

Learning From GPS Trajectories of Floating Car for CNN-Based Urban Road Extraction With High-Resolution Satellite Imagery

Ju Zhang¹, Qingwu Hu¹, Jiayuan Li¹, and Mingyao Ai

Abstract—Deep learning has achieved great success in recent years, among which the convolutional neural network (CNN) method is outstanding in image segmentation and image recognition. It is also widely used in satellite imagery road extraction and, generally, can obtain accurate and extraction results. However, at present, the extraction of roads based on CNN still requires a lot of manual preparation work, and a large number of samples can be marked to achieve extraction, which has to take long drawing cycle and high production cost. In this article, a new CNN sample set production method is proposed, which uses the GPS trajectories of floating car as training set (GPSTasST), for the multilevel urban roads extraction from high-resolution remote sensing imagery. This method rasterizes the GPS trajectories of floating car into a raster map and uses the processed raster map to label the satellite image to obtain a road extraction sample set. CNN can extract roads from remote sensing imagery by learning the training set. The results show that the method achieves a harmonic mean of precision and recall higher than road extraction method from single data source while eliminating the manual labeling work, which shows the effectiveness of this work.

Index Terms—Convolution neural network (CNN), GPS trajectories of floating car, high-resolution satellite imagery, road extraction.

I. INTRODUCTION

DEEP learning is one of the important techniques to classify terrains and extract objects from remote sensing imageries[1]–[3]. As a geographical element closely related to people’s life, road is characterized by complex and changeable. Based on deep learning technology, especially the convolutional neural network (CNN) model, the urban road extraction from remote sensing imagery has been a research focus, which is useful, effective, and efficient. For example, Kahraman *et al.* [4] proposed a road detection model based on neural network from remote sensing imagery, mainly using a multilayer sensor to identify road based on

Manuscript received October 17, 2019; revised April 27, 2020; accepted June 1, 2020. This work was supported by the Open Research Fund of Key Laboratory for National Geography State Monitoring (National Administration of Surveying, Mapping and Geoinformation) under Grant 2018NGCMZD02. (Corresponding author: Qingwu Hu.)

The authors are with the School of Remote Sensing Information Engineering, Wuhan University, Wuhan 430079, China (e-mail: zhangju@whu.edu.cn; huqw@whu.edu.cn).

This article has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the authors.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2020.3003425

the RGB eigenvalues of the image; Xia *et al.* [5] divided road into different categories, then used pixel-level estimation for remote sensing imagery by deep convolutional network, predicted all probabilities of different categories, and finally connected missing or nonsmooth roads; Saito *et al.* [6] take the original pixel values in the spatial image as input and output of the predicted three-channel label image (building-road-background), which can effectively train a single CNN model, extract multiple objects at the same time, and improve the performance of prediction. Combining the advantages of residual operation and U-Net, Zhang *et al.* [7] proposed a semantic segmentation road extraction neural network, which is constructed using the units and has a similar architecture to U-Net, which make the model use fewer parameters but achieve better performance.

In road extraction, a road extraction network model is constructed based on a CNN, and road objects in the image are segmented and extracted. The commonly used CNN networks for road extraction include LinkNet, D-LinkNet, D-LinkNet-1D, and so on. The LinkNet [8] architecture is similar to a ladder network architecture in which the feature map from the encoder is added to the up-sampled feature map from the decoder. Each encoder is coupled to a decoder such that the input of the encoder is coupled to the output of the corresponding decoder. ResNet18 is used as the backbone of the network structure, which reduces the parameters of the network and simplifies the structure. The D-LinkNet [9] model uses LinkNet and a pretraining encoder as its backbone and adds an additional expansion convolution to the center. The basic structure of D-LinkNet is mainly composed of three parts: encoder, center part, and decoder. D-LinkNet uses ResNet34 as an encoder. The D-LinkNet center part uses several expansion convolutions and jump connections to expand the receptive field while preserving the resolution and spatial information of the image. The D-LinkNet decoder is identical to the original LinkNet, which guarantees training speed. D-LinkNet-1D is a network model specifically used for road extraction. Roads are thin and long, and the 1-D filters are more aligned with road shapes. D-LinkNet-1D replaces the 3×3 transposed convolution in decoder of D-LinkNet with 1-D transposed convolution.

In the process of road extraction, the neural network model should be trained first by the training set composed of the remote sensing imageries that have been marked out of roads, and the model is fitted. After the training is completed,

the trained model can be used to extract the road from the remote sensing imagery. There are many open source road extraction sample sets in the field of deep learning. However, influenced by climate, topography, economic, and culture, roads in different regions vary widely. Especially in some cities, roads are divided into many grades, and high-grade roads are generally wide and easy to be identified unless large areas are blocked. While some low-grade roads, such as roads inside the community, tend to very narrow, some are blocked by vehicles, trees, and buildings, which make them difficult to be identified. Existing data sets do not satisfy road extraction in all regions. If the characteristics of the roads in training set and in the area to be predicted have a big difference, the large generalization error will be produced and the prediction result will be affected [10]. Therefore, it is necessary to manually generate the sample set for road extraction, which is more reliable, and the obtained extraction result has high precision, but the initial investment is large, and the practicality is low in actual production [11]. From the current road extraction research, it is difficult that the extraction accuracy and efficiency reach a good balance. We need to explore an easy-to-access and reliable production method of road extraction sample set, which will make the road extraction based on the CNN model from imagery to ensure the accuracy and integrity, while the preparation time is shortened and the overall efficiency is improved.

High-resolution remote sensing imagery can reflect surface features in more detail, making remote sensing imagery data into one of the main data sources for road extraction. Gruen and Li [12] combined wavelet decomposition with road sharpening and proposed a model-driven linear feature extraction algorithm based on dynamic programming, which can successfully extract a complete road network from a single SPOT satellite imagery or an aerial imagery; Cheng-Li *et al.* [13] used Gaussian probability iterative segmentation to process the road area to extract road information, and then used Hough transform to detect the straight edge of the road. The test proved that good branch road information can be obtained. The IHS transform is used by Yan and Zhao for true color fused images for threshold segmentation, then the connected regions are processed, and the spurs are removed to obtain the road network [14]. Liu *et al.* [15] used improved directional consistency to segment the high-resolution imagery to obtain the road vectorgraph. The ground cover information of remote sensing imagery is comprehensive, abundant, and accurate. However, the objects in remote sensing imagery are redundant, and the objects on the road and along the road will interfere with the extraction. Many roads are blocked by trees or buildings or are in the shadows, which is difficult to identify them. Road extraction from remote sensing imagery has the problem of different objects have the same spectrum and the same objects have different spectrum in remote sensing imagery, and it is difficult to reflect the topological relationship and geometric structure of the road in the image.

Floating cars such as express and special cars provided by platforms and taxis are equipped with GPS

positioning equipment. Most civilian private cars are also equipped with GPS equipment for navigation. These floating cars generate massive GPS trajectories during driving, and these floating car data the included spatial location information can directly reflect the geometry of the road traffic network, which is a good data source for extracting road information. Based on the characteristics of GPS trajectories of floating car, the method of extracting road from floating car data is real time and available. A lot of research results have been obtained by extracting road geometry information from floating car data. Chen and Cheng [16] used image processing technology to convert the GPS trajectories points into a binary image, and use the morphological method to extract the road skeleton line to construct the road network map. The nuclear density function was used by Biagioni and Eriksson [17], multiscale raster maps are obtained by setting different thresholds, and finally the skeleton extraction algorithm was used to obtain the road centerline. Liu *et al.* [18] combined the kernel density estimation method with the *K*-Means clustering method to extract road network data. Xie *et al.* [19] proposed a method to infer road networks from GPS trajectories of floating car. They detected potential intersections by clustering the turning points on the GPS trajectories of floating car and inferred the geometry of the road segments between intersections by aligning GPS tracks point by point using a “stretch and then compress” strategy. Qiu and Wang [20] proposed a point segmentation and grouping method to generate road maps from GPS trajectories of floating car. Guo *et al.* [21], Wu *et al.* [22], and Wang *et al.* [23] used clustering, machine learning, probability statistics, and other methods to study the geometry extraction of road intersections.

Compared with satellite imagery, the continuity of the GPS trajectories of floating car and good geometric features can help solve the problem of remote sensing imagery and can reflect the road topological relationship lacking in remote sensing imagery. To put it in another way, the ground cover information of remote sensing imagery is more comprehensive, abundant, and accurate. It can compensate for the problem of extracting from floating car data to a certain extent, including the uneven GPS trajectories distribution, reflecting incomplete information and poor GPS positioning accuracy. The use of remote sensing imagery combined with floating car data for road network extraction can fully exploit their respective advantages and more comprehensively and accurately extract road information under complex scenes. This article will combine the floating car data and high-resolution remote sensing imagery in the city to study the fusion method that can give full play to their respective advantages. Use deep neural network learning methods, which is difficult to obtain sample data during network training. The method of quickly labeling and obtaining sample data based on floating car data is proposed, and network training is carried out to realize accurate, complete, and effective urban road extraction.

In this article, we proposed a method for automatically generating urban road extraction sample sets using floating car data. The floating car data containing only road information will be used instead of the prior knowledge of artificial road extraction, which has high efficiency and small investment.

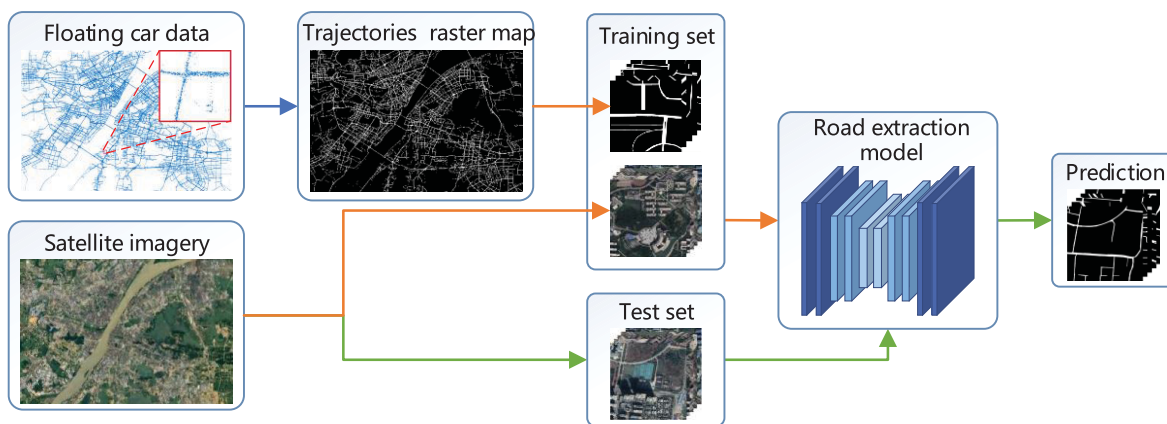


Fig. 1. Technical route of road extraction.

It can realize the rapid extraction and update of roads and the extraction of large-scale and complex scenes. With the prior knowledge of the urban road extraction, the results are comprehensive and accurate, thus achieving a good balance between efficiency and effect.

II. METHODOLOGY

At the beginning of the extraction, the filtered floating car data are rasterized. Considering the problem of uneven distribution of GPS trajectories of different grades of roads, the GPS trajectories of floating car are rasterized according to the points density. The GPS trajectories of the points dense road area are directly rasterized, and the GPS trajectories of the sparse area are rasterized according to the order according to the timing. The two raster maps are combined to get a raster image of all the GPS trajectories. Then, a series of mathematical morphology processing is performed to obtain a smooth GPS trajectories raster map. Then, the resolution of the GPS trajectories raster map and the remote sensing imagery is unified, and the corresponding road is marked by the floating car data to generate a training sample set. Using a CNN as a basic structure to establish an urban road extraction network model, and learning training data including floating car data marked. The network model of road extraction is obtained, and the roads in the high-resolution remote sensing imagery of the test set are extracted. The road extraction result is finally obtained.

The overall technical route of road extraction by CNN based on floating car data from the high-resolution imagery is shown in Fig. 1.

A. Road Extraction From GPS Trajectories of Floating Car

Floating car data are one of the important data sources for urban road extraction. It is a direct representation of the road shape. Due to the huge number of GPS trajectories of floating car, the method of extracting roads from GPS trajectories vectorgraph has huge computation and storage space, which is difficult to achieve in actual production. The image morphology method is used to extract the urban road network information from the GPS trajectories of floating car, which greatly reduces the time complexity and space complexity, and can meet the needs of actual production. This article

will elaborate on the method and process of extracting roads from trajectories. Firstly, the expression form of the trajectory data and the conditions of data cleaning will be described. Then, the method of GPS trajectories rasterization will be discussed. An algorithm of floating car data rasterization based on GPS trajectories density is proposed. A smooth GPS trajectories raster image of research region is obtained based on mathematical morphology.

1) *GPS Trajectories Rasterization*: The GPS trajectories of floating car data are the GPS information continuously collected and recorded by a large number of floating car over a period of time. It has the characteristics which are large data volume and uneven data quality. The floating car data used in this article are the floating car data of the Wuhan express car provided by the Drip Express. Each piece of data contains order information, longitude, latitude, speed, direction, time stamp, and horizontal accuracy factor information. Due to the interference of the accuracy of the GPS trajectories of floating car and the occlusion of the road segment, the accuracy of the GPS trajectory data is different. The trajectory data with high precision and good quality are screened for the next step.

In this article, the speed, acquisition interval, and horizontal component precision factor are used as the screening conditions for the floating car data. If the speed of a floating car at a certain position is very slow or even 0, its GPS data will drift around it. Such GPS trajectories points will seriously affect the positioning accuracy and need to be removed. At the same time, the trajectory speed is limited to the range of [5 m/s, 25 m/s] according to the general driving speed of floating car on the urban road. The acquisition interval of the floating car data is mostly less than or equal to 5 s. To improve the accuracy of the data, the floating car data with an acquisition interval greater than 5 s are removed. The floating car data also contain the horizontal component precision factor information, which is used to indicate the error of the trajectory in the latitude and longitude direction. It is generally considered that the horizontal precision factor is greater than 7 is the error, and less than or equal to 3 is the data with high positioning accuracy [24]. In this article, the amount of floating car data is large. In order to obtain higher precision extraction results, the GPS trajectories with horizontal precision factor less than or equal to 3 are selected. The floating car data satisfying the

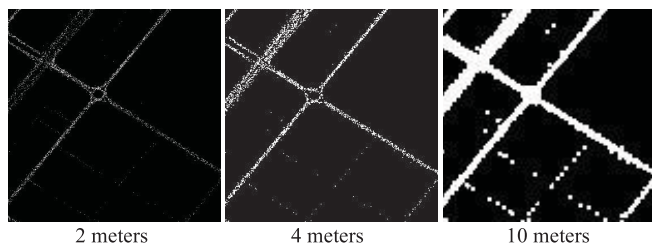


Fig. 2. Raster map with different resolution.

above three conditions are taken as the data to be processed, and the remaining data are all removed.

Rasterizing floating car data is the mapping of GPS trajectories-points into a 2-D space. The image obtained by rasterizing the point directly is a binary image, the grid pixel value of trace points is 1, and the grid pixel value of without trace point is 0. Therefore, the floating car data can be rasterized by calculating the latitude and longitude coordinates (lon_p , lat_p) of the trace point corresponding to the spatial coordinates (x , y) of the grid. The calculation method is

$$x = \text{round}((lon_p - lon_{\min}) \times \text{num}_{\text{pix}}) \quad (1)$$

$$y = \text{num}_y - \text{round}((lat_p - lat_{\min}) \times \text{num}_{\text{pix}}) \quad (2)$$

where round is a rounding function. lon_{\min} and lat_{\min} are the minimum latitude and longitude of the research area. num_{pix} is the number of pixels per degree. This variable is directly determined by the resolution, which is the ratio of the distance to the resolution per degree. num_y is the number of pixels in the y -direction of the research region. Since the origin of the screen coordinates is the upper left corner of the screen, the coordinate value in the y -direction needs to be subtracted from num_y . The calculation of num_y is

$$\text{num}_y = \text{round}((lat_{\max} - lat_{\min}) \times \text{num}_{\text{pix}}) \quad (3)$$

where lat_{\max} is the maximum longitude of the research area?

The resolution of the raster image is also one of the factors that determine the road extraction result: if the resolution is too high, the pixel rasterization points of the traces will be too scattered. Although the geometric features of the road can be more clearly and accurately reflected, the sparse trace points are removed more easily as noise when they are denoised. At the same time, the amount of high-resolution raster data will be too large, which will affect the efficiency of post processing. If the resolution is too low, the proportion of pixels occupied by the track will increase, which will result in inaccurate road information presented by the raster image. For example, two adjacent roads may be misjudged as one. Therefore, choosing the suitable resolution is one of the keys to rasterize floating car data.

Fig. 2 shows an image generated by rasterizing the trajectory with different resolutions. Figure 2(a) is a GPS trajectories rasterization image with 2 m resolution. Although the effect of reflecting the geometric features of the road is good, some GPS trajectories of road segments are too sparse, and GPS trajectories in the entire road may be removed during denoising. Fig. 2(c) is an image with 1 m resolution. In this image, the two roads on the left side are rasterized to one.

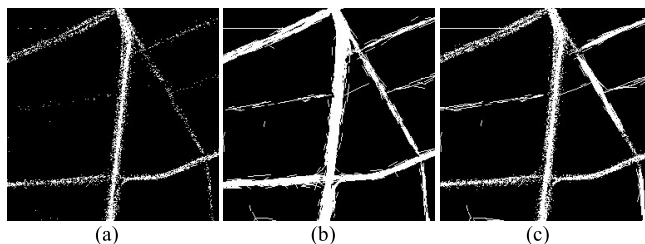


Fig. 3. Results with different rasterization methods. (a) Result of GPS trajectories points rasterization. (b) Result of rasterizing traces using linear interpolation. (c) Result of rasterizing the floating car data according to the traces density.

And it cannot reflect the geometric characteristics of the roundabout. Fig. 2(b) uses the 4 m resolution, it reflects the geometry of the road better. This article uses the 4 m resolution to rasterize the GPS trajectories.

To put it in another way, the method of rasterizing GPS trajectories points is also a problem to be discussed. For the GPS trajectories of floating car with higher acquisition frequency, the method of rasterizing GPS trajectories points generally has two kinds of algorithms: direct rasterization and linear interpolation. Direct rasterizing means that only the pixels of GPS trajectories points are assigned a value of 1, while the linear interpolation algorithm is to connect two GPS trajectories points with a straight line, and the pixels through which the line passes are also assigned a value of 1. It can be seen that for Fig. 3(a), there may be only a few discrete GPS trajectories points in the sparse road segment, which can be easily removed as salt-pepper noise during late denoising, resulting in loss of road information; (b) interpolation of the GPS trajectories by linear interpolation in the section with dense points will result in data redundancy and, at the same time, increase edge burrs and aggravate noise.

Therefore, we rasterize the floating car data according to the GPS trajectories density. The result is shown in Fig. 3(c). First, we split road into sparse and dense road segments based on floating car data for one week. Then, according to the density of the road trace points, the GPS trajectories of floating car are processed separately by different algorithms. The linear interpolation method is adopted in the road with sparse GPS trajectories. On the road with heavy traffic, only the trace points are rasterized, avoiding redundant operations and ensuring good shape of the roads.

2) *GPS Trajectories Raster Map Mathematical Morphology*: Although we have used reasonable conditions to clean the floating car data and rasterize it with appropriate algorithm. Due to hardware and other conditions, the rasterized image will still have noise, holes, edge burrs, and so on. Adopting the mathematical morphology algorithm, designing a reasonable process to de-noise and smooth the rasterized image, can obtain images with better road geometry.

After obtaining the rasterized image of the GPS trajectories of floating car, first select the appropriate template according to the actual situation and use the median filtering to denoise. The salt-pepper noise on the image due to GPS trajectories drift can be eliminated by median filtering. The algorithm can effectively remove discrete noise points and can maintain the

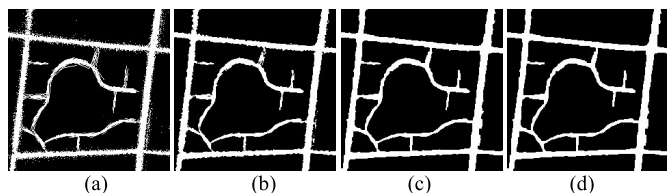


Fig. 4. Mathematical morphology processing results of GPS trajectories raster map. (a) Result of rasterization. (b) Result of denoising. (c) Result of removing holes. (d) Result of deburring.

original shape of the image well. Fig. 4(a) is an image after rasterization, and Fig. 4(b) is the result of denoising by median filtering. The noise points have been removed substantially.

The denoised image still has some holes and gaps that need to be processed. Using the appropriate structural elements for the dilation operation eliminates the hole gap in the target object. Then, restore the edge to the original shape by erosion operation with the same structural element, that is, the closing operation. The algorithm which dilates firstly and then erodes the target object can effectively eliminate the holes in the raster map. Fig. 4(c) shows the result of removing the holes after the closing operation.

In opening operation, the first erosion operation eliminates edge protrusions smaller than the structural elements and results in smooth images; the subsequent dilation operation can expand the nonburred part back to the original shape, achieving the effect of both eliminating the burr and maintaining the original shape of the target object. Process the raster map by opening operation to remove burrs on the edge while ensuring the original geometry of the road. In Fig. 4(d), after processing with the opening operation, the smoothed result is obtained.

B. Establishment of a Road Extraction Sample Set

At present, deep learning technology is developing rapidly and becomes more mature, but the sample data production in deep learning are often carried out by means of manual labeling, which consumes a lot of time and energy. The floating car data are a direct reflection of the road network. This article chooses to use the floating car data to automatically label the training data to quickly create a deep learning sample set.

This article designs a reasonable experimental process to make a road extraction sample set. Since the resolution of the satellite imagery selected in this article is higher than the resolution of the GPS trajectories raster map, this article will first improve the resolution of the GPS trajectories raster map to the resolution of the satellite imagery. In addition, since the high-resolution image takes up a large space, which is difficult to further processing on the image, the two images are first divided according to the size required by the final sample set, which facilitates image processing operations and avoids memory leaks. Then, this article proposes an appropriate algorithm to adjust the width of the floating car data so that it can match the road on the remote sensing imagery well. Finally, the test set and training set data will be distributed in a reasonable proportion to complete the production of the road extraction sample set. The specific process of road extraction sample set production is shown in Fig. 5.

1) *Composition of the Sample Set*: In this article, the smooth GPS trajectories raster map is used to label the roads in the satellite imagery. The sample set for general deep learning consists of a training set and a test set. The Google Earth satellite imagery of research region with a resolution of 0.54 m is used. The training data set is composed of remote sensing imageries with 1024×1024 size and road binary images with the same size corresponding to them. The road raster in the road binary map takes 1 and the background value is 0. The remote sensing imagery is divided into subimages with 1024×1024 size, and the total number is 72×48 . The pixel value at the edge of the image, which size is less than 1024, is assigned a value of zero.

The GPS trajectories of floating car coordinate system in this article and the Google Earth satellite image coordinate system use the same coordinate system, and no coordinate conversion operation is required. The image rasterized by the GPS trajectories of floating car is superimposed with the satellite imagery, and most of the GPS trajectories grid falls into the road area of the image. The high-resolution imageries are processed by geometric correction, radiation correction, atmospheric correction, and so on.

2) *Matching of GPS Trajectories to Roads on Imagery*: The resolution of the high-resolution imagery is about 0.54 m, while the resolution of the GPS trajectories raster map is 4 m. To use the floating car data to label the road on the remote sensing imagery, firstly, the resolution of the images should be unified. The resolution of the GPS trajectories raster map is increased to match the high-resolution imagery. Then, the satellite imagery and the GPS trajectories of floating car raster image are, respectively, divided into 1024×1024 subimages by the grid segmentation tool.

In theory, the GPS trajectories points of floating car should fall into and cover the entire road area, so that the width of the GPS trajectories of floating car raster is the width of the corresponding road. In reality, the various scenes are very complex, and this ideal situation is often difficult to achieve. In the main roads, especially in the busy roads of the city center, there are more vehicles, and the GPS trajectories drift phenomenon is more common. This makes the GPS trajectories drifting to the sides of the road no longer discrete grid points but connected to the normal GPS trajectories. Many connected raster areas are unable to be removed by denoising operations. As shown in Fig. 6(a), the GPS trajectories grid represented by the yellow area has covered the buildings and other objects on both sides of the road. Such a training sample that mistakes the ground objects on both sides of the road for the road will cause the neural network model to extract the wrong information during training. A large number of such samples will affect the extraction results of the network model.

Conversely, in some roads with less traffic, there may be only a small number of vehicles passing through the time range of experimental data acquisition. The GPS trajectories cannot cover the entire road surface. Even some GPS trajectories widths are too little to reflect the true characteristics of the corresponding roads. The width of a road label is required in the network learning in order to extract features. The yellow

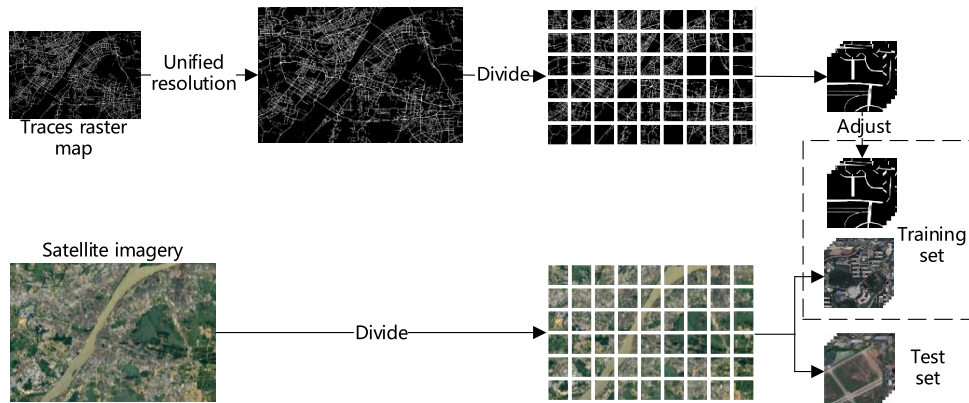


Fig. 5. Sample set production route.



Fig. 6. (a) GPS trajectories raster is narrower than the road width. (b) GPS trajectories raster overflows road range.

GPS trajectories grids in Fig. 6(b) occupy only a small part of the road surface.

For the abovementioned case, the width of the trajectory should be appropriately adjusted to reflect the actual road width in the satellite imagery. The first thing to do is to judge the area where the GPS trajectories raster is too wide and too narrow. According to the actual situation of urban road, the width of higher-grade roads is generally greater than 30 m. The lowest level road displayed on the general map is the residential group level road, which is connecting the residential communities in the community, and the road width is generally not less than 5 m. Since the road center line is finally extracted in this article, the width of the labeled road area can be slightly smaller than the actual road width. Therefore, this article assigns that the GPS trajectories grid with a width greater than 30 m may have the problem of GPS trajectories drift, and it is necessary to erode the appropriate width inward. In this article, the erosion width is set to 5 m, and if the width is less than 30 m after erosion, it will erode to 30 m. A 5- to 30-m wide GPS trajectories raster is considered to be an appropriate GPS trajectories raster which is not need to process, while a GPS trajectories raster less than 5 m wide is considered to be too narrow and needs to be dilated to 5 m width.

The resolution of the GPS trajectories raster map is already the same as the resolution of the high-resolution satellite imagery, which is about 0.54 m. When screening a GPS trajectories grid with 5 m width, a circular-like structural element with a diameter of 9 as shown by B in Fig. 7 is selected. Use it to scan the GPS trajectories object in the GPS trajectories raster image. The structural object proceeds along the center line of the GPS trajectories object and determines whether the structural element B is completely covered by

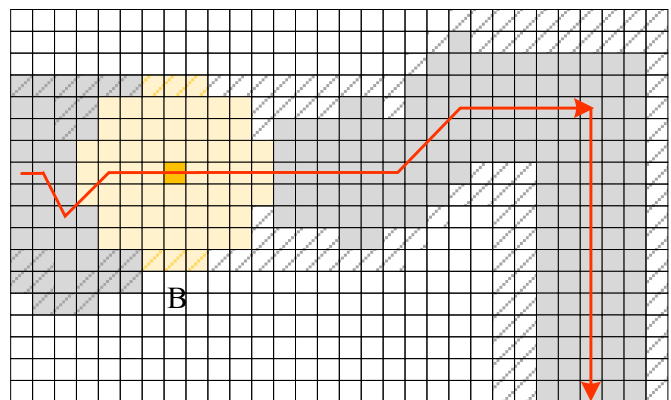


Fig. 7. Schematic of the expansion of a narrow track object.

the GPS trajectories object during the scanning process; if it is, it proceeds to the next position of the center line without processing and, if it is not, the pixel with value of 0 in the structure element is assigned a value of 1, then continue scanning until the end of the entire image is complete. The schematic diagram is shown in Fig. 7. The structural element B proceeds along the centerline of the GPS trajectories object represented by the red line. The gray grids are the GPS trajectories, and the grids filled with oblique line are the grid of which value is changed from 0 to 1 by dilation.

Similarly, for roads with a GPS trajectories raster width greater than the actual, a structural element B_1 with 55 pixels diameter is used to proceed along the center line of the GPS trajectories object, and for a GPS trajectories raster with a width larger than the structural element B_1 , a structural element B_2 with a side length of 30 m, that is, 9×9 , is used for erosion. Determine if the erosion will be less than B_1 , if it is not, continue, if it is, then erode to the end of the B_1 edge. Through this algorithm, a track grid with good matching with the road can be obtained. The width of road raster W_r is

$$W_r = \begin{cases} D_B, & (0 < W_t < D_B) \\ W_t, & (D_B \leq W_t \leq D_{B_1}) \\ W_t - D_{B_2}, & (W_t > D_{B_1}) \end{cases} \quad (4)$$

where D_B is the diameter of structural element B, W_t is the width of GPS trajectories raster, and D_{B_1} and D_{B_2} are the diameters of structural elements B_1 and B_2 , respectively.

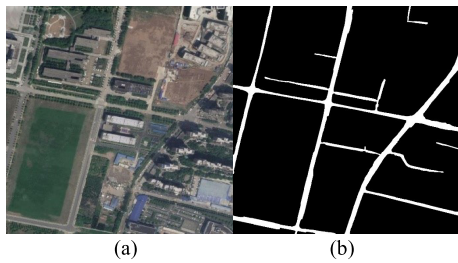


Fig. 8. Example of training set. (a) One of the satellite imageries. (b) Label data with the roads labeled based on the traces data.

Finally, a test set consisting of 600 satellite imageries and a training set consisting of 2856 sets of satellite imageries and label data with the roads labeled are obtained. Fig. 8 shows a set of training set data. The use of GPS trajectories data to label remote sensing imagery saves a lot of labors and time costs and is convenient and effective.

III. EXPERIMENTAL RESULTS

A. Data Sets and Models

1) *Data Sets*: The floating car data used in this article are the 19-day express floating car data from June 20, 2018 to July 8, 2018. The average daily data are about 85 million, and the total data size is 140 G. The research area is the urban area of Wuhan, with a longitude range of 114.125° – 114.500° , a latitude range of 30.417° – 30.667° , and an area of approximately 900 km^2 . The research region is located in the Wuhan Fourth Ring Road and contains most of the area within the Third Ring Road.

The Google Earth satellite imagery of research region with a resolution of 0.54 m is used. The image size is 22G, and the image acquisition time is of May 2018. There is little difference between the acquisition time of the floating car data and the satellite imagery. In order to classify the CNN model for the high-resolution imagery, its size needs to satisfy the receptive field of the model. At the same time, the connectivity and complexity of the road should be ensured. Therefore, the high-resolution imagery should be segmented and divided into subimages of appropriate size.

In this article, a rectangular area with an area of about 150 km^2 is selected as the research area. The area is located on the lower reaches of the Han River, mainly crossing the Jiangnan District, Qiaokou District, and Hanyang District of Wuhan. The specific location is the yellow area, as shown in Fig. 9. The area includes some urban centers and suburbs of Wuhan. The types of roads include highways, urban roads of various grades, like main roads, internal roads, factories and mines roads, and rural roads. They meet the complexity of urban roads. It facilitates detection of the effect of the road extraction results. There are 600 images in 30×20 in the test set, and the remaining 2856 images are labeled to the training set data.

The DeepGlobe Road Extraction Data set [25] and Road and Building Detection Data set (the Massachusetts Roads Data set) [26] are used for training. The data set includes 6226 pieces of high-resolution remote sensing imageries with label data, and the image size is 1024×1024 pixels. The captured countries involve Thailand, India, and Indonesia,

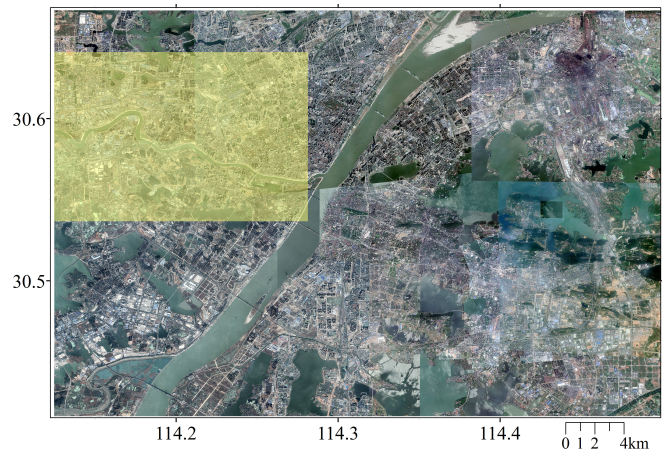


Fig. 9. Areal map of research area.

including urban, rural, coastal, and tropical rainforests. The main roads are labeled, but some side roads are ignored. The size and format of each image is the same as the data set based on GPS trajectories. The Massachusetts Roads Data set consists of 1171 aerial images of the state of Massachusetts. Each image is 1500×1500 pixels in size, covering an area of 2.25 km^2 . Because the image size is different from the GPS data set, we clip the image to a size of 1024×1024 pixels. The data set covers a wide variety of urban, suburban, and rural regions and covers an area of over 2600 km^2 . The label data are a binary image in which the road area pixel value corresponding to the high-resolution imagery is 1 and the background value is 0. The information of the data sets used in this article is shown in Table I.

2) *Models*: In order to validate the effectiveness of the training set labeled by GPS trajectories of floating car, LinkNet, D-LinkNet, and D-LinkNet-1D are used for road extraction. These models all have an encoder–decoder architecture. Models' structure is light and has high precision and efficient. They are image semantic segmentation neural networks which are widely used for road extraction from remote sensing imagery [27]–[30].

The LinkNet achieves a good balance between speed and accuracy. The residual operation of the jump link between the encoder and the decoder is introduced, so that the accuracy of the model is improved, while the speed is improved, and the most advanced precision level of image segmentation is achieved. The expansion convolution can expand the receptive field of the convolution kernel while keeping the number of parameters unchanged, and it can ensure that the size of the feature map of the output remains unchanged. There are usually two forms of expansion convolution, cascaded and parallel, both of which have the power to improve segmentation accuracy. D-LinkNet adds an expansion convolution to the center of LinkNet, which combines the advantages of both forms, and using a shortcut connection to combine the two forms. D-LinkNet-1D changes the shape of the transposed convolution in the decoder based on D-LinkNet to make it more suitable for road extraction. They are all recent network models used for road extraction, with high accuracy and fast training speed. We use DeepGlobe Road

TABLE I
INFORMATION OF THE DATA SETS

Dataset	Image size (pixels)	Number of images with labels	Location	Coverage area type
RBDD	1500×1500	1171	Massachusetts, USA	Urban, suburban, and rural
DeepGlobe	1024×1024	6226	Thailand, India, Indonesia	Urban, rural, coastal, and tropical rainforests
GPSTasTS	1024×1024	3456	Wuhan, China	Urban, suburban, and rural

TABLE II
IOU OF DIFFERENT TRAINING SET AND MODEL COMBINATIONS

Training set	Method	IoU on Wuhan test set	IoU on their own test set
RBDD	Link	0.287	0.686
	D-LinkNet	0.281	0.727
	D-LinkNet-1D	0.291	0.716
RBDD + GPSTasTS	Link	0.735	/
	D-LinkNet	0.756	/
DeepGlobe	D-LinkNet	0.475	0.750
	D-LinkNet-1D	0.478	0.758
DeepGlobe + GPSTasTS	Link	0.729	/
	D-LinkNet	0.764	/
	D-LinkNet-1D	0.795	/

Extraction Data set (DeepGlobe) and Road and Building Detection Data set (RBDD) for training and testing to evaluate the performance of the models, respectively; as shown in the last column of Table II, from IoU, they have a good effect of extracting roads.

B. Experimental Setup

Each CNN model is separately trained by four training sets to obtain two road extraction models. First, the DeepGlobe road extraction data set and Road and Building Detection Data set are used for training. After the training is completed, a satellite imagery road extraction model is obtained. Then, based on this, the model trained by 2856 pieces of sample data which are labeled by GPS trajectories raster, and the second road extraction model is obtained. The label data are a binary image in which the road area pixel value corresponding to the high-resolution imagery is 1 and the background value is 0.

The test set data account for about 20% of all sample data, which conform to the principle of data distribution for general deep learning. Cross-validation was not used during the training to prevent over-fitting, but data are increased through data transformation and color transformation. Morphological transformation methods for data increase include random horizontal folding, vertical folding, diagonal folding, image displacement, and image scaling. The color transformation is that an image transforms in the HSV space. Because data have increased, the validation set was not used. The CNN

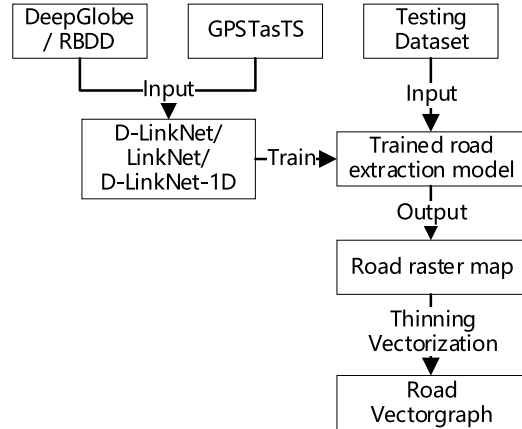


Fig. 10. Experimental setup of road extraction.

models are trained by different training sets to extract roads from satellite imagery.

LinkNet uses ResNet34 pretrained on the ImageNet[31] data set as an encoder for road extraction. D-LinkNet adds an expansion convolution of size 3×3 between the encoder and the decoder based on LinkNet. D-LinkNet-1D changes the transconvolution of the D-LinkNet decoder to a 1-D shape. After the training is completed, the test set data are used for testing, and the road extraction result of the test area is obtained.

The models are, respectively, predicted the Wuhan test set, and binary images with road information having the same format as the label are obtained. Then, use the road raster map refinement and vectorization method to generate the road vectorgraph of the testing area. Experimental setup of road extraction is shown in Fig. 10.

C. Result Analysis

To evaluate the performance of different models on the Wuhan test set, we calculated the mean IoU, as shown in the third column of Table II. In addition, we add an experiment that the three neural networks trained from scratch on “the DeepGlobe Road Extraction Data set” and “Road and Building Detection Data sets,” which evaluate the performance of networks and the quality of available labeled training data, as shown in the last column of Table II. It can be seen that the IoU using the open source data set training test is greater than 0.6, indicating that the network performance and data set quality are good. However, the IoU on the Wuhan test set of the models without the Wuhan training set is all less than 0.5. After adding the Wuhan training data, the performance has increased significantly.

TABLE III
ROAD EXTRACTION RESULTS EVALUATION

Training data	Model	Precision	Recall rate			F
			all roads	main roads	side roads	
RBDD	Link	0.785	0.403	0.560	0.279	0.533
	D-LinkNet	0.861	0.415	0.586	0.279	0.560
	D-LinkNet-ID	0.882	0.474	0.653	0.332	0.617
RBDD+GPSTasTS	Link	0.993	0.626	0.858	0.442	0.768
	D-LinkNet	0.992	0.647	0.879	0.463	0.783
	D-LinkNet-ID	0.994	0.642	0.872	0.460	0.780
DeepGlobe	Link	0.870	0.615	0.710	0.519	0.721
	D-LinkNet	0.895	0.591	0.703	0.478	0.712
	D-LinkNet-ID	0.898	0.617	0.738	0.494	0.731
DeepGlobe+GPSTasTS	Link	0.977	0.740	0.914	0.564	0.842
	D-LinkNet	0.990	0.745	0.917	0.571	0.850
	D-LinkNet-ID	0.991	0.655	0.883	0.473	0.788

The label data used in this article are obtained by GPS trajectories of floating car; it is not real road network data. This article uses the real road vectorgraph as the truth data for the evaluation of the results. The lowest urban road level in the true value is the residential group level road, that is, the road connecting the residential community in the community. The road is a linear element and its length is usually used as an index for evaluation. The extracted road results are in linear format with some slight deviation between the data. Therefore, this article refers to the linear element accuracy measurement algorithm base on buffer proposed in reference[32]. And the precision, the recall rate, and the comprehensive evaluation index F value are selected as the evaluation indexes.

The precision (P), the recall (R), and the F value (F) are defined as

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F = \frac{(\alpha^2 + 1)P \times R}{\alpha^2(P + R)}. \quad (7)$$

Among them, TP is to predict the positive class as a positive class, FP is to predict the negative class as a positive class, and FN is to predict the positive class as a negative class. The value of α in (3) is 1, which is also the most common F value

$$F = \frac{2 \times P \times R}{P + R}. \quad (8)$$

The extraction result obtained by direct vectorization of the trajectory raster data, the extraction results with only learning the available labeled training data, and the extraction result with learning the available labeled training data and the GPS trajectories labeled data set are set as the evaluation objects. The buffer is established according to width in the field of the true value. It is compared with the road vectorgraph to be evaluated. The length of the true positive TP, false positive FP, which is the length of the background that is mistaken

for the road, and false negative FN, which is the length of the road that is mistaken for the FN and the length of the road that is mistaken for the background is calculated. When calculating the recall rate, for more detailed analysis and evaluation, the true value data are also classified according to the road type field. It is divided into main roads including highway, first, second, and third grade road, and side roads including internal roads, factories and mines roads, and rural roads. Table III is an evaluation index of these road extraction results.

The precisions of the road results extracted by the three models with learning GPS trajectories raster labels are higher than 90%, while the precision of results extracted by models without GPS trajectories as training data is significantly reduced in precision, which is less than 90%. Without learning the floating car data, the extraction result will have many false positives. As the green arrows point to in Fig. 11(a), the railways in the satellite imagery are identified as a road mistakenly.

In terms of the recall of the road extraction results, the overall recall rates of the extractions with learning the GPS trajectories label are the highest, while the road extraction recall rates without the learning GPS trajectories label are the lowest. For the recall rate of the all roads, the recall of extraction base on the GPS trajectories of floating car is highest, which has little difference from the results by CNN models with learning floating car data. The extraction recall of the side roads is low. Relatively, the recalls of extractions based on CNN models with learning floating car data are higher than the others, while the recall of result extracted directly by GPS trajectories of floating car performs the worst. As far as the recall of the main road extraction, vehicles are mainly driven on the main road, so the integrity of extraction based on only GPS trajectories is the highest. The CNN models are able to extract most of the roads covered by buildings or trees after learning to the labels obtained by floating car data. The roads pointed by red arrows in Fig. 11(b)–(d) are in the shadow but is still extracted.

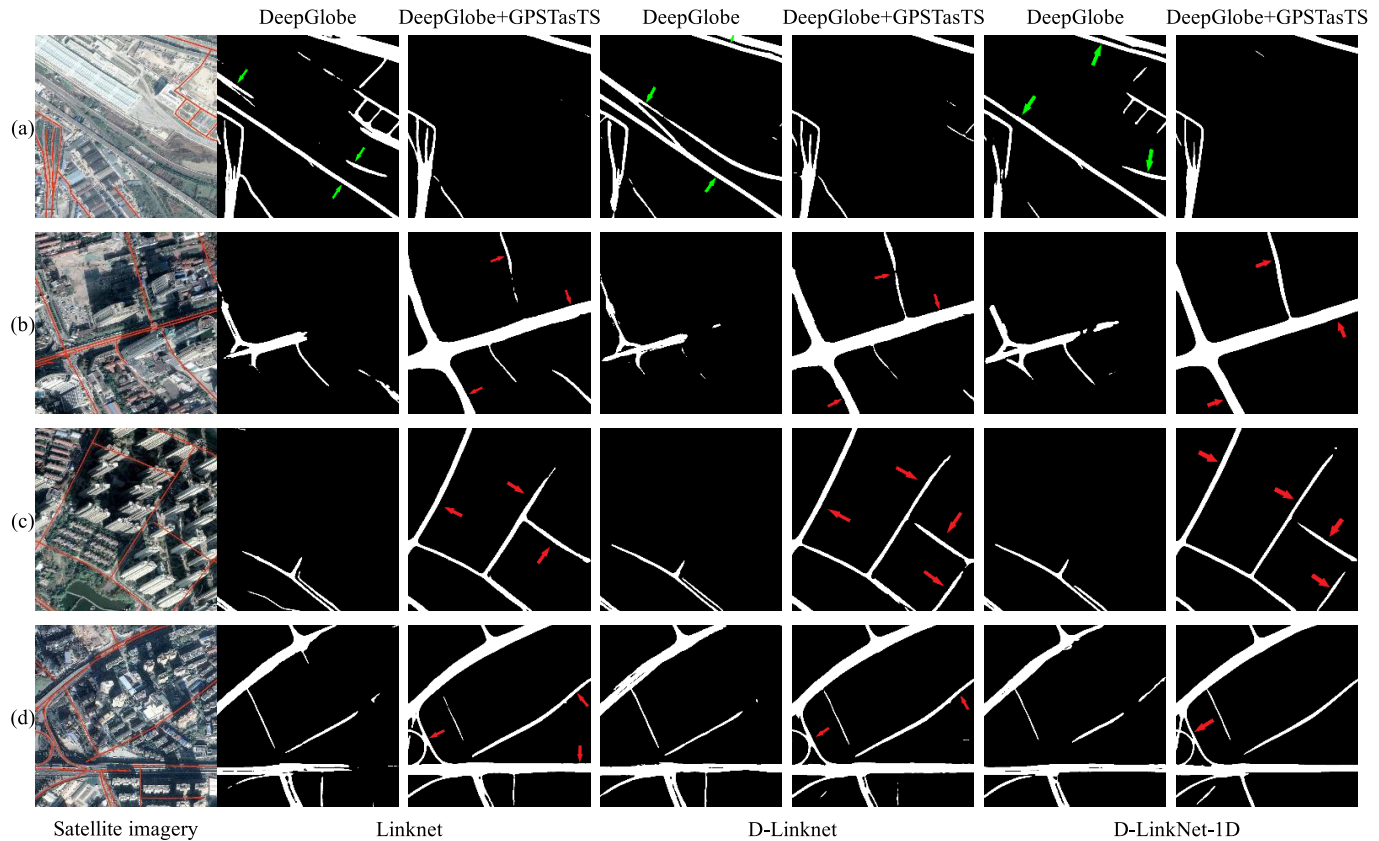


Fig. 11. Road extraction results using different methods and data sets (Take DeepGlobe as an example).

It can be seen from the comprehensive evaluation results that the result of each model with learning floating car data is better than those without learning. And the extractions obtained from satellite imagery based on models with learning floating car data are relatively good in all aspects and the comprehensive evaluation index is higher than the result extracted by only GPS trajectories.

IV. DISCUSSION

In this article, the advantages of floating car data and high-resolution satellite imagery are fully utilized, and a reasonable process is designed. The method of effectively combining the two to extract urban roads is proposed, and the extraction results are relatively good.

In terms of precision, the precision of road extraction results from only floating car data is the highest. This is because the floating car data do not contain other features, and interference information is less than satellite imagery. It is only the representation of road geometric features. The false positive in the extraction result from only GPS trajectories is mainly because the GPS trajectories have drifted away from the road. The road extraction results based on the CNN models with learning GPS trajectories label from satellite imagery are also excellent. Since the geometric features and spectral features of the railways, ditches, and river banks are very similar to the roads, a small number of linear features are identified roads mistakenly. In contrast, the results by the models without learning GPS trajectories label are significantly reduced in terms of precision. The models misjudge a large number of railways as roads, and the learning GPS trajectories labels

can effectively avoid such false positives and improve the precision.

In terms of the integrity of the road extraction results, for the main road, because the driving range of the floating cars is mainly concentrated on the main road, almost only the main road in the suburban area did not pass the vehicles of the experimental platform within the experimental time range, resulting in false negatives. Therefore, the integrity of extraction from only GPS trajectories is the best. The models can extract most of the main roads that are blocked by buildings and trees and located in the shadow area after learning to the GPS trajectories labels. However, some of the main roads are missed because the ground objects are severely blocked or the whole section is in the shadow. At the same time, due to the limitation of the receptive field size of the network, remote sensing imagery is divided into subimages which sizes are 1024×1024 . So that some roads are cut too short on the subimages. As a result, the geometric features are not obvious and the false negatives occur. Models that do not learn floating car data have poor integrity of the road extraction, and most of the roads which are obscured by buildings or trees, and within the shadows, are missing. For the side roads, the results extracted from only GPS trajectories are the worst, which is because the floating GPS trajectories of floating car have low coverage to the internal roads, rural roads, and factory mine roads. Roads without GPS trajectories cannot be extracted, which reduces the integrity of the extraction. The recall of results extracted from imagery by the neural network model is all higher than the extraction results from the GPS trajectories. Like the main roads, the integrity of results obtained by models

with learning floating car data is higher than that without learning GPS trajectories. However, due to the width of side roads is small and severely obscured by objects, the geometric features and regional features are not obvious on the imagery, resulting in low extraction integrity of the side roads.

Overall, the extractions from imagery based on CNN models with learning GPS trajectories have the best extraction effect, both in terms of extraction precision and recall. The precision of result and the recall of main roads extraction from only floating car data are the highest, while the integrity of the side road extraction is the worst. Models that do not learn floating car data have a poor effect on road extraction. It is mainly due to the difference in the sample set of training compared with the models with learning GPS trajectories. Without floating car data, the models only learn the available labeled training data. The DeepGlobe road extraction data set reflects the surface environment of Southeast Asia and Road and Building Detection Data set reflects North America. There are certain differences in the features of natural landscapes and humane landscapes, which affect the road extraction effect. It can be seen that adding floating car data to the training set in the process of neural networks training can improve the accuracy and integrity of the road extraction.

V. CONCLUSION

For the huge cost generated by hand-label training data in deep learning, a method of labeling targets with existing information is proposed. For starters, urban roads are extracted based on the GPS trajectories of floating car. The floating car are preprocessed with appropriate conditions, and then a GPS trajectories rasterization algorithm combining GPS trajectories points rasterization and linear interpolation rasterization is proposed to ensure data integrity and avoid data redundancy. After the GPS trajectories raster map is matched with the remote sensing imagery, the roads on the remote sensing imagery are automatically marked. It saves a lot of time and labor costs for manual labeling and avoids subjective judgment errors, making road extraction more accurate and efficient. At the last, CNN models are trained by the data set with GPS trajectories to extract roads from satellite imagery.

Make full use of the complementary characteristics of floating car data and high-resolution remote sensing imageries, focusing on the fusion of the two types of data. Make the advantages of both are fully utilized, and a road network with high accuracy and complete integrity is obtained. A good balance has been achieved in the overall efficiency and effectiveness of the complex road extraction in a large area of the city.

REFERENCES

- [1] T. Ishii, R. Nakamura, H. Nakada, Y. Mochizuki, and H. Ishikawa, "Surface object recognition with CNN and SVM in Landsat 8 images," in *Proc. 14th IAPR Int. Conf. Mach. Vis. Appl. (MVA)*, May 2015, pp. 341–344.
- [2] I. Sevo and A. Avramovic, "Convolutional neural network based automatic object detection on aerial images," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 5, pp. 740–744, May 2016.
- [3] H. Wu, H. Zhang, J. Zhang, and F. Xu, "Fast aircraft detection in satellite images based on convolutional neural networks," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 4210–4214.
- [4] I. Kahraman, M. K. Turan, and I. R. Karas, "Road detection from high satellite images using neural networks," *Int. J. Model. Optim.*, vol. 5, no. 4, p. 304, 2015.
- [5] W. Xia, N. Zhong, D. Geng, and L. Luo, "A weakly supervised road extraction approach via deep convolutional nets based image segmentation," in *Proc. Int. Workshop Remote Sens. Intell. Process. (RSIP)*, May 2017, pp. 1–5.
- [6] S. Saito, T. Yamashita, and Y. Aoki, "Multiple object extraction from aerial imagery with convolutional neural networks," *Electron. Imag.*, vol. 2016, no. 10, pp. 1–9, Feb. 2016.
- [7] Z. Zhang, Q. Liu, and Y. Wang, "Road extraction by deep residual U-Net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, May 2018.
- [8] A. Chaurasia and E. Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4.
- [9] L. Zhou, C. Zhang, and M. Wu, "D-LinkNet: LinkNet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 182–186.
- [10] T. Sun, Z. Di, P. Che, C. Liu, and Y. Wang, "Leveraging crowd-sourced GPS data for road extraction from aerial imagery," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7509–7518.
- [11] R. Jones, A. McCallum, K. Nigam, and E. Riloff, "Bootstrapping for text learning tasks," in *Proc. IJCAI Workshop Text Mining, Found., Techn. Appl.*, 1999, vol. 1, no. 7.
- [12] A. Gruen and H. Li, "Road extraction from aerial and satellite images by dynamic programming," *ISPRS J. Photogramm. Remote Sens.*, vol. 50, no. 4, pp. 11–20, Aug. 1995.
- [13] J. Cheng-Li, J. Ke-Feng, J. Yong-Mei, and K. Gang-Yao, "Road extraction from high-resolution SAR imagery using Hough transform," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2005, p. 4.
- [14] D. Yan and Z. Zhao, "Road detection from Quickbird fused image using IHS transform and morphology," in *Proc. IGARSS. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 6, Jul. 2003, pp. 3967–3969.
- [15] R. Liu, J. Song, Q. Miao, P. Xu, and Q. Xue, "Road centerlines extraction from high resolution images based on an improved directional segmentation and road probability," *Neurocomputing*, vol. 212, pp. 88–95, Nov. 2016.
- [16] C. Chen and Y. Cheng, "Roads digital map generation with multi-track GPS data," in *Proc. Int. Workshop Edu. Technol. Training Int. Workshop Geosci. Remote Sens.*, Dec. 2008, pp. 508–511.
- [17] J. Biagioni and J. Eriksson, "Map inference in the face of noise and disparity," in *Proc. 20th Int. Conf. Adv. Geographic Inf. Syst. (SIGSPATIAL)*, 2012, pp. 79–88.
- [18] X. Liu, J. Biagioni, J. Eriksson, Y. Wang, G. Forman, and Y. Zhu, "Mining large-scale, sparse GPS traces for map inference: Comparison of approaches," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2012, pp. 669–677.
- [19] X. Xie, K. B.-Y. Wong, H. Aghajan, P. Veelaert, and W. Philips, "Inferring directed road networks from GPS traces by track alignment," *ISPRS Int. J. Geo-Inf.*, vol. 4, no. 4, pp. 2446–2471, Nov. 2015.
- [20] J. Qiu and R. Wang, "Automatic extraction of road networks from GPS traces," *Photogramm. Eng. Remote Sens.*, vol. 82, no. 8, pp. 593–604, Aug. 2016.
- [21] T. Guo, K. Iwamura, and M. Koga, "Towards high accuracy road maps generation from massive GPS traces data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2007, pp. 667–670.
- [22] J. Wu, Y. Zhu, T. Ku, and L. Wang, "Detecting road intersections from coarse-gained GPS traces based on clustering," *J. Comput.*, vol. 8, no. 11, pp. 2959–2965, Nov. 2013.
- [23] J. Wang, J. Song, M. Chen, and Z. Yang, "Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine," *Int. J. Remote Sens.*, vol. 36, no. 12, pp. 3144–3169, Jun. 2015.
- [24] R. Langley, "Nmea 0183: A GPS Receiver," *GPS World*, vol. 6, no. 7, pp. 54–57, 1995.
- [25] I. Demir *et al.*, "DeepGlobe 2018: A challenge to parse the Earth through satellite images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 172–181.
- [26] V. Mnih, "Machine learning for aerial image labeling," Ph.D. dissertation, Dept. Comput. Sci., Univ. Toronto, Toronto, ON, Canada, 2013.
- [27] F. Wulff, B. Schaufele, O. Sawade, D. Becker, B. Henke, and I. Radusch, "Early fusion of camera and lidar for robust road detection based on U-Net FCN," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1426–1431.

- [28] O. Filin, A. Zapara, and S. Panchenko, "Road detection with EOSResUNet and post vectorizing algorithm," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 211–215.
- [29] T. Sun, Z. Chen, W. Yang, and Y. Wang, "Stacked U-Nets with multi-output for road extraction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 202–206.
- [30] A. Buslaev, S. Seferbekov, V. Iglovikov, and A. Shvets, "Fully convolutional network for automatic road extraction from satellite imagery," in *Proc. CVPR Workshops*, vol. 207, 2018, p. 210.
- [31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [32] M. F. Goodchild and G. J. Hunter, "A simple positional accuracy measure for linear features," *Int. J. Geographical Inf. Sci.*, vol. 11, no. 3, pp. 299–306, Apr. 1997.



Ju Zhang received the B.S. degree from the College of Nature Resources and Environment, Northwest Agriculture and Forestry University, Yangling, China, in 2016, and the M.S. degree from the School of Remote Sensing Information Engineering, Wuhan University, Wuhan, China, in 2019, where she is pursuing the Ph.D. degree with the School of Remote Sensing Information Engineering.

Her research interests include image processing, trajectory data mining, imagery classification, and deep learning.



Qingwu Hu received the B.Eng. and M.Eng. degrees in photogrammetry and remote sensing from the Wuhan Technical University of Surveying and Mapping, Wuhan, China, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, in 2007.

He was with Tennessee State University, Nashville, TN, USA. He is a Professor with the School of Remote Sensing Information Engineering, Wuhan University. He is a Guest Researcher with the State Key Laboratory of Geographic Information Engineering, Wuhan University, and the State Key Laboratory of Informatization of Rail Transit Engineering, Xi'an, China. He has authored more than 60 peer-reviewed articles in international journals from multiple domains, such as remote sensing and photogrammetry. His research interests include methods, techniques and applications of remote sensing, GIS and GPS integration, and photogrammetry.

Dr. Hu is a member of the Academic Committee of the Qingdao Hailu Geographic Information Integration and Application Key Laboratory. He has won two second prizes for national scientific and technological progress and two scientific and technological progress awards for surveying and mapping. In 2017, he was awarded the Talbert Abrams Award by the American Society for Photogrammetry and Remote Sensing (ASPRS). He is a National Natural Science Foundation Reviewer and a Reviewer of several international academic journals.



Jiayuan Li received the B.Eng., M.Eng., and Ph.D. degrees from the School of School of Remote Sensing Information Engineering, Wuhan University, Wuhan, China.

He is an Associate Researcher with Wuhan University. He has authored more than ten peer-reviewed articles in international journals. His research interests lie in remote sensing and computer vision. His research is mainly focused on image matching and feature registration.

Dr. Li was awarded the Talbert Abrams Award by ASPRS in 2017. He is a Reviewer for international journals such as the IEEE TRANSACTIONS ON IMAGE PROCESSING and the *International Journal of Remote Sensing*.



Mingyao Ai received the B.Eng. and M.Eng. degrees from the School of School of Remote Sensing Information Engineering, Wuhan University, Wuhan, China, where he is currently pursuing the Ph.D. degree.

He is a Technician with Wuhan University. His research interests include remote sensing and photogrammetry. His research is mainly focused on SLAM and 3-D reconstruction.