certain, the larger the width of the detected crack, the higher the reliability of the crack recognition based on camera imaging; when the width of the detected crack is certain, it is recommended to choose a camera with more than 12 million pixels for processing and analysis.

(4) As can be seen from Figure 22, comparing and analyzing the error rates of crack width detection with three different pixel resolution correction coefficients $\cos\theta$, it is found that the smaller the pixel resolution correction coefficient $\cos\theta$ is, the larger the error rate of crack detection is. Therefore, the measured plane and the imaging plane are overlapped as much as possible in the actual shooting process, and the angle between the two planes is reduced to improve the crack detection accuracy.

(5) By analyzing the shooting angles α_i, β_i, and the angle θ_i between the measured plane and the imaging plane, the results show that $\alpha_i + \theta_i \approx 90°$ ($i = a$, b, c), so in the process of identifying the cracks in the tunnel tube sheet, by measuring the angle of α between the vertical axis line where the midpoint of the lens is located and the line connecting the target of the shooting with the center of the lens, we can roughly calculate the

angle θ between the imaging plane and the measured plane, which can provide validation for the correction factor of pixel resolution.

Author Contributions: Writing—original drafting and conceptualization, X.X.; drafting, editing and methodology, Q.L.; drafting, editing and formal analysis, S.L.; visualization and drafting, F.K.; methodology and supervision, G.W.; software and visualization, T.W.; validation, writing—review and editing, S.W. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: Author Shue Li, Fengyi Kang were employed by the company Nantong Highway Development Center. 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. Wang, F.W.; Feng, A.J. 2022 China Urban Rail Transit Data Statistics and Development Analysis. *Tunn. Constr.* **2023**, *43*, 521–528, rear insertion 1-rear insertion 8.
2. Li, Z.H.; Tang, C. Analysis of Intelligent Identification Algorithm for Shield Tunnel Cracks Based on High Definition Industrial Camera. *Surv. Mapp. Bull.* **2021**, *8*, 83–87+101.
3. Jiang, S.S. *Image Detection of Subway Tunnel Crack Diseases Based on Deep Learning*; East China Normal University: Shanghai, China, 2022.
4. Wang, J.H.; Jie, Q.Y.; Liu, J. Junichiro Koizumi. Current Situation of Railway Tunnel Diseases and Operation Management in Japan and Suggestions for the Development of Tunnel Operation and Maintenance Technology in China. *Tunn. Constr.* **2020**, *40*, 1824–1833.
5. Xue, Y.D.; Gao, J.; Li, Y.C.; Huang, H.W. optimization of subway tunnel lining defect detection model based on deep learning. *J. Hunan Univ. (Nat. Sci. Ed.)* **2020**, *47*, 137–146.
6. Huang, H.; Sun, Y.; Xue, Y.; Wang, F. Inspection equipment study for subway tunnel defects by grey-scale image processing. *Adv. Eng. Inform.* **2017**, *32*, 188–201. [CrossRef]
7. Li, Y.; Yang, N. An Improved Crack Identification Method for Asphalt Concrete Pavement. *Appl. Sci.* **2023**, *13*, 8696. [CrossRef]
8. Nhat-Duc, H.; Nguyen, Q.; Tran, V. Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network. *Autom. Constr.* **2018**, *94*, 203–213. [CrossRef]
9. Lee, T.; Yoon, Y.; Chun, C.; Ryu, S. CNN-Based Road-Surface Crack Detection Model That Responds to Brightness Changes. *Electronics* **2021**, *10*, 1402. [CrossRef]
10. Wu, Q.; Song, Z.; Chen, H.; Lu, Y.; Zhou, L. A Highway Pavement Crack Identification Method Based on an Improved U-Net Model. *Appl Sci.* **2023**, *13*, 7227. [CrossRef]
11. Li, P.; Xia, H.; Zhou, B.; Yan, F.; Guo, R. A Method to Improve the Accuracy of Pavement Crack Identification by Combining a Semantic Segmentation and Edge Detection Model. *Appl. Sci.* **2022**, *12*, 4714. [CrossRef]
12. Li, D.; Xie, Q.; Gong, X.; Yu, Z.; Xu, J.; Sun, Y.; Wang, J. Automatic defect detection of metro tunnel surfaces using a vision-based inspection system. *Adv. Eng. Inform.* **2021**, *47*, 101206. [CrossRef]
13. Cha, Y.J.; Choi, W.; Suh, G.; Mahmoudkhani, S.; Büyüköztürk, O. Autonomous Structural Visual Inspection Using Region-Based Deep Learning for Detecting Multiple Damage Types. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 731–747. [CrossRef]
14. Kim, B.; Cho, S. Image-based concrete crack assessment using mask and region-based convolutional neural network. *Struct. Control Health Monit.* **2019**, *26*, e2381. [CrossRef]
15. Vazquez-Nicolas, J.M.; Zamora, E.; González-Hernández, I.; Lozano, R.; Sossa, H. PD+SMC Quadrotor Control for Altitude and Crack Recognition Using Deep Learning. *Int. J. Control Autom. Syst.* **2020**, *18*, 834–844. [CrossRef]
16. Gonthina, M.; Chamata, R.; Duppalapudi, J.; Lute, V. Deep CNN-based concrete cracks identification and quantification using image processing techniques. *Asian J. Civ. Eng.* **2023**, *24*, 727–740. [CrossRef]
17. Altabey, W.A.; Noori, M.; Wang, T.; Ghiasi, R.; Kuok, S.; Wu, Z. Deep Learning-Based Crack Identification for Steel Pipelines by Extracting Features from 3D Shadow Modeling. *Appl. Sci.* **2021**, *11*, 6063. [CrossRef]
18. Kim, H.; Lee, S.; Ahn, E.; Shin, M.; Sim, S. Crack identification method for concrete structures considering angle of view using RGB-D camera-based sensor fusion. *Struct. Health Monit.* **2020**, *20*, 500–512. [CrossRef]

19. Kim, I.; Jeon, H.; Baek, S.; Hong, W.; Jung, H. Application of Crack Identification Techniques for an Aging Concrete Bridge Inspection Using an Unmanned Aerial Vehicle. *Sensors* **2018**, *18*, 1881. [CrossRef] [PubMed]
20. Terven, J.; Córdova-Esparza, D.; Romero-González, J. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Mach Learn. Know Extr.* **2023**, *5*, 1680–1716. [CrossRef]
21. Yang, H.; Yang, L.; Wu, T.; Meng, Z.; Huang, Y.; Wang, P.S.; Li, P.; Li, X. Automatic Detection of Bridge Surface Crack Using Improved YOLOv5s. *Int. J. Pattern Recognit.* **2022**, *36*, 2250047. [CrossRef]
22. Yu, G.; Zhou, X. An Improved YOLOv5 Crack Detection Method Combined with a Bottleneck Transformer. *Mathematics* **2023**, *11*, 2377. [CrossRef]
23. Ling, L.; Xue, F.; Zhang, J.; Huan, Y.; Chen, F. Intelligent detection of fine cracks on sleepers based on improved YOLOv5 model of cascade fusion. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2023**, *237*, 1273–1283. [CrossRef]
24. Wang, G.; Chen, Y.; An, P.; Hong, H.; Hu, J.; Huang, T. UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. *Sensors* **2023**, *23*, 7190. [CrossRef] [PubMed]
25. Wu, Y.; Han, Q.; Jin, Q.; Li, J.; Zhang, Y. LCA-YOLOv8-Seg: An Improved Lightweight YOLOv8-Seg for Real-Time Pixel-Level Crack Detection of Dams and Bridges. *Appl. Sci.* **2023**, *13*, 10583. [CrossRef]
26. Rahman, S.; Rony, J.H.; Uddin, J.; Samad, M.A. Real-Time Obstacle Detection with YOLOv8 in a WSN Using UAV Aerial Photography. *J. Imaging* **2023**, *9*, 216. [CrossRef] [PubMed]
27. Design Standards for Shield Tunnel Engineering. China—National Standards—State Administration for Market Regulation CN-GB. 2022. Available online: https://www.mohurd.gov.cn/gongkai/zhengce/zhengcefilelib/202110/20211025_762616.html (accessed on 1 December 2023).
28. Qian, Y.X.; Cheng, X.W.; Gao, X.D.; Liang, W.; Li, X.Y. Analysis of the Impact of Vibration on the Imaging Quality of Aviation CCD Cameras. *Electro Opt. Control* **2008**, *15*, 55–58, 66.
29. Jiang, G.Y. Collected Works on Evaluation and Inspection of Optical System; Imaging Quality Collection of Documents on Evaluation and Inspection of Imaging Quality of Optical Systems. 1988. Available online: https://xueshu.baidu.com/usercenter/paper/show?paperid=d3b96107521f449f88d0ba9030d19c93&site=xueshu_se (accessed on 1 December 2023).
30. Zhong, X.G.; Peng, X.; Shen, M. Feasibility Study on Bridge Crack Width Identification Based on Unmanned Aerial. *Veh. Imaging J. Civ. Eng.* **2019**, *52*, 52–61.
31. Pan, X.D. *Research on a Deep Learning Based System for Detecting and Evaluating Apparent Defects in Subway Tunnels*; Shandong University: Jinan, China, 2023.
32. Yin, G.S.; Gao, J.G.; Shi, M.H.; Jin, M.Z.; Tuo, H.L.; Li, C.; Zhang, B. A method for identifying tunnel cracks under image segmentation. *J. Transp. Eng.* **2022**, *22*, 148–159.
33. Geng, H.T.; Liu, Z.Y.; Jiang, J.; Fan, Z.C.; Li, J.X. An embedded road crack detection algorithm based on improved YOLOv8. *Comput. Appl.* **2023**, 1–8. Available online: http://www.joca.cn/CN/10.11772/j.issn.1001-9081.2023050635 (accessed on 1 December 2023).
34. Cheng, C.X.; Qiao, Q.Y.; Luo, X.L.; Yu, S.J. Target detection algorithm for UAV aerial images based on improved YOLOv8. *Radio Eng.* **2023**. Available online: https://link.cnki.net/urlid/13.1097.TN.20231127.1157.002 (accessed on 1 December 2023).
35. Peng, X. Research on Bridge Crack Shape and Width Recognition Based on Unmanned Aerial Vehicle Imaging. Master's Thesis, Hunan University of Science and Technology, Xiangtan, China, 2017.