

## Research Article

# Bridge Extraction Algorithm Based on Deep Learning and High-Resolution Satellite Image

Wenbing Yang,<sup>1</sup> Xiaoqi Gao,<sup>2</sup> Chunlei Zhang,<sup>2</sup> Feng Tong,<sup>3</sup> Guantian Chen,<sup>2</sup> and Zhijian Xiao <sup>4</sup>

<sup>1</sup>Yiwu Industrial and Commercial College, Yiwu, Zhejiang 322000, China

<sup>2</sup>Zenghe Packaging Co., Ltd., Wenzhou, Zhejiang 325000, China

<sup>3</sup>Wenzhou Heshun Packaging Machinery Co., Ltd., Wenzhou, Zhejiang 325000, China

<sup>4</sup>College of Digital Engineering, Zhejiang Dongfang Polytechnic, Wenzhou, Zhejiang, China

Correspondence should be addressed to Zhijian Xiao; xiaozhijian@126.com

Received 20 March 2021; Revised 23 April 2021; Accepted 7 May 2021; Published 2 June 2021

Academic Editor: Shah Nazir

Copyright © 2021 Wenbing Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a novel method of extracting roads and bridges from high-resolution remote sensing images based on deep learning. Edge detection is performed on the images in the road area along with the road skeleton line, and the result of the detected binary edge is vectorized. The interference of protective belts on both sides of the road, road vehicles, road green belts, traffic signs, etc. and the shadow interference of the bridge itself are eliminated to determine the parallel sides of the road. The bridge features on the road are used to locate the detected bridge and obtain information such as the location, length, width, and direction of the bridge, verifying the experimental results of the Shaoguan Le point images. In addition, in order to learn higher-level road feature information, the algorithm in this paper introduces the hollow convolution and multicore pooling modules. Secondly, the residual refinement network further refines the output of the prediction network to improve the ambiguity of the prediction network results. In addition, in view of the small proportion of road pixels in remote sensing images, the network also integrates binary cross entropy, structural similarity, and intersection ratio loss function to reduce road information loss. The applicability of the proposed study was tested, and the results show that the algorithm is very effective for the extraction of road and bridge targets.

## 1. Introduction

The bridge is one of the important artificial buildings. Its detection and identification are of great significance for GIS data acquisition, map update, and auxiliary supervision of bridge construction. Bridges are divided into water bridges and road bridges. At present, most of the existing algorithms are carried out for the target of bridges on the water. These algorithms mainly adopt a top-down knowledge-driven recognition method, that is, first use prior knowledge to establish a target recognition model. According to the hypothesis, segmentation, marking, and feature extraction are carried out on a purposeful basis, and then, further target detection is carried out [1–3]. These algorithms have certain reference significance for the extraction of road bridges on high-resolution images [4], but the surface conditions of the

images of bridges on the road are complex and changeable. Buildings, protective belts on both sides of the road, vehicles on the road, road green belts, traffic signs, signs, billboards, and other factors will interfere with the extraction of bridges. At present, there is no complete algorithm for bridge extraction on the road [5–7].

In order to effectively identify the road bridge target in the remote sensing image, the key is how to extract the characteristics of the road bridge in the remote sensing image [8, 9]. It is concluded that the bridge on the road mainly has the following characteristics: (1) the bridge body is suspended in the air, except for the supporting piers and cables, and the bridge body has no contact with other objects on the ground. What is shown in the image is that the bridge body and the surrounding area have a clear boundary line. Even the overpass between the upper and lower highways

has a difference in brightness due to the influence of light. (2) There is basically no big change in the gray value inside the bridge, and it has local averageness, and the gray level of the whole bridge does not change much and it is the same. (3) The shape of the bridge imaged on the image is a long rectangle, with a pair of approximately parallel sides that intersect the road but are not necessarily perpendicular to the road. (4) The bridge has a certain width on the high-resolution image, usually a few pixels. (5) Across the road. Among them, (1) and (2) are the radiation characteristics of road bridges; (3) and (4) are the geometric characteristics of on-road bridges; and (5) is the functional characteristics of on-road bridges. We make full use of these features, formulate different criteria pertinently, and propose a reasonable method of extracting bridges on the road [10].

The organization of this paper is as follows: Section 2 briefly depicts the algorithm and realization of road bridge recognition. Section 3 shows the details of the deep-learning-based method. The experiments of the proposed study are presented in Section 4. The conclusion and future work are given in section 5 of this paper.

## 2. The Algorithm and Realization of Road Bridge Recognition

In this paper, aiming at the imaging characteristics and resolution of high-resolution visible light images [11] and the main characteristics of road bridge targets, a method for extracting road bridges based on multiple information is proposed. First, a certain size of sliding window is used to track the road skeleton line for Canny edge detection, and the edge information within the road domain is detected. Then, the detected edges are vectorized to obtain a set of vector lines, the criteria are used to remove invalid line segments, and finally, the bridge parallel lines are detected, the bridge target is identified, and the parameters of the bridge are determined. The overall framework of the algorithm is shown in Figure 1.

**2.1. Edge Detection.** The edge of the image is the most basic feature of the image. The most notable feature of the bridge is a pair of parallel lines. This article attempts to analyze and determine the outer edge line of the bridge body from edge detection [12, 13]. Edges are usually detected by using first-order and second-order derivatives to calculate the abruptness of the brightness value. Sobel, Roberts, LOG, and Canny are mainstream operators for calculating the abruptness of the brightness value.

Most of the bridges on the road have low contrast with the road, and the edges are relatively weak. Compared with Sobel, Roberts [14], and LOG, Canny edge detection uses high and low thresholds to extract edges and then connects the edges in the high-threshold image into contours. When the end point of the contour is reached when connecting, the edges that can be connected on the low-threshold image are found and collected until all the gaps are connected. As shown in Figure 2, the Canny operator can accurately locate the abrupt signal of the bridge edge line. Therefore, the

Canny operator is selected as the edge detection operator in this paper.

In order to avoid the impact of buildings, vegetation, water system, etc. other than the road and because the road bridge is located within the road area, we can track the existing road vector data through the sliding window method and perform edge detection on the image within the road area. To ensure that bridges on the road can be fully detected, the length and width of the sliding window here are 1.5 times the width of the road. This not only further restricts the scope of bridge detection but also detects the edge of the bridge more quickly and accurately.

**2.2. Edge Vectorization.** In order to better analyze the geometric feature lines of the bridge, the binarized edge image can be vectorized to obtain a series of edge vector lines. For the removal of invalid line segments later, the straight line fitting of bridge lines, etc. provide edge point collections of edge line segments. The edge vectorization in this article is based on the ArcGIS Engine development environment, using the GP tool under development to vectorize the detected edge binarization image to obtain the corresponding vector file [15].

**2.3. Invalid Line Segment Removal.** The types of image features in the road area are more complicated. After edge extraction, a large number of invalid edge lines that interfere with the identification of bridges on the road will be obtained. Therefore, those invalid line segments that affect bridge detection must be removed before detecting bridge parallel lines. Specific interference factors include vehicles, trees, traffic signs, signs, billboards, railings on individual bridges, and shadows of the bridge itself.

**2.3.1. Interference Removal Caused by Factors such as Vehicles, Trees, Traffic Signs, and Billboards.** In the image after edge extraction, the vehicles and trees on the road mostly correspond to relatively small discontinuities and relatively short discontinuous edge lines, and the edges are mostly rough and the local curvature is larger. According to this feature, in order to avoid the impact of vehicles, road green belts, traffic signs, billboards, etc. on bridge recognition, this paper formulates guidelines to remove those very short or relatively large curvature edge lines. The specific algorithm guidelines are as follows:

Rule 1: all the edge lines are recorded as set  $N$ , each edge vector line of the edge vector file is traversed, and the curve length  $l_i$  of the edge vector line is calculated. Suppose the length threshold in the rule of removing line segments is  $l_{min}$ , then the set after removing the too short broken lines is  $T = \{t_i | t_i \in N, l_i > l_{min}\}$ .

Criterion 2: for the edge line in the set  $T$ , the distance between the end points of the line segment is  $d_i$ , the curve length of the line segment is  $l_i$ , then the curve degree parameter  $c_i = d_i/l_i$  of the line segment is set, and  $c_{min}$  is the curvature threshold. Then, most of the

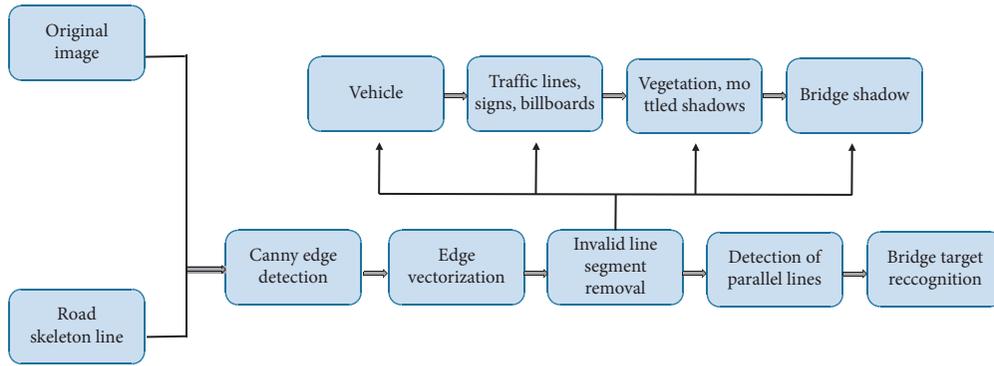


FIGURE 1: Overall framework of the bridge to identify.

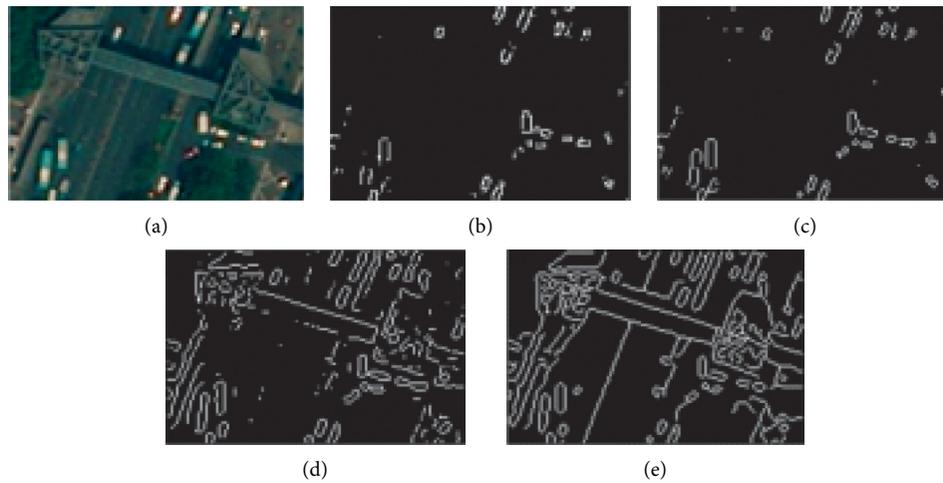


FIGURE 2: Result of edge detection. (a) Original image. (b) Roberts. (c) Sobel. (d) Log. (e) Canny.

invalid edge lines corresponding to the larger curvature of the vehicles on the road are removed, and the set  $T' = \{ti | ti \in T, li > lmin, ci > cmin\}$  is obtained.

2.3.2. Removal of Interference Caused by Traffic Signs.

The types of traffic signs on the road include guide lane lines, lane dividing lines, center double solid lines, sidewalk lines, and stop lines. Among them, the guide lane line, the lane dividing line, the center double solid line, and the sidewalk line are imaged on the high-resolution remote sensing image. After edge detection, the edge line formed is almost parallel to the road centerline. In the image, the bridge and the road center line intersect. Figure 3(a) shows a schematic diagram of a traffic sign line, and Figure 3(b) shows it on a remote sensing image with a resolution of 0.2 m. The spatial relationship between the traffic sign line and the road centerline is approximately parallel and does not intersect, and the bridge edge line is close to the road centerline. Therefore, in order to obtain more reliable results before identification, the article adopts criterion 3 to restrict.

Criterion 3: due to the imaging method of the central projection, the bridge and the road across it cannot be parallel. By setting the angle between the sideline of the bridge and the centerline of the road to be greater than a

certain threshold  $\alpha$ , after experiments, setting the threshold  $\alpha = 30^\circ$  can remove most of the virtual scene information caused by the traffic sign line. The interference caused by the stop line needs to be further eliminated by the following parallel line detection criterion.

2.4. Interference Removal of the Shadow of the Bridge Itself.

For shadows on remote sensing images, it is mainly caused by insufficient illumination, which leads to low brightness in shadow areas and lack of color information [7].

After the Canny edge detection, the shadows of the bridge and itself often form three approximately parallel edge lines. Based on the intersection of the line in the middle and the center line of the road, a certain range of data is taken on both sides to take the average. The side with the smaller mean value is judged to be the shadow part, and the corresponding edge line is regarded as the shadow edge line and deleted to remove the shadow interference of the bridge itself. The schematic diagram of shadow interference removal is shown in Figure 4.

2.4.1. Interference Removal Caused by Factors such as Bridge Railings.

If there is a bridge on the road, there must be intersections between the center line of the road and the edge

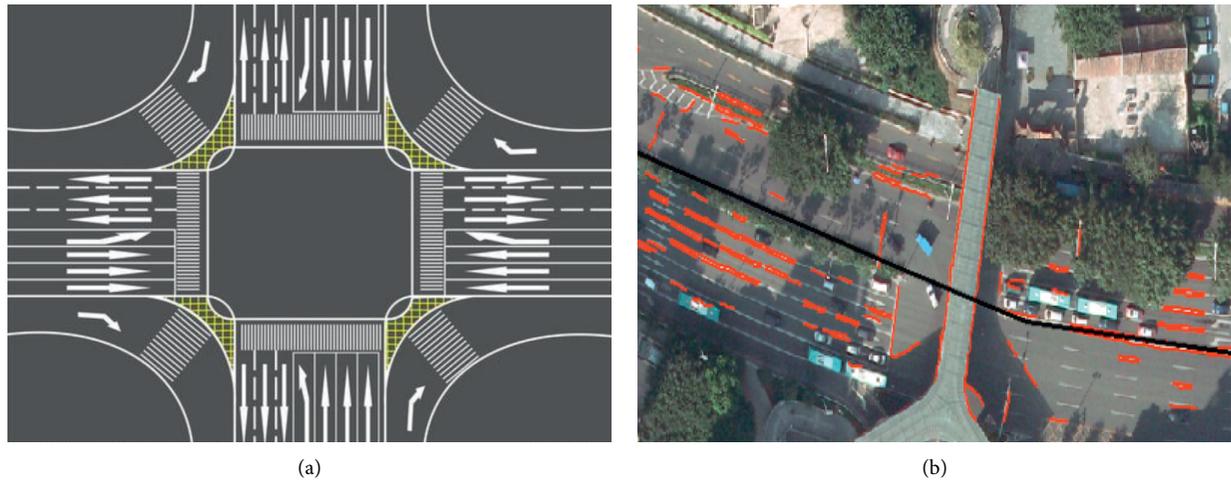


FIGURE 3: Traffic signs line. (a) Schematic diagram of traffic lines. (b) Spatial relationship among traffic sign lines, bridge edges, and road center lines on remote sensing images.

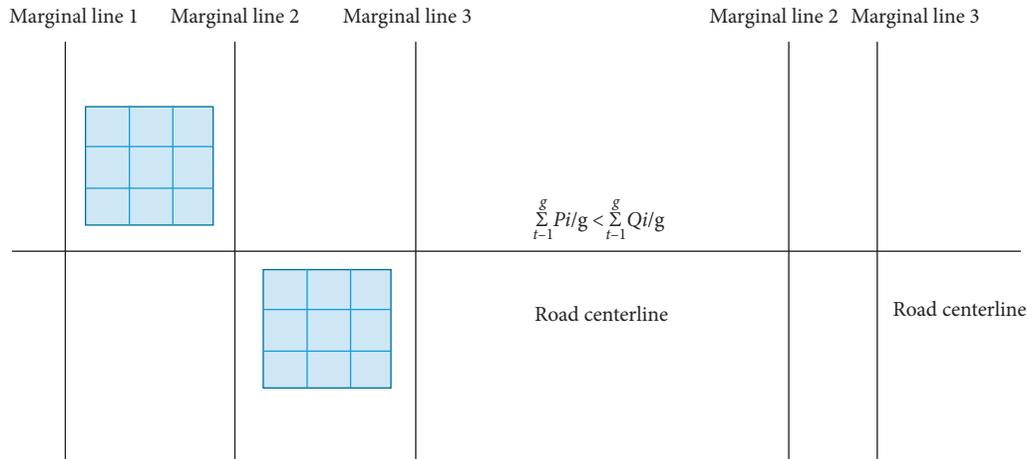


FIGURE 4: Schematic of bridge shadow interference removal.

of the bridge [3], and these intersections appear in pairs under normal circumstances. However, due to the angle between the normal of the sun and the angle of incidence, the image of the railing of the bridge may fall within the range of the bridge, and the intersection point with the center line of the road will also be recorded. Therefore, two or more points may be recorded at and near a certain edge of the bridge when detecting along the center line. Therefore, redundant points need to be deleted to avoid interference caused by factors such as bridge railings. The processing method is to traverse all the intersections, and if there is an intersection within the 8 neighborhood of another point, one of them is recorded [3].

**2.5. Detect Parallel Edges.** Through the deletion of invalid line segments, we obtain potential bridge parallel line pairs, but there may still be virtual scenes in these lines, so the obtained edge lines need to be further confirmed before

identification. Because the two edges of the bridge are not ideal parallel lines and the bridge has a certain width, the virtual scene can be further eliminated by extracting approximately parallel edge line segments within a certain range [6], as shown in Figure 5.

The line segments AB and CD are the suspected edge line segment pairs of the bridge. A1 and B1 are the projections of AB on CD. C1 and D1 are the projections of CD on AB, respectively. d1–d4 are the distances from the end of the straight line to another straight line. The angle between the two straight lines is  $\theta$ .

Criterion 4: (1) the bridge has a certain width, so d1–d4 is less than a specified threshold; (2) the intersection of the detected bridge edge and the road centerline is taken as the starting point, and the road width is used as the limit to track the edge line of the bridge, record the coordinates of the tracking point, and perform least-squares fitting on it to obtain the slope of the edge. Because the bridge has two

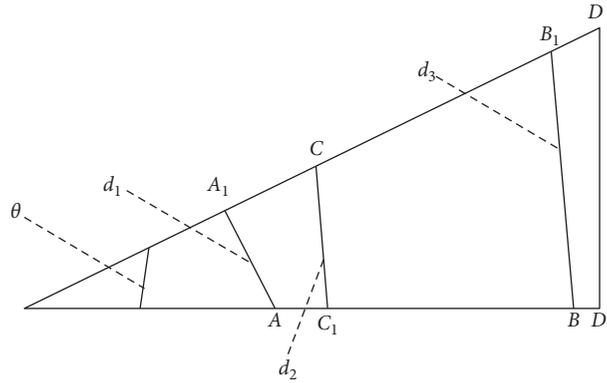


FIGURE 5: Criteria for detection of parallel lines.

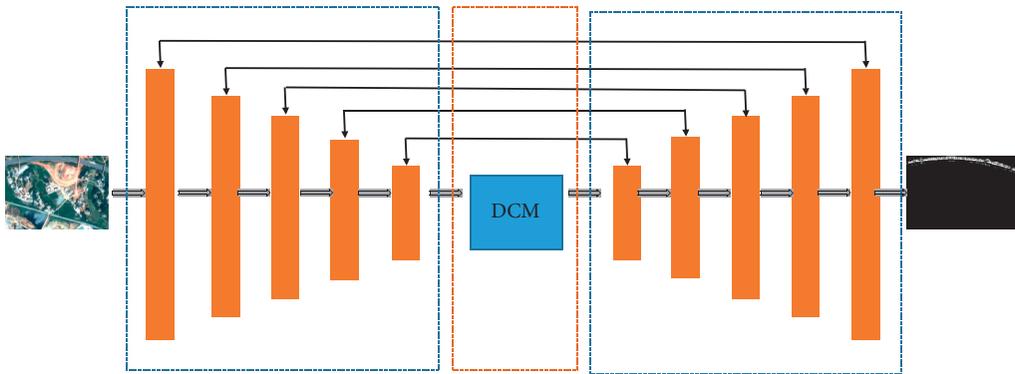


FIGURE 6: Encoder-decoder structure overall frame diagram.

sides, two slope values can be obtained. According to the angle formula, the angle  $\theta$  corresponding to the two sides can be obtained. If  $\theta$  is less than a certain threshold, it is considered as a pair of sides of the bridge.

**2.6. Bridge Recognition.** After the invalid line segments are removed, the edges of the bridge can be paired. Starting from the recorded first edge line, two adjacent lines can determine a bridge [16], and the intersection points with the center line of the road are  $B1(x1, y1)$  and  $B2(x2, y2)$ .

- (1) Determination of the location of the bridge:

$$B_c(x, y) = \left( \frac{(x_1 + x_2)}{2}, \frac{(y_1 + y_2)}{2} \right), \quad (1)$$

- (2) Determination of the bridge position: the azimuth angle of the bridge sideline pair is averaged as the bridge position
- (3) Determination of the width of the bridge: after obtaining the slope of the side line of the bridge, a straight line equation of two parallel lines is constructed according to the midpoint of the two side lines of the bridge, and the distance between the two lines is the width of the bridge
- (4) Determination of the length of the bridge: the calculation principle of the length of the bridge is to take

the position of the bridge as the starting point and grow on the image between the two bridge edges according to the consistency of the area

### 3. Deep-Learning-Based Method

This paper also uses deep-learning methods [16, 17] to assist road and bridge extraction tasks. In the image segmentation task, the more common structure is the encoder-decoder structure. The encoder is responsible for downsampling and reducing the image dimensions, and the decoder is responsible for upsampling and reducing the image dimensions. As shown in Figure 6, the method proposed in this paper chooses to load the ResNet-34 pretraining model. The depth of the entire network is designed as five layers, including one input convolutional layer and four encoder layers. When using the Massachusetts dataset for training, the image size of the input into the network is  $500 \times 500 \times 3$ ; then, it will undergo a bilinear interpolation operation to interpolate the image size to  $512 \times 512 \times 3$  in the input convolutional layer, and the feature map of the size of  $512 \times 512 \times 3$  that will pass the step size is 2, the convolution kernel is a  $7 \times 7$  convolution process, and the output feature map size at this time is  $2_{56} \times 2_{56} \times 64$ ; then, through the maximum pooling process with a convolution kernel of  $2 \times 2$  and a step size of 2, the feature map obtained the size  $128 \times 128 \times 64$ , and then, it is used as the input of the first-

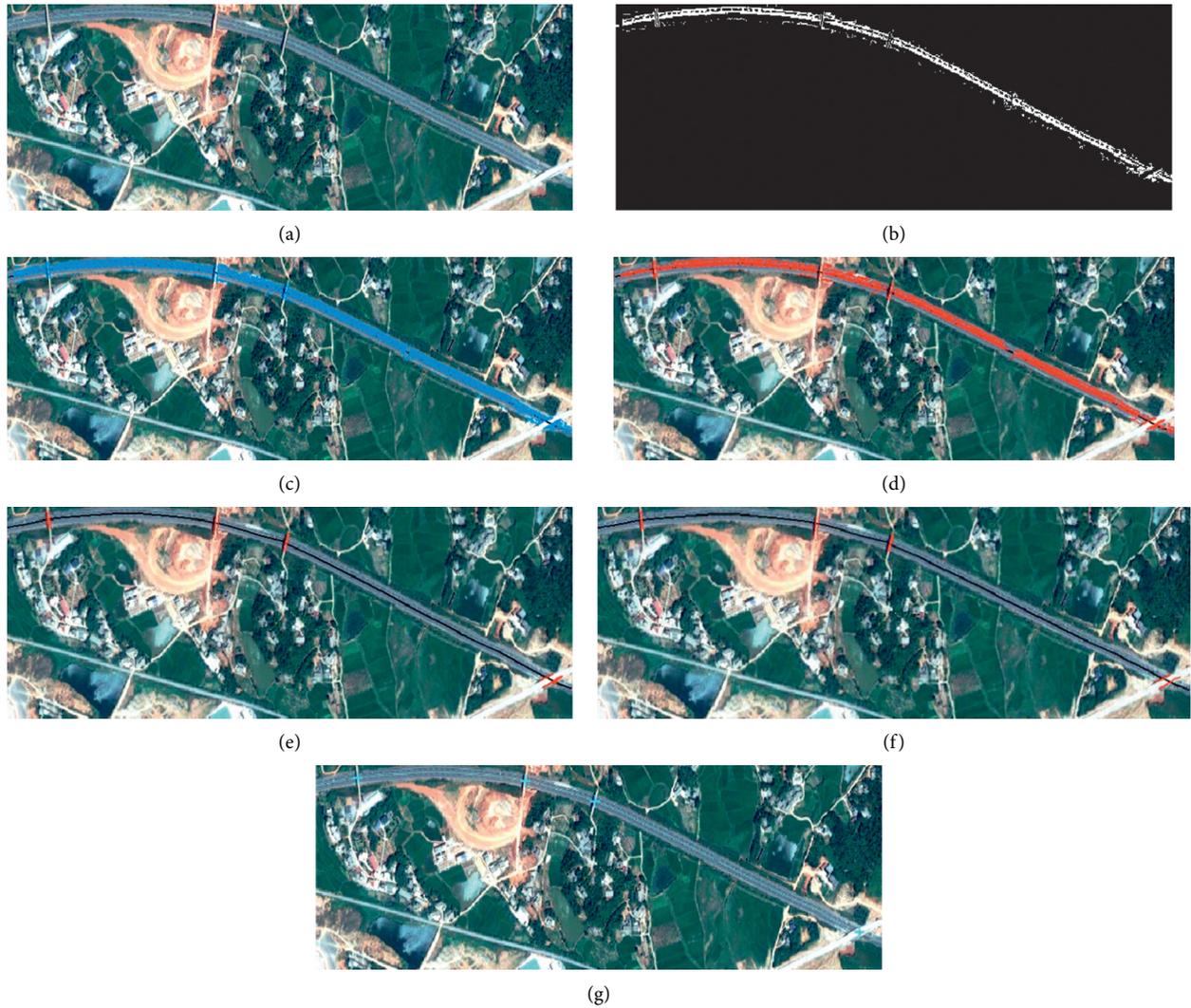


FIGURE 7: Original and extraction results of the experimental zone. (a) Original image. (b) Binary image after edge detection in the road area. (c) Edge vectorization results within the road area. (d) Result after removing interference from vehicles and road green belts. (e) The result of removing interference from traffic sign lines. (f) The result after removing the shadow. (g) Bridge extraction results.

layer decoder. In the abovementioned four-layer encoder layer, each layer of encoder is composed of a different number of residual blocks, and the residual block includes two  $3 \times 3$  convolution kernels and a combined jump branch. The output generated by the four-layer encoder through the downsampling convolution operation is  $128 \times 128 \times 64$ ,  $64 \times 64 \times 128$ ,  $32 \times 32 \times 256$ , and  $16 \times 16 \times 512$ , and the output of the last encoder layer will be used as the input of the transition layer. Figure 6 represents the encoder-decoder structure overall frame diagram.

#### 4. Experiment

The recognition method proposed in this paper is implemented in the VS2008 + ArcGIS Engine environment. Figure 7 shows the recognition results of part of the experimental area of Pleiades remote sensing image in Shaoguan.

Figure 7 shows part of the recognition results and intermediate results. Figure 7(a) is the original image; Figure 7(b) is the binarized image after the edge is extracted along the center line of the road using the Canny operator; Figure 7(c) is the edge raster image converted into vector lines by ArcGIS Engine, for the subsequent removal of invalid line segments and rapid topological calculation for detecting parallel lines; Figure 7(d) is the removal of interference factors such as vehicles and green belts, the threshold of curvature is set to 0.9, and the threshold of length is set to 4 pixels, and good results are obtained; Figure 7(e) is the removal of other interference factors such as traffic sign lines, and the threshold  $a$  of the angle with the road centerline is  $30^\circ$ ; Figure 7(f) is the result of removing the shadow interference of the bridge itself, and further removing the virtual scene by detecting the parallel lines, the bridge linear edge pair is basically recognized [16]. According to that the pixel value of the bridge shadow is

TABLE 1: Bridge parameter identification results of the experimental zone (from left to right).

Serial number	Bridge center position (latitude and longitude)	Bridge length (m)	Bridge width (m)	Bridge orientation (°) (angle with true north)
1	(112.557886, 27.907211)	39.63964	4.951638	-7.569795
2	(112.562436, 27.907155)	39.63964	8.250913	9.733668
3	(112.564332, 27.906753)	39.63964	5.392012	17.764025
4	(112.571607, 27.904112)	39.63964	9.117343	69.819042

smaller than the pixel value of the bridge itself, the shadow effect is eliminated. When detecting parallel lines, the threshold of the angle  $\theta$  between the two sides is set to  $5^\circ$ , and the distance threshold is set to 40 m; Figure 7(g) is the final recognition result, and the light blue cross represents the bridge target. Table 1 shows the position, orientation, length, width, and other parameters of the bridge detection results from left to right in Figure 7(g), and it can be seen that bridges on the road can be better identified. It can be seen from Table 1 that the position, orientation, length, and width of the bridge can be obtained relatively correctly using the algorithm of this article.

## 5. Conclusions and Further Work

This paper proposes a set of high-resolution remote sensing image road bridge extraction technologies based on multiple information. Based on the main characteristics of the bridge, the algorithm establishes a criterion method for bridge recognition, using the road area as the recognition area and detecting along the center line of the road. Limiting the detection range to the vicinity of the center line of the road is conducive to the improvement of detection efficiency and results; the vectorization of the edge after Canny edge detection is beneficial to the deletion calculation of the virtual scene line, and the bridge edge line is quickly and accurately determined. The identification of bridges on the road provides direct and rapid dynamic updates for the map and provides auxiliary reference for municipal planning.

The algorithm in this paper performs detection along the center line of the road, which depends on the accuracy of the road network information provided. If the road network is not accurately extracted, it will lead to misdetection and missed detection of bridge extraction; in addition, the detection algorithm in this paper is aimed at parallel bridges on the road, but it cannot be extracted correctly for ring-shaped and irregular bridges. This requires us to conduct further research.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding this paper.

## References

- [1] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 4, pp. 612–616, 2018.
- [2] J. Nie, S. Qu, Y. Wei, L. Zhang, and L. Deng, "An infrared small target detection method based on multiscale local homogeneity measure," *Infrared Physics & Technology*, vol. 90, pp. 186–194, 2018.
- [3] Y. Shi, Y. Wei, H. Yao, D. Pan, and G. Xiao, "High-boost-based multiscale local contrast measure for infrared small target detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 1, pp. 33–37, 2017.
- [4] H. Zhao, X. Qi, X. Shen, J. Shi, and J. Jia, "Icnet for real-time semantic segmentation on high-resolution images," in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 405–420, Ithaca, NY, USA, August 2018.
- [5] C. Tan, N. Uddin, E. J. O'Brien, P. J. McGetrick, and C. W. Kim, "Extraction of bridge modal parameters using passing vehicle response," *Journal of Bridge Engineering*, vol. 24, no. 9, Article ID 04019087, 2019.
- [6] T. Sun, Z. Di, and Y. Wang, "Combining satellite imagery and gps data for road extraction," in *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, pp. 29–32, New York, NY, USA, November 2018.
- [7] D. Wang, X. Meng, C. Gao, S. Pan, and Q. Chen, "Multipath extraction and mitigation for bridge deformation monitoring using a single-difference model," *Advances in Space Research*, vol. 60, no. 12, pp. 2882–2895, 2017.
- [8] X. Zhang, Y. Yang, Z. Li, X. Ning, Y. Qin, and W. Cai, "An improved encoder-decoder network based on strip pool method applied to segmentation of farmland vacancy field," *Entropy*, vol. 23, no. 4, p. 435, 2021.
- [9] W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools and Applications*, vol. 80, no. 7, pp. 11291–11312, 2021.
- [10] Z. Shi and N. Uddin, "Extracting multiple bridge frequencies from test vehicle—a theoretical study," *Journal of Sound and Vibration*, vol. 490, Article ID 115735, 2021.
- [11] A. Lichtenegger, D. J. Harper, M. Augustin et al., "Spectroscopic imaging with spectral domain visible light optical coherence microscopy in Alzheimer's disease brain samples," *Biomedical Optics Express*, vol. 8, no. 9, pp. 4007–4025, 2017.
- [12] Y. Liu, M. M. Cheng, X. Hu, K. Wang, and X. Bai, "Richer convolutional features for edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 8, pp. 3000–3009, 2017.
- [13] P. Ganesan and G. Sajiv, "A comprehensive study of edge detection for image processing applications," in *Proceedings of the 2017 International Conference on Innovations in*

- Information, Embedded and Communication Systems (ICI-IECS)*, pp. 1–6, IEEE, Coimbatore, India, March 2017.
- [14] H. A. Sholeh, Y. Mulyani, and H. D. Saptama, “Studi perbandingan pengenalan karakter aksara lampung dengan metode deteksi tepi roberts dan sobel,” *Inovasi Pembangunan: Jurnal Kelitbangan*, vol. 6, no. 3, pp. 261–272, 2018.
- [15] B. Dahlberg and M. Versen, “Loop transformation algorithm for test vector accessing at high speed,” in *Proceedings of the 45th International Symposium for Testing and Failure Analysis (ISTFA)*, p. 434, ASM International, Chardon, Ohio, December 2019.
- [16] X. Ning, W. Li, and W. Liu, “A fast single image haze removal method based on human retina property,” *IEICE Transactions on Information and Systems*, vol. E100.D, no. 1, pp. 211–214, 2017.
- [17] X. Ning, X. Wang, S. Xu et al., “A review of research on co-training, concurrency and computation: practice and experience,” *Concurrency and Computation Practice and Experience*, p. e6276, 2021, in Press.