

# RESEARCH ON HIGH RESOLUTION REMOTE SENSING IMAGE CLASSIFICATION BASED ON SEGNET SEMANTIC MODEL IMPROVED BY GENETIC ALGORITHM

Ming Xue<sup>1,3</sup>, Bo Wei<sup>1,3,\*</sup>, Lu Yang<sup>1,2</sup>

<sup>1</sup> Guangxi Key Laboratory of Spatial Information and Geomatics, Guilin University of Technology, Jiangan Road, Guilin, China - 67142046@qq.com, (superweibo, yll1222)@glut.edu.cn

<sup>2</sup> Zhuhai ORBITA Aerospace Science & Technology CO., LTD, Baisha Road, Zhuhai, China

<sup>3</sup> College of Geomatics and Geoinformation, Guilin University of Technology, Jiangan Road, Guilin, China

## Commission VI, WG VI/4

**KEY WORDS:** Genetic Algorithm, Deep Learning, Convolution Neural Network, Full Convolution Neural Network, SegNet Semantic Model, Weight Optimization, Remote Sensing Image Classification

### ABSTRACT:

SegNet model is an improved model of Full Convolutional Networks (FCN). Its encoder, i.e. image feature extraction, is still a convolutional neural network (CNN). Aiming at the problem that most traditional CNN training uses error back propagation algorithm (BP algorithm), which has slow convergence speed and is easy to fall into local optimum solution, this paper takes SegNet as the research object, and proposes a method of extracting partial weights by using genetic algorithm (GA) to select features of SegNet model, and to alleviate the problem that SegNet is easy to fall into local optimal solution. In the training process of SegNet model, the weight of convolution layer of SegNet model used to extract features is optimized through selection, crossover and mutation of genetic algorithm, and then the improved SegNet semantic model (GA-SegNet model) is obtained by GA. In order to verify the image classification effect of the proposed GA-SegNet model, the same high-resolution remote sensing image data are used for experiments, and the model is compared with maximum likelihood (ML), support vector machine (SVM), traditional CNN and SegNet semantic model without GA improvement. The experimental results show that the proposed GA-SegNet model has the best classification accuracy and effect, which GA overcomes the problem of premature convergence of BP random gradient descent to a certain extent, and improves the classification performance of SegNet semantic model.

## 1. INTRODUCTION

In recent years, with the rapid development of remote sensing technology, the spatial resolution of remote sensing image is getting higher and higher, and the spatial information of image is also more rich and detailed. At the same time, the correlation between information in high-resolution remote sensing images is more complex, and the key to extract the information of remote sensing images is to select a reasonable and effective image classification method. High-resolution remote sensing images have rich texture features and obvious geometric structure. Designing a more efficient and accurate classification model of remote sensing images can quickly and accurately grasp the number and distribution of various types of objects in the region. It has important practical significance for environmental protection, urbanization construction and sustainable development.

At present, the commonly used classification algorithms in the field of remote sensing image classification are maximum likelihood (ML) (Milas et al., 2017, Peng et al., 2018), ISO clustering (Tao et al., 2018), support vector machine (SVM) (Wang et al., 2017, Li et al., 2018), random forest (RF) (Liu et al., 2012, Shi et al., 2018), neural network (NN) algorithms (Nogueira et al., 2016, Ma et al., 2018). However, these commonly used methods basically rely on the spectral characteristics of images for classification. When the spectral characteristics of remote sensing images are low, the

classification effect of these methods will be greatly affected. At this stage, deep learning algorithms have shown immeasurable potential in image classification (Li et al., 2016, Ma et al., 2016), target detection (Druzhkov and Kustikova, 2016, Han et al., 2018), speech recognition (Dai, 2017). Since image semantics segmentation algorithm has strong spatial feature extraction ability, an increasing number of scholars have applied it to remote sensing image classification, such as Markov random field algorithm in document (Nishii, 2003, Yang et al., 2013), Bayesian algorithm in document (Bruzzone, 2000, Li and Yin, 2013), and conditional random field algorithm in document (Zhong et al., 2011, Zhong et al., 2014, Guo et al., 2016). However, these traditional segmentation methods have a large amount of parameters, and the efficiency of image segmentation is relatively low. Literature (Long et al., 2014) proposes a classical full convolutional networks (FCN) based on semantic segmentation network. Classification based on FCN remote sensing imagery is an end-to-end architecture. The network can restore image size by up-sampling, which can not only recognize the category of pixels, but also restore the original pixels. Pixel-level classification of image is realized by locating in the original image. FCN abandons the full connection layer of traditional convolutional neural network, reduces the parameters of the neural network, reduces the complexity of the network, and improves the efficiency of segmentation. Literature (Chen et al., 2018) constructs a remote sensing image classification framework using FCN, and realizes dense pixelwise classification of high resolution remote sensing

\* Corresponding author

images. Literature (Badrinarayanan et al., 2017a) proposed the SegNet semantic segmentation model, the SegNet model is an improved model of FCN, It inherits the idea of FCN image semantic segmentation. The network combines the features of encoder-decoder structure and hopping network to make the model more Accurate output feature maps provide more accurate classification results with limited training samples. Literature (Yang et al., 2019) proposed the use of SegNe semantic model for high-resolution remote sensing imagery rural construction land extraction, and finally formed a better classification model. However, whether it is the traditional FCN or the SegNet semantic model, the encoder image extraction part is CNN. The most commonly used model training method of CNN is the back propagation algorithm. The core idea of this algorithm is random gradient descent, slow convergence speed, easy to fall into local optimal solution, which affects the image segmentation efficiency of the model. Designing a semantics segmentation model that can make the weight parameters converge globally effectively is still the research focus of using semantics segmentation model to classify high-resolution remote sensing images at present.

Therefore, in view of the above problems, and considering that genetic algorithm can inspire adaptive global search, this paper proposes to use genetic algorithm to select the weight of convolution layer in the encoder part of SegNet Semantic Model, in order to improve the BP algorithm convergence speed is slow, easy to fall into the problem of local optimal solution, and then improve the classification efficiency and accuracy of high-resolution remote sensing images.

## 2. EXPERIMENTAL DATA SET

The data set used in this experiment is a high-resolution remote sensing image of a region in southern China in 2015. It contains five large-scale RGB remote sensing images with spatial resolution of sub-meter and size ranging from 3000×3000 pixels to 6000×6000 pixels. Four types of objects are marked in the image, namely water body (Mark 1), vegetation (Mark 2), building (Mark 3) and road (Mark 4). Among them, grassland, cultivated land and woodland are classified as vegetation. Because the memory of computer is limited and the size of remote sensing image varies, the image can not be directly input into the neural network for training. Therefore, firstly, the remote sensing image is cut randomly, and the  $x$  and  $y$  coordinates are generated randomly on the image. Then, the 256×256 pixels small image is cut out under the coordinates. Thus, a small part of the training samples whose width and height are all 256 pixels are obtained. Figure 1 shows two original maps and corresponding label maps in the data set.

Because there are fewer data for network training and

verification, the data is enhanced, and the corresponding enhancement function is written using the OpenCV library. The enhancement function can rotate the original image and the label image by 90 degrees, 180 degrees, and 270 degrees, and Perform mirroring along the  $y$ -axis, then blur the original image, adjust the illumination, and increase the noise operation. After data enhancement operation, a large training set of 15,000 256×256 pixels pictures were obtained as input data for model training.

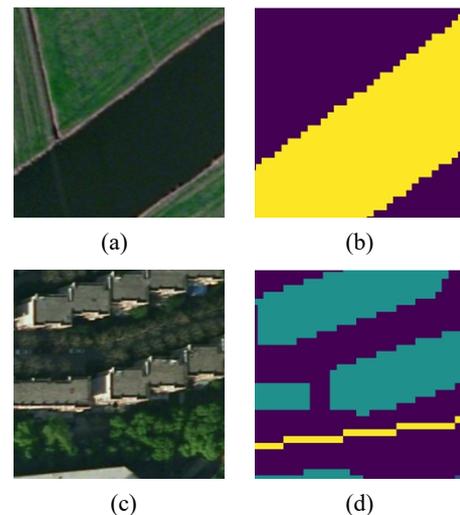


Figure 1. Original maps (a) and (c), and corresponding label maps (b) and (d)

## 3. GA-SEGNET SEMANTIC MODEL

### 3.1 SegNet Model

SegNet semantic segmentation model follows the design idea of FCN and improves FCN. The difference between them lies in the different technologies used in the encoder and decoder parts of the network structure. The whole framework of SegNet is shown in Figure 2 (Badrinarayanan et al., 2017b). The left half of the framework is the convolution feature extraction part, which enlarges the local receptive field and reduces the size of the picture by pooling operation. This process becomes Encoder. The Encoder part uses the first 13-layer convolution network of VGG16; The right half is the deconvolution and upsampling operation. The features of the classified image are reproduced by the deconvolution operation. The input image is restored to the original size by upsampling. The process is Decoder. Finally, through the Softmax layer, the maximum probability of different categories is output, and the final image segmentation image is obtained.

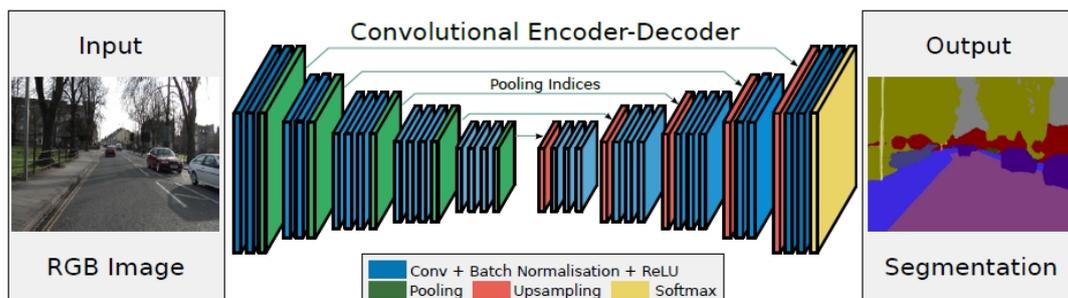


Figure 2. SegNet framework

Compared with FCN, SegNet semantic model has been improved in pooling and upsampling operations. Pooling in CNN can reduce the size of a picture by half. Pooling usually operates in two ways: max-pooling (Figure 3) and mean-pooling. Unlike the traditional pooling operation, the pooling operation in SegNet has an additional index function, that is, after each pooling operation, the corresponding position of the weight selected by max-pooling operation in  $2 \times 2$  filter is saved, such as the number "5" in figure 3, index starting from 0, the position of "5" in red  $2 \times 2$  filter is (1, 1), and the index corresponding to blue "3" is (0, 0). Upsampling operation is the inverse process of pooling operation. Upsampling operation can increase the size of the image. SegNet model uses index information to directly put data back to the corresponding location in the operation of upsampling, and then convolution layer training and learning. In addition to occupying some storage space, this upsampling operation does not need training and learning. In contrast, FCN uses deconvolution strategy, i.e., feature is deconvoluted to get upsampling after deconvolution operation. This process requires training and learning. Figure 4 illustrates the difference between FCN and SegNet models in upsampling operations (Badrinarayanan et al., 2017b).

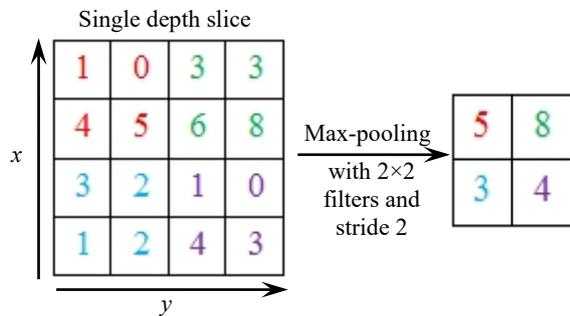


Figure 3. Max-pooling

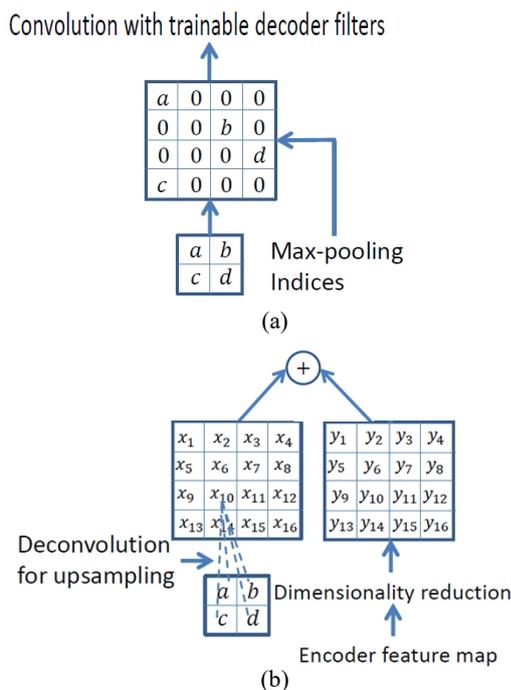


Figure 4. Upsample operation comparison between SegNet (a) and FCN (b)

### 3.2 Genetic Algorithm

GA is a random search algorithm based on natural selection and natural mechanism of biology. It is very suitable for dealing with complex and non-linear optimization problems which are difficult to be solved by traditional search algorithms. The genetic algorithm starts from the initial solution generated randomly, and generates new solution by iterating through certain selection, crossover and mutation operations step by step. The following are the concrete steps of the algorithm implementation:

1. Population initialization. Because the genetic algorithm can not directly deal with the parameters of the problem space, it is necessary to transform the feasible solution of the problem into chromosomes in the genetic space by coding. Common coding methods include bit string coding, Grey coding, real number coding, etc.
2. Fitness calculation. Fitness function is a criterion for judging the quality of individuals in a population. The fitness function is the only basis for natural selection. The fitness function is usually transformed by the objective function.
3. Select operation. Selection operation chooses good individuals from the old population with certain probability to form a new population, and further generates the next generation population. Individual fitness value is related to the probability of individual being selected. The larger the fitness value of individual is, the higher the probability of being selected is, and the probability of individual  $i$  being selected is as follows:

$$p_i = \frac{F_i}{\sum_{j=1}^N F_j} \quad (1)$$

where  $F_i$  = the fitness value of individual  $i$   
 $N$  = the number of individual population

4. Cross operation. Crossover operation refers to the random selection of two individuals from a population and the exchange of two chromosomes to transfer the excellent genes of the father generation to the offspring, thus producing new excellent individuals. The cross-manipulation method of the  $k$  chromosome  $r_k$  and  $l$  chromosome  $r_l$  at the  $j$  position is as follows:

$$r_{kj} = r_{ij}(1-b) + r_{lj}b \quad (2)$$

$$r_{lj} = r_{ij}(1-b) + r_{kj}b \quad (3)$$

where  $b$  = a random number in the  $[0,1]$  interval

5. Mutation operation. The main purpose of the mutation operation is to maintain the diversity of the population. The mutation operation randomly selects an individual from the population and mutates a point in the individual to produce a better individual. The method of mutation of the  $j$  gene  $a_{ij}$  of the  $i$  individual is as follows:

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max}) \times f(g), & r \geq 0.5 \\ a_{ij} + (a_{\min} - a_{ij}) \times f(g), & r < 0.5 \end{cases} \quad (4)$$

$$f(g) = r_2(1 - g / G_{\max})^2 \quad (5)$$

where  $a_{\max}$  = the upper bound of gene  $a_{ij}$   
 $a_{\min}$  = the lower bound of gene  $a_{ij}$

$r_2$  = a random number  
 $g$  = the current iteration number  
 $G_{\max}$  = the largest evolution number  
 $r$  = the random number of [0,1] interval

### 3.3 Design of GA-SegNet Model

The GA-SegNet model is implemented in Python language. Figure 5 introduces the main technical process of using genetic algorithm to improve the semantics segmentation model SegNet.

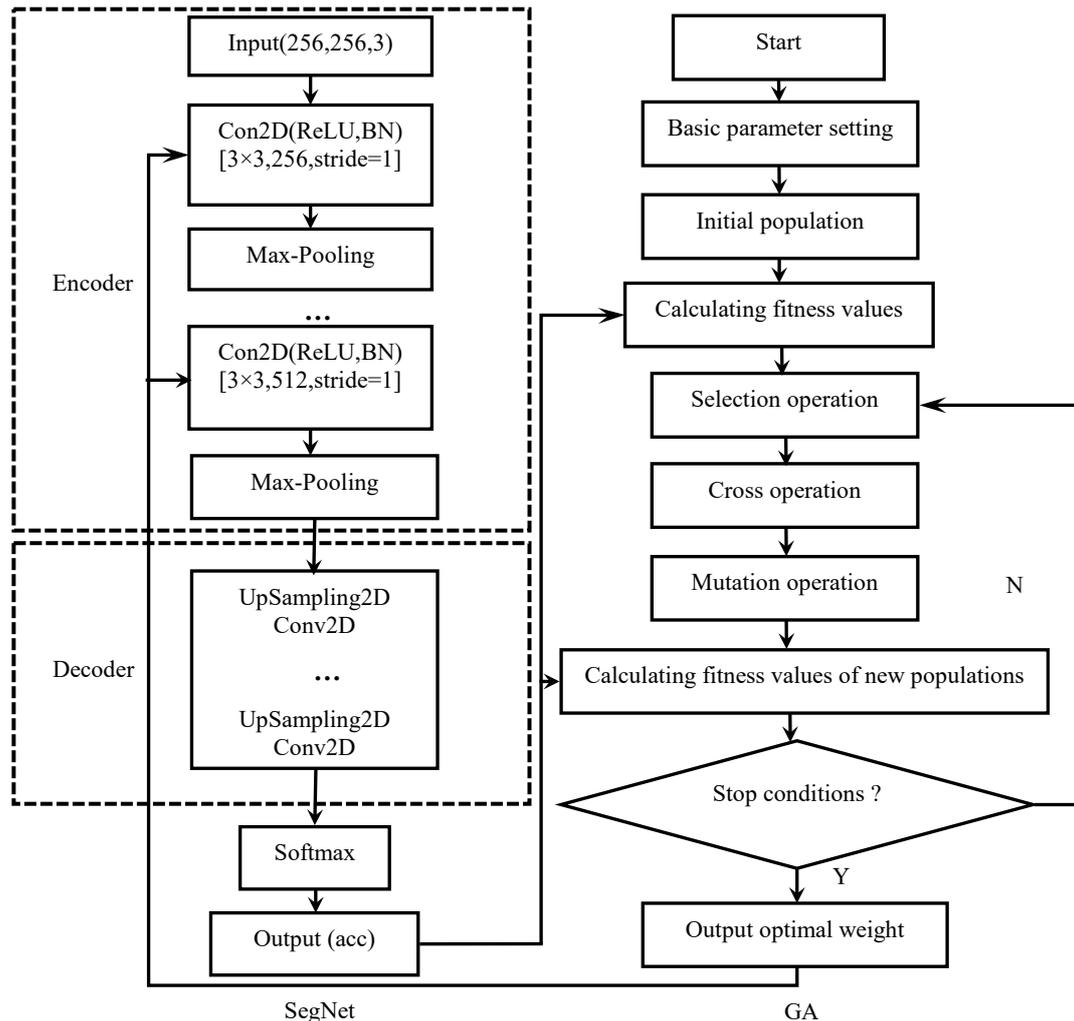


Figure 5. Technical flow chart of GA-SegNet model

The main design idea of the model GA-SegNet is to first randomly generate chromosomes corresponding to the convolution layer weights of the SegNet encoder part of the semantic model using the genetic algorithm. Each chromosome gene contains the total number of weights, and then the randomly generated chromosomal genes are used as the weight is assigned to the corresponding encoder convolution layer, the input training and verification samples are trained on the SegNet model, and the validation accuracy (acc) of the final output sample is used as the fitness function of genetic algorithm. The fitness value is used to evaluate the chromosome quality of genetic algorithm. Finally, the population is updated iteratively by genetic operation selection, crossover and mutation, and the new population is calculated. The fitness of the semantics model SegNet encoder leaves a chromosome with

high fitness in the iteration, and then gets the optimal weight of the convolution layer of the semantics model SegNet encoder.

The key point of GA-SegNet model design is how to determine the number of chromosome nodes, and convert it into a weight matrix consistent with the form of convolution layer to give the network convolution layer. Finally, the fitness function of genetic algorithm is selected. In order to determine the number of nodes (variables) of each chromosome in genetic algorithm, the weight matrix of each convolution layer in SegNet coding part of the semantic model is obtained, and the number of elements of the weight matrix is calculated. The total weight number of all convolution layers, i.e. the number of nodes of chromosomes is obtained. After the population initialization of the genetic algorithm is performed, the gene string of each

chromosome in the population is sliced, the number of segment nodes of the slice is equal to the weight of each convolution layer, and is set to the convolution layer weight matrix. In a completely consistent form, the matrix obtained by gene conversion is directly assigned to the convolutional layer of the encoder to train the network. As for the fitness function determination of GA, the accuracy of the final output of the validated samples in the training process of the model can directly explain the selection of the weight parameters of the whole network and indirectly evaluate the merits and demerits of each chromosome gene, so the accuracy of the final output of the validated samples in the network is regarded as the fitness function of GA.

#### 4. EXPERIMENTAL RESULTS AND ANALYSIS

##### 4.1 Training and Validation Sample Organization

In order to input sample size smoothly into the semantic model SegNet and ensure the best training effect, and maximize the classification accuracy of the model, the cutting size is  $256 \times 256$  pixels to ensure the accuracy of classification. Finally, when the model reads the data set, the size of the validation set selected is 25% of the training set. Finally, the model forms 11,250 training samples with  $256 \times 256$  pixel size and 3,750 validation samples with  $256 \times 256$  pixel size.

##### 4.2 Model Parameter Setting

Before training GA-SegNet model, some basic parameters of the model need to be set. Learning rate is used to control the global learning rate of the model. Excessive and low learning rate will lead to slow divergence and convergence of the model, respectively. In this experiment, the initial learning rate is set to 0.01; momentum can be used to control the global learning rate.

Accelerating the convergence speed of the model is set to 0.8; the learning rate change index (gamma) determines the acceleration of learning rate and sets it to 0.1; the weight decay value (weight decay) and the stepsize evaluation rate (stepsize) are respectively set to 0.0004, 1500; considering the batch sizes that computer memory can withstand. After testing with different batch sizes, the final settings are 4, and the number of model epochs is 15. In the basic parameter setting of GA, the number of population is set to 60, the number of genetic iterations is set to 5, the probability of genetic crossover is set to 0.8, and the probability of variation is set to 0.01.

In terms of parameter setting of comparative experiments, the ML, SVM, traditional CNN and model SegNet were used to classify high-resolution remote sensing images. When using the SVM for classification, the maximum number of examples for each category is set to 500; in the image classification using the traditional CNN of the two layers of the convolutional pooling layer and the unmodified semantic model SegNet, the settings of learning rate, gamma, momentum value, weight decay value, stepsize frequency and iteration times are the same as that of GA-SegNet model.

##### 4.3 Experimental Results

After the training of GA-SegNet model is completed, the GA-SegNet.h5 file of the model is saved, which contains the structure and weight parameters of the model. The test remote sensing image is input into the trained model for prediction and classification. At the same time, the same test is carried out using ML, SVM, traditional CNN and the SegNet model. The remote sensing image is classified and predicted. Figure 6 is the result of remote sensing image classification of the method and the comparison method in this paper.

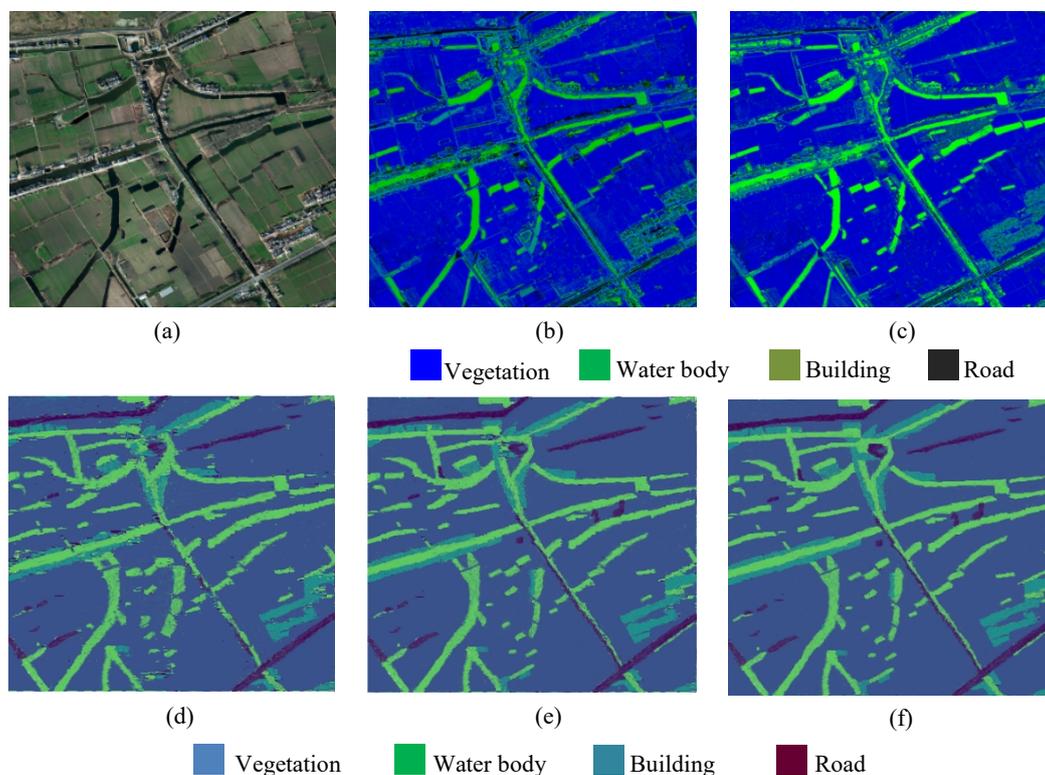


Figure 6. Classification results of remote sensing image (a) by ML (b), SVM (c), CNN (d), SegNet (e) and GA-SegNet (f)

#### 4.4 Accuracy Analysis

Accuracy evaluation of classification results is an important work in the process of remote sensing image classification. Usually, confusion matrix, classification accuracy of various categories, overall classification accuracy (OA), Kappa coefficient are used as the evaluation index of image classification accuracy. In order to verify the validity of GA-SegNet model, and based on the real image and classification

result of high resolution remote sensing images, 1000 random points are generated on the real image using ArcMap, and the confusion matrices of classification results between the proposed method and the comparison method are calculated respectively (Tables 1-5). According to the confusion matrix of each classification result, the corresponding classification accuracy of single object category, overall classification accuracy and Kappa coefficient can be obtained. The accuracy comparison of classification results is shown in Table 6.

Category	Water body	Vegetation	Building	Road	Line total
Water body	191	31	13	11	246
Vegetation	26	423	22	12	483
Building	6	7	150	20	186
Road	6	7	10	62	85
Column total	229	468	195	105	1000

Table 1. ML classification obfuscation matrix

Category	Water body	Vegetation	Building	Road	Line total
Water body	202	25	10	9	246
Vegetation	19	439	17	8	483
Building	5	7	155	19	186
Road	5	5	8	67	85
Column total	231	476	190	103	1000

Table 2. SVM classification obfuscation matrix

Category	Water body	Vegetation	Building	Road	Line total
Water body	211	20	8	7	246
Vegetation	14	450	13	6	483
Building	3	4	168	11	186
Road	2	3	5	75	85
Column total	230	477	194	99	1000

Table 3. CNN classification obfuscation matrix

Category	Water body	Vegetation	Building	Road	Line total
Water body	214	19	8	5	246
Vegetation	10	461	9	3	483
Building	3	3	172	8	186
Road	1	2	3	79	85
Column total	228	485	192	95	1000

Table 4. SegNet classification obfuscation matrix

Category	Water body	Vegetation	Building	Road	Line Total
Water body	227	11	5	3	246
Vegetation	7	469	5	2	483
Building	2	3	175	6	186
Road	1	1	2	81	85
Column Total	237	484	187	92	1000

Table 5. GA-SegNet classification obfuscation matrix

Accuracy	Category	ML	SVM	CNN	SegNet	GA-SegNet
Single category classification accuracy	Water body	0.7764	0.8211	0.8577	0.8699	0.9228
	Vegetation	0.8758	0.9089	0.9317	0.9546	0.9710
	Building	0.8065	0.8333	0.9032	0.9247	0.9409
	Road	0.7294	0.7882	0.8824	0.9294	0.9529
OA		82.60%	86.30%	90.40%	92.60%	95.20%
Kappa		0.7412	0.7553	0.8564	0.8889	0.9279

Table 6. Accuracy comparison of classification results of different classification methods

Using ML, SVM, CNN, SegNet semantic model and GA-SegNet model to classify remote sensing images, the above five different obfuscation matrices are obtained, and the accuracy of single category classification, OA and Kappa coefficient are obtained by further calculation of obfuscation matrices. Through the analysis of the data in Table 6, it can be seen that the classification accuracy of water body has increased from 0.7764 of ML to 0.9228 of this method, and the accuracy has increased by 0.2164. The classification accuracy of vegetation, buildings and roads has increased by 0.0952, 0.1344 and 0.2235 respectively from ML to this method. In terms of OA, the accuracy of the method in this paper and the comparison method is more than 82%. Compared with the ML and SVM the overall classification accuracy of the method in this paper is significantly improved. Compared with the 90.40% and 92.60% of the traditional CNN method and SegNet semantic model, it is also slightly improved, reaching 95.20%. In terms of Kappa coefficients, the Kappa coefficients of the comparison methods are below 0.9. The Kappa coefficients of the ML and the SVM are below 0.8. The traditional CNN method and SegNet semantic model are 0.8564 and 0.8889, respectively. The highest Kappa coefficients of this method is 0.9279. Whether from the single category classification accuracy, OA or Kappa coefficient, it can be seen that the method in this paper is superior to the other four comparison methods, showing better classification ability.

## 5. CONCLUSION

Aiming at the problem that the error back propagation algorithm used in SegNet semantic model encoder is easy to fall into local optimal solution, this paper proposes to introduce genetic algorithm into the weight optimization of SegNet semantic model encoder, so that the original SegNet model has the advantage of global convergence of genetic algorithm. According to the experiments of high-resolution remote sensing image classification, compared with the ML, SVM, traditional CNN and SegNet semantics segmentation model, the GA-SegNet model presented in this paper shows the best classification effect and accuracy in single category classification accuracy, OA and Kappa coefficient. To some extent, genetic algorithm overcomes the problem that BP random gradient descent easily falls into premature convergence, and improves the classification performance of SegNet semantic model. In the future, the work will focus on how to set the optimal basic parameters of genetic algorithm to play its maximum role.

## ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under grant number 41461085; the Natural Science Foundation of Guangxi Province under grant number 2016GXNSFAA380035; the Foundation of Guangxi Key Laboratory of Spatial Information and Geomatics under grant number 16-380-25-04; the “BaGui Scholars” Special Funds of Guangxi Province under grant number 2019-79; and the Doctoral Foundation of Guilin University of Technology under grant number 1996015.

## REFERENCES

- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017a. SegNet: a deep convolutional encoder-decoder architecture for scene segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 2481-2495.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017b. SegNet: a deep convolutional encoder-decoder architecture for scene segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 99, 1-1.
- Bruzzzone, L., 2000. An approach to feature selection and classification of remote sensing images based on the Bayes rule for minimum cost. *IEEE Transactions on Geoscience & Remote Sensing*, 38(1), 429-438.
- Chen, G. S., Chen, S. Y., Jing, W. P., Li, C., 2018. Pixel-level classification of remote sensing images based on genetic algorithm and full convolution network. *Journal of Chinese Mini-Micro Computer Systems*, 39(7), 216-221.
- Dai, L. R., Zhang, S. L., Huang, Z. Y., 2017. Status and prospect of speech recognition technology based on in-depth learning. *Journal of Data Acquisition & Processing*, 32(2), 221-231.
- Druzhkov, P. N., Kustikova, V. D., 2016. A survey of deep learning methods and software tools for image classification and object detection. *Pattern Recognition & Image Analysis*, 26(1), 9-15.
- Guo, C., Zhou, L., Li, Y., 2016. A new change-detection method in high-resolution remote sensing images based on a conditional random field model. *International Journal of Remote Sensing*, 37(5), 1173-1189.

- Han, J. W., Zhang, D. W., Cheng, G., Liu, N., Xu, D., 2018. Advanced deep-learning techniques for salient and category-specific object detection: a survey. *IEEE Signal Processing Magazine*, 35(1), 84-100.
- Li, C., Yin, J., 2013. A multispectral remote sensing data spectral unmixing algorithm based on variational bayesian ICA. *Journal of the Indian Society of Remote Sensing*, 41(2), 259-268.
- Li, N., Zhu, X. F., Pan, Y. Z., Zhan, P., 2018. SVM remote sensing image classification optimized by artificial bee colony algorithm. *Journal of Remote Sensing*, 22(4), 559-569.
- Li, W. J., Fu, H. H., Yu, L., 2016. Stacked autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping. *International Journal of Remote Sensing*, 37(23), 5632-5646.
- Liu, Y., Du, P. J., Xia, J. S., Liu, S. C., 2012. Classification of small satellite remote sensing images based on random forests. *Science of Surveying and Mapping*, 37(4), 194-196.
- Long, J., Shelhamer, E., Darrell, T., 2014. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 39(4), 640-651.
- Ma, X., Wang, H., Wang, J., 2016. Semisupervised classification for hyperspectral image based on multi-decision labeling and deep feature learning. *ISPRS Journal of Photogrammetry & Remote Sensing*, 120, 99-107.
- Ma, Y. F., Zhao, J. B., Li, J., Qiao, X. S., 2018. Classification of remote sensing images based on artificial neural networks. *Standardization of Surveying and Mapping*, 34 (4), 34-37.
- Milas, A. S., Arend, K., Mayer, C., Simonson, M. A., Mackey, S., 2017. Different colours of shadows: classification of UAV images. *International Journal of Remote Sensing*, 38(8-10), 3084-3100.
- Nishii, R., 2003. A Markov random field-based approach to decision-level fusion for remote sensing image classification. *Geoscience & Remote Sensing IEEE Transactions on*, 41(10), 2316-2319.
- Nogueira, K., Penatti, Otávio, A. B., Jefersson, A. D. S., 2016. Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognition*, 61, 539-556.
- Peng, J., Li, L., Tang, Y. Y., 2018: Maximum likelihood estimation-based joint sparse representation for the classification of hyperspectral remote sensing images. *IEEE Transactions on Neural Networks and Learning Systems*, 99, 1-13.
- Shi, C. X., Zhao, C. G., Pang, L., 2018. Random forest remote sensing image classification based on superpixel statistics. *Application Research of Computers*, 35(12), 284-288.
- Tao, J., Dan, H., Yu, X., 2018. Enhanced IT2FCM algorithm using object-based triangular fuzzy set modeling for remote-sensing clustering. *Computers & Geosciences*, 118, 14-26.
- Wang, H. L., Guo, W., Shi, Y., Qiao, H. B., 2017 Research on hyperspectral remote sensing image classification method based on Grey-Sigmoid kernel function support vector machine. *Journal of Henan Agricultural University*, 51(3), 402-407.
- Yang, H. L., Peng, J. H., Xia, B. R., Zhang, D. X., 2013. Remote sensing classification using fuzzy C-means clustering with spatial constraints based on Markov random field. *European Journal of Remote Sensing*, 46(1), 305-316.
- Yang, J. Y., Zhou, Z. X., Du, Z. R., Xu, Q. Q., Yin, H., Liu, R., 2019. Extraction of rural construction land from high resolution remote sensing images based on SegNet semantic model. *Transactions of the CSAE*, 35(5), 259-266.
- Zhong, P., Liu, F., Wang, R., 2011. Learning sparse conditional random fields to select features for land development classification. *International Journal of Remote Sensing*, 32(15), 4203-4219.
- Zhong, Y., Ji, Z., Zhang, L., 2014. A hybrid object-oriented conditional random field classification framework for high spatial resolution remote sensing imagery. *IEEE Transactions on Geoscience & Remote Sensing*, 52(11), 7023-7037.