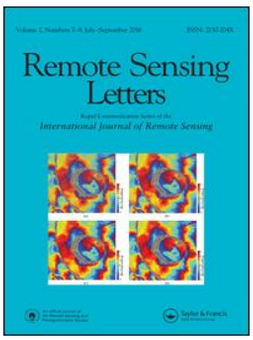

Remote Sensing Letters

ISSN: 2150-704X (Print) 2150-7058 (Online) Journal homepage: http://www.tandfonline.com/loi/trsl20

# Estimating Road Widths From Remote Sensing Images

Lun Luo, Yu Zang, Xiaofang Wang, Cheng Wang, Jonathan Li, Sheng Wu & Yuelei Liu

**To cite this article:** Lun Luo, Yu Zang, Xiaofang Wang, Cheng Wang, Jonathan Li, Sheng Wu & Yuelei Liu (2018) Estimating Road Widths From Remote Sensing Images, Remote Sensing Letters, 9:9, 819-828, DOI: 10.1080/2150704X.2018.1484957

**To link to this article:** https://doi.org/10.1080/2150704X.2018.1484957

Published online: 03 Jul 2018.

Submit your article to this journal

Article views: 104

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=trsl20

REMOTE SENSING LETTERS
2018, VOL. 9, NO. 9, 819–828
https://doi.org/10.1080/2150704X.2018.1484957

Taylor & Francis
Taylor & Francis Group

Check for updates

# Estimating Road Widths From Remote Sensing Images

Lun Luo[a,b], Yu Zang[a], Xiaofang Wang[a], Cheng Wang[a], Jonathan Li[a,c], Sheng Wu[d] and Yuelei Liu[e]

[a]Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen, China; [b]China Transport Telecommunications and Information Center, Bei Jing, China; [c]Department of Geography and Environmental Management, University of Waterloo, Waterloo, Canada; [d]Spatial Information Research Center of Fujian, Fuzhou University, Fuzhou, China; [e]School of Electronics and Control Engineering, Chang'An University, Xi'an, China

**ABSTRACT**

In this paper, we propose a learning-based approach for road width estimation from high resolution satellite or aerial images. With the road centerline given, a proposed pixel-wise descriptor describes the distribution of the road-like pixels. Then, in the predictive phase, we demonstrate that performance is significantly improved by considering the spatial coherence of the output labels. Experimental results show that the proposed approach provides high quality width estimation results on roads in various terrains.

**ARTICLE HISTORY**
Received 13 February 2018
Accepted 25 May 2018

## 1. Introduction

Road width estimation from satellite images is a fundamental problem in remote sensing image processing. Accurate road width information makes it easier to survey road networks in vast rural areas. In the past decade, road width estimation, relying on a steady road extraction method, has received much attention. Recent methods employ multi-stage (Unsalan and Beril 2012; Shi, Miao, and Debayle 2014; Zang et al. 2016) schemes to further improve road extraction performance, while the road width is estimated by some empirical schemes. Using spectral, shape, and gradient features, Ünsalan et al. (Unsalan and Beril 2012) constructed a graphical description of the road primitives and then proposed a graph-based topology analysis scheme to refine the road map. Based on a pre-trained spectral-spatial classifier, Shi et al. (Shi, Miao, and Debayle 2014) developed a multi-step road centerline extraction scheme, which significantly improves detection robustness. In these works, road width is estimated by an empirical approach (Unsalan and Beril 2012) or a spectral intensity approximation (Shi, Miao, and Debayle 2014).

Using an aperiodic directional structure measurement, Zang et al. (Zang et al. 2016, 2017) proposed a structure aware road extraction scheme by exploiting the aperiodic directional structure measurement along with a guided remote sensing image enhancing scheme. However, road width is simply estimated by an adaptive thresholding scheme.

On the other hand, the development of deep neural network (Hinton and Salakhutdinov 2006) provides a new idea for road network extraction. Mnih and Hinton (Volodymyr and

**CONTACT** Yu zang zangyu7@126.com Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen, FJ 361005, China

© 2018 Informa UK Limited, trading as Taylor & Francis Group

Hinton 2010) first proposed a multi-level network to assign each pixel a label to denote whether it belongs road region or not. Cheng (Cheng et al. 2017) proposed a cascaded end-to-end convolutional neural network (CasNet) to address the road segmentation and centerline extraction tasks, such approach work well for the urban roads with explicit spectral features. Wei et al. (Wei, Wang, and Mai 2017) modified traditional convolutional neural network to refine the road structure from aerial images. Saito et al. (Saito, Yamashita, and Aoki 2016) use the convolutional neural network to extract terrestrial objects from aerial imagery and got pretty good results. These learning based methods have provided interesting inspirations for the road width estimation.

Based on these works, we found that few of the road network extraction approaches proposed a systematic analysis of road properties, such as road width estimation or pavement material detection, etc. In this work, we propose a learning based approach to provide the accurate estimation of road width; the basic assumption is that, in previous works (Shi, Miao, and Debayle 2014; Zang et al. 2016)the road centerline was appropriately extracted. The main contribution of this paper relies on the following: 1) We propose a simple, yet effective, descriptor to describe the intensity distribution of the context at the road centerline pixel. Such a descriptor is then employed to train a CNN to provide a pixel-wise probability of the width at each road pixel. 2) In the predictive phase, we propose an optimization based label scheme, which takes into account the spatial coherence of the data. Such a scheme demonstrates significant improvement of the estimation performance in our experiments.

## 2. Methodology

The overall framework of our approach is shown in Figure 1. With the given road centerline and a set of manually labeled samples, we first propose a road width descriptor to describe the road pixel distribution on a local patch centered on the target pixel. Then, the

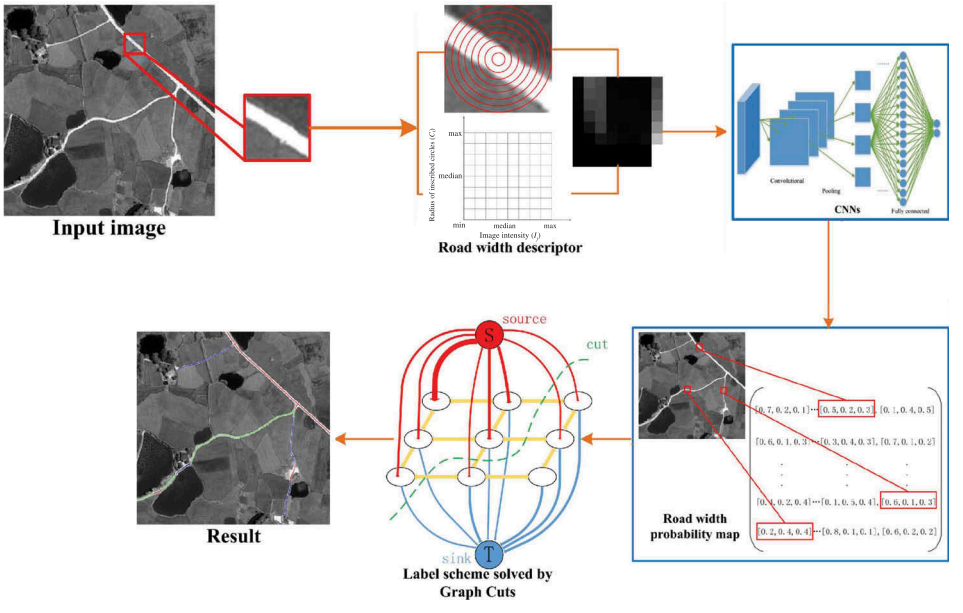


Figure 1. The workflow of our approach.

established features, along with the labels, are employed to train a CNN model to generate a probability map describing roads with different widths. For the prediction, by considering the spatial coherence of data, the labeling process is modeled by a proposed optimization problem, which can then be solved by classic graph cut methods.

## 2.1. *Road width descriptor*

Our descriptor is established based on the following two observations: 1) The spectral performances of road pixels and background are visually distinguishable; 2) For roads with different widths, the pixel intensity distribution at local areas can be very different. Based on these two observations, it is desired that the descriptor have identifiable responses to roads with different widths and be designed simply to acquire a satisfactory timely performance.

The descriptor is actually a two-dimensional histogram of the road pixel intensity distribution for a certain patch. One of the dimensions is the radius; the other dimension is the intensity. To cover the maximum width of a road, we first cut a patch around each road pixel. Then, to count the intensity distribution of the pixels around the center, we build an $8 \times 8$ 2D-histogram, the first dimension denotes the radius of the circle and the second dimension denotes the intensity of the image. The range of a road intensity (common range is [0, 255]) is quantified into eight equal subintervals–$I_j$ ($j \in \{1..8\}$). The radius of the maximum inscribed circle in a patch is also divided into eight equal subintervals to obtain eight concentric circles–$C_i$ ($i \in \{1..8\}$).

Figure 2 shows an example of the descriptor structure. Figure 2 (*a*) shows a typical patch center at a road centerline pixel. Around the center pixel, eight concentric circles (highlighted in red) are established; the radius of the maximum circle is half the patch length. Figure 2 (*b*) shows the structure of the descriptor; the radius of the maximum inscribed circle and the maximum intensity of the pixels in the patch are divided into eight equal subintervals to form an $8 \times 8$ matrix.

The value of each element, denoted as $N_{i,j}$, represents, in the circle, $C_i$, the ratio between the number of pixels (with intensity in the range $I_{j-1}I_j$) and the total number of pixels. Mathematically, the matrix is expressed as follows:

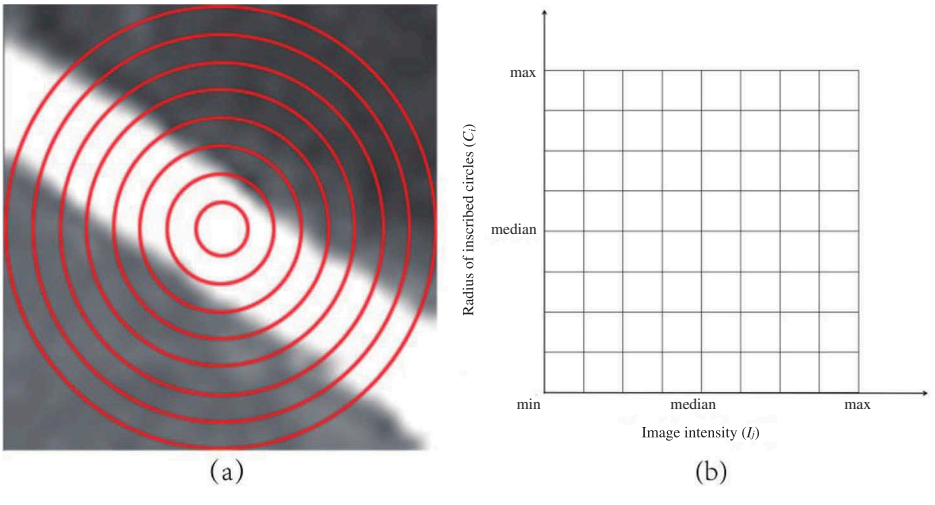


**Figure 2.** Example of a descriptor. (*a*) is the visualized result of the constructed concentric circles; (*b*) is the structure of the descriptor.

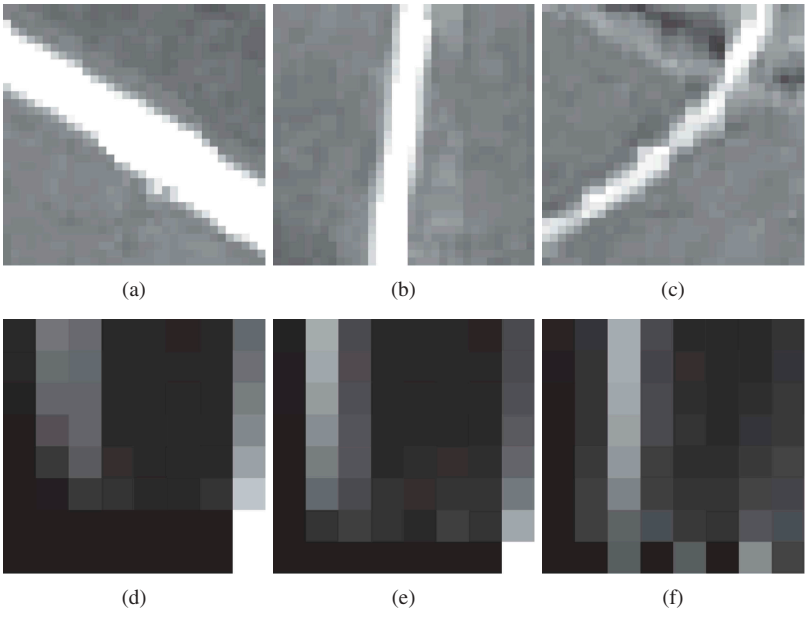


Figure 3. The diagram of the descriptor. Where (*a*)-(*c*) are road patches with different widths, (*d*)-(*f*) are the visualized results of the proposed descriptor.

$$V(j,i) = \frac{N_{i,j} \times 255}{K_i} \tag{1}$$

where $V(j,i)$ is the visualized value of each element in the matrix; $K_i$ is the total number of pixels in circle, $C_i$. Then, to form an image for visualizing and training, all the values in our matrix are re-scaled to [0, 255].

Figure 3 shows a group of examples of our descriptor. The patches of roads with various widths are shown in Figure 3 (*a*) to Figure 3 (*c*). The response of the proposed descriptor is shown in Figure 3 (*d*) to Figure 3 (*f*). With a change in road width, the intensity distribution of the patches is very different. For the wider road, the pixel intensities in the innermost circle tend to be distributed in the brightest (or darkest) area, depending on the spectral performance of the road. By contrast the pixel intensities for a narrow road appear to distribute uniformly.

## 2.2. *Road width estimation via cnns*

Using the designed width descriptor along with a set of manually labeled training samples, CNNs (Krizhevsky, Sutskever, and Hinton 2012) are employed to provide the initial probability that each road centerline pixel belongs to each width category. Training a set of weights, W, and bias, *b*, we find the minimum error between the test data and the samples, which is written as follows:

$$\arg\min_{\mathrm{W},b}\left(\sum_{l_i \in \mathrm{L}} |P_{l_i}(\mathrm{I}(i,j), \mathrm{N_I}(i,j)) - P_{l_i}(\mathrm{S}(i,j), \mathrm{N_S}(i,j)|\right) \tag{2}$$

where $P_{l_i}(\cdot)$ describes the probability that the road centerline pixel belongs to the width class, $l_i$. L is the set of all the width labels. $\mathrm{N_I}(i,j)$ and $\mathrm{N_S}(i,j)$ are certain neighborhoods of the input image $\mathrm{I}(i,j)$ and training samples $\mathrm{S}(i,j)$, respectively, centered at position, $(i,j)$.

As shown in Figure 4, the structure of the network contains three convolutional layers and one fully connected layer. The kernel number is 32 for the first and second convolutional layers, and 64 for the third layer. The kernel sizes of all the three convolutional layers are $5 \times 5$, and the kernel size of the pooling layer is $3 \times 3$. The convolutional, local response normalization, pooling, and fully connected layers are denoted by Conv, Lrn, Pool, and Fc, respectively.

At training time, a neural network is trained by minimizing the cross-entropy loss between the predicted sample and the training sample. Denoting label domain, $\mathrm{L} = \{l_1, l_2, \cdots l_n\}$, and the score function of the $i$th sample, $x_i$, under weight, $w_{l_j}$, as:

$$S_i^{l_j}(x_i, w_{l_j}) = x_i \times w_{l_j} + b \tag{3}$$

Then, for each sample, $x_i$, the cost function is written as:

$$E_i = -S_i^{l_T} - \log_{10} C_i + \log_{10} \sum_{j=l_1}^{l_n} e^{S_i^{l_j} + \log_{10} C_i} \tag{4}$$

where $l_T$ is the manually labeled width class of sample, $x_i$; $C_i$, a constant for program implementation, is set in our system at the value of $\log_{10} C_i = -\max S_i^{l_j}$.

With these definitions and employing the stochastic gradient descent algorithm, we find the set of optimal weights, W, and bias, $b$. Specifically, the classification dataset is divided into three parts: training, validation, and test. Each time, we use 256 examples of the training and validation dataset to train the model. The weights, $w_{i+1}$, are updated by a momentum value, $\delta$, as follows:

$$w_{i+1} := w_i + \delta_{i+1} \tag{5}$$

where $i$ is the iteration index and $\delta_{i+1}$ is the momentum variable. The whole training process is repeated until either of the following two conditions is satisfied: (1) The

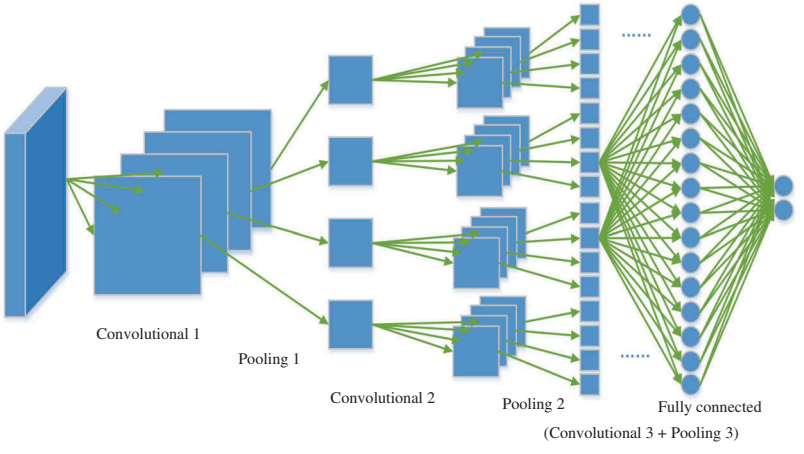


Figure 4. The structure of our learning network.

accuracy of prediction in the test dataset is greater than 95%; (2) There are a total of 65,000 iterations. The update rule for weight, $w$, is written as:

$$\delta_{i+1} := 0.9 \times \delta - 0.0005 \times \xi \times w_i - \xi \times \left\langle \frac{\partial L}{\partial w_i} \right\rangle_{D_i} \tag{6}$$

Here the momentum is set at 0.9, and weight decay is set at 0.0005; the learning rate, $\xi$, is set at 0.0001. $\left\langle \frac{\partial L}{\partial w_i} \right\rangle_{D_i}$ is the average over the $i$th batch, $D_i$, of the derivative of the objective with respect to W, evaluated at $w_i$. After the model is trained, a pixel-wise probability map on road centerlines is constructed. Each element of the map is a $n$-dimensional vector denoting the probability that the pixel belongs to each width category -$l_i, i \in [1, n]$. With this map, a labeling scheme is designed to provide the final estimation of the road width.

## 2.3. *Spatial coherent width labeling*

Using the initial probability distribution, we assign a label to each road centerline pixel to denote which width class it belongs to. A straight forward way to achieve this is to confirm the label according to the maximum element of the probability vector. However, there is a problem inherent in such a simple way. As shown in Figure 5, different spectral bands represent different widths. The widest and narrowest roads are denoted by red and blue, respectively. The zoomed-in patches show some typical 'misrecognition' points. Such cases occur because adjacent pixels may acquire different labels.

Obviously, a more reasonable way to assign the label is to not only consider the probability value, but also take into account the spatial coherence of the pixels. In other words, two adjacent pixels should have the same label. Based on such a consideration, we propose an optimization scheme to model the labeling problem. Specifically, denoting the label set $L = \{l_1 \cdots l_n\}$, the cost function is written as:

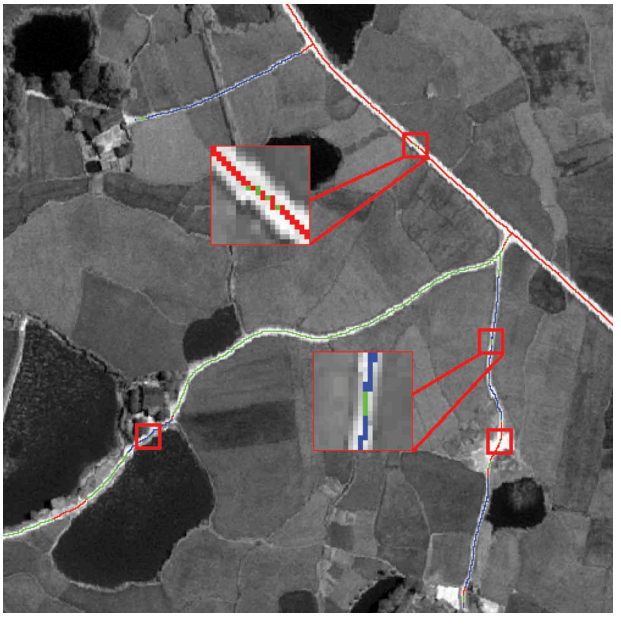

Figure 5. A classic result by using the straightforward labeling scheme, in which the widest, narrowest and middle width roads are denoted by red, blue and green, respectively.

$$\arg\min_{\mathrm{L}}(D(P) + \alpha S(\mathrm{L})) \tag{7}$$

where $D(\cdot)$ is the data term defined by the probability map, $P$; $S(\cdot)$ is the smoothing term defined by the label set, L. The data term is written as:

$$D(P) = 1 - \mathrm{prob}(p, l_i); \tag{8}$$

where $\mathrm{prob}(p, l_i)$ is the probability that pixel $p$ belongs to the width class $l_i$. $\alpha$ is a balancing factor to determine how the spatial coherence affects the results. In our experiments, the typical range of $\alpha$ is [0.1, 1].

The smoothing term, $S(\cdot)$, designed to indicate the influence from its neighbors when labeling the pixel, $p$, is defined as:

$$S(l_p, l_q) = \sum_{q \in \mathrm{N}_p} \frac{\mathrm{sign}(l_p, l_q)}{||p - q||} \tag{9}$$

where $\mathrm{N}_p$ is a certain neighborhood of pixel $p$. Here $\mathrm{N}_p$, containing twelve nearest neighbors of pixel, $p$, is defined in a one-dimensional domain. The width classes of the road patch centered at pixels, $p$ and $q$, are indicated by $l_p$ and $l_q$, respectively. The sign function $\mathrm{sign}(l_p, l_q)$ equals zero if $l_p = l_q$; otherwise, it equals one. $||p - q||$ indicates the distance between pixels $p$ and $q$.

Previous work (Boykov and Funka-Lea 2006), demonstrated that such an optimization model, having the classical form of Markov Random Fields (MRF), is solvable by classical graph cut methods. The essence of the smoothing term is to add a penalty for two adjacent pixels with different labels. The closer the two pixels are, the larger the penalty, thereby attaining coherent spatial labeling results.

## 3. Results and analysis

Our approach is implemented based on C++ programming on a PC with a 3.2GHz Intel Core i5–3200 CPU and 4G RAM. Several experiments were designed to comprehensively evaluate the algorithm. First, applied for testing were images from different satellite sensors, such as Pleiades-1A, Gaofen2 and Aerial, which involved various motorways, arterial highways and rural roads. We also evaluated our approach on the Pleiades-1A remote sensing image of Shaoshan City, China, where reference data was obtained by ground survey and provided by the China Transport Telecommunications & Information Center.

### 3.1. *Evaluation on images from various satellites*

*Road width descriptor evaluation*. To evaluate the observations of our road width descriptor, we apply the simple *k*-means algorithm to divide 5000 samples into different classes. These samples are collected from various satellites, such as Pleiades-1A, Gaofen2 satellites and Aerial images. The test region include the city of Shaoshan and Yiyang, in Hunan province, China. All the samples are downsampled to a uniform resolution (1 m). In the experiment, the number of class categories is chosen from 2–5, and corresponding classification correctness are 85.7%, 83.2%,74.5% and 67.3% respectively. Overall, the result is consistent with our

expectation when the number of class categories is appropriately set. While as the increase of the number of class categories, the size of the descriptor should also change accordingly.

*Road width estimation evaluation*. To comprehensively evaluate the performance of our approach, inputs from various sensors, such as the Pleiades-1A, Gaofen2 satellites, and Aerial image (Pleiades-1A satellite and Aerial images are with resolution of 0.5 m, Gaofen2 satellite image is with resolution 1 m), are applied for testing. Figure 6 shows the results of our road width estimation approach. Manually labeled ground truths, corresponding to the Pleiades-1A, Gaofen2 satellites, and Aerial images are shown in Figure 6 (*a*), (*b*), (*c*), respectively. Roads with different widths are marked with different spectral bands. Table 1 shows the relationship between the spectral bands and road widths. The roads in the Pleiades-1A image are basically from a rural area, and the widths range from 3m to 6m; the roads in the Gaofen2 and Aerial images are from a rural-urban fringe area, and the widths range from 4 m to 25 m.

Our estimated results are shown in Figure 6 (*d*), (*e*), (*f*). As shown, most of the errors occur in the crossroads, or areas where the roads are mixed with the adjacent buildings. In these cases the intensity distribution of the road patches is significantly changed, thus leading to the errors, as shown in the zoomed in patches. However, these errors are easily corrected in the following road network vectorization process. The final performance, in practical applications, is even better.

## 3.2. *Evaluation on various regions*

We also select three typical road width estimation works including the proposed approach, and evaluate the kilometre statistics of these works on the remote sensing image of Shaoshan City. The road centerline of this paper and approach (Xia et al. 2017) are extracted by (Zang et al. 2016). The image is recorded by a Pleiades-1A satellite with resolution 0.5 m and size

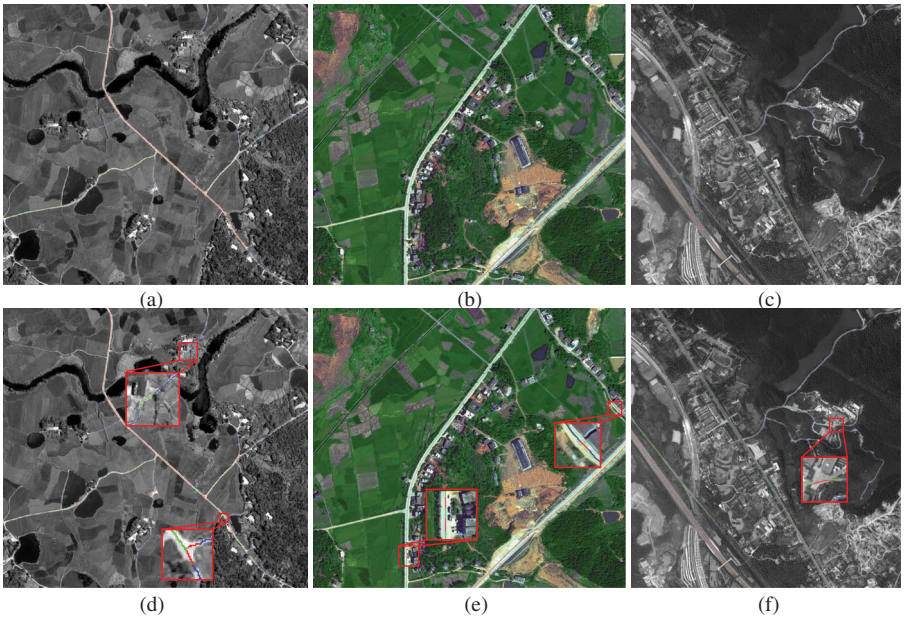


Figure 6. Our road width estimation results. (*a*)-(*c*)are the manually labeled ground truthes corresponding to the Pleiades-1A, Gaofen2 satellites and Aerial images; (*d*)-(*f*) are our results. In these figures, red, blue and green indicate the widest, narrowest and middle width roads, respectively.

**Table 1.** Correspondence relationship between various spectral bands and road width ranges.

| Sensor | Road width in red band (m) | Road width in green band (m) | Road width in yellow band (m) | Road width in blue band (m) |
|---|---|---|---|---|
| Pleiades-1A | 3.8–5.5 | 3.1–3.8 | N/A | 2.5-3.1 |
| Gaofen2 | 14.0–20.0 | 8.0–14.0 | 4.5–8.0 | 3.0–4.5 |
| Aerial image | 18.0–26.0 | 9.0–18.0 | N/A | 4.0–9.0 |

$28,648 \times 37,929$ pixels. The ground truth of the kilometres of roads with different widths are provided by the China Transport Telecommunications & Information Center. In such experiment, we employ 3000 samples to train the whole CNNs, and the training time is less than two hours due to the created width descriptor. Then in the test phase, each road centerline pixel is assigned a width label, and it takes about 4 hours for the entire Shaoshan image. The statistical results are shown in Table 2. Where all the roads, according to the technical standard, are quantified in four levels, as shown in Row 1; the ground truth along with the results of the three methods are shown in Rows 2 to 5, respectively.

As the results show, most of the roads in Shaoshan City are country roads with widths less than 10 m, and the proposed approach got the best ratio of the relative error (less than 10%). Also the proposed approach got the best performance for the motorways and arterial highways, which have widths ranging from 20 m to 25 m, with relative error only 3.6%, which is satisfactory. The relative errors of all the three methods for the feeder highways, ranging from 10 m 20 m, are fairly high because the total length of these roads is rather small.

To evaluate the estimation accuracy, we divide the image into small patches (with size $1000 \times 1000$ pixels), and collect the statistics of the approaches (Xia et al. 2017; Guan, Wang, and Yao 2010) and our method. The images are collected by Pleiades-1A and Gaofen2 satellites and Aerial image for Shaoshan and Yiyang city, in Hunan province, China. The image sizes are $28,648 \times 37,929$, $17,656 \times 28,413$ and $22,231 \times 29,874$ pixels, and the resolutions are 0.5 m, 1 m and 0.5 m respectively. All the samples are rescaled to the same resolution (1 m), and use the same road width classification standard, as shown in Table 2. We also compare the results to the latest deep neural networks Residual Network (ResNet) (He et al. 2016) and VGG (Simonyan and Zisserman 2015), and for both of the networks, we also add our labeling scheme. The results are shown in Table 3. From the results, it is viewed that the performance of the learning based approaches are rather impressive, while the statistics between ResNet, VGG and our approach are rather close, and in this paper we choose simple CNNs for its efficiency and simple network structure.

**Table 2.** Evaluation results of recent works for lengths of roads with various widths.

| Road width range (m) | 3–5 | 5–10 | 10–20 | 20–25 |
|---|---|---|---|---|
| The true lengths of roads with different widths(km) | 591 | 211 | 33 | 167 |
| Xia et al. (2017) road length (km) and error (%) | 493 (16.6%) | 279 (32.2%) | 20 (39.4 %) | 190 (13.8%) |
| Guan, Wang, and Yao (2010) road length (km) and error (%) | 517 (12.5 %) | 273 (29.3%) | 42 (27.3%) | 215 (28.7%) |
| The proposed method road length (km) and error (%) | 543 (8.1%) | 227 (7.5%) | 51 (54.5%) | 161 (3.6%) |

**Table 3.** Road width estimation correctness of various methods.

| | Results of Xia et al. (2017) | Results of Guan, Wang, and Yao (2010) | Results of ResNet | Results of VGG | Results of proposed method |
|---|---|---|---|---|---|
| Average classification correctness | 69.25% | 72.82% | 83.24% | 82.00% | 81.60% |

## 4. Conclusions

In this letter, we presented a learning based approach for accurate and stable road width evaluation from remote sensing images. The main contribution of this paper relies on the following: (1) a simple, yet effective, descriptor to describe the context of the road region distribution at the center pixel, and (2) an optimization based label scheme, which takes into account the spatial coherence of the data. Experiments on images from various satellites and statistics about Shaoshan city show that our approach provides stable evaluation results for various kinds of roads.

## Funding

This work was supported by the National Natural Science Foundation of China [61501387];

## References

Boykov, Y., and G. Funka-Lea. 2006. "Graph Cuts and Efficient N-D Image Segmentation." International Journal of Computer Vision 70 (2): 109–131. doi:10.1007/s11263-006-7934-5.

Cheng, G., Y. Wang, X. Shibiao, H. Wang, S. Xiang, and C. Pan. 2017. "Automatic Road Detection and Centerline Extraction via Cascaded End-to-End Convolutional Neural Network." *IEEE Transactions on Geoscience and Remote Sensing PP* 99: 1–16.

Guan, J., Z. Wang, and X. Yao. 2010. "A New Approach for Road Centerlines Extraction and Width Estimation." In *IEEE International Conference on Signal Processing*, Bangalore, India, 924–927.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. "Deep Residual Learning for Image Recognition." In *Computer Vision and Pattern Recognition*, IEEE, 770–778.

Hinton, G. E., and R. Salakhutdinov. 2006. "Reducing the Dimensionality of Data with Neural Networks." *Science* 313 (5786): 504. doi:10.1126/science.1127647.

Krizhevsky, A., I. Sutskever, and G. E. Hinton. 2012. "ImageNet Classification with Deep Convolutional Neural Networks." In *International Conference on Neural Information Processing Systems*, Doha, Qatar, 1097–1105.

Saito, S., T. Yamashita, and Y. Aoki. 2016. "Multiple Object Extraction from Aerial Imagery with Convolutional Neural Networks." *Journal of Imaging Science and Technology* 60 (1): 10402–1/10402–9. doi:10.2352/J.ImagingSci.Technol.2016.60.1.010402.

Shi, W., Z. Miao, and J. Debayle. 2014. "An Integrated Method for Urban Main-Road Centerline Extraction from Optical Remotely Sensed Imagery." *IEEE Transactions on Geoscience and Remote Sensing* 52 (6): 3359–3372. doi:10.1109/TGRS.2013.2272593.

Simonyan, K., and A. Zisserman. 2015. "Very Deep Convolutional Networks for LargeScale Image Recognition." In *International Conference on Learning Representations*, San Diego, USA.

Unsalan, C., and S. Beril. 2012. "Road Network Detection Using Probabilistic and Graph Theoretical Methods." *IEEE Transactions on Geoscience and Remote Sensing* 50 (11): 4441–4453. doi:10.1109/TGRS.2012.2190078.

Volodymyr, M., and G. E. Hinton. 2010. "Learning to Detect Roads in High-Resolution Aerial Images." In *European Conference on Computer Vision*, Springer, 210–223.

Wei, Y., Z. Wang, and X. Mai. 2017. "Road Structure Refined CNN for Road Extraction in Aerial Image." *IEEE Geoscience and Remote Sensing Letters* 14 (5): 709–713. doi:10.1109/LGRS.2017.2672734.

Xia, Z., Y. Zang, C. Wang, and L. Jonathan. 2017. "Road Width Measurement from Remote Sensing Images." In *2017 IEEE International Geoscience and Remote Sensing Symposium*, Fort Worth, USA, 902–905.

Zang, Y., C. Wang, L. Cao, and Y. Yao. 2016. "Road Network Extraction via Aperiodic Directional Structure Measurement." *IEEE Transactions on Geoscience and Remote Sensing* 54 (6): 3322–3335. doi:10.1109/TGRS.2016.2514602.

Zang, Y., C. Wang, Y. Yao, L. Luo, K. Yang, and L. Jonathan. 2017. "Joint Enhancing Filtering for Road Network Extraction." *IEEE Transactions on Geoscience and Remote Sensing* 55 (3): 1511–1525. doi:10.1109/TGRS.2016.2626378.