International Journal of Innovative Computing, Information and Control
Volume **18**, Number **3**, June **2022**

ICIC International ©2022 ISSN 1349-4198
pp. **851–866**

AN IMPROVED WATERSHED ALGORITHM ON MULTI-DIRECTIONAL EDGE DETECTION FOR ROAD EXTRACTION IN REMOTE IMAGES

Kang Lv[1], Weixing Wang[2,*], Zhen Zhou[3,*] and Xueying Wang[4]

[1]Teacher Education and Training Department
Henan Finance University
No. 76, Zhengkai Avenue, Zhengzhou 451464, P. R. China
lvkang@hafu.edu.cn

[2]School of Information Engineering
Chang'an University
Middle-section of Nan'er Huan Road, Xi'an 710064, P. R. China
*Corresponding author: wxwang@chd.edu.cn

[3]School of Information Technology
Luoyang Normal University
No. 6, Jiqing Road, Yibin District, Luoyang 471934, P. R. China
*Corresponding author: zhouzhen2001@126.com

[4]Quality Monitoring and Appraisal Station for Traffic Construction Project
of Inner Mongolia Autonomous Region
Harbin Institute of Technology
No. 92, West Dazhi Street, Nangang District, Harbin 150001, P. R. China
343693555@qq.com

Received August 2021; revised February 2022

Abstract. *Road detection on aerial and remote sensing vague images is a hard task. In this study, an automatic road detection algorithm is proposed. The algorithm utilizes the multi-directional edge detection operators to obtain a gradient magnitude image, in which, an eight-neighborhood template is considered as a processing unit in that a threshold based on entropy is used for all the pixels in all directions to obtain the final gradient magnitude image; then, the gradient magnitude image is roughly segmented on the principle of Watershed algorithm; and finally, a number of morphological functions are utilized to finalize the roads. In experiments, a number of aerial and remote sensing road images in a public dataset were selected for testing, by comparing to several traditional algorithms and semantic segmentation methods, the testing results show that the studied algorithm is satisfactory for rural road images, and the road detection accuracy can be up to 93.2%, and the recall rate is up to 89.3%.*

Keywords: Remote sensing image, Road detection, Gradient magnitude, Watershed, Adaptive thresholding

1. **Introduction.** It is important to retrieve the earth surface information from aerial images. High spatial resolution aerial images provide the more accurate information, which offers the advantages over urban planning, geometrics, military monitoring, object extraction, changes monitoring and GIS data updating. The fast and accurate useful information extraction from aerial images has become a hot research topic, and the road detection in vague aerial images is one of the hardest tasks in this research field.

DOI: 10.24507/ijicic.18.03.851

In the past three decades, a lot of efforts have been made for developing road detection algorithms. Many research institutions have done a lot of research work in this area, and the research content includes the traditional image processing and segmentation algorithms [1] and the recent semantic segmentation methods such as deep learning method. No matter what kind of algorithms and methods is applied, the automatic road detection is still difficult for aerial and remote sensing images. A road is a linear object which is similar to the pavement crack in the road construction, since the pavement surface is rough and illumination is uneven sometimes, the studied algorithms for crack detection are similar to that for road detection. For the road construction, a lot of new algorithms, such as the algorithms based on level set and hydrodynamics, have been studied [2-4], which can also be applied for road detection.

In general, the surface of the road is uniform with similar colors in certain length of road segment, which is good for road detection in some environments. As traditional classification, the road detection algorithms can be divided into two types: semi-automatic and full-automatic algorithms. At present, there is no breakthrough with the full-automatic algorithms for aerial and remote sensing images, but some successes on the semi-automatic ones have been achieved [5,6]. In 2010, [7] presented an algorithm for road-tracing, it is valuable for normal road detection, but it might fail in some road images. After then, many algorithms were studied based on spectral or textures [8,9], such as the algorithm for image segmentation based on a multi-resolution application of a wavelet transform and feature distribution [10]; and a number of the recent road detection algorithms based on color information [11], but not for the images under extreme environment that make image vague much [12,13]. In addition to the above various image information, [14] preliminarily used road shape features and other road features to detect road. It is good for the binary images in some cases, but it is lacking of versatility.

For a complex aerial or remote sensing image, before road detection, an image should be enhanced for sharpening roads and smoothing image for reducing noise, which is needed indeed. One of examples for road detection in a vague aerial image is to use fractional differential to sharpen the road edges [15], which is useful for the image with many textures in a blur image, but it might be not effective for very dark and vague road images. For the poor quality remote sensing images, some defog and dehaze algorithms may be useful, but this kind of algorithms [16,17] is not discussed in this paper. Anyhow, the new method/ algorithm study for automatic road detection in a vaguer aerial image is important.

In recent years, different semantic segmentation methods have been applied into road detection. Abdollahi et al. made a review for deep learning used into remote sensing datasets of road detection [18]. They did a systematic survey for deep learning techniques utilized in remote sensing benchmarks for road extraction. The review mainly includes four kinds of deep learning methods, namely, the FCNs, deconvolutional networks, the GANs model and patch-based CNNs models. They also compared different deep learning models for remote sensing datasets to show which one performs satisfactorily in detecting roads from high-resolution images. Moreover, they described the future research directions and research contents.

To detect roads automatically and accurately in a complex aerial and remote sensing image, and to overcome the deficiencies of the previous algorithms for vaguer aerial images, a new method is proposed in this paper. The new method includes different algorithms, and the main algorithms are improved Watershed, multi-directional edge detection, and adaptive thresholding. The main procedure is that for the vague aerial image, it firstly does image enhancement based on an oriented edge detection, then makes road detection roughly on the improved Watershed, and subsequently accurately finalizes road by applying

the road shape features. The studied method is compared to about 10 different traditional image processing algorithms and several semantic segmentation methods.

2. **Methodology.** The method includes three main algorithms. The main algorithms are 1) multi-directional gradient magnitude operator; 2) adaptive thresholding algorithm; and 3) improved Watershed algorithm. The details of the algorithms are described as the following sections respectively.

2.1. **Multi-directional gradient magnitude operator.** The traditional method of obtaining gradient image is to use the difference algorithm to obtain approximate derivative of gray level image. The equation is shown in Equation (1).

$$G(x,y) = |f(x,y) - f(x+1,y)| + |f(x,y) - f(x,y+1)| \tag{1}$$

where $f(x,y)$ and $G(x,y)$ are the gray level image and its gradient magnitude image, respectively.

The process for obtaining the gradient magnitude image is to convert the color image into gray level image, then to use Equation (1) to calculate the gradients, and finally to get the result (gradient magnitude image). The traditional algorithm for obtaining gradient magnitude image can only process a single gray level image but cannot directly process the color image, which will lose some information; because the edge direction of the image is changeable, the difference operation should be rotation invariant. Hence, the first-order operators are generally applied to deriving images. The Sobel operator is widely utilized because of its simple template, fast running speed and sensitive to edge information. However, the template has only two directions, which can only make calculation in the horizontal and vertical directions, and the center of the template cannot be aligned with the pixel, resulting in the inaccurate position of the edge, resulting in misjudgment and lack of information.

The depth field image is obtained to enhance the road information and part of the noise. Therefore, in the process for getting the gradient magnitude image, the noise should be denoising, and the enhanced road information should not be lost. Thereby, this paper uses the improved multi-directional Sobel operator to obtain the gradient magnitude image, because the Sobel operator itself uses the weighted sum method to smooth the noise. By increasing the number of templates from two to eight and changing the direction from only two to eight, the gradient information is collected from eight directions, so that the image information will not be lost easily and the positioning is more accurate.

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}$$

0° **45°** **90°** **135°**

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$

180° **225°** **270°** **315°**

FIGURE 1. Improved direction templates

As shown in Figure 1, the direction number of the template is increased from two to eight, which ensures the rotation invariance during the operation, and the gradient magnitude value obtained by convolution of the image from multiple directions is more accurate. Assume that the gray level image is $f(x, y)$, and the eight directional gradient magnitude value in pixel (x, y) can be calculated respectively as $F1, F2, \ldots, F8$:

$$\begin{aligned} F1 = {} & f(x+1, y-1) + 2 * f(x+1, y) + f(x+1, y+1) - f(x-1, y-1) \\ & - 2 * f(x-1, y) - f(x-1, y+1) \end{aligned} \tag{2}$$

$$\begin{aligned} F2 = {} & f(x+1, y) + 2f(x+1, y+1) + f(x, y+1) - f(x, y-1) \\ & - 2 * f(x-1, y-1) - f(x-1, y) \end{aligned} \tag{3}$$

$$\cdots$$

$$\begin{aligned} F7 = {} & f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1) - f(x-1, y+1) \\ & - 2 * f(x, y+1) - f(x+1, y+1) \end{aligned} \tag{4}$$

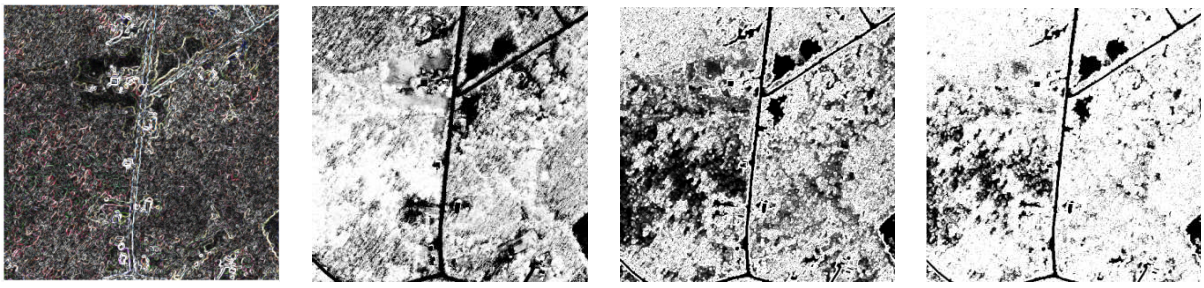
$$\begin{aligned} F8 = {} & f(x, y-1) + 2f(x+1, y-1) + f(x+1, y) - f(x-1, y) \\ & - 2 * f(x-1, y+1) - f(x, y+1) \end{aligned} \tag{5}$$

Finally, the gradient image is obtained by using Equation (6).

$$G(x, y) = \sqrt{\sum_{i=1}^{8} Fi^2} \tag{6}$$

where $G(x, y)$ is the gradient image, and Fi $(i \in (1, 8))$ is for convolution results in different directions. After the modular operation, the information of multiple directions is integrated to get a complete gradient image $G(x, y)$, which cannot only smooth the image, but also further enhance the interested target information.

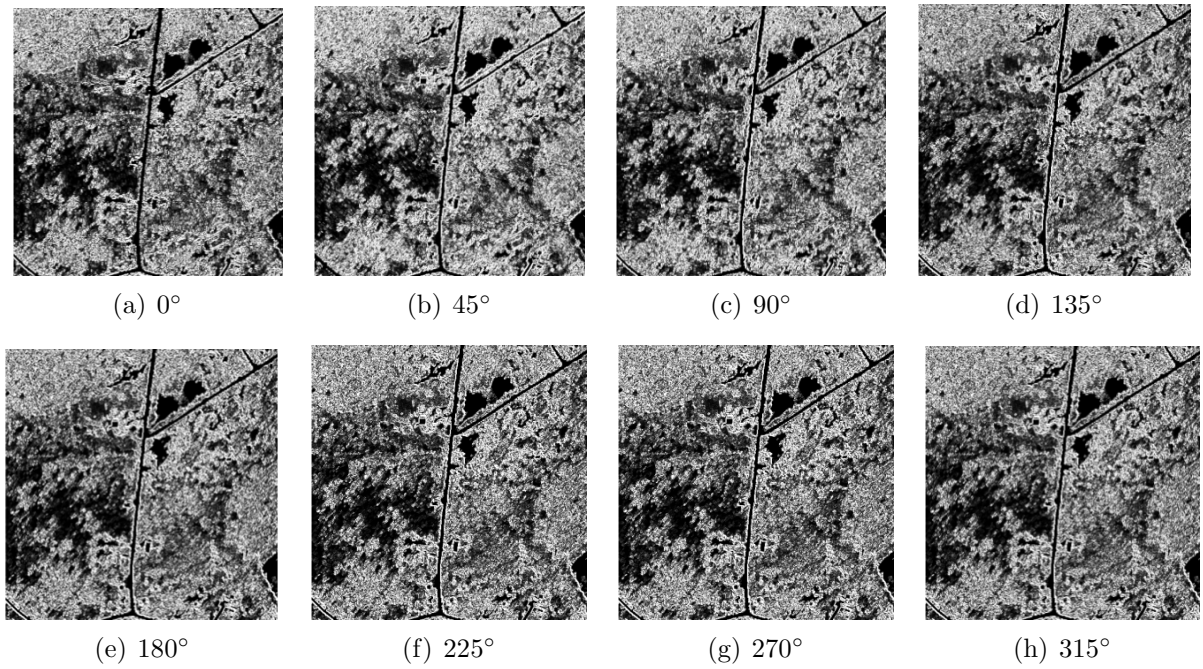
The result of obtaining the gradient magnitude image by using the improved multi-directional gradient operator is shown in Figure 2 and the multiple directional gradient magnitude results are shown in Figure 3, where the comparison between the images is synthesized by eight directional information and the gradient magnitude images are obtained by the traditional Sobel operator.



(a) Sobel result (b) Depth image (c) Traditional gradient (d) Improved gradient

FIGURE 2. Gradient magnitude results by different algorithms

In Figure 2 and Figure 3, it can be seen that the gradient magnitude image obtained by the improved multi-directional Sobel operator is clearer and the edges are more complete. It shows that some information is lost in the convolution result of each direction, but the image information is complemented by the modulus operation, so that the road information in the image is preserved completely. Moreover, because the Sobel operator itself is sensitive to the edge information, the edge information of the road in the gradient



(a) 0° (b) 45° (c) 90° (d) 135°

(e) 180° (f) 225° (g) 270° (h) 315°

FIGURE 3. Results by multi-directional gradient magnitude operators

magnitude image is more prominent, and the part of the noise in the image is also filtered out, which tells us that the multi-directional gradient magnitude operator can make the Watershed algorithm get better segmentation effect.

For the eight Sobel edge images, for each pixel, find out its maximum gradient magnitude value as the value in a new gradient magnitude image in the same position. In this way, the final gradient magnitude image can be got.

2.2. **Adaptive thresholding algorithm.** Because the small change of the pixel gray level in the same region of the image will be detected by the Watershed algorithm, producing in the incomplete regions in the image, resulting in the over-segmentation phenomenon, the gradient magnitude image is thresholded to remove the enhanced noise and pseudo minimum points generated by non-road objects, to extract the foreground objects, and to improve the image segmentation results. So the selection of the threshold is the key, using adaptive threshold algorithm can produce different thresholds in different images, it can reduce the influence of human factors, because manual threshold setting needs prior knowledge, the extraction results need a lot of statistics and calculation, and the robustness of the algorithm is not good, and will produce large errors. The new adaptive threshold algorithm is a two-dimensional maximum entropy method, and its calculation principle is in [19,20].

Machine learning is playing a major role in understanding complex human activity patterns related to the problems of human activity recognition (HAR). In this paper, we propose a machine learning-based method that works in two phases. In the first phase, the prediction probabilities are combined by soft voting. Later, the method is periodically capable of training itself. All the simulations are performed using R machine learning and a statistical platform. All the experiments are performed on the Dell Precision M4800 workstation with Intel Core i7, which has eight cores and 16 GB RAM. The multiple cores are used for parallel processing to speed up the training of the models.

Suppose that the gradient magnitude image I has L gray levels, the total number of pixels is N, and the frequency of each gray level is expressed as $f_1, f_2, f_3, \ldots, f_L$ respectively, in I, for the pixels with gray level i, if its neighboring pixel gray level is j, which can be expressed as f_{ij}, assume the histogram of I is $h(i,j) = p_{ij}$, and the level range of i and j is $0 \le i \le L-1$, $0 \le j \le L-1$. Then the histogram of the two-dimensional gradient magnitude image can be expressed by

$$p_{ij} = \frac{f_{ij}}{N}, \quad p_{ij} \to \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} p_{ij} = 1 \tag{7}$$

For the thresholding range (s,t), s is the threshold of i, and t is the threshold of j. Object and background accumulative probabilities p_1 and p_2 are respectively

$$p_1 = \sum_{i=0}^{s} \sum_{j=0}^{t} p_{ij}, \quad p_2 = \sum_{i=s+1}^{L-1} \sum_{j=t+1}^{L-1} p_{ij}$$

The two-dimensional entropy of object and background of image I is defined as

$$H_1 = -\sum_{i=0}^{s} \sum_{j=0}^{t} p_{ij} \lg p_{ij} \tag{8}$$

$$H_2 = -\sum_{i=s+1}^{L-1} \sum_{j=t+1}^{L-1} p_{ij} \lg p_{ij} \tag{9}$$

The total entropy of the gradient magnitude image is defined as

$$H(s,t) = H_1 + H_2 = \lg \left(p_1(1-p_1) + \frac{H_1}{p_1} + \frac{H_L - H_1}{1-p_1} \right) \tag{10}$$

where $H_L = -\sum_{i=0}^{L} \sum_{j=0}^{L} p_{ij} \lg p_{ij}$. The final threshold $H(*\_s, *\_t)$ satisfies the following conditions:

$$H(*\_s, *\_t) = \max\{H(s,t)\} \tag{11}$$

Based on the above threshold, the gradient magnitude image is thresholded, and the pixel's value less than the threshold will be set as 0.

2.3. **Improved Watershed algorithm.**

2.3.1. *Traditional Watershed.* Watershed is a segmentation method based on topological theory and mathematical morphology, and many researchers have made the modification and improvement for this kind of algorithms [21-24]. Its idea is to treat the gray level of each point on the image as height to form terrain. The minimum value and the area effects are called catchment basin. Image segmentation is realized by simulating immersion process [21,22]. The immersion process is simulated, that is, if there is water pushing in or out, the image will be divided into different areas, and the water will flow to the catchment basin. When the basin is full of water, the edge of each basin can be clearly seen, which is called Watershed.

According to the simulated immersion process, the Watershed formation model can be divided into two types: top-down, and the simulated precipitation model. The simulated precipitation model is that when the rain falls on the surface of the terrain, although it passes through different paths, it will eventually flow into the same place. Combining all the paths together is a complete catchment basin, and the boundary formed at this time is the Watershed line, which may be the dam of a pool [23].

Whether it is to simulate precipitation or flood, it highlights an important property of the Watershed algorithm, that is, the Watershed lines obtained by the Watershed algorithm are connected, and the boundaries between regions are continuous. Finally, the set of closed segmentation regions is obtained. The purpose of the Watershed segmentation algorithm is to segment and extract the pixels in the image according to the nature of similarity. The pixels in each region are similar, and there is a minimum point in each region. The Watershed segmentation algorithm is very sensitive to the subtle gray level changes in the image; therefore, the edges of the objects can be accurately located in the image [22]. The results of the traditional Watershed algorithm are shown in Figure 4.

(a) Original image (b) Watershed (c) Smooth + Watershed

FIGURE 4. Effect of traditional Watershed method

Although Watershed segmentation algorithm has many advantages, the effect of traditional Watershed segmentation algorithm for road detection is not ideal. Because it has the strong sensitivity to small gray level changes, it usually acts on the gradient image, but there are gray level changes in the flat area of the gradient image, which will also be responded to produce the over segmentation phenomenon. A 3D road image simulated by gray levels is shown in Figure 5.

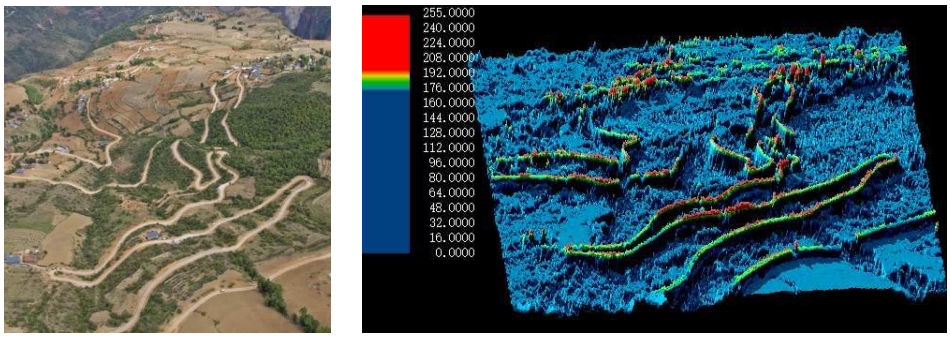


FIGURE 5. Road image and its pixel relief map

In Figure 5, the image topology and geomorphology map formed after the height represented by gray levels can be seen that there are many catchment basins in the image. After image segmentation with the Watershed algorithm, many segmented areas will be generated, and the target cannot be effectively extracted. The result of Figure 4 shows that there are too many minima to extract the effective road area, which may result in over segmentation. Therefore, many researchers have put forward many solutions to solve the problem of image segmentation by Watershed, which can be roughly divided into image filtering, modified gradient image and region merging.

2.3.2. *Improved Watershed.* Based on the above analysis and consideration, we improved the Watershed algorithm in the following procedure:

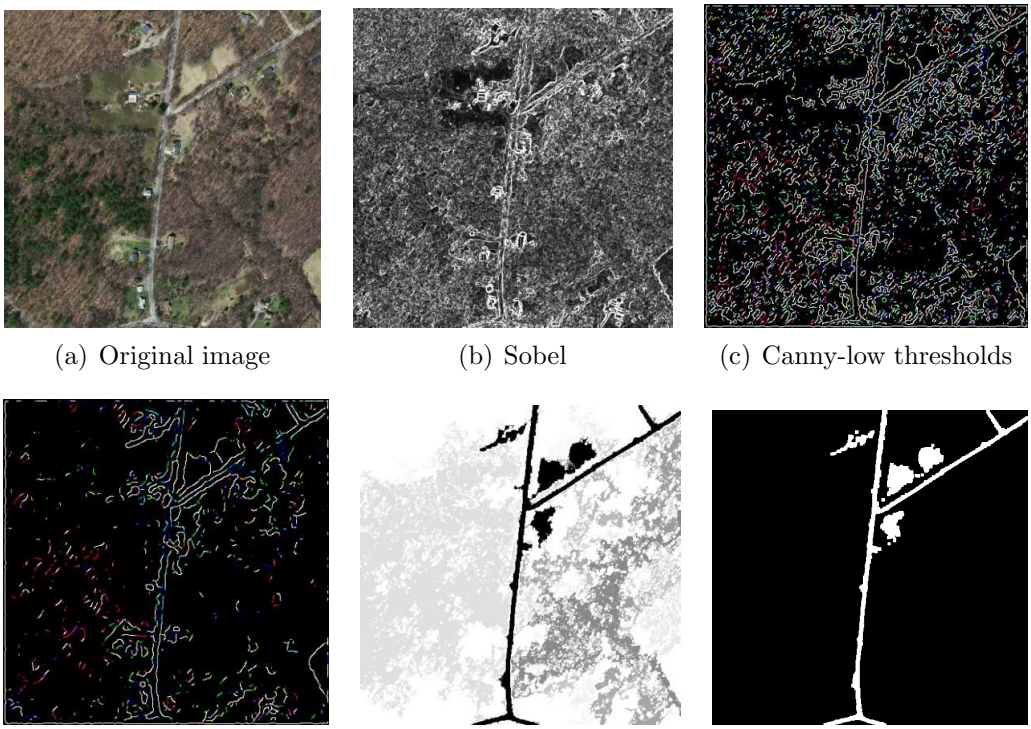
1) Firstly, the original image is enhanced by using a Gaussian filter;

2) The gradient magnitude image is obtained by using the multi-directional gradient magnitude operator to enhance the road information and filter out the image noise;

3) Adaptive thresholding based on maximum entropy is applied on the gradient magnitude image to removing un-necessary pixels;

4) Watershed is operated on the thresholded gradient magnitude image to get the output.

Figure 6(a) shows the original image of rural road; Figures 6(b)-6(d) show the Sobel, Canny edge detector with low thresholds and Canny edge detector with high thresholds, respectively; Figure 6(e) presents the new edge detection result; and Figure 6(f) shows the binarization result of 6(e).



(a) Original image (b) Sobel (c) Canny-low thresholds

(d) Canny-high thresholds (e) New gradient (f) Binarizing

FIGURE 6. Comparison results by different edge detection algorithms: (a) Original image; (b) Sobel; (c) Canny-low thresholds; (d) Canny-high thresholds; (e) new algorithm; and (f) binarizing for (e)

In Figure 6, by comparing different edge detectors, it can be seen that the edge information of the enhanced gradient magnitude image $f(x, y)$ by the new algorithm is more obvious, and the image is brighter and richer in detail. However, the pseudo minimum value in the gradient magnitude image needs to be removed. Therefore, the adaptive threshold algorithm is used to obtain the gradient magnitude image threshold H. The regions larger than the threshold are retained and marked, and the parts smaller than the threshold are discarded to obtain the marked image $mark(x, y)$. After labeling the gradient magnitude image $f(x, y)$ by $mark(x, y)$, $f_{mark}(x, y)$ is obtained. In $f_{mark}(x, y)$, only the pixels whose gray level is not 0 in $mark(x, y)$ have the minimum value, and the pixels in other positions no longer have the minimum value. One of the modified gradient magnitude images is shown in Figure 6(e).

For the post-processing of segmentation results, it can be seen in Figure 7 that the improved Watershed algorithm has extracted the road main body information, but there are still non road areas in the road. After analysis, it is found that the reason for this part of the area is that the characteristics of the open space in the non-road area in the rural road scene are similar to those of the road area, and are enhanced by the depth field image; some of the houses on both sides of the road are mistakenly identified as the road area after enhanced. Although most of the noise is removed when the adaptive threshold is applied, this part of the regions is also enhanced. Because the intensity of this part of the regions is greater than that in most of the regions, when the threshold is utilized, it will be classified as the road region, resulting in the segmentation error.

(a) Binary image (b) Edges on (a) (c) Original + (a) (d) Original + (b)

FIGURE 7. Optimization results on the images Figure 6(a) and Figure 6(f)

Hence, this part of the region can be removed by a morphological algorithm, because the region itself is not connected, the algorithm uses the template of 3×3 structure operator for morphological reconstruction to make the edge more smooth, de-burring, de-adhesion, and then calculates the area of each region to remove these separated small areas of non-road regions.

3. Results and Analyses.

3.1. **Source of experimental dataset.** The dataset of this experiment uses the Massachusetts dataset, and selects the images which are divided into rural road scenes for processing. The selected images are identified as road scenes, and the standard is whether the manually labeled images in the dataset contain road information.

3.2. **Evaluation criteria.** In the field of image processing, there are two kinds of image evaluation criteria. One is the qualitative analysis of visual contrast; and the second type is the quantitative analysis of indicators based on standards [25,26].

The visual contrast qualitative analysis is to compare the results of test image according to different processing algorithms. The algorithm used in this study is visual contrast qualitative analysis. The quantitative analysis is to evaluate the performance of the image processing results by applying the computer-aided detection system. The calculation algorithm for comparison is based on the similarity between the image extraction results and the truth map, to evaluate whether the algorithm model is effective or not. According to this standard, the area where the road is located is set as positive, and the background area is set as negative. Hence, there are the following situations in the evaluation indicators.

In this study, the performance of the algorithm is evaluated quantitatively by three indexes: Precision (P), Recall (R) and F1-Score (F1). They are defined as

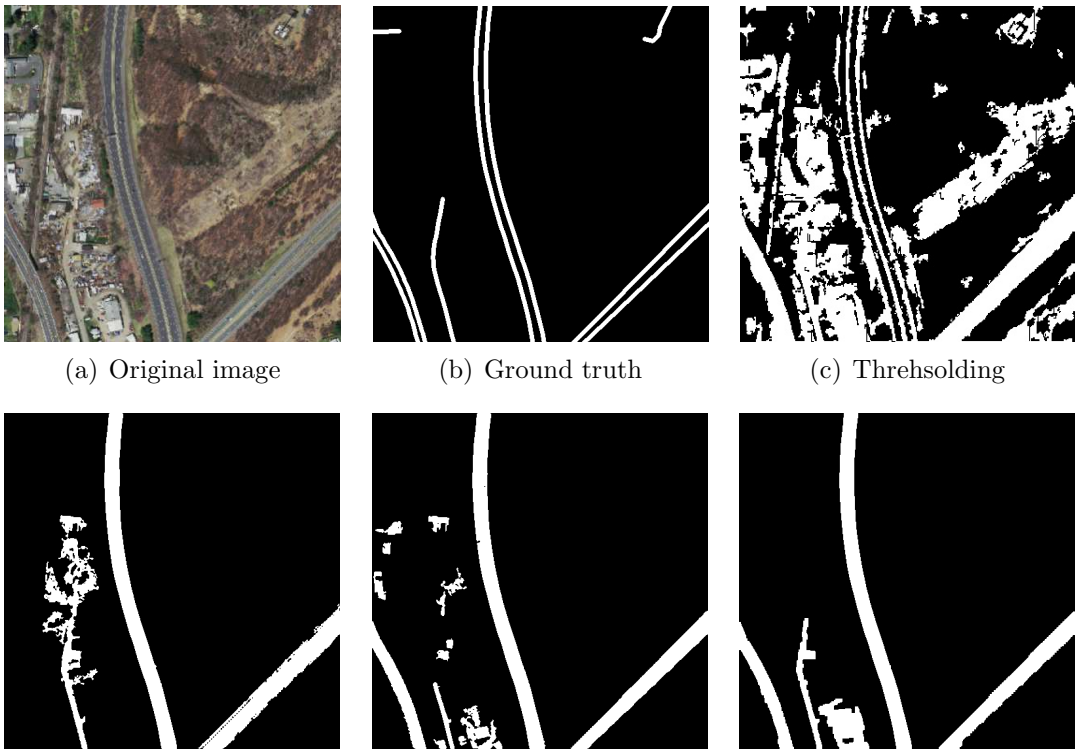
$$P = TP/(TP + FP) \tag{12}$$

$$R = TP/(TP + FN) \tag{13}$$

$$F1 = \left[\left(1 + a^2\right) PR\right] / \left[a^2(P + R)\right] \tag{14}$$

P represents how many pixels in the detected road area are accurate, and R is recall, which represents how many pixels of road are detected in the original image. $F1$ is the harmonic mean value of P and R, which is used to evaluate the overall performance of the algorithm. Generally, the value of a is set to 1, that is $F1$, and its maximum value is also 1, the larger the result is, the better the performance of the algorithm model is.

3.3. **Result analysis by comparing traditional image processing algorithms.** There are two kinds of reference algorithms in the comparative experiments in this study. One is the improved Watershed algorithm combined with region growing proposed in [27]. The other is the object-oriented method suggested in [28] to optimize the Watershed algorithm. Figure 8 shows the image processing results of different algorithms. In order to better compare with manual marking and traditional methods/algorithms obviously, the following examples are presented.



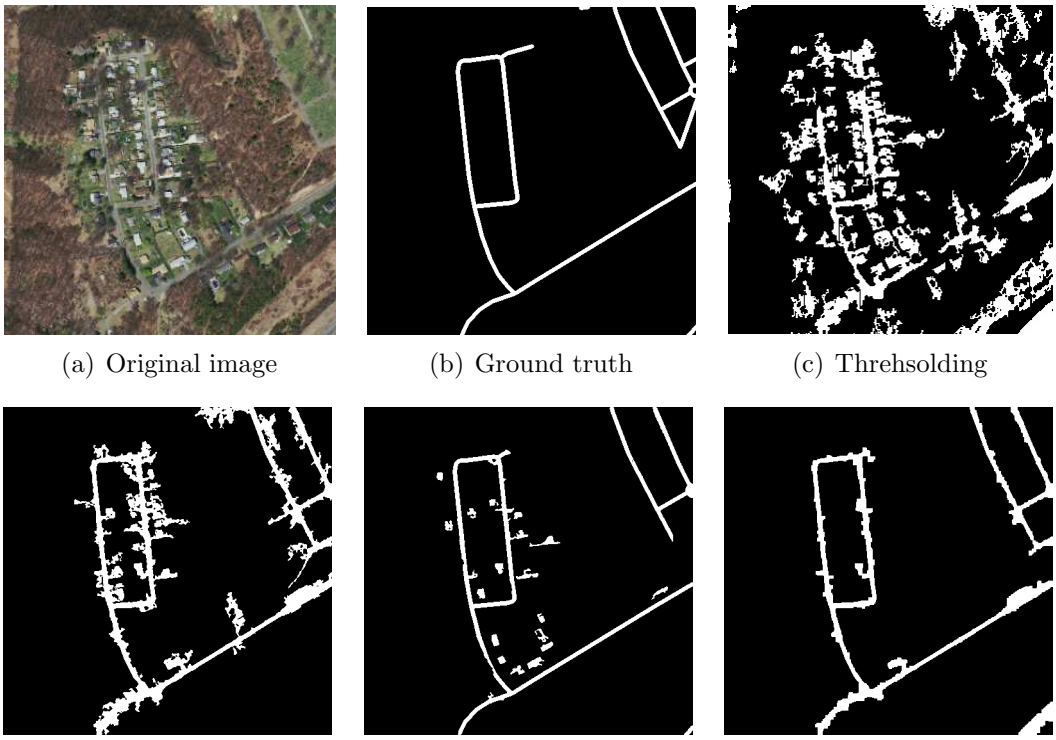
(a) Original image (b) Ground truth (c) Threhsolding

(d) Watershed-1 (e) Watershed-2 (f) New algorithm

FIGURE 8. Image #1 segmentation results of different algorithms: (a) Original image; (b) ground truth; (c) threhsolding; (d) Watershed-1; (e) Watershed-2; and (f) new algorithm

In Figure 8(a), the image includes three main roads and three small roads as marked in Figure 8(b), the colors are different in different regions, so the thresholding algorithm is difficult to use for extracting the roads as shown in Figure 8(c). The Watershed-1 (Watershed + Region Growing) can extract two main roads compared to the ground truth, as shown in Figure 8(d). If the Watershed is applied and a number of post functions are added (Watershed-2), the three main roads and a small road can be detected in Figure 8(e). The new algorithm in this study can get the better result in Figure 8(f) than that in Figure 8(d) and Figure 8(e).

Figure 9(a) shows a remote sensing image including the roads in a building region, the simple thresholding algorithm makes image result mesh by comparing the ground truth in Figure 9(b), and the roads cannot be seen clearly, as shown in Figure 9(c). The improved Watershed algorithm can extract a part of roads (Figure 9(d)), after a number of post processing, the roads can be seen clearly, and by removing all the spot regions, the final result is obtained satisfactorily.



(a) Original image (b) Ground truth (c) Threhsolding

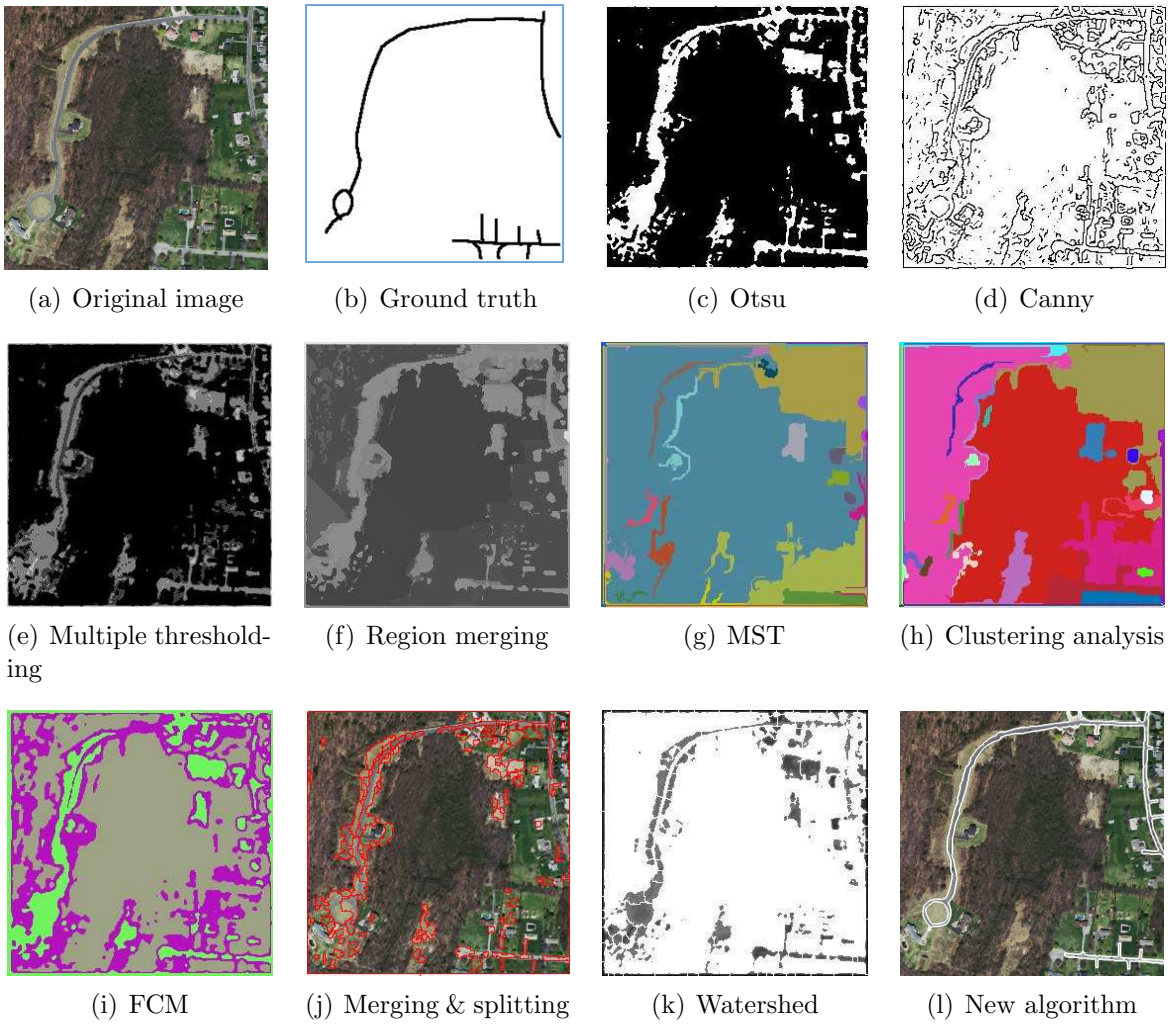
(d) Improved Watershed (e) Watershed + Post processing (f) New algorithm

FIGURE 9. Image #2 segmentation results of different algorithms: (a) Original image; (b) ground truth; (c) threhsolding; (d) improved Watershed; (e) Watershed + Post processing; and (f) new algorithm

TABLE 1. Evaluation results of different algorithms in Figure 8 and Figure 9

Image	Algorithm	Recall	Precision	F1-score
Figure 8(a)	Thresholding based	65.1%	25.2%	38.3%
	Improved Watershed + Region Growing	71.3%	81.4%	76.0%
	Watershed + Post functions	80.3%	85.6%	82.9%
	Algorithm in this study	81.3%	86.7%	83.9%
Figure 9(a)	Thresholding based	64.3%	20.4%	36.1%
	Improved Watershed + Region Growing	88.3%	79.8%	83.8%
	Watershed + Post functions	92.1%	88.8%	90.4%
	Algorithm in this study	89.3%	93.2%	91.2%

Because the image will be scaled when it is displayed in the text, the effect is as shown in Figure 10, in order to better highlight the extracted edges, the extracted edge information is enhanced, with the edge, one pixel is added to the left and right, so that the edges in the image can be seen clearly after being scaled, and the edge position does not change.



(a) Original image (b) Ground truth (c) Otsu (d) Canny

(e) Multiple thresholding (f) Region merging (g) MST (h) Clustering analysis

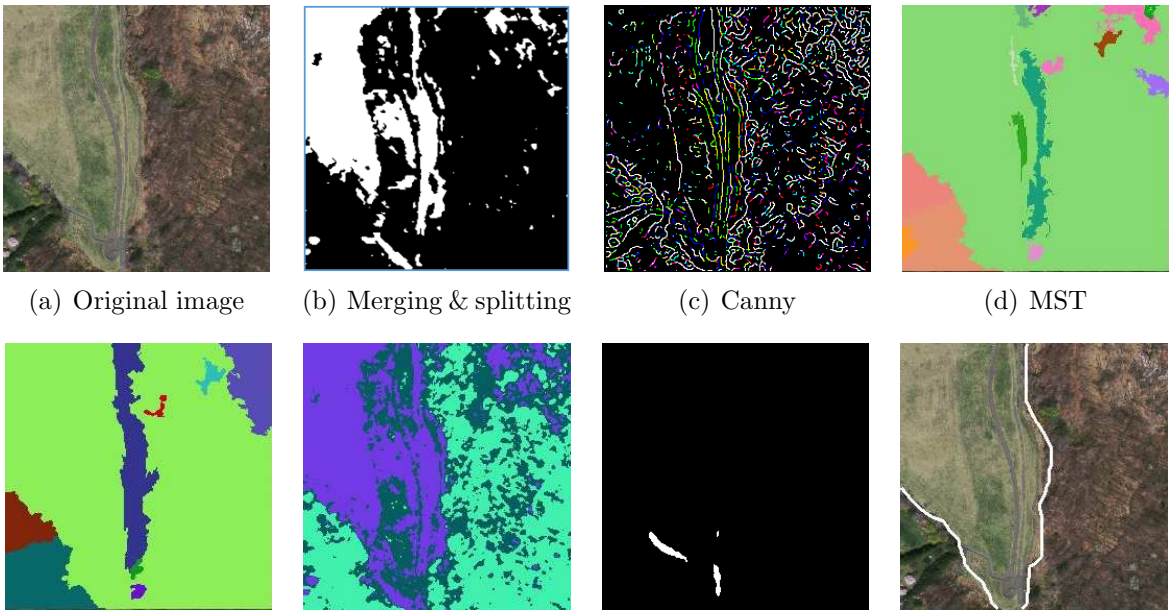
(i) FCM (j) Merging & splitting (k) Watershed (l) New algorithm

FIGURE 10. Road extraction results of rural scene 1

For the road extraction of other rural road scenes in the dataset, the enhanced results are shown in Figure 10.

In Figures 10(a) and 10(b), the original image seems to be simple, just including three long roads and some short roads, 9 different basic algorithms are tested for road detection, but none of them is successful for the detection. Otsu thresholding can output the regions including the main roads, but non-road regions are intersected to the roads, and the roads are difficult to extract completely, as shown in Figure 10(c). Since the image includes a lot of grass and forests, the edge based algorithms such as Canny edge detector will produce a lot of noise edges, see Figure 10(d). The multiple thresholding in Figure 10(e) and region merging in Figure 10(f) give out the better results than that in Figure 10(c), but they still have the same problems as that in Figure 10(c). The MST (minimum spanning tree), clustering analysis and FCM (fuzzy c-means) might be not suitable for this kind of images as shown in Figures 10(g)-10(i). The results by merging & splitting Figure 10(j) and Watershed Figure 10(k) are better than that by the above algorithms, but roads are segmented into different small areas which are mixed with non-road areas, so the further image segmentation is needed. The new algorithm studied in this paper can give a satisfactory result referring to that in Figure 10(l); hence it is successful for this kind of images.

3.4. **Result analysis by comparing semantic segmentation methods.** In addition to comparing the traditional image processing algorithms, the new method is also compared to the semantic segmentation methods [29], such as U-net [30] and SegNet [31]. In Figure 11(a), the image may be cut into 4-6 regions by colors, and it includes 4 roads. The similarity based algorithm only detects high gray scale areas, not roads, as shown in Figure 11(b); the discontinuity based algorithm can detect road boundaries, but many noise edges; hence the roads are hard extracted in the further processing, as shown in Figure 11(c); MST and clustering algorithms can extract the main road region including gaps and touching areas, as shown in Figures 11(d) and 11(e); FCM (Figure 11(f)) has the similar result as shown in Figure 11(b); and when using the semantic method for the image, the U-net [30] segmentation method only detects 10% of roads, as shown in Figure 11(g). The new algorithm can detect at least two main roads, more than 60% (Figure 11(h)).



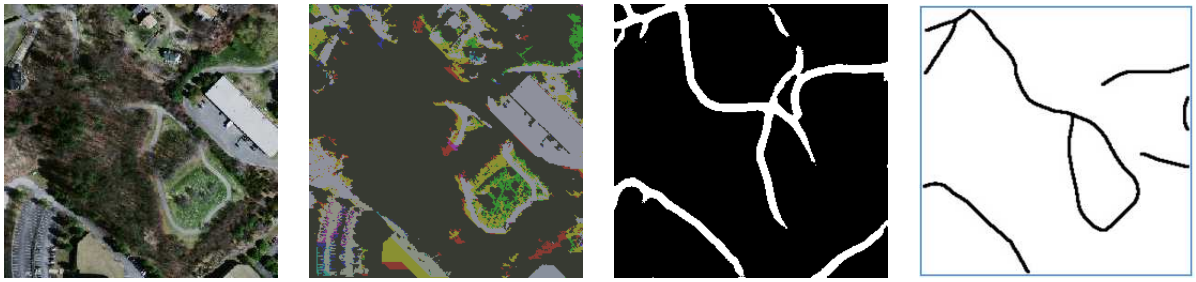
(a) Original image (b) Merging & splitting (c) Canny (d) MST

(e) Clustering analysis (f) FCM (g) U-net [30] (h) New algorithm

FIGURE 11. Road extraction results in rural scene 2

In Figure 12(a), the original image includes a large part of forests and buildings, three main roads are surrounded by the forests and buildings, and they are difficult to detect. By using the traditional Watershed algorithm, no road can be clearly detected, if the semantic segmentation method SegNet [31] is applied, the result image has over-segmentation and under-segmentation phenomena, as shown in Figure 12(c), and the best result is obtained by the new algorithm, as shown in Figure 12(d).

4. **Conclusions.** In this study, a road extraction algorithm in rural road scene is studied, and the improved Watershed algorithm is researched based on the multi-directional edge detection operators, in which, eight templates are considered to obtain the final gradient magnitude image; then, the gradient magnitude image is roughly segmented on the principle of Watershed algorithm; and finally, a number of morphological functions are utilized to finalize the roads. Firstly, the common image segmentation algorithm and the morphological functions are applied, and then, the improved Watershed algorithm is studied. Before the Watershed algorithm is used, the gradient magnitude image is obtained by using multi-directional gradient magnitude operators on the enhanced image, the image



(a) Original image (b) Merging & splitting (c) SegNet [31] (d) New algorithm

FIGURE 12. Road extraction results in rural scene 3

noise is smoothed out, and the entropy thresholding algorithm is utilized to obtain the threshold; then, the gradient magnitude image is modified, and the image segmentation results are finalized. After image segmentation, the morphological functions are utilized to remove the interference of road surrounding information without destroying the useful information, and the relatively complete road areas are obtained. The experimental results show that the studied method can obtain ideal and meaningful road image segmentation results.

In the future research of the method, since the image number in this study is limited, we should test more images to validate the method. Some parameters used in this study may be changed as image type changes. The method can be used for the post processing after semantic segmentation.

Acknowledgments. This research is financially supported by scientific and technological projects of Henan Province, China (grant no. 202102210172), Inner Mongolia Autonomous Region Transportation Development Open Fund (2019KFJJ-003), research on intelligent data processing and resource scheduling of Industrial Internet of Things based on edge computing and the Natural Science Foundation of Zhejiang Province (No. LY18F030016).

REFERENCES

[1] W. Wang, N. Yang et al., A review of road extraction from remote sensing images, *Journal of Traffic and Transportation Engineering (English Edition)*, vol.3, no.3, pp.271-282, 2016.

[2] W. Wang, M. Wang et al., Pavement crack image acquisition methods and crack extraction algorithms: A review, *Journal of Traffic and Transportation Engineering (English Edition)*, vol.6, no.6, pp.535-556, 2019.

[3] W. Wang, L. Li and Y. Han, Crack detection in shadowed images on gray level deviations in a moving window and distance deviations between connected components, *Construction Building Materials*, vol.271, 121885, 2021.

[4] W. Wang, W. Chen et al., Extraction of tunnel centerline and cross sections on fractional calculus and 3D invariant moments and best-fit ellipse, *Optics & Laser Technology*, vol.128, 106220, 2020.

[5] S. Udomhunsakul, S. P. Kozaitis and U. Sritheeravirojana, Semi-automatic road extraction from aerial images, *Proceedings of the SPIE*, vol.5239, pp.26-32, 2004.

[6] T. Yoon, W. Park and T. Kim, Semi-automatic road extraction from IKONOS satellite image, *Proceedings of the SPIE*, vol.4545, pp.320-328, 2002.

[7] S. Movaghati, A. Moghaddamjoo and A. Tavakoli, Road extraction from satellite images using particle filtering and extended Kalman filtering, *IEEE Transactions on Geoscience and Remote Sensing*, vol.48, no.7, pp.2807-2817, 2010.

[8] D. Haverkamp and R. Poulsen, Complementary methods for extracting road centerlines from IKONOS imagery, *Proceedings of the SPIE*, vol.4885, pp.501-511, 2003.

[9] W. Wang, H. Li et al., Pavement crack detection on geodesic shadow removal with local oriented filter on LOF and improved level set, *Construction Building Materials*, vol.237, 117750, 2020.

[10] B. G. Kim, J. I. Shim and D. J. Park, Fast image segmentation based on multi-resolution analysis and wavelets, *Pattern Recognition Letters*, vol.24, no.16, pp.2995-3006, 2003.
[11] G.-S. Hong, B.-G. Kim, D. P. Dogra and P. P. Roy, A survey of real-time road detection techniques using visual color sensor, *Journal of Multimedia Information System*, vol.5, no.1, pp.9-14, DOI: https://doi.org/10.9717/JMIS.2018.5.1.9, 2018.
[12] A. Chandra, B. Acharya and M. I. Khan, Retinex image processing: Improving the visual realism of color images, *International Journal of Information Technology and Knowledge Management*, vol.4, no.2, pp.371-377, 2011.
[13] K. He, J. Sun and X. Tang, Single image haze removal using dark channel prior, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.33, no.12, pp.2341-2353, 2011.
[14] T. Tang, X. Wang, J. Carbonara and Z. Shi, Feature shape and elevation based road classification and extraction on high spatial resolution remote sensing imageries, *2009 17th International Conference on Geoinformatics*, pp.1-6, 2009.
[15] W. Wang, W. Li and X. Yu, Fractional differential algorithms for rock fracture images, *The Imaging Science Journal*, vol.60, pp.103-111, 2012.
[16] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi and N. Mazur, A novel fast single image dehazing algorithm based on artificial multiexposure image fusion, *IEEE Transactions on Instrumentation and Measurement*, vol.70, pp.1-23, 2021.
[17] W. Li, H. Wei, G. Qi, H. Ding and K. Li, A fast image dehazing algorithm for highway tunnel based on artificial multi-exposure image fusion, *IOP Conference Series: Materials Science and Engineering*, vol.741, no.1, 012038, 2020.
[18] A. Abdollahi, B. Pradhan, N. Shukla et al., Deep learning approaches applied to remote sensing datasets for road extraction: A state-of-the-art review, *Remote Sensing*, vol.12, no.9, 2020.
[19] S. J. Phillips, R. P. Anderson and R. E. Schapire, Maximum entropy modeling of species geographic distributions, *Ecological Modelling*, 2006.
[20] R. Kikuchi and B. H. Soffer, Maximum entropy image restoration. I. The entropy expression, *Journal of the Optical Society of America*, vol.67, no.12, 1977.
[21] F. K. Siddiqui and V. Richhariya, An efficient image segmentation approach through enhanced watershed algorithm, *Computer Engineering & Intelligent Systems*, vol.5, no.3, 2013.
[22] G. N. Girish, A. Kothari and J. Rajan, Marker controlled watershed transform for intra-retinal cysts segmentation from optical coherence tomography B-scans, *Pattern Recognition Letters*, 2017.
[23] M. L. Silvoster, R. Mathusoothana and S. Kumar, Watershed based algorithms for the segmentation of spine MRI, *Int. J. Inf. Tecnol.*, DOI: https://doi.org/10.1007/s41870-021-00644-8, 2021.
[24] M. H. Ghaznavi, S. Ghaderi and K. Ghaderi, Using marker-controlled watershed transform to detect Baker's cyst in magnetic resonance imaging images: A pilot study, *Journal of Medical Signals & Sensors*, vol.12, no.1, 2021.
[25] C. Heipke, H. Mayer, C. Wiedemann et al., Evaluation of automatic road extraction, *International Archives of Photogrammetry and Remote Sensing*, vol.32, pp.151-160, 1997.
[26] C. Wiedemann, C. Heipke, H. Mayer et al., Empirical evaluation of automatically extracted road axes, *Empirical Evaluation Techniques in Computer Vision*, pp.172-187, 1998.
[27] T. Pan and C. Li, Road extraction method of satellite image based on region-growth watersheds algorithm, *Computer Engineering and Design*, vol.29, no.19, pp.4987-5013, 2010.
[28] G. Xu, J. Bi, X. Wang et al., An experiment of automatic road extraction from high-resolution remote sensing image based on object-oriented technology, *Remote Sensing Information*, no.2, pp.108-111, 2019.
[29] Y. Shin, M. Kim, K.-W. Pak and D. Kim, Practical methods of image data preprocessing for enhancing the performance of deep learning based road crack detection, *ICIC Express Letters, Part B: Applications*, vol.11, no.4, pp.373-379, 2020.
[30] O. Ronneberger, P. Fischer and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science*, N. Navab, J. Hornegger, W. Wells and A. Frangi (eds.), Cham, Springer, 2015.
[31] V. Badrinarayanan, A. Kendall and R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.39, no.12, pp.2481-2495, 2017.

Author Biography

Kang Lv received the B.Sc. degree in computer application technology from Henan Normal University, China, 2004; the M.Sc. degree in network engineering from Zhengzhou University, China, 2008. He is currently an associate professor at the Teacher Education and Training Department, Henan Finance University, China. His main research interests include the computer application technology, computer data mining technology, development and utilization of information technology resources.

Weixing Wang received the Ph.D. degree in computer vision from Royal Institute of Tech., Sweden, in 1994. He is currently a doctoral supervisor in Royal Institute of Tech., Sweden, and a visiting professor with the Department of Computer Science and Engineering, Chang'an University, Xi'an, China. His current research interests include pattern recognition and intelligence, image and graph processing and analysis, and computer vision.

Zhen Zhou received the B.Sc. degree in computer application technology from Henan Normal University, China, 2001; the M.Sc. degree in computer application technology from Northwestern Polytechnical University, China, 2008. He is currently a full-time professor at the School of Information Technology, Luoyang Normal University, China. His main research interests include the image processing technology, cloud computing and IoT edge calculation.

Xueying Wang received the B.Sc. degree from Shandong University of Technology, China in 2010; the M.Sc. degree from Chongqing Jiaotong University, China in 2013. He is a senior engineer at Quality Monitoring and Appraisal Station for Traffic Construction Project of Inner Mongolia Autonomous Region, and he is studying at Harbin Institute of Technology for the Ph.D. degree, China. His main research interests are new materials and testing methods of asphalt pavement.