A Deep Learning Approach to Nightfire Detection based on Low-Light Satellite

Yue Wang, Ye Ni, Xutao Li and Yunming Ye

Department of Computer Science,
Harbin Institute of Technology, Shenzhen, China

ABSTRACT

Wildfires are a serious disaster, which often cause severe damages to forests and plants. Without an early detection and suitable control action, a small wildfire could grow into a big and serious one. The problem is especially fatal at night, as firefighters in general miss the chance to detect the wildfires in the very first few hours. Low-light satellites, which take pictures at night, offer an opportunity to detect night fire timely. However, previous studies identify night fires based on threshold methods or conventional machine learning approaches, which are not robust and accurate enough. In this paper, we develop a new deep learning approach, which determines night fire locations by a pixel-level classification on low-light remote sensing image. Experimental results on VIIRS data demonstrate the superiority and effectiveness of the proposed method, which outperforms conventional threshold and machine learning approaches.

KEYWORDS

Night fire detection, pixel segmentation, low-light satellite image

1. Introduction

Wildfire is a severe threat to forests and human estate, which often takes place frequently due to the increase of dry fuel and the impact of extreme climatic conditions. A small wildfire could grow into a serious and big one if it is not detected in time. The problem is especially fatal at night, when the wildfires are less likely to be noticed. Hence, how to accurately detect night fires becomes an very important issue.

Low-light satellite, e.g., National Polar-Orbiting Partnership (NPP)/Visible Infrared Imaging Radiometer Suite (VIIRS), Defense Meteorological Satellite Program (DMSP)/ Operational Lines can System (OLS) and Luojia-1A Satellite, which can take remote sensing pictures at night, offers an opportunity to identify night fires. The low-light satellite image is a multichannel matrix, and its spectral bands span visible light, near infrared, short wave infrared and medium wave infrared. With the rich band information, night fires can be identified. Previous studies solve the problem based on threshold or conventional machine learning approaches, which are not robust and accurate enough. Recently, deep learning techniques have shown promising performance on conventional computer vision tasks. Hence, in this paper, we aim to develop a new deep learning approach, which can accurately detect night fires based on low-light satellite images.

David C. Wyld et al. (Eds): CoSIT, AIAPP, SIGL, CRIS, NLPML - 2021
pp. 01-09, 2021. CS & IT - CSCP 2021 DOI: 10.5121/csit.2021.110401

In particular, we propose a pixel-level classification method based on recursive convolutional neural network for night fire detection. In our method, the spatial context is effectively exploited by the recursive convolution mechanism. Moreover, a squeeze excitation (SE) module is introduced to model the channel correlations for night fire detection. To validate the effectiveness of the proposed method, we conduct experiments on VIIRS satellite data. The experimental results show that the developed method performs better than conventional machine learning approaches, including Light GBM [1], random forest [2]. Moreover, it also outperforms existing deep learning techniques, e.g., multi-scale convolutional network and UNet [3].

2. RELATED WORK

Most of previous studies focus on the fire detection in small scale range. For example, in [4], a fire detection algorithm is developed based on the video data. In the method, the videos are from city monitors, where the covered area is quite small. Due to observation limitation, some studies attempt to detect fires based on low-light satellite images. The studies address the problem via a threshold segmentation. For example, in [5], an active fire detection algorithm is developed based on VIIRS. The method mainly analyzes the data characteristics and adopts some threshold rules to detect fires. However, the threshold based methods are not robust enough.

In addition to the threshold based methods, some machine learning approaches are also leveraged for fire detection upon low-light satellite data. For example, by extracting the multi-channel features, light GBM and random forest are applied to fire identification. However, the performance of the methods significantly relies on the manually constructed features.

Recently, with the rapid development and great success of deep learning techniques, its performance often beats that of the conventional machine learning algorithms. Moreover, it works in an end-to-end manner and does not need any manually constructed features. Though some deep learning algorithms have been developed for remote sensing tasks, none has ever touched the fire detection on low-light satellite. In this paper, we aim to develop a new fire detection algorithm based on the low-light satallite images.

3. THE PROPOSED APPROACH

In this section, we introduce the proposed deep learning approach to fire detection. Our notion is treating the fire location detection as a pixel level binary classification task. Hence, we build a deep learning approach to solve the problem. There are three key issues to be taken into account: (i) the spatial contexts are very important and should be effectively exploited; (ii) the channels contribute differently to the identification; (iii) the positive (fire pixels) and negative (non-fire pixels) examples are totally imbalanced. Next, we elaborate the proposed approach in the following four subsections. First, we introduce the developed multi-scale recursive convolution neural network unit. Second, the detection architecture is built upon the unit. Third, a squeeze excitation scheme is incorporated into the architecture. Finally, we apply the focus loss to learn the parameters in the architecture.

3.1. Multi-Scale Recursive Convolution Neural Network Unit

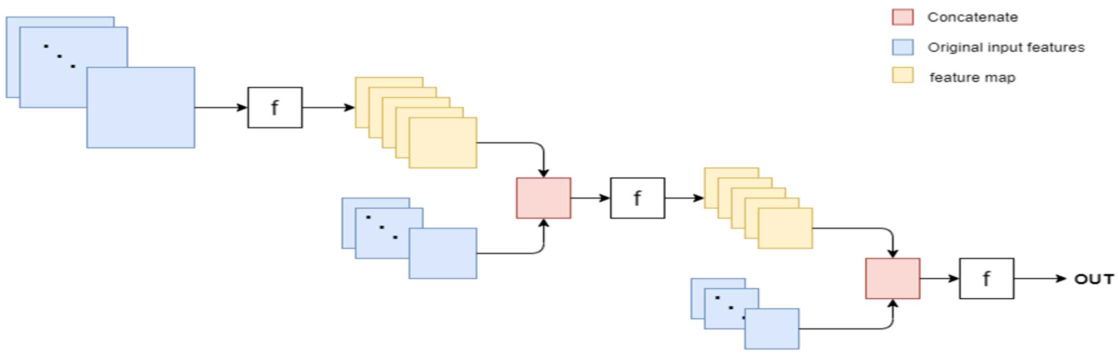

Figure 1. The Structure of Multi-scale Recursive Convolution Network Unit.

To detect the fire locations, we consider the task as a pixel-level classification problem. One of the key issues of the pixel-level classification is to effectively exploit the spatial contexts around the pixel. Inspired by [6], we develop the multi-scale recursive convolutional network unit. The main idea of the new unit can be illustrated as Fig. 1. We can see that given a multi-channel image patch, the new unit leverages three scales of convolutions to compute its output feature maps. In the first scale, a convolution filter is first utilized to compute the feature maps. Then, we crop the corresponding multi-channel input into the same size as the feature maps, and then combined it with the corresponding feature maps. In the second scale, the fused results are processed by the similar procedures to produce the output. Finally, the results in the second scale are further convolved with a filter to compute the third scale output, which is also the final output of the new unit.

The new unit can effectively exploit the spatial context due to its multi-scale recursive convolution. Moreover, in the new unit the contexts are gradually shrunk to the center point of each image patch. By doing so, the contexts can be exploited and the noise information is also effectively controlled at the same time, which is especially important for the pixel level based classification.

3.2. The Fire Detection Network Architecture

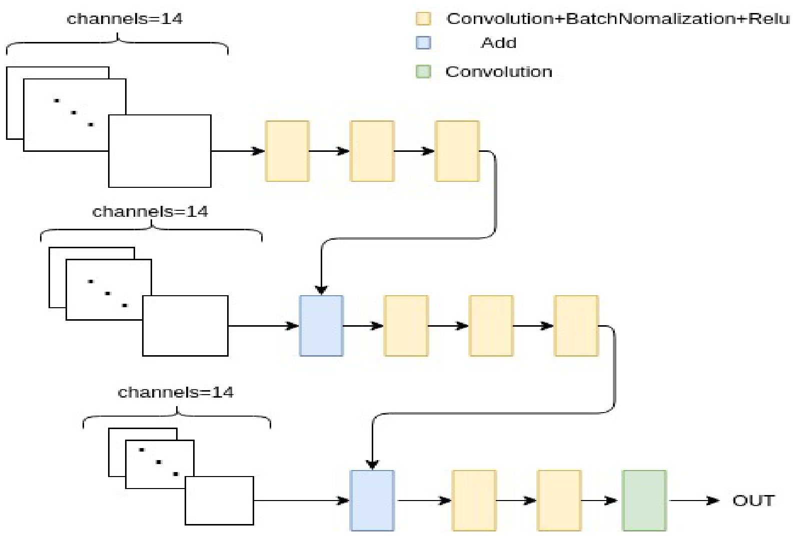

Figure 2. The Architecture of Fire Detection Network.

Upon the multi-scale recursive convolution network unit, we develop the fire detection network architecture. In our approach, we treat the fire location detection from low-light satellite images as an pixel based classification problem. For each pixel, we determine whether it is fire or not by extracting a small patch center at the pixel. With the small patch as input, a binary classification neural network architecture is established based on the multi-scale recursive convolution neural network unit, which is shown as in Fig. 2. We can see that given a 14-channel image patch as input, the patch is first convolved with three filter layers appended by a batch normalization layer and a rectified linear unit (ReLU) activation function. The input patch is cropped into appropriate size with the output feature maps and then added together in the second scale. In the third scale, the similar procedure is performed. With the output feature maps from the scale, a classification decision is made to determine whether the center point of the input patch is fire or not.

3.3. Incorporation of Squeeze Excitation Scheme

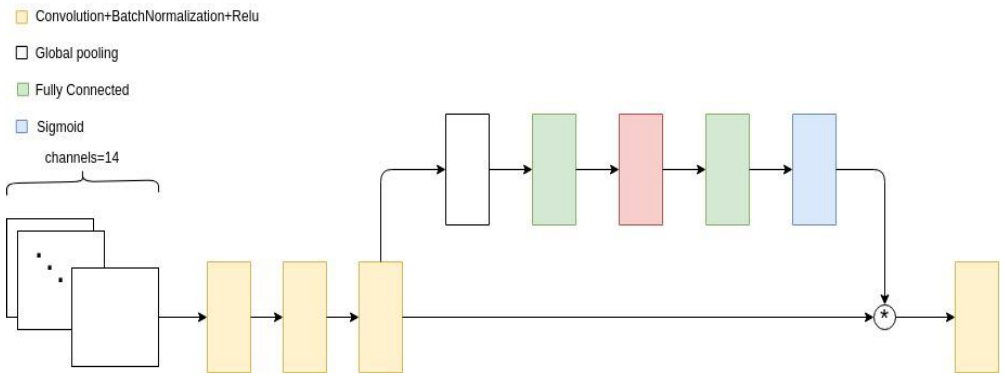

Figure 3. The structure of SE-block.

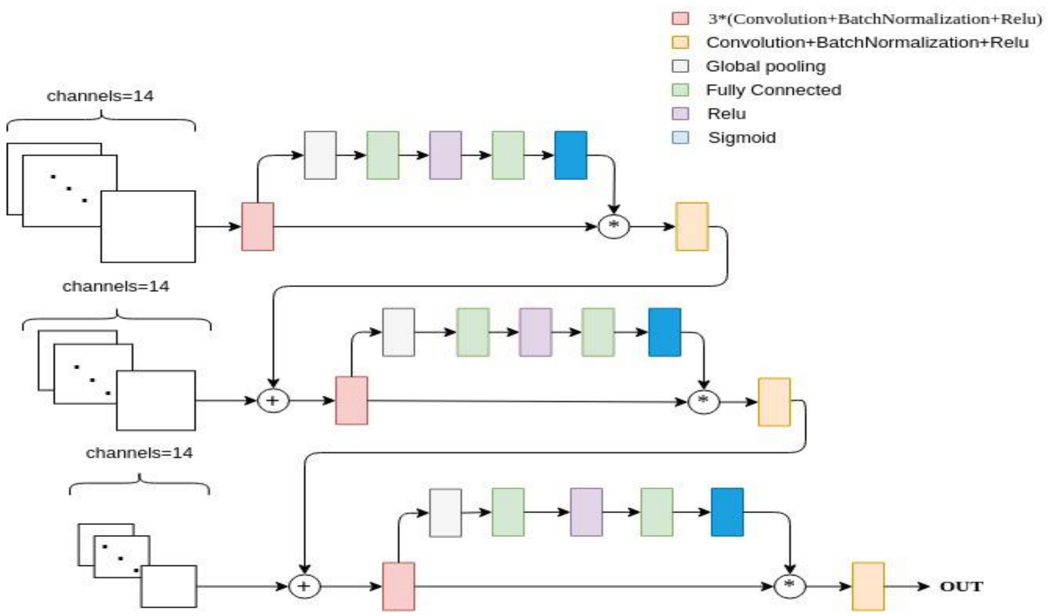

Figure 4. The structure of the multi-scale recursive convolution network model.

As shown in Fig. 3, the input low-light satellite image patch has multiple channels. In the channels, some are helpful for the fire detection while some are less important. Hence, the detection approach should effectively model the importance of different channels. Inspired by [7], we incorporate the squeeze excitation module into our fire detection network architecture.

The modified network architecture with squeeze excitation (SE) module is shown in Fig. 4. We can see the architecture is still a three-scale recursive convolution structure. Some modifications are made to incorporate the SE module. First, we use multiple convolution layers and activation function layers to extract the features. Then with the feature maps in each layer, we use a global pooling for sampling, and use multiple fully connected layers to characterize the contributions of different channels. The SE-module part outputs a multi-channel weighting matrix, which reserves the same size as feature maps. Finally, we perform the dot product with the feature maps and computed weighting matrix. As a result, the produced feature maps utility the weights computed, and channel importances are effectively modeled and fused.

Table 1 summarizes the detailed parameter settings of our fire detection network. We can see that the proposed fire detection network is mainly divided into three stages, and each stage consists of a similar network structure. In the stages, the model leverage inputs of different scales. The feature maps generated in the previous stage are combined with the low-light remote sensing image of the same scale to compute the network input of this stage. Finally, through this network we can obtain the fire detection results at the central point.

Table 1. The Structure of the Fire Detection Network.

Stage	Name	Layer	Filter	Stride	Output Size
	Input_1				31 * 31 * 14
First	Output_1	Conv1	5 * 5 / 32	1	27 * 27 * 64
	Output_2	Conv2	3 * 3 / 64	1	25 * 25 * 64
	Output_3	Conv3	3 * 3 / 128	1	23 * 23 *128
	Output_4	AvgPool	23 * 23		1 * 1 * 128
	Output_5	Fc1	128 * 8		1 * 1 * 8
	Output_6	Fc2	8 * 128		1 * 1 * 128
	Output_7 = Output_3 * Output_6				23 * 23 * 128
	Output_8	Conv4	3 * 3 / 64	1	21 * 21 * 64
	Output_9	Conv5	1 * 1 / 2	1	21 * 21 * 2
	Input_2				21 * 21 * 14 + 21 * 21 * 2
Second	Output_10	Conv6	5 * 5 / 32	1	17 * 17 * 64
	Output_11	Conv7	3 * 3 / 64	1	15 * 15 * 64
	Output_12	Conv8	3 * 3 / 64	1	13 * 13 * 64
	Output_13	AvgPool	13 * 13		1 * 1 * 128
	Output_14	Fc1	128 * 8		1 * 1 * 8
	Output_15	Fc2	8 * 128		1 * 1 * 128
	Outpu_16 = Output_12 * Output_15				13 * 13 * 128
	Output_17	Conv9	3 * 3 / 64	1	11 * 11 * 64
	Output_18	Conv10	1 * 1 / 2	1	11 * 11 * 2
	Input_3				11 * 11 * 14 + 11 * 11 * 2
Third	Output_19	Conv11	5 * 5 / 32	1	7 * 7 * 64
	Output_20	Conv12	3 * 3 / 64	1	5 * 5 * 64
	Output_21	Conv13	3 * 3 / 64	1	3 * 3 * 64
	Output_22	AvgPool	3 * 3		1 * 1 * 128
	Output_23	Fc1	128 * 8		1 * 1 * 8
	Output_24	Fc2	8 * 128		1 * 1 * 128
	Output_25 = Output_21 * Output_24				3 * 3 * 128
	Output_26	Conv14	3 * 3 / 64		1 * 1 * 64
	Output_27	Conv15	1* 1 / 2		1 * 1 * 2

3.4. The Loss Function

As a binary classification problem, one remainder issue is that the positive and negative samples are extremely imbalanced. In reality, the night fires take place at very few locations and time points. Hence, most of the low-light satellite images do not contain any positive examples. Only some of them include positive examples. The negative examples are very prevalent. Hence, we need to address the imbalance issue appropriately.

To tackle the issue, we adopt the focal loss [9] as our objective function. By a carefully modified cross entropy loss, the focal local loss can nicely deal with the skew positive and negative sample ratio. Specifically, it is formally computed as follows:

$$L_{fl} = \begin{cases} -\alpha(1-y')^{\gamma} logy' \;, & y=1 \\ 1-(1-\ \alpha)y'^{\gamma}\log(1-y')\,, & y=0 \end{cases} \quad (1)$$

Here α and γ are two positive parameters, and y and y' is the ground-truth class label and the prediction probability delivered by models, respectively. Α ∈ (0,1)controls the class weights, which is utilized to adjust the imbalanced ratio between positive and negative examples. γ is a tunable parameter to adjust the loss. When γ is approaching 0, the loss resembles the cross entropy. When it becomes larger, the loss pays more attentions to indistinguishable part. The loss function is equivalent to cross entropy if we set α = 1/2 and γ = 0.

4. EXPERIMENTS

4.1. Experiments Setup

Data Sets. In the experiment, we leverage the M-band data from Suomi NPP VIIRS to test the models. The data includes 16 channels, which are emissive, reflective and temporal brightness channels. Each pixel in the image denotes a resolution of 750m. Missing value is very prevalent in Suomi satellite data. To tackle the missing values, we replace them with the average value of each channel. The ground truth fire locations are obtained as night fire product introduced in [9].

In the VIIRS Nightfire (VNF) the occurrence time and locations are recorded for each fire point at night. We leverage the information as our ground-truth labels and correspond them to the Suomi data. In the experiment, we utilize the VIIRS night low-light monitoring data from June 5 to 6, 2019 and October 2 to 6, 2019. As for a comparison, we leverage conventional machine learning approaches light BGM and random forest as our baselines. In addition, the deep learning techniques convolutional neural network (CNN), UNet are also compared.

Evaluation Metrics. To validate the performance of different methods, we leverage the widely utilized precision, recall and F1-score. The three metrics are computed as follows:

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

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$F_1 = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (4)$$

Here *TP*, *TN*, *FP* and *FN* denote the numbers of true positive, true negative, false positive and false negative samples, respectively. In the three metrics, both precision and recall are biased while F1-score is a more comprehensive measure. Hence, F1-score is utilized to denote the overall evaluation. The higher the F1-score is, the better the performance is.

4.2. Experiment Results

Table 2 reports the experimental results of different models. We can see from the Table that UNet performs the worst, because the positive and negative examples are totally imbalanced. UNet fails to produce a promising segmentation. CNN is better but performs worse than conventional machine learning approaches light GBM and random forest. This is because: (i) the CNN model does carefully exploit the spatial contexts but utilizes them directly, where noisy contexts may hurt the performance; (ii) the imbalance issue of samples is not carefully considered. When equipped with the focal loss, the CNN model delivers better performance than lightBGM and random forest. Our proposed method yields better result than the CNN with focal loss, and its result is further improved when combined with the focal loss. All the results demonstrate the superiority and effectiveness of the proposed method.

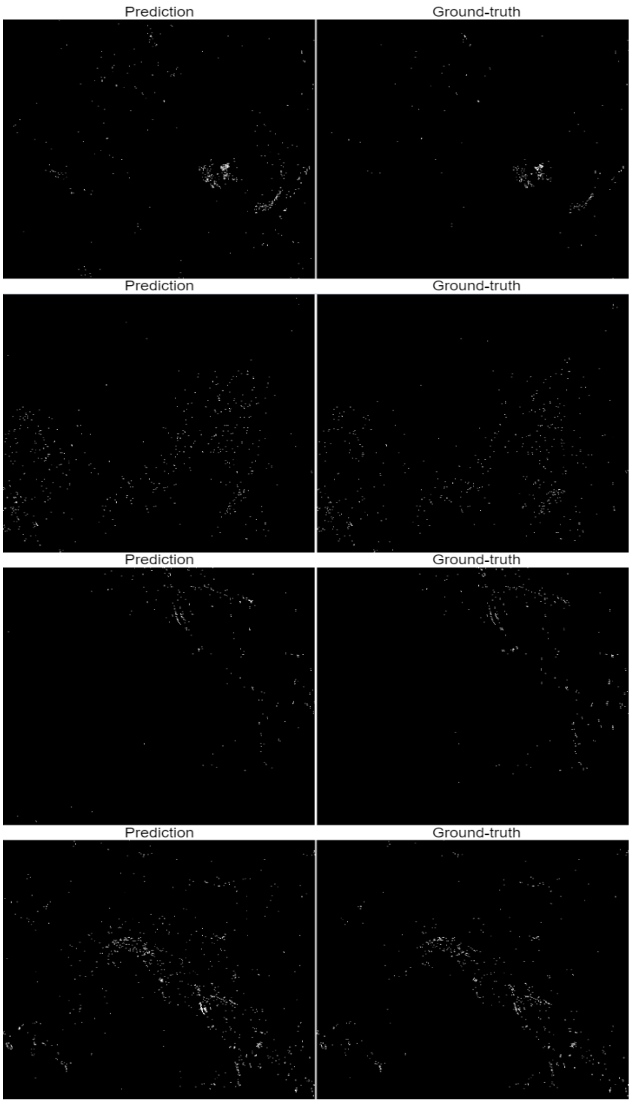

Figure 5. The night fire detection results of our method.

To visually examine the performance of our proposed method, we depict in Fig. 5 the detected locations of our model and the ground truth in four examples. We can see that the fire points detected by our method are quite consistent with the ground-truth locations, which further validates the effectiveness of the proposed method.

Table 2. The experimental results of different models on fire point detection tasks, where the symbol with * is the best result.

Method	Precision	Recall	F1-score
Random forest	0.756	1.0*	0.861
LightGBM	0.907	0.797	0.848
CNN	0.955	0.563	0.708
CNN + Focal Loss	0.893	0.871	0.882
UNet	0.294	0.165	0.211
Our method	0.983*	0.820	0.894
Our method + Focal Loss	0.949	0.929	0.939*

5. CONCLUSION

In this paper, we propose a new deep learning method to detect night fires based on the low-light satellite images. To effectively exploit the spatial contexts, a multi-scale recursive neural network unit is developed. Squeeze excitation module is incorporated in our method to characterize the channel importance. A focal loss objective function is adopted to tackle the sample imbalance issue. Experimental results on VIIRS low-light data set demonstrate the effectiveness and superiority of our method over existing techniques.

ACKNOWLEDGEMENT

The research was supported by the Shenzhen Science and Technology Program under Grant JCYJ20170811160212033.

REFERENCES

[1] black/white or gKe G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree[C]//Advances in neural information processing systems. 2017: 3146-3154.

[2] Liaw A, Wiener M. Classification and regression by randomForest[J]. R news, 2002, 2(3): 18-22.

[3] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.

[4] Günay O, Taşdemir K, Töreyin B U, et al. Video based wildfire detection at night[J]. Fire Safety Journal, 2009, 44(6): 860-868.

[5] Schroeder W, Oliva P, Giglio L, et al. The New VIIRS 375 m active fire detection data product: Algorithm description and initial assessment[J]. Remote Sensing of Environment, 2014, 143: 85-96.

[6] Pinheiro P, Collobert R. Recurrent convolutional neural networks for scene labeling[C]//International conference on machine learning. 2014: 82-90.

[7] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.

[8] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.

[9] Elvidge C D, Zhizhin M, Hsu F C, et al. VIIRS nightfire: Satellite pyrometry at night[J]. Remote Sensing, 2013, 5(9): 4423-4449.

AUTHORS

Yue Wang received a bachelor's degree from South China University of Technology in Guangzhou, China in 2019, and started studying for a master's degree in Harbin Institute of Technology Shenzhen campus in 2019. His research direction is the application of semantic segmentation technology and satellite remote sensing.

Ye Ni received the B.Sc degree from Harbin Institute of Technology, Weihai, China, in 2019, and he is currently pursuing the M.Sc. degree in Harbin Institute of Technology, Shenzhen, China. His research interests include computer vision and remote sensing.

Xutao Li is now an Associate Professor in the Shenzhen Graduate School, Harbin Institute of Technology. He received the Ph.D. and Master degrees in Computer Science from Harbin Institute of Technology in 2013 and 2009, and the Bachelor from Lanzhou University of Technology in 2007. His research interests include data mining, machine learning, graph mining and social network analysis, especially tensor based learning and mining algorithms.

Yunming Ye received the Ph.D. in Computer Science from Shanghai Jiao Tong University. He is now a professor in the Shenzhen Graduate School, Harbin Institute of Technology. His research interests include data mining, text mining, and ensemble learning algorithms.