


ITC 1/50 Information Technology and Control Vol. 50 / No. 1 / 2021 pp. 89-101 DOI 10.5755/j01.itc.50.1.27987	A GF-3 SAR Image Dataset of Road Segmentation	
	Received 2020/11/09	Accepted after revision 2020/11/23
	 http://dx.doi.org/10.5755/j01.itc.50.1.27987	

HOW TO CITE: Sun, Z., Zhao, M., Jia, B. (2021). A GF-3 SAR Image Dataset of Road Segmentation. *Information Technology and Control*, 50(1), 89-101. <https://doi.org/10.5755/j01.itc.50.1.27987>

A GF-3 SAR Image Dataset of Road Segmentation

Zengguo Sun

Key Laboratory of Modern Teaching Technology, Ministry of Education, Xi'an Shaanxi 710062, China
School of Computer Science, Shaanxi Normal University, Xi'an 710119, China

Mingmin Zhao, Bai Jia

School of Computer Science, Shaanxi Normal University, Xi'an 710119, China

Corresponding author: sunzg@snnu.edu.cn

We constructed a GF-3 SAR image dataset based on road segmentation to boost the development of GF-3 synthetic aperture radar (SAR) image road segmentation technology and make GF-3 SAR images be applied to practice better. We selected 23 scenes of GF-3 SAR images in Shaanxi, China, cut them into road chips with 512×512 pixels, and then labeled the dataset using LabelMe labeling tool. The dataset consists of 10026 road chips, and these road images are from different GF-3 imaging modes, so there is diversity in resolution and polarization. Three segmentation algorithms such as Multi-task Network Cascades (MNC), Fully Convolutional Instance-aware Semantic Segmentation (FCIS), and Mask Region Convolutional Neural Networks (Mask R-CNN) are trained by using the dataset. The experimental result measures including Average Precision (AP) and Intersection over Union (IoU) show that segmentation algorithms work well with this dataset, and the segmentation accuracy of Mask R-CNN is the best, which demonstrates the validity of the dataset we constructed.

KEYWORDS: GF-3 SAR image dataset; road segmentation; deep learning; Mask R-CNN.

1. Introduction

Synthetic aperture radar (SAR) is a microwave sensor that uses an active mode of operation. The radar sensor itself emits a pulse of energy to the ground, and at the same time receives signals scattered from the

ground surface for ground detection. Due to the active imaging mode, SAR emits microwaves that can penetrate clouds and fogs, so it has the ability to obtain ground information all day and all weather. Therefore,

SAR is an indispensable remote sensing sensor in the field of ground observation, and plays an important role in both military and civilian fields [9, 13, 32, 46].

The GF-3 satellite, launched on August 10, 2016, is China's first C-band multi-polarized SAR satellite with the 1-meter resolution. The GF-3 satellite has the most imaging modes in the world (12 types in 5 categories), and its main technical indicators have reached or exceeded the level of similar international satellites. Therefore, the GF-3 SAR images have been widely used in agriculture, disaster monitoring, geology, lakes, ocean monitoring, navigation and other fields [18, 37].

As road segmentation in SAR images is very important for national economy and people's livelihood, such as transport system, urban development, residential life, and industrial distribution [19], it has been a research hotspot in the field of SAR image interpretation. Yang et al. addressed the fusion of image and point cloud data for road detection [45]. An proposed a method for extracting roads in complex scenes using the optimized Hough algorithm [1], Fu et al. proposed high-resolution remote sensing images road extraction based on wavelet transform and Hough transform [15], Geman et al. presented a new approach to tracking roads from satellite images [17], Tsutsui et al. presented an approach to road segmentation that only requires image-level annotations at training time [39], and Cheng et al. proposed a novel method of fusing geometric and appearance cues for road surface segmentation [5]. However, SAR imaging is not the same as optical imaging, and its characterization is not intuitive. The phenomena such as overlays during imaging will interfere with the target interpretation, and due to the interference of speckle, bright and dark grainy noise appears in images, which seriously affects the interpretation of images and the extraction of road features [47]. In deep learning, there is less research on road segmentation because road segmentation is much more difficult than road detection. The aim of road detection is to draw a border around the road, predict its label, and only output the presence or absence of the road and position information; while the road segmentation outputs the position information of the road, predicts the contour and area of the road, and further outputs the shape of the road.

With the development of deep learning technology, there have been many segmentation algorithms us-

ing deep neural network models. Commonly used network models are convolutional neural networks. Basic networks such as AlexNet [24], VGG [35], GoogleNet [38], and ResNet [21] have successively appeared, and segmentation models based on this structure have appeared, including Multi-task Network Cascades (MNC) [10], Fully Convolutional Instance-aware Semantic Segmentation (FCIS) [28], and Mask Region Convolutional Neural Networks (Mask R-CNN) [20]. MNC is mainly the application of multi-task learning in the case perception segmentation. FCIS is the first full convolution, end-to-end solution for image instance segmentation task. Without using any skills, Mask R-CNN is superior to all existing single-model networks, and it is recognized as a flexible, efficient, and universal segmentation architecture at present [20]. Deep learning methods often require a large amount of training sample data as a support. There are more public sample datasets in computer vision, such as ImageNet [11], PASCAL VOC [14], and COCO [27], and the data size reaches thousands of objects and millions of chips. In the field of optical remote sensing, the public datasets mainly include NWPU-RESISC45 [3], DIOR [25], NWPU VHR-10 [2, 4, 6], DOTA [43], HRRSD [31], and RSOD [30, 44]. In the research field of SAR image, there are mainly the AIR-SARShip-1.0 [36] dataset and the dataset mentioned in [41]. Above datasets have facilitated numerous studies. However, due to the extremely high accuracy required for sample labeling in road segmentation, public and free datasets of road segmentation for SAR images are very scarce, which seriously affects the development of deep learning technology of road segmentation for GF-3 SAR images. Only by constructing a set of road sample databases suitable for road segmentation can we promote research on SAR road segmentation and make GF-3 SAR images better serve the national road planning, urban construction and other aspects [29, 42].

A GF-3 SAR image dataset of road segmentation is constructed in this paper, named SARroad. This dataset contains 10026 image chips, which are from 23 scenes of GF-3 SAR images with size of 512×512 pixels each chip. The imaging modes include Spotlight (SL), Ultra-Fine Strip (UFS), Fine Strip I (FSI), and Fine Strip II (FSII), and the corresponding resolution is 1m, 3m, 5m, and 10m, respectively. The three segmentation algorithms of MNC, FCIS, and Mask

R-CNN are trained by our dataset, and the testing results demonstrate the validity of the dataset we constructed. It is necessary to note that our dataset is publicly available on the website [33, 34].

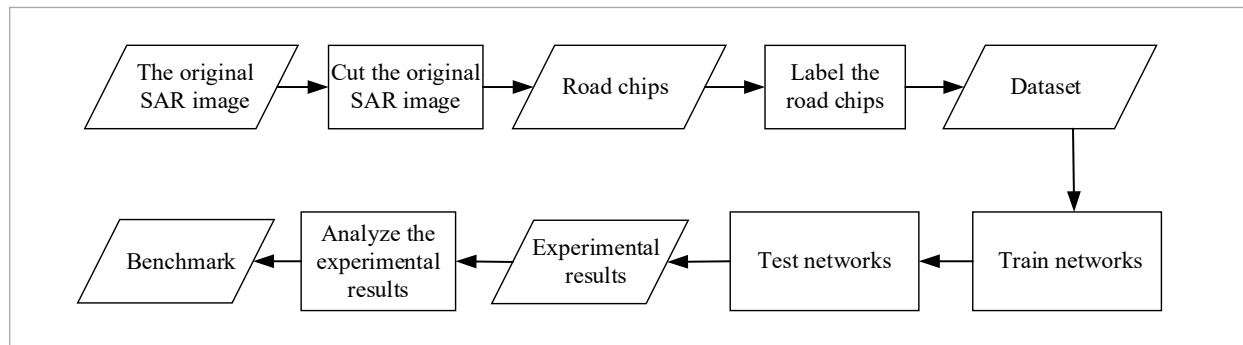
2. Methodology

As shown in Figure 1, we firstly selected 23 original GF-3 SAR images, cut them into 10026 road chips with 512×512 pixels, then we used the LabelMe label-

ing tool to segment and label the roads, and finally we constructed the dataset. We used the training dataset to train MNC, FCIS, and Mask R-CNN, and then used the testing dataset to test these three networks. We used a combination of quantitative analysis and qualitative analysis to analyze the experimental results in detail. Furthermore, we evaluated the accuracy of the three models, and the segmentation accuracy of Mask R-CNN is the best. Finally, we formed a benchmark through experimental comparison analysis.

Figure 1

Methodological steps



3. Construction of SAR Image Dataset of Road Segmentation

3.1. The Original SAR Image

In this paper, 23 scenes of GF-3 SAR images of Shaanxi Province in China were used to construct the dataset. All the GF-3 SAR images are obtained from China Centre for Resources Satellite Data and Application, which is the national remote sensing satellite data center with the largest number of civil land observation satellites in orbit in China. The GF-3 SAR

images in FSI mode and FSII mode were taken in 2017. The GF-3 SAR images in SL mode were taken in 2018. The GF-3 SAR images in UFS mode were taken in 2019. Table 1 gives the detailed information of the GF-3 SAR image dataset. The images have resolutions of 1m, 3m, 5m, and 10m, so there is a big difference among the images. Figure 2 is a GF-3 SAR image in SL mode, with various terrains of road, lake, farmland, and town. Lakes usually have lower gray values due to specular scattering, towns usually have higher gray values due to dihedral or trihedral scattering, and roads have the gray values in between.

Table 1

Details of the road dataset of GF-3 SAR image

Imaging Mode	Resolution (m)	Incident Angle (°)	Swath (km)	Polarization	Number of Images	Number of Road Chips
SL	1	20-50	10	Single	6	4775
UFS	3	20-50	30	Single	7	2193
FSI	5	19-50	50	Dual	3	1471
FS II	10	19-50	100	Dual	7	1587

Figure 2

GF-3 SAR image in SL mode



3.2. Construction Strategy of Road Segmentation Dataset

The construction process of the dataset is shown in Figure 3.

Firstly, 23 scenes of original GF-3 SAR images have a size of pixels around 13200×24300 . Sub-images containing roads in GF-3 SAR images are selected. The sub-images should contain clear road edges and moderate background complexity, with various imaging modes.

Secondly, we developed an image cropping software with Python and Open Source Computer Vision Library (OpenCV). When we input the original SAR image, the software can crop the original image with the size predefined, and the cropped image is not distorted. The software's crop function is not affected by the

resolution and size of the input image. Since we have set the size of the cropped image chip to 512×512 pixels in the software, a 512×512 pixels window is slid along the sub-image's row with a step of 256 pixels, and then the road chips are cut with a size of 512×512 pixels. The chips contain roads with various shapes as well as related background information such as farmland, houses, airports, green belts and cities, which meets the requirements of actual road segmentation. The 512×512 pixels size images can be used to clearly extract the type, shape and background information of the roads.

Thirdly, we use the LabelMe labeling tool to segment and label the roads. In the labeling process, the road edge is labeled with a series of points. After labeling, each road image corresponds to a JSON file, such as Figure 4(c). Since the shape of the roads in our dataset is rich, the average time to label a road chip is about 1 minute. We used a total of about 167 working hours to label 10026 road chips with the LabelMe labeling tool. The operating system of the labeling computer is Windows 10, with Intel Core i7-9750H CPU.

As shown in Figure 4(b), there are seven marked points, which are denoted as 1234567. Taking the upper left corner of the image as the coordinate origin, the horizontal axis is the X axis, and the vertical axis is the Y axis. The coordinate value of each labeled point is its actual pixel position in the image.

Finally, the entire dataset is randomly divided into a training dataset (70%), a validation dataset (20%), and a test dataset (10%). Table 2 gives the allocation details. Since our dataset is tens of thousands of lev-

Figure 3

The construction process of the dataset proposed as open public repository

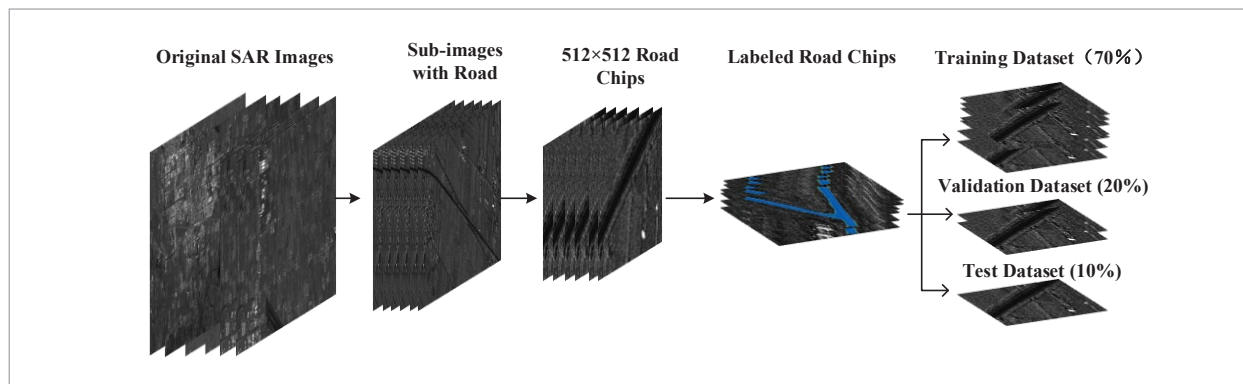
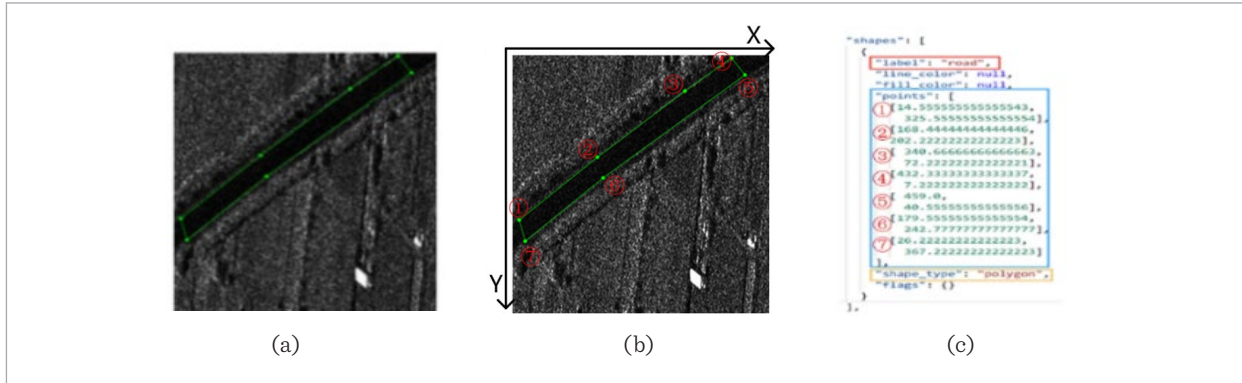


Figure 4

Road labeling details: (a) example of road image; (b) labeled image; (c) JSON file after labeling



els, we used the 7: 2: 1 ratio to divide the entire dataset into a training dataset, a validation dataset, and a test dataset. We need to ensure that enough data (70%) is used as the training dataset to avoid overfitting of road segmentation in deep learning. We selected a small amount of data (20%) as the validation dataset to adjust the model and selected the fewest data (10%) as the test dataset to evaluate the model indicators.

Table 2

Numbers of training set, validation set and testset

	SL	UFS	FS I	FS II	Total
Training set	3591	1157	988	1311	7047
Validation set	758	683	300	264	2005
Test set	426	199	183	166	974

3.3. Details of Dataset

Table 1 provides details on the road chips in different imaging modes. Figure 5 shows the road chips in different imaging modes. Figure 6 shows the roads with different shapes in the same imaging mode (SL), including airport runway, highway, country road, fork road, cross road, and curved road. The selection criteria of the road segment type used in our research are that the roads have a high probability of appearance, different shapes, and a large degree of difference. The background of roads in the dataset is also very diverse, which is shown in Figure 7.

Due to various shapes and backgrounds of roads in the SARroad dataset, the overfitting of road segmentation in deep learning can be avoided effectively.

Figure 5

Roads in different imaging modes: (a) SL (resolution of 1m); (b) UFS (resolution of 3m); (c) FS I (resolution of 5m); (d) FS II (resolution of 10m).

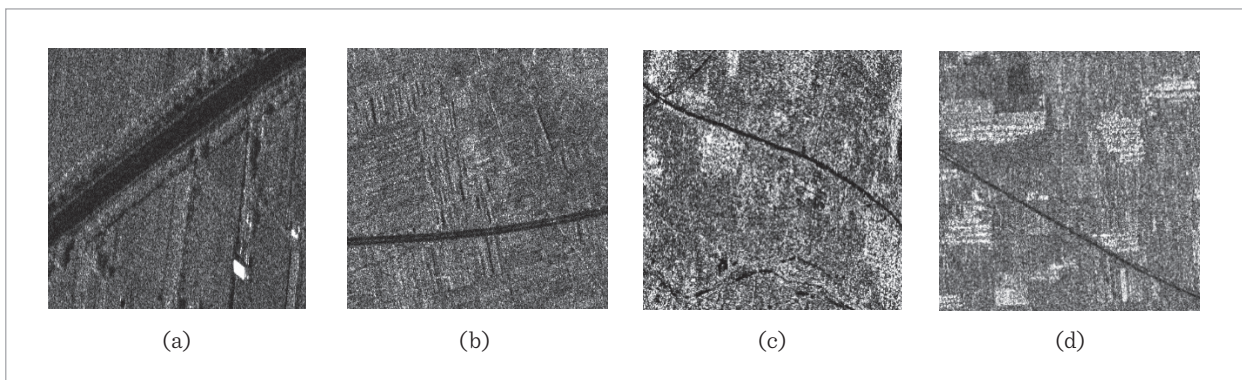
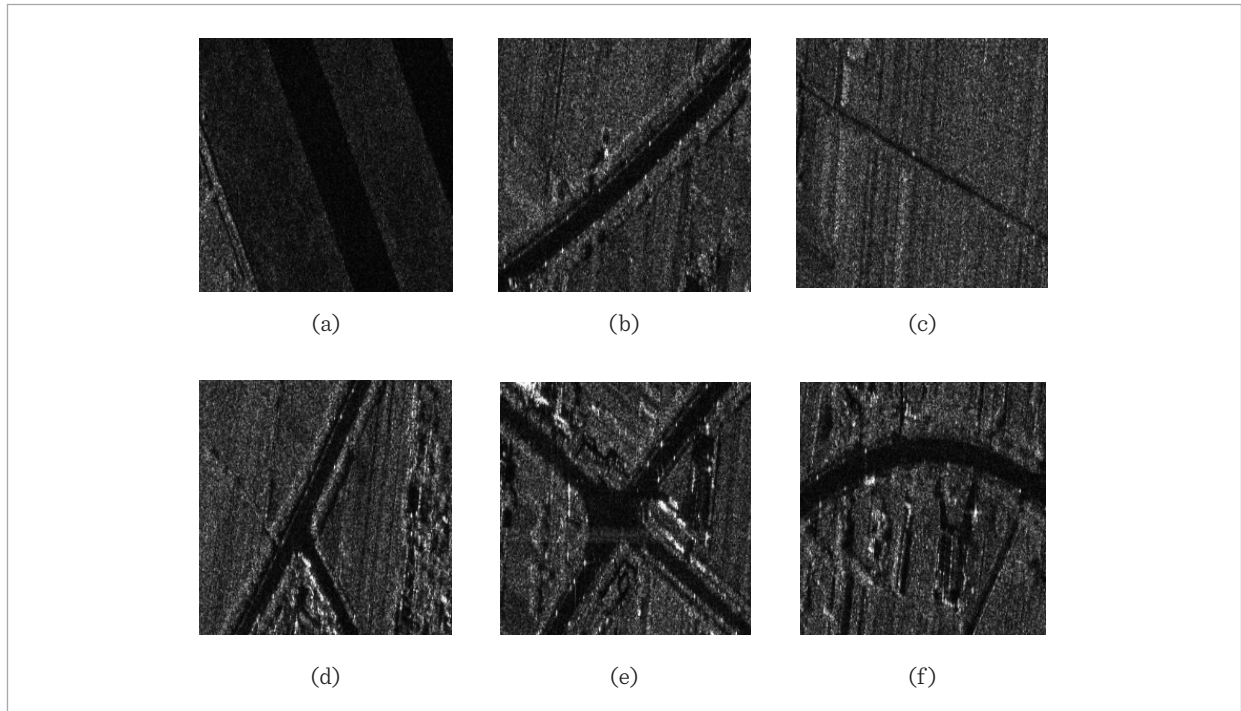
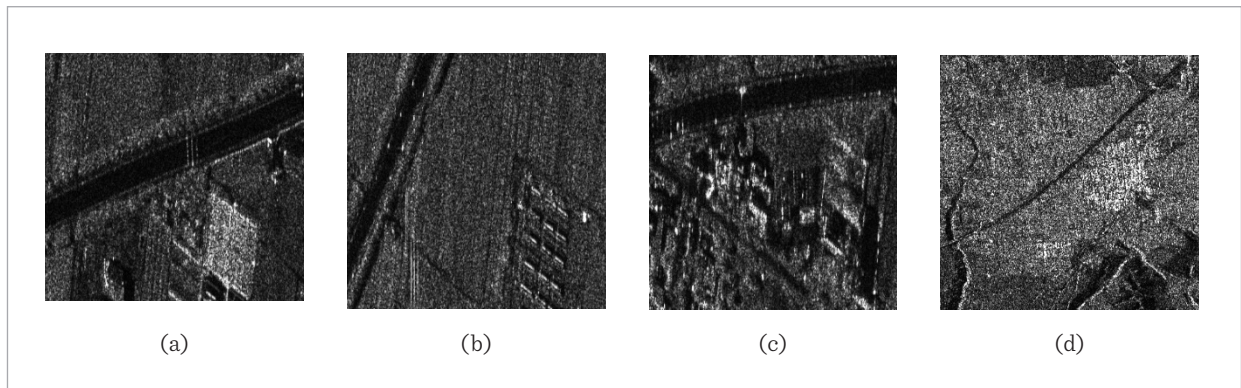


Figure 6

Roads with different shapes in SL mode (resolution of 1m): (a) airport runway; (b) highway; (c) country road; (d) fork road; (e) cross road; (f) curved road

**Figure 7**

Roads with different backgrounds: (a) rural house; (b) farmland; (c) urban area; (d) ridge

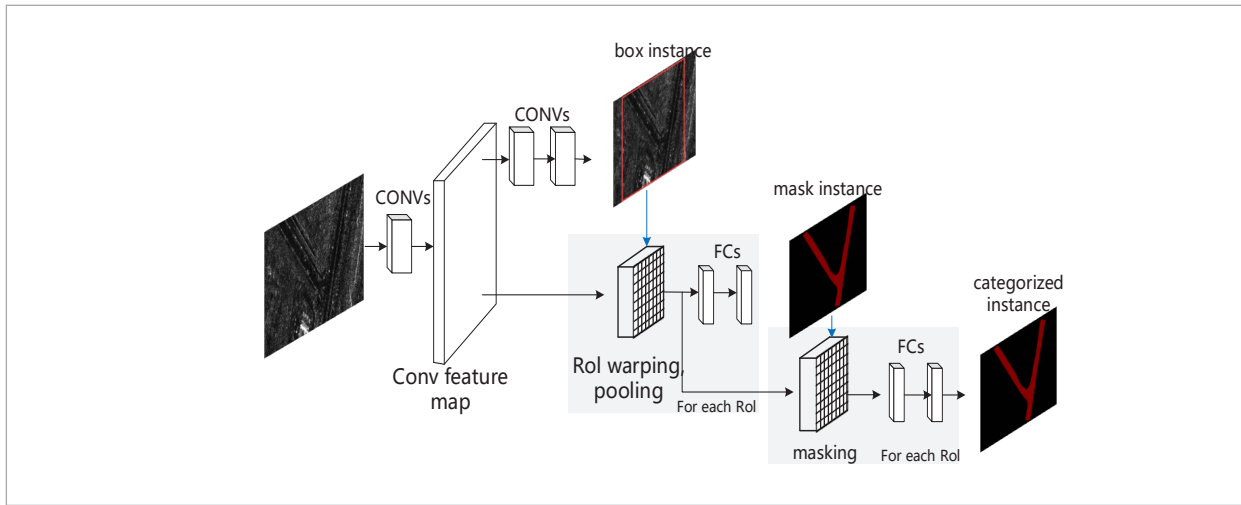


4. Segmentation Algorithms

There are three representative segmentation algorithms of MNC, FCIS, and Mask R-CNN [16], the most famous of which is Mask R-CNN. Above segmentation algorithms in computer vision have similar basic networks,

such as FPN [26] and ResNet. In this paper, ResNet50 is selected as the basic network for road feature extraction. The MNC, FCIS, and Mask R-CNN are briefly introduced respectively in the following Sections.

Figure 8
MNC network structure



4.1. MNC

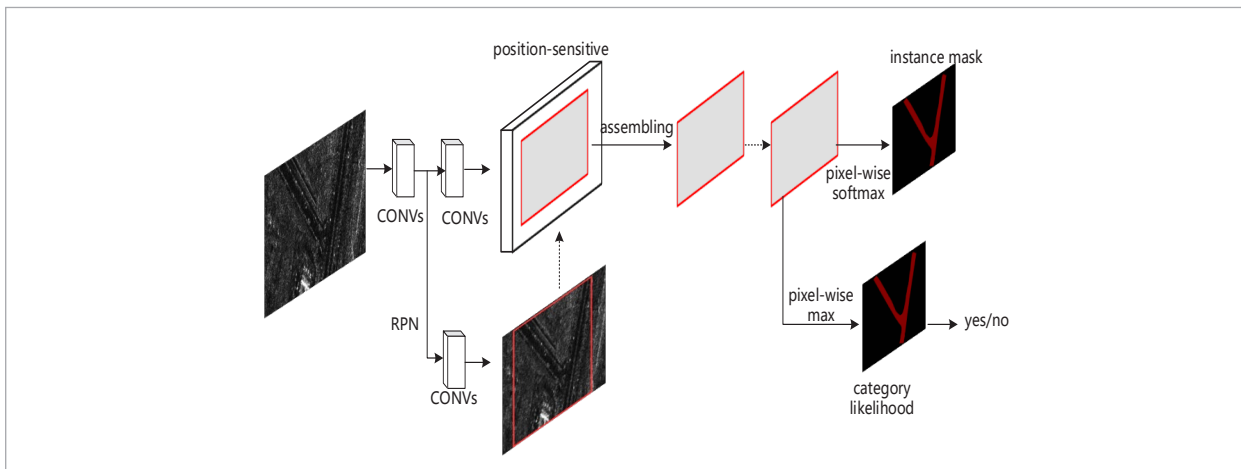
Figure 8 shows the network structure of MNC. MNC is mainly composed of three phases [10]. The first stage task determines the road region of interest (ROI) based on the road features extracted by the convolutional neural network. The second stage task uses ROI Warping and ROI Pooling and forms a preliminary road physical mask through two fully-connected layers. The third stage task uses the two-layer fully-connected network to further determine the road mask based on the second stage road segmentation.

Different from the FCIS and Mask R-CNN methods, the biggest feature of MNC is its segmentation method. On the basis of shared features, each task is performed simultaneously, and each task does not interfere with each other. On the basis of shared features, the next task depends on the previous task and shared features, forming a hierarchical multi-tasking structure.

4.2. FCIS

Figure 9 shows the network structure of FCIS. FCIS is mainly composed of three parts [28]. One is the convolu-

Figure 9
FCIS network structure



tional neural network part for extracting feature maps. The second is ROI region generation network part, using ROI Warping and ROI Pooling to extract corresponding features and generate road and background. The third is fully-connected layer part, which is used for road and background division to generate the final mask.

The preliminary features of road image are extracted by the convolution layer, and then these features are used to extract the ROI region through the RPN (Region Proposal Network) network, and to generate a feature map through the convolution layers. After the generated ROI region and feature map are aggregated, each ROI and its corresponding feature map are sent to the final segmentation discriminator for classification to obtain an accurate road mask.

4.3. Mask R-CNN

Figure 10 shows the network structure of Mask R-CNN. Mask R-CNN is mainly composed of three parts of the network: the RPN network part using the convolutional neural network to extract the feature map, the network part to generate the target classification using regional proposals, and the network part for semantic segmentation and mask generation [20]. Firstly, images are fed into a deep convolutional network to obtain a feature map, and then a set of rectangular target frames and their corresponding target scores are obtained using RPN. After that, the region of interest is further processed using the method of

ROI. Finally, these converted proposed regions are passed to the classifier to output the bounding boxes of the corresponding roads, while the semantic segmentation network part generates road masks in parallel.

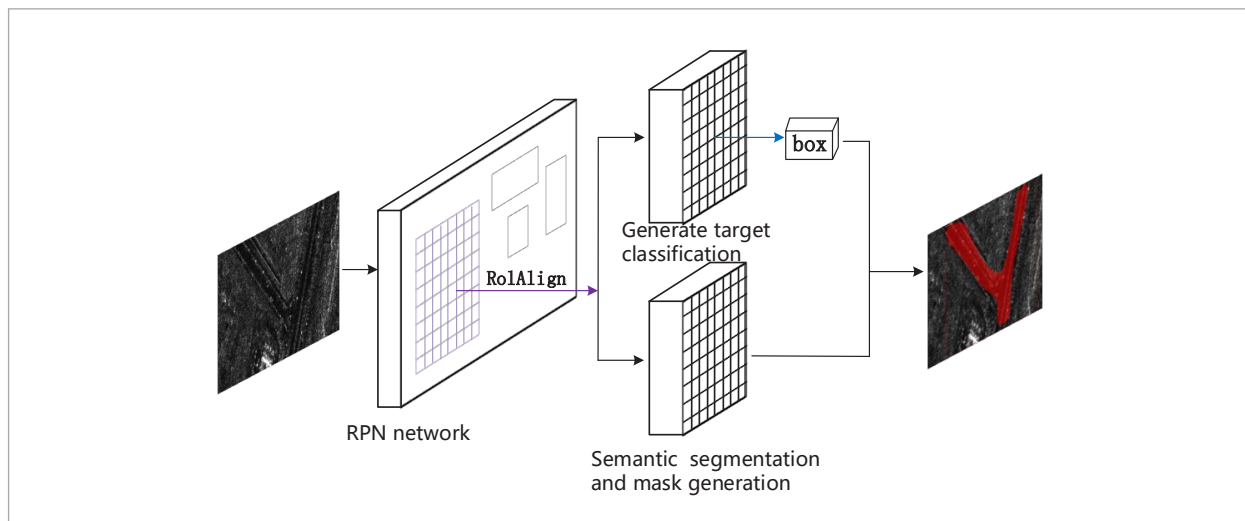
5. Analysis of Experimental Results

Based on our dataset, we compare the performance of the three aforementioned deep learning segmentation algorithms to verify the effect of deep learning methods on road segmentation and give specific analysis. The operating system of the experimental machine is Ubuntu 16.04, and the GPU is NVIDIA 2080tiGPU.

In deep learning, precision (p) is often used to reflect the correct rate of a category being correctly predicted, and recall (r) is used to reflect the proportion of correctly predicted samples among all predicted samples. The calculation method is shown in (1) and (2), where TP represents the number of detection frames whose detection result is true and the true value is true, FP represents the number of detection frames whose detection result is true but the true value is false, and FN represents the number of detection frames whose detection result is false but the true value is true [36]. AP (Average Precision) reflects the trade-off between the algorithm's recognition accuracy of positive examples and its ability to cover positive examples. AP is defined as the area under the preci-

Figure 10

Mask R-CNN network structure



sion-recall curve (PR curve). The x-axis is recall, and y-axis is precision. However, due to relatively difficulty of such integration, AP is often calculated by (3), where $p(\tilde{r})$ is the measured precision at recall \tilde{r} [14, 30]. The higher the AP value, the better the algorithm performance. IoU (Intersection over Union) reflects the degree of area overlap between the mask predicted by the algorithm and the label frame [22], and it is calculated by (4), where B_p is the predicted road mask, B_{gt} is the road label frame, $area(B_p \cap B_{gt})$ is the area of the intersection of the predicted and ground truth bounding boxes, and $area(B_p \cup B_{gt})$ is the area of their union. If the prediction is perfect, $IoU = 1$, and if it completely misses, $IoU = 0$. The higher the IoU value, the better the algorithm performance.

When IoU is greater than 0.5, the detection is considered to be successful, which is marked as TP . When IoU is smaller than 0.5, a false alarm is considered to appear, which is marked as FP . Since this dataset only contains roads, mAP (average value of AP for all classes) is same as AP itself.

$$p = \frac{TP}{TP + FP}, \quad (1)$$

$$r = \frac{TP}{TP + FN}, \quad (2)$$

$$AP = \sum_0^1 (r_{n+1} - r_n) \max_{\tilde{r}: \tilde{r} \geq r_{n+1}} p(\tilde{r}), \quad (3)$$

$$IoU = \frac{area(B_p \cap B_{gt})}{area(B_p \cup B_{gt})}. \quad (4)$$

We train the models using the images from the training dataset, and adjust the models using the images from the validation dataset. After these two steps are finished, we finally test the performance of the models using the images from the test dataset. The testing results of each segmentation model are shown in Table 3, where each algorithm is measured by AP value and IoU. When IoU is greater than 0.5, the classification is considered to be correct. The input image size is 512×512 pixels. It can be seen that the performance of Mask R-CNN is the best, while the performance of MNC is the worst. One of the reasons is that, in Mask R-CNN, RoI Pooling is replaced by RoIAlign, leading to more accurate features extracted [20].

Table 3

Performance benchmarking of three deep learning algorithms for road segmentation

Algorithms	AP	IoU
Mask R-CNN	86.5%	88.2%
FCIS	79.2%	85.7%
MNC	73.1%	80.0%

Figure 11, Figure 12, and Figure 13 show the segmentation results of the Mask R-CNN algorithm for various testing images, respectively. Three testing images are straight road, fork road and "V" type road. Each time 500 training samples are randomly selected from the training set and added into the training data. The network model after the last training is used as the initial network model for the next training. It can be seen that with the increase of training samples,

Figure 11

Segmentation results of Mask R-CNN for straight road: (a) test image; (b) 500 training samples; (c) 1000 training samples; (d) 1500 training samples; (e) 2000 training samples

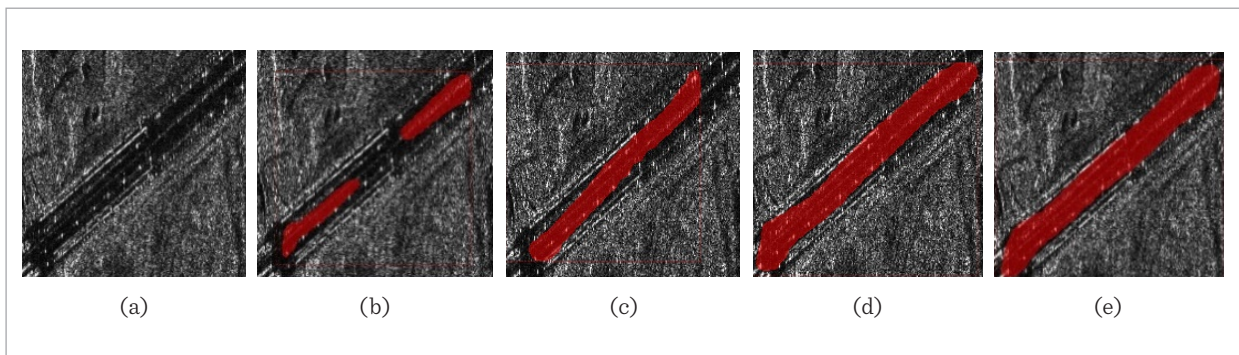
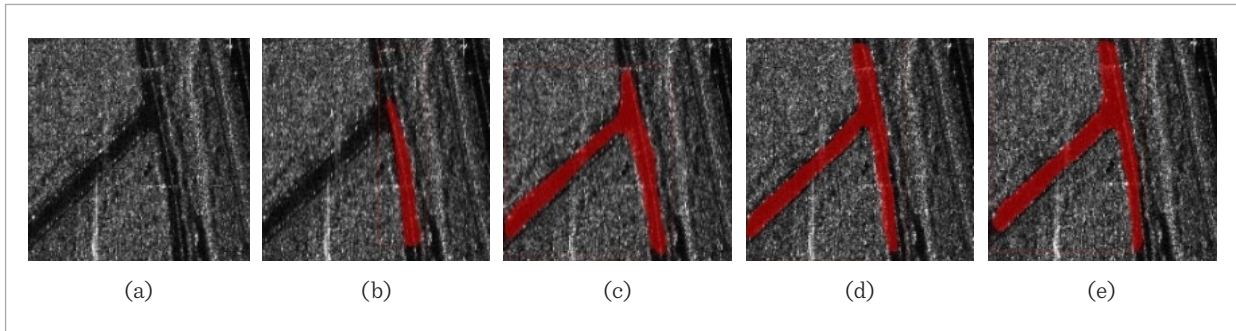
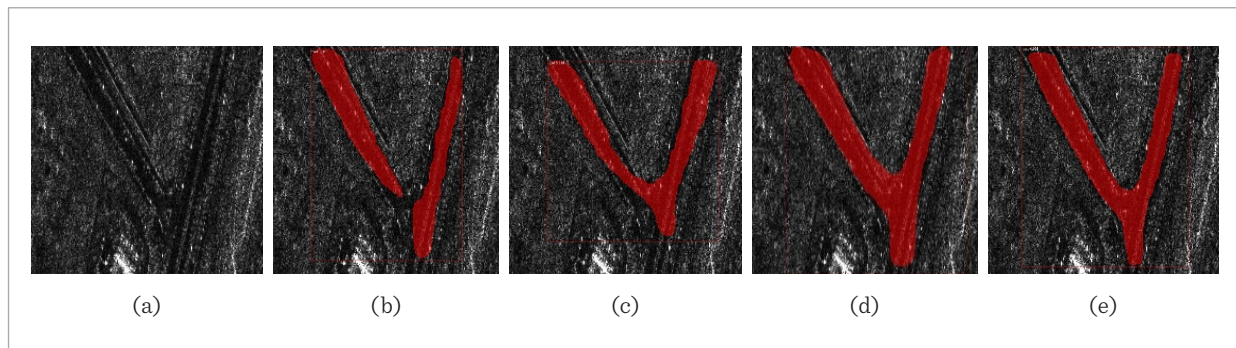


Figure 12

Segmentation results of Mask R-CNN for fork road: (a) test image; (b) 500 training samples; (c) 1000 training samples; (d) 1500 training samples; (e) 2000 training samples

**Figure 13**

Segmentation results of Mask R-CNN for "V" type road: (a) test image; (b) 500 training samples; (c) 1000 training samples; (d) 1500 training samples; (e) 2000 training samples



Mask R-CNN segmentation is more accurate. Overall, the Mask R-CNN algorithm performs relatively well. However, there are still some problems of insufficient semantic segmentation, and the performance can be further improved due to use of more training samples [7, 8, 12, 23, 40, 48].

6. Conclusions

We have constructed a GF-3 SAR image dataset for road segmentation named SARroad. Three deep-learning based segmentation algorithms of MNC, FCIS and Mask R-CNN are used to perform experiments on the SARroad dataset. The results show that the Mask R-CNN performs best. The three

algorithms are trained by using the SARroad dataset, and the testing results demonstrate the validity of the SARroad dataset. The experimental results establish a performance benchmark for the SARroad dataset, which is convenient for other researchers to carry out related research on road segmentation of GF-3 SAR images. It should be noted that the current segmentation results are not accurate enough. This is because the accuracy of road labeling is not enough and there is not enough training data. In future research, we will improve the accuracy of road labeling, increase the capacity of the dataset, and enrich the shape of roads. We believe that these further improvements make the constructed SARroad dataset more efficient for road segmentation of GF-3 SAR images.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61102163), the Fundamental Research Funds for the Central Universities (No. GK201903085), the State Key Laboratory of Geo-Information Engineering (No. SKL-

GIE2019-M-3-5) and the Key Laboratory of Land Satellite Remote Sensing Application Center, Ministry of Natural Resources of the People's Republic of China (No. KLSMNR-202004). The authors would like to thank China Centre for Resources Satellite Data and Application for providing GF-3 SAR images.

References

1. An, L. Road Extraction from High Resolution Remote Sensing Image based on Hough Transform. *Geomatics & Spatial Information Technology*, 2018, 41(3), 116-117.
2. Cheng, G., Han, J. W. A Survey on Object Detection in Optical Remote Sensing Images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 117, 11-28. <https://doi.org/10.1016/j.isprsjprs.2016.03.014>
3. Cheng, G., Han, J. W., Lu, X. Q. Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proceedings of the IEEE*, 2017, 105(10), 1865-1883. <https://doi.org/10.1109/JPROC.2017.2675998>
4. Cheng, G., Han, J. W., Zhou, P. C., Guo, L. Multi-class Geospatial Object Detection and Geographic Image Classification based on Collection of Part Detectors. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2014, 98, 119-132. <https://doi.org/10.1016/j.isprsjprs.2014.10.002>
5. Cheng, G., Qian, Y., Elder, J. H. Fusing Geometry and Appearance for Road Segmentation. *Proceedings of IEEE International Conference on Computer Vision Workshops, (ICCVW 2017), Venice, October 22-29, 2017*, 166-173. <https://doi.org/10.1109/ICCVW.2017.28>
6. Cheng, G., Zhou, P. C., Han, J. W. Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(12), 7405-7415. <https://doi.org/10.1109/TGRS.2016.2601622>
7. Chen, G., Li, C., Wei, W., Jing, W., Woźniak, M., Blažauskas, T., Damaševičius, R. Fully Convolutional Neural Network with Augmented Atrous Spatial Pyramid Pool and Fully Connected Fusion Path for High Resolution Remote Sensing Image Segmentation. *Applied Sciences*, 2019, 9(9), 1816. <https://doi.org/10.3390/app9091816>
8. Chouhan, V., Singh, S. K., Khamparia, A., Gupta, D., Tiwari, P., Moreira, C., Damaševičius, R., de Albuquerque, V. H. C. A Novel Transfer Learning based Approach for Pneumonia Detection in Chest X-ray Images. *Applied Sciences*, 2020, 10(2), 559. <https://doi.org/10.3390/app10020559>
9. Cui, B., Zhang, Y. H., Li Yan, Wei, J. J., Huang, Q. L. A SAR Change Detection Method based on The Consistency of Single-Pixel Difference and Neighbourhood Difference. *Remote Sensing Letters*, 2019, 10(5), 488-495. <https://doi.org/10.1080/2150704X.2018.1562256>
10. Dai, J. F., He, K. M., Sun, J. Instance-Aware Semantic Segmentation via Multi-Task Network Cascades. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (CVPR 2016), Las Vegas, NV, June 27-30, 2016*, 3150-3158. <https://doi.org/10.1109/CVPR.2016.343>
11. Deng, J., Dong, W., Socher, R., Li, L., Li, K., Li, F. F. Imagenet: A Large-Scale Hierarchical Image Database. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (CVPR 2009), Miami, FL, June 20-25, 2009*, 248-255. <https://doi.org/10.1109/CVPR.2009.5206848>
12. Dirvanauskas, D., Maskeliūnas, R., Raudonis, V., Damaševičius, R., Scherer, R. HEMIGEN: Human Embryo Image Generator based on Generative Adversarial Networks. *Sensors*, 2019, 19(16), 3578. <https://doi.org/10.3390/s19163578>
13. Enderli, C., Savy, L., Refregier, P. Application of the Deflection Criterion to Classification of Radar SAR Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, 29(9), 1668-1672. <https://doi.org/10.1109/TPAMI.2007.1121>
14. Everingham, M., Gool, L. V., Williams, C. K. I., Winn, J., Zisserman, A. The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 2010, 88(2), 303-338. <https://doi.org/10.1007/s11263-009-0275-4>
15. Fu, J. Z., Yang, S. M., Liu, H. High-Resolution Remote Sensing Images Road Extraction based on Wavelet

- Transform and Hough Transform. *Journal of Geomatics*, 2015, 40(4), 48-50. <https://doi.org/10.14188/j.2095-6045.2015.04.014>
16. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villeina-Martinez, V., Garcia-Rodriguez, J. A Review on Deep Learning Techniques Applied to Semantic Segmentation. *arXiv e-prints*, 2017. <https://doi.org/10.1016/j.asoc.2018.05.018>
 17. Geman, D., Jedynek, B. An Active Testing Model for Tracking Roads in Satellite Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18(1), 1-14. <https://doi.org/10.1109/34.476006>
 18. Guo, Y. J., Chen, E., Guo, Y., Li, Z. Y., Li, C. G., Xu, K. P. Deep Highway Unit Network for Land Cover Type Classification with GF-3 SAR Imagery. *Proceedings of SAR in Big Data Era: Models, Methods and Applications, (BIG-SAR DATA 2017)*, Beijing, November 13-14, 2017, 1-6. <https://doi.org/10.1109/BIGSAR DATA.2017.8124926>
 19. Han, Z. S., Wang, C. P., Fu, Q. Arbitrary-Oriented Target Detection in Large Scene SAR Images. *Defence Technology*, 2020, 16(4), 933-946. <https://doi.org/10.1016/j.dt.2019.11.014>
 20. He, K. M., Georgia, G., Piotr, D., Ross, G. Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 42(2), 386-397. <https://doi.org/10.1109/TPAMI.2018.2844175>
 21. He, K. M., Zhang, X. Y., Ren, S. Q., Sun, J. Deep Residual Learning for Image Recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (CVPR 2016)*, Las Vegas, NV, June 27-30, 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
 22. Jaszewski, M., Parameswaran, S., Hallenborg, E., Bagnall, B. Evaluation of Maritime Object Detection Methods for Full Motion Video Applications Using the PAS-CAL VOC Challenge Framework. *Proceedings of Video Surveillance and Transportation Imaging Applications*, San Francisco, California, United States, March 4, 2015, 9407. <https://doi.org/10.1117/12.2175585>
 23. Juočas, L., Raudonis, V., Maskeliūnas, R., Damaševičius, R., Woźniak, M. Multi-Focusing Algorithm for Microscopy Imagery in Assembly Line Using Low-Cost Camera. *The International Journal of Advanced Manufacturing Technology*, 2019, 102, 3217-3227. <https://doi.org/10.1007/s00170-019-03407-9>
 24. Krizhevsky, A., Sutskever, I., Hinton, G. Imagenet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 2017, 60(6), 84-90. <https://doi.org/10.1145/3065386>
 25. Li, K., Wan, G., Cheng, G., Meng, L. Q., Han, J. W. Object Detection in Optical Remote Sensing Images: A survey and a New Benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 159, 296-307. <https://doi.org/10.1016/j.isprsjprs.2019.11.023>
 26. Lin, T. S., Dollár, P., Girshick, R., He, K. M., Hariharan, B., Belongie, S. Feature Pyramid Networks for Object Detection. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (CVPR 2017)*, Honolulu, HI, July 21-26, 2017, 936-944. <https://doi.org/10.1109/CVPR.2017.106>
 27. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Zitnick, C. L. Microsoft COCO: Common Objects in Context. *Proceedings of European Conference on Computer Vision, (ECCV 2014)*. Springer International Publishing, September 6-12, 2014, 740-755. https://doi.org/10.1007/978-3-319-10602-1_48
 28. Li, Y., Qi, H. Z., Dai, J. F., Ji, X. Y., Wei, Y. C. Fully Convolutional Instance-Aware Semantic Segmentation. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (CVPR 2017)*, Honolulu, HI, July 21-26, 2017, 4438-4446. <https://doi.org/10.1109/CVPR.2017.472>
 29. Liu, J. Y., Qiu, X. L., Han, B., Xiao, D. J. Study on Geo-Location of Sliding Spotlight Mode of GF-3 Satellite. *Proceedings of IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar, (APSAR 2015)*, Singapore, September 1-4, 2015, 417-420. <https://doi.org/10.1109/APSAR.2015.7306238>
 30. Long, Y., Gong, Y. P., Xiao, Z. F., Liu, Q. Accurate Object Localization in Remote Sensing Images based on Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(5), 2486-2498. <https://doi.org/10.1109/TGRS.2016.2645610>
 31. Lu, X. Q., Feng, Y. C., Yuan, Y., Zhang, Y. L. Hierarchical and Robust Convolutional Neural Network for Very High-Resolution Remote Sensing Object Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(8), 5535-5548. <https://doi.org/10.1109/TGRS.2019.2900302>
 32. Marques, R. C. P., Medeiros, F. N., Santos Nobre, J. SAR Image Segmentation Based on Level Set Approach and $\{\text{cal G}\}_A^0$ Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(10), 2046-2057. <https://doi.org/10.1109/TPAMI.2011.274>
 33. SARroad Available online: <https://github.com/liekkas966/SARroad>.
 34. SARroad (Training Set of SL) Available online: https://github.com/liekkas966/SARroad_train_1m.

35. Simonyan, K., Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *Computer Science*, 2014.
36. Sun, X., Wang, Z. R., Sun, Y. R., Diao, W. H., Zhang, Y., Fu, K. AIR-SARShip-1.0: High Resolution SAR Ship Detection Data Set. *Journal of Radars*, 2019, 8(6), 852-862. <https://doi.org/10.12000/JR19097>
37. Sun, Z. G., Song, Y. J. Structural Information Detection based Filter for GF-3 SAR Images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2018, XLII-3, 1617-1621. <https://doi.org/10.5194/isprs-archives-XLII-3-1617-2018>
38. Szegedy, C., Liu, W., Jia, Y. Q., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. Going Deeper with Convolutions. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, (CVPR 2015)*, Boston, MA, June 7-12, 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
39. Tsutsui, S., Kerola, T., Saito, S. Distantly Supervised Road Segmentation. *Proceedings of IEEE International Conference on Computer Vision Workshops, (ICCVW 2017)*, Venice, October 22-29, 2017, 174-181. <https://doi.org/10.1109/ICCVW.2017.29>
40. Urbonas, A., Raudonis, V., Maskeliunas, R., Damaševičius, R. Automated Identification of Wood Veneer Surface Defects Using Faster Region-Based Convolutional Neural Network with Data Augmentation and Transfer Learning. *Applied Sciences*, 2019, 9(22), 4898. <https://doi.org/10.3390/app9224898>
41. Wang, Y. Y., Wang, C., Zhang, H., Dong, Y. B., Wei, S. S. A SAR Dataset of Ship Detection for Deep Learning under Complex Backgrounds. *Remote Sensing*, 2019, 11(7), <https://doi.org/10.3390/rs11070765>
42. Wang, Z. Z., Jiao, J., Zeng, Q. M., Liu, J. Y. A Three-Hierarchy Evaluation of Polarimetric Performance of GF-3, Compared with ALOS-2/PALSAR-2 and RADAR-SAT-2. *Sensors*, 2019, 19(7). <https://doi.org/10.3390/s19071493>
43. Xia, G. S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J. B., Datcu, M., Pelillo, M., Zhang, L. P. DOTA: A Large-Scale Dataset for Object Detection in Aerial Images. *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition, (CVPR 2018)*, Salt Lake City, UT, June 18-23, 2018, 3974-3983. <https://doi.org/10.1109/CVPR.2018.00418>
44. Xiao, Z. F., Liu, Q., Tang, G. F., Zhai, X. F. Elliptic Fourier Transformation-Based Histograms of Oriented Gradients for Rotationally Invariant Object Detection in Remote-Sensing Images. *International Journal of Remote Sensing*, 2015, 36(2), 618-644. <https://doi.org/10.1080/01431161.2014.999881>
45. Yang, F., Wang, H., Jin, Z. A Fusion Network for Road Detection via Spatial Propagation and Spatial Transformation. *Pattern Recognition*, 2019, 100, 107139. <https://doi.org/10.1016/j.patcog.2019.107141>
46. Yang, T., Zhu, J. B., Liu, J. Y. SAR Image Target Detection and Recognition based on Deep Network. *Proceedings of SAR in Big Data Era: Models, Methods and Applications, (BIGSAR DATA 2019)*, Beijing, China, August 15-16, 2019, 1-4. <https://doi.org/10.1109/BIGSAR.2019.8858465>
47. Yu, Z., Wang, W. Q., Li, C. S., Liu, W., Yang, J. Speckle Noise Suppression in SAR Images Using a Three-Step Algorithm. *Sensors*, 2018, 18(11), 1-13. <https://doi.org/10.3390/s18113643>
48. Zhou, B., Duan, X., Ye, D., Wei, W., Woźniak, M., Połap, D., Damaševičius, R. Multi-Level Features Extraction for Discontinuous Target Tracking in Remote Sensing Image Monitoring. *Sensors*, 2019, 19(22), 4855. <https://doi.org/10.3390/s19224855>

