

## Study on the Classification of Gaofen-3 Polarimetric SAR Images Using Deep Neural Network

ZHANG Jiaqi<sup>1</sup>, ZHANG Jixian<sup>1,2</sup>, ZHAO Zheng<sup>1,3</sup>

<sup>1</sup> Chinese Academy of Surveying and Mapping, Beijing 100830, China – sddmzhang@163.com

<sup>2</sup> National Quality Inspection and Testing Center for Surveying and Mapping Products, Beijing 100830, China – zhangjx@casm.edu.cn

<sup>3</sup> Beijing Key Laboratory of Urban Spatial Information Engineering, Beijing 100830, China – 20324790@qq.com

Commission VI, WG III/2

**KEY WORDS:** SAR Classification, Deep Learning, GoogLeNet, Landcover, Gaofen-3 satellite

### ABSTRACT:

Polarimetric Synthetic Aperture Radar (POLoSAR) imaging principle determines that the image quality will be affected by speckle noise. So the recognition accuracy of traditional image classification methods will be reduced by the effect of this interference. Since the date of submission, Deep Convolutional Neural Network impacts on the traditional image processing methods and brings the field of computer vision to a new stage with the advantages of a strong ability to learn deep features and excellent ability to fit large datasets. Based on the basic characteristics of polarimetric SAR images, the paper studied the types of the surface cover by using the method of Deep Learning. We used the fully polarimetric SAR features of different scales to fuse RGB images to the GoogLeNet model based on convolution neural network Iterative training, and then use the trained model to test the classification of data validation. First of all, referring to the optical image, we mark the surface coverage type of GF-3 POLoSAR image with 8m resolution, and then collect the samples according to different categories. To meet the GoogLeNet model requirements of  $256 \times 256$  pixel image input and taking into account the lack of full-resolution SAR resolution, the original image should be pre-processed in the process of resampling. In this paper, POLoSAR image slice samples of different scales with sampling intervals of 2m and 1m to be trained separately and validated by the verification dataset. Among them, the training accuracy of GoogLeNet model trained with resampled 2-m polarimetric SAR image is 94.89%, and that of the trained SAR image with resampled 1 m is 92.65%.

### 1. INTRODUCTION

Polarization Synthetic Aperture Radar (PolSAR) image classification is one of the hot research directions. It can provide basic support for land use surveys, geographic national conditions monitoring, urban and rural planning and many other fields. Traditional image classification methods often use a specific calculation method to classify specific features. As a result, their accuracy will exhibit greater instability because of the effect by different features.

Because the polarimetric SAR image has much random speckle noise and the lower resolution, The features of different types of objects also easily show similar features in PolSAR images. So the traditional image classification method will be subject to greater interference. The manual interpretation method not only consumes a lot of manpower and time, but the accuracy rate is usually also influenced by the subjective factors.

Considering the above problems and the existing deficiencies of the existing classification methods, and after studying the outstanding achievements of the deep learning method in the field of image processing, we introduce the idea of deep learning into the study of polarimetric SAR image classification. So we use a convolutional neural network to classify the sample sets of polarized SAR images at different scales.

The main idea of the research is to construct a sample set based on different scales by re-sampling the original resolution GF-3 PolSAR data with different resolutions. Then each type of sample is divided into a training sample set and a validation sample set according to a certain. At the same time, multi-scale convolutional neural networks are built for samples of different

scales, and the models are trained and parameters are optimized to obtain classification models for SAR images of different scales. Finally, we will use the validation sample set to verify the model classification accuracy.

### 2. DEEP LEARNING

The deep learning model is inspired by the human brain structure. It relies on the coupling of multiple neurons to perform layer-by-layer abstraction on the input data. So it can display its powerful capabilities in areas such as images, text, and voice, and then be used in more fields. Due to its powerful learning ability and generalization ability, deep neural networks have gradually replaced other machine learning methods and become the most important technology in the field. More and more scholars in the field of machine learning begin to get involved and apply deep learning.

In recent years, computer computing capabilities have advanced rapidly and in 2006 Hinton et al. proposed a layer-by-layer greedy algorithm for training neural networks, marking the third wave of deep learning and the beginning of a real era of deep learning. There are already many existing deep learning models, and new models have emerged on the frequency of months or even weeks. However, the basic frameworks of these models mainly include deep confidence neural networks (DBN), convolutional neural networks (CNN), Self-encoder (AE) and so on. Among them, CNN is the most widely used in image processing.

## 2.1 Convolutional Neural Networks

Convolutional neural networks (CNN) is a neural network designed to process data with a grid structure, such as time series data (which can be viewed as one-dimensional grid data arranged along the time axis) and image data (two-dimensional Grid data). Next, we only discuss its application in the field of image data.

The core algorithm of convolutional neural networks is convolution and pooling.

Convolution is usually a mathematical operation of two real-variable functions, such as:

$$s(t)=(x*\omega)(t)$$

Among them, the function  $x$  is often referred to as an input, and the function  $\omega$  is referred to as a kernel function. When the parameter is a negative value, the value of  $\omega$  is 0.  $S(t)$ , the output, is commonly referred to as a feature map.

In the convolutional neural network model, it first performs partial perceptions on the lower layers of the network, and gradually synthesizes low-level information at a higher level to complete the overall feeling and obtain overall information. This can greatly reduce the amount of calculations without having to perceive global information. At the same time, the convolutional neural network can obtain different characteristics by changing different convolution kernels.

Because the convolutional neural network model has a small convolution kernel, the feature dimension obtained through the convolution operation is still extremely large. In order to avoid over-fitting of the network, pooling is a good solution. Pooling functions use the overall statistical characteristics of adjacent outputs at a location instead of the network output. The max pooling function is usually the most common, replacing the pixel value with the largest value in the adjacent rectangular area.

After Lecun's first use of convolutional neural networks for the identification of handwritten digits, different convolutional neural network models have mushroomed and the number of model layers has increased rapidly. LeNet, which was originally used for hand-written digital recognition, has only 8 layers of neurons. But its structure is very complete, including the input layer, convolution layer, pooling layer, full connection layer and output layer. Small but complete. The later models such as Alexnet, VGG, and GoogLeNet have also been based on LeNet to deepen the hidden layer and adjust the structure, but the basic structure remained the same. Among them, GoogLeNet is a new model proposed by Google for participating in ILSVRC 2014. With its superior computing efficiency and excellent network structure, GoogLeNet has gradually become one of the most commonly used basic models of deep learning research in image processing.

## 2.2 GoogLeNet

GoogLeNet first appeared in the ILSVRC 2014 contest. Prior to this convolutional neural network model, the most direct way to improve computing performance was to increase the number and width of network layers. But the ensuing increase in the huge number of parameters makes it easy for the network to enter a dead-end of overfitting and greatly increase the amount of computation. To resolve this problem, the Google

team proposed a new structure called Inception to maintain network sparsity while also making better use of the computational performance of dense matrices.

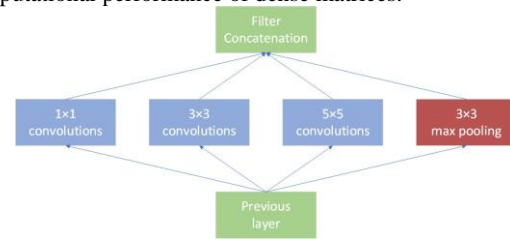


Figure 1. Inception structure

In the GoogLeNet V1 and V2, each new  $1 \times 1$  convolution kernel is again added to form a new basic structure. The GoogLeNet network model is composed of these basic structures combined with multiple softmax classifiers. This structure effectively avoids the over-fitting caused by overly dense network models, and solves the problems of traditional neural network models that use random sparse links to reduce the computational performance for improving learning ability. It provides a good structural basis for the development of the convolutional neural network model.

## 3. EXPERIMENTAL RESULTS

In order to satisfy the determination of PolSAR image sample categories at different scales, we created 2m and 1m training sample sets to construct a multi-scale GoogLeNet network model. Data is divided into 30 epochs during training. In order to speed up the fitting speed of the convolutional neural network model to the data and prevent overfitting, the learning rate was chosen to be 0.01 in the first ten Epochs and then decreased to 0.001 and 0.0001. In the final network model, the trained training accuracy of the GoogLeNet network model with two different resampling resolutions is 94.89% and 92.65%, respectively. It can be seen that as the sample size decreases, the training accuracy of the model decreases. It is not difficult to find the reason after analyzing the sample and convolutional neural network structure. The upsampling of the original image does not increase the information contained in the image itself. Under the same pixel number standard, the higher the resampling resolution, the less information it contains. The working method of convolutional neural network is to continuously abstract the image to higher level features. Therefore, the model classification performance depends on the amount of effective information contained in the image, and the richer the effective information contained in the image itself. Under the same network training conditions, the better the model classification performance. The prediction experiment of the validation set also confirmed the rule of decreasing precision.

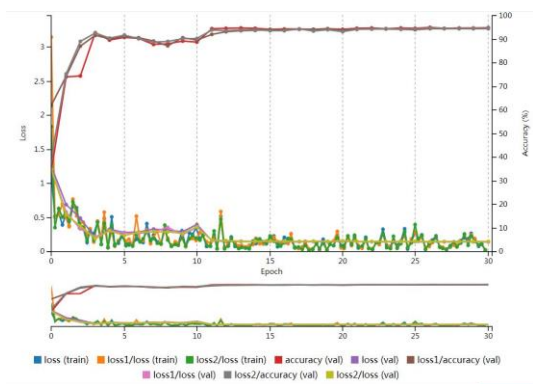


Figure 2. Model training accuracy with 2m resolution samples

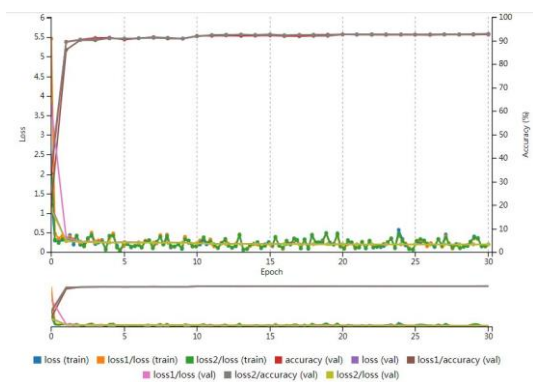


Figure 3. Model training accuracy with 1m resolution samples

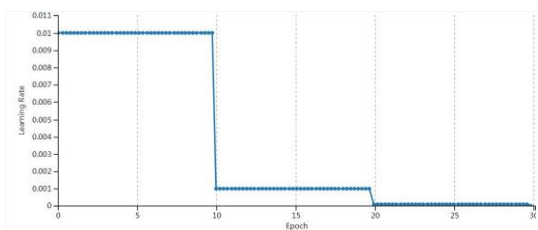


Figure 4. Learning rate

From the training accuracy chart above, we can see that the training accuracy of the two models tends to be stable after 10 epochs, and the learning rate decreases from 0.01 to 0.001.

Resolution	Type		
	Water	Building	Vegetation
2m	96%	96%	89.43%
1m	97.19%	94.98%	80.65%

Table 1. GoogLeNet model validation accuracy at different resampling resolutions

After training the GoogLeNet network model with different scale training sets, the classification accuracy of the three different types of land features has presented the law of higher recognition accuracy of land objects and buildings and lower recognition accuracy of vegetation types, which means that The well-trained model can achieve better classification performance for water bodies and buildings, while there are

many samples of the type of vegetation that are wrongly divided into building types. After analyzing the misclassified samples, it was found that during the sampling process, due to some obvious boundaries between the vegetation coverage areas, some of the samples contained some regular textures resembling building edges, and the convolutional neural network itself It is more sensitive to texture information, which makes the model erroneously judge it as a building category based on this feature when predicting the sample category. At the same time, in the course of the experiment, it was found that not only large-scale water samples were found in the water body samples, but also some of the samples were paddy fields or ponds that contained not only water but also some regular roads or ground areas. This part of the sample also has a very high classification accuracy.

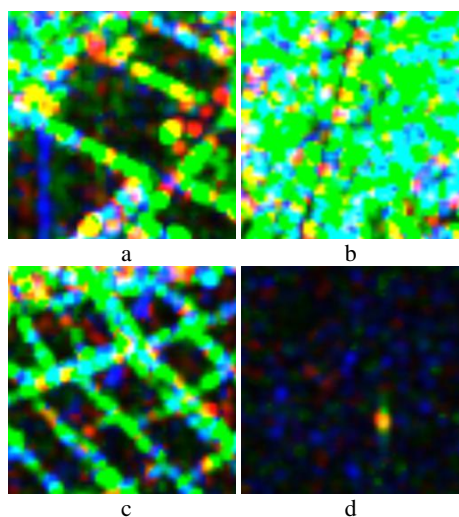


Figure 5. Sample diagram of 2m resolution. a)Building, b)Vegetation, c) Paddy field, d)Water

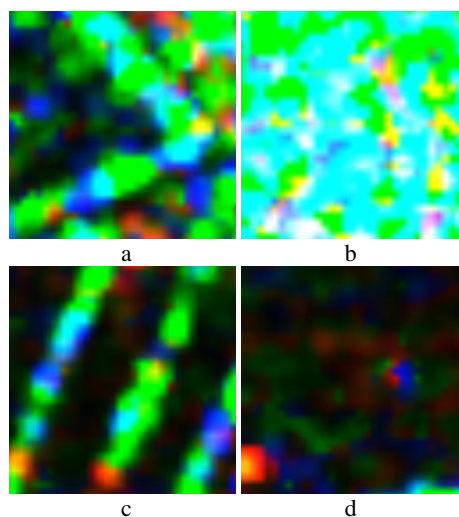


Figure 6. Sample diagram of 1m resolution. a)Building, b)Vegetation, c) Paddy field, d)Water

Among them, paddy field samples and water samples are classified as water samples.

It can be clearly seen from the above sample diagram that as the sample size is getting smaller and smaller, the amount of information contained in a single sample is smaller, and the

ability of the GoogLeNet network model to distinguish between different classes is also weaker.

#### 4. SUMMARY

After making some research on the existing methods for the classification of polarimetric SAR imagery, we use the deep learning image classification method of convolutional neural network to study the classification of Gaofen-3 PolSAR images. In order to study the ability of convolutional neural networks to classify the SAR samples under different scales, we use different scale samples to train multiple parallel GoogLeNet network models. And then Input different scale samples into corresponding scale network models for class prediction. Analyze the causes of differences in training accuracy at different scales, and the reasons for the differences in verification accuracy of different types of terrain.

At the same time, the classification method proposed in this paper also has its shortcomings. For example, for a small-scale classification sample, the fewer effective feature information it contains, the influence of the GoogLeNet convolutional neural network model structure characteristics, the accuracy of the model using small-scale sample training will be reduced accordingly. However, it is not considered that because the number of model layers is too large and the depth of image features is too deep, it is not as good as a shallower convolutional neural network model. Therefore, in future research work, it may be considered to use different network models with different layers for different scales to classify samples to improve the classification accuracy of samples at different scales.

#### ACKNOWLEDGEMENTS

This work was supported in part by the CASM Fundamental research funds (No.7771715).

#### REFERENCES

GUO Yanhe. Polarimetric SAR Classification Based on Depth Learning[D]. Xi'an University of Electronic Science and Technology,2015.

Zhong Z, Jin L, Xie Z. High Performance Offline Handwritten Chinese Character Recognition Using GoogLeNet and Directional Feature Maps[J]. 2015:846-850.

Jin X, Chi J, Peng S, et al. Deep image aesthetics classification using inception modules and fine-tuning connected layer[C]// International Conference on Wireless Communications & Signal Processing. IEEE, 2016:1-6.

Cerentini A, Welfer D, Cordeiro M D, et al. Automatic Identification of Glaucoma Using Deep Learning Methods.[C]// Medinfo 2017 Precision Healthcare through Informatics. 2017.

Xu Y, Scott K A. Sea ice and open water classification of SAR imagery using cnn-based transfer learning[C]// IGARSS 2017 - 2017 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2017:3262-3265.

Hao Guo, Danni Wu, Jubai An. Discrimination of Oil Slicks and Lookalikes in Polarimetric SAR Images Using CNN[J]. Sensors, 2017,08,1837

Ballester P, Araujo R M. On the performance of GoogLeNet and AlexNet applied to sketches[C]// Thirtieth AAAI Conference on Artificial Intelligence. AAAI Press, 2016:1124-1128.

An Q, Pan Z, You H. Ship Detection in Gaofen-3 SAR Images Based on Sea Clutter Distribution Analysis and Deep Convolutional Neural Network[J]. Sensors, 2018, 18(2):334.

Hao Dong ,Xin Xu,Lei Wang,Fangling Pu .Gaofen-3 PolSAR Image Classification via XGBoost and Polarimetric Spatial Information[J]. Sensors, 2018, 18(2), 611

Han Y, Li Y, Yu W. SAR target segmentation based on shape prior[C]// Geoscience and Remote Sensing Symposium. IEEE, 2014:3738-3741.

Cao F, Hong W, Pottier E. An improvement for the unsupervised Wishart Freeman classification with fully polarimetric SAR data[C]// Geoscience and Remote Sensing Symposium. IEEE, 2011:320-322.

Yang J, Yang R. Unsupervised classification of polarimetric SAR images using complex Wishart distribution based on H/ $\alpha$  decomposition and algorithm evaluation[J]. 2007, 6790

Deng J, Zhang Z, Marchi E. Sparse Autoencoder-Based Feature Transfer Learning for Speech Emotion Recognition[C]// Affective Computing and Intelligent Interaction. IEEE, 2013:511-516.

Chen S, Wang H. SAR target recognition based on deep learning[C]// International Conference on Data Science and Advanced Analytics. IEEE, 2015:541-547.

Mason E, Yazici B. Deep learning for SAR image formation[C]// Society of Photo-Optical Instrumentation Engineers. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2017:1020104.

Liu C, Yin J, Yang J. Application of deep learning to polarimetric SAR classification[C]// Iet International Radar Conference. 2015:4. -4.

Gong M, Yang H, Zhang P. Feature learning and change feature classification based on deep learning for ternary change detection in SAR images[J]. Isprs Journal of Photogrammetry & Remote Sensing, 2017, 129:212-225.

An Q, Pan Z, You H. Ship Detection in Gaofen-3 SAR Images Based on Sea Clutter Distribution Analysis and Deep Convolutional Neural Network[J]. Sensors, 2018, 18(2):334.