QUALITY EVALUATION OF LAND-COVER CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

Dang Yu[1],Zhang Jixian[1*], Zhao Yousong[1],Luo Fujun[1],Ma Wei[1],Yu Fan[2]

[1]National Quality Inspection and Testing Center for Surveying and Mapping Products,
[2]Chinese Academy of Surveying and Mapping

Commission VI, WG III/IVb

KEY WORDS: Remote Sensing, Land Cover classification, Deep Learning, Convolutional Neural Network, Quality evaluation

ABSTRACT:

Land-cover classification is one of the most important products of earth observation, which focuses mainly on profiling the physical characters of the land surface with temporal and distribution attributes and contains the information of both natural and man-made coverage elements, such as vegetation, soil, glaciers, rivers, lakes, marsh wetlands and various man-made structures. In recent years, the amount of high-resolution remote sensing data has increased sharply. Accordingly, the volume of land-cover classification products increases, as well as the need to evaluate such frequently updated products that is a big challenge. Conventionally, the automatic quality evaluation of land-cover classification is made through pixel-based classifying algorithms, which lead to a much trickier task and consequently hard to keep peace with the required updating frequency. In this paper, we propose a novel quality evaluation approach for evaluating the land-cover classification by a scene classification method Convolutional Neural Network (CNN) model. By learning from remote sensing data, those randomly generated kernels that serve as filter matrixes evolved to some operators that has similar functions to man-crafted operators, like Sobel operator or Canny operator, and there are other kernels learned by the CNN model that are much more complex and can’t be understood as existing filters. The method using CNN approach as the core algorithm serves quality-evaluation tasks well since it calculates a bunch of outputs which directly represent the image’s membership grade to certain classes. An automatic quality evaluation approach for the land-cover DLG-DOM coupling data (DLG for Digital Line Graphic, DOM for Digital Orthophoto Map) will be introduced in this paper. The CNN model as an robustness method for image evaluation, then brought out the idea of an automatic quality evaluation approach for land-cover classification. Based on this experiment, new ideas of quality evaluation of DLG-DOM coupling land-cover classification or other kinds of labelled remote sensing data can be further studied.

1. REMOTE SENSING IMAGE CLASSIFICATION BY CNN

1.1 Instructions of Convolutional Neural Network

Convolutional Neural Network(CNN) as an important model of machine learning have been proposed for a long time, it delivers an innovative approach of convolutional structure to solve the problem of computer vision. The CNN named by convolution in reason of including convolution layers, each convolution layer contents a number of kernels which will be random initialized in the start, then being trained by training-use data set using gradient descent back propagation algorithm, every kernel would apply to input images with a sliding strategy. At the bottom of net structure, there will be at least one fully-connected layer which serves as the out-put classifier. With a depth neural network structure, nowadays Deep Convolutional Neural Networks(DCNN) outperforms all existed classic models with better accuracy and robustness.

Including convolution layers to the neural network brought out a few advantages. Firstly, it works the way closer to the human eye does, which provide some biomimetic clues in computer vision task then those existed machine learning methods. Secondly it brings the feature of a sparsely connected structure, which will cut down a considerably amount of computation complexity compare to traditional matrix multiplication. Thirdly, convolution kernels inside the structure will apply through the whole image being inputted both training phase and testing phase, which means that the weights in the network will be shared the whole time which helps to achieve a better robustness of feature extraction with lower performance demand.

In recent years, CNN has become a trend to improve the accuracy of tasks like classification, segmentation and recognition. But most of those works are targeting close range images which provides a large number of redundant features, applying and modify those CNN method to fit the characteristics of remote sensing image, and base on that, to brought up a land-cover classification quality evaluation approach, is the idea of this paper.

1.2 Comparison experiment of model and optimal function choices

In order to achieve automatic accuracy evaluation of land cover classification vector, an important part is to build a DCNN that can identify various kinds of single-class images with efficiency. In this experiment, a series of comparative experiments will be carried out to verify whether the trained network has the ability to identify the land cover classification images of remote sensing.

[1] Corresponding author: Zhang Jixian (zhangjx@sbsm.gov.cn)

https://doi.org/10.5194/isprs-archives-XLII-3-257-2018

The experiment is using a data-set that contains 10150 training images, plus about 1500 verification images and about 1100 test images. The size of all those images is 256 * 256 pixels, and the scale is 1:10000 in resolution of 1M. After training AlexNet and GoogLeNet, a weight system after training was obtained. Finally, 70 images which not used by neither training, verifying or testing data-set are using for error evaluation, including 7 classes of 10 images per class. The definition of all single-class images is that of all 256 * 256 pixels, there are more than 85% pixels belongs to the major class.

Forecast / Fact	Meadow	Building	Farmland	Forest	Water	Excavated ground	Garden plot
Meadow	9	0	0	1	0	0	0
Building	0	10	0	0	0	0	0
Farmland	0	0	10	0	0	0	0
Forest	0	0	0	10	0	0	0
Water	0	0	0	0	10	0	0
Excavated ground	1	0	1	0	0	8	0
Garden plot	1	0	0	2	0	0	7

Table 1. The confusion matrix of AlexNet

Forecast / Fact	Meadow	Building	Farmland	Forest	Water	Excavated ground	Garden plot
Meadow	10	0	0	0	0	0	0
Building	0	10	0	0	0	0	0
Farmland	0	0	10	0	0	0	0
Forest	0	0	0	10	0	0	0
Water	0	0	0	0	10	0	0
Excavated ground	1	0	1	0	0	8	0
Garden plot	1	0	0	2	0	0	7

Table 2. The confusion matrix of GoogLeNet

Obviously, for AlexNet :

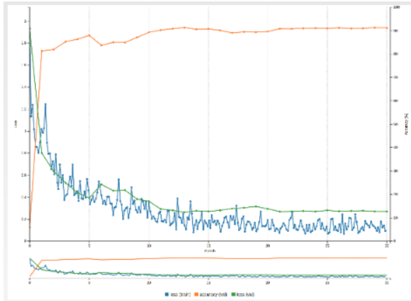
$$p_0 = \frac{9+10+10+10+10+8+7}{70} = 0.9143$$

$$p_e = \frac{11\times 10 + 10\times 10 + 10\times 11 + 10\times 13 + 10\times 10 + 10\times 8 + 10\times 7}{70\times 70} = 0.14$$

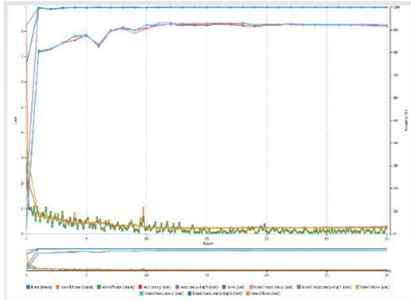
$$k = \frac{p_0 - p_e}{1-p_e} = \frac{0.9143-0.14}{1-0.14} = 0.9$$

The kappa coefficient of confusion matrix of AlexNet is 0.9. For the GoogLeNet, p_0 is 0.9286， p_e is 0.16， the kappa coefficient of confusion matrix of GoogLeNet is 0.915.

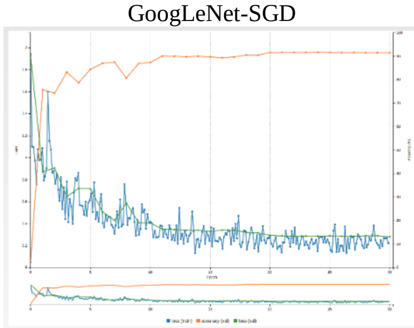
After that, in purpose of finding the most fitted optimization algorithm for the remote sensing images, a series of experiments has been carried out. The performance of AlexNet and GoogLeNet using SGD, NAG, AdaGrad, Adam and RMSProp algorithm is compared, and the training efficiency has been analysed with training curves.



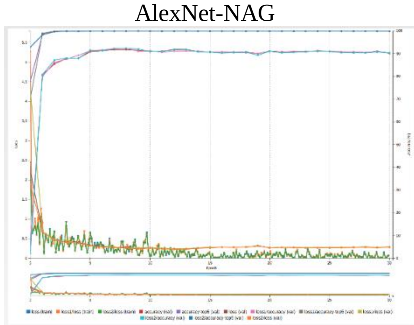
AlexNet-SGD



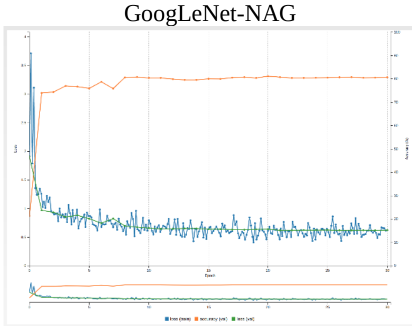
GoogLeNet-SGD



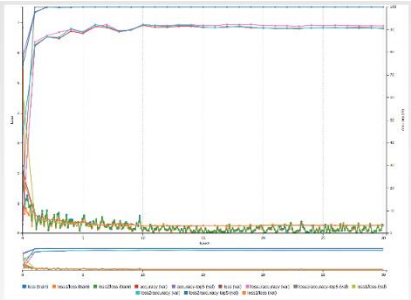
AlexNet-NAG



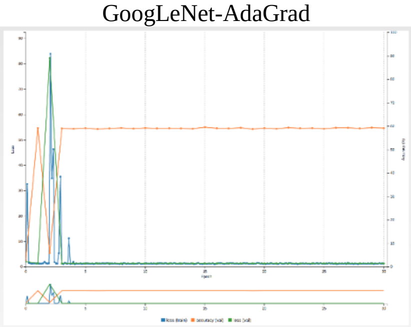
GoogLeNet-NAG



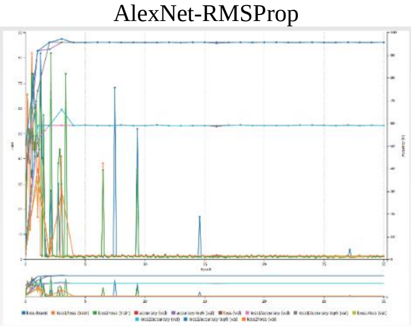
AlexNet-AdaGrad



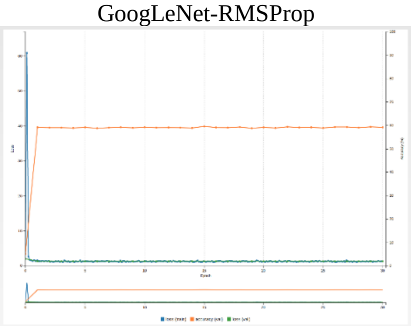
GoogLeNet-AdaGrad



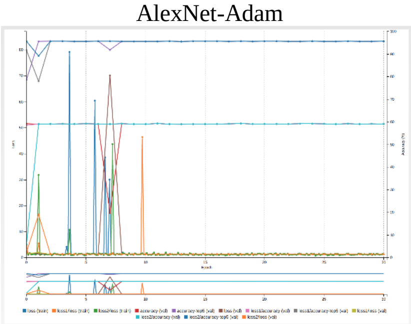
AlexNet-RMSProp



GoogLeNet-RMSProp



AlexNet-Adam



GoogLeNet-Adam

Figure 1. Comparison of the optimization algorithms

It can be found that, on the basis of the SGD, the NAG's addition improves the network performance better. But through AlexNet observation, after AdaGrad and RMSProp application, due to the gradient accumulation problem, the learning rate has been greatly restricted. Although the generalization performance (training loss and verification loss overlapping) has been improved, the overall accuracy has decreased. The results of RMSProp and Adam algorithm in the training is a failure.

1.3 The model using for quality evaluation

After the comparison experiment, we concluded that the 22 layers GoogLeNet with NAG algorithm is the one that outperforming other choices.

Forecast / Fact	Meadow	Building	Farmland	Forest	Water	Excavated ground	Garden plot
Meadow	10	0	0	0	0	0	10
Building	0	10	0	0	0	0	0
Farmland	0	0	10	0	0	0	0
Forest	0	0	0	10	0	0	0
Water	0	0	0	0	10	0	0
Excavated ground	1	0	1	0	0	8	1
Garden plot	1	0	0	2	0	0	1

Table 3. The confusion matrix of GoogLeNet-NAG

The kappa coefficient of this trained model is 0.915.

2. QUALITY EVALUATION METHOD OF REMOTE SENSING LAND-COVER CLASSIFICATION

2.1 Pixel-based automatic quality evaluation versus CNN method

The existing accuracy evaluation usually stared from two images comparing, one is the remote sensing classification image being evaluated, the other is the more accurate reference map. There is no doubt that the accuracy of reference map is the basis of all accuracy evaluation. If the accuracy of reference map or reference information as a basis for evaluation is not guaranteed, a series of accuracy evaluation based on this is also of no significance.

The surface information is constantly changing with time and place, so as the "true value" reference map / reference information to ensure its accuracy also need to meet as check information updating. However, in the reality accuracy evaluation task, the demand often focuses on the evaluation of absolute accuracy, that is, the reference information as the evaluation standard needs to be higher than the accuracy of the checked class diagram, which is also a issue difficult to solve in many projects. The general solutions are obtaining the same phase with the inspected products of high precision reference / reference map, or make a field inspection to the questioned area, these two methods are time-consuming, and is difficult to put into practice sometimes.

At the evaluation phase, the neural network are able to output the classification results directly which in order to achieve a series of parameters of image evaluation results, and the evaluation does not need any reference map to make a comparison, but expressing by distributed internal information of the network which trained by reference information through data-set.

2.2 Single class and mixed classes image evaluation strategy

In the single-class of scenario, the strategy of evaluation is relatively straightforward, these images can be evaluated directly and output the results of the evaluation.

For mixed classes of images, the result of land classification is often broken. Therefore, the single class image in size of 256 * 256 pixels is difficult to achieve in practical applications. At the same time, an important feature of land-cover classification is there are no overlap in it. Therefore, it provides a realistic basis for classifying and extracting single class of land cover classification results based on single class. By using the vector as mask extraction of the image, a bunch of single class images can be extracted from the mixed classes image. After that, the same strategy for single class image can be apply to the mixed classes images as well.

Figure 2. Original image and its classification results

Figure 3. A single map layer of grass(up) and house(down)

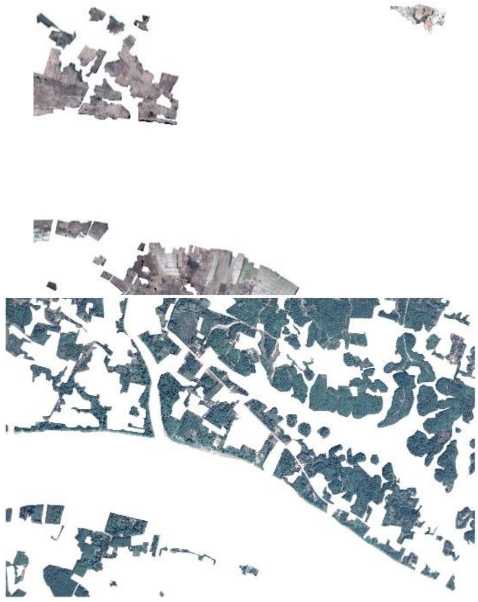


Figure 4. A single map layer of Farmland (up) and woodland (down)

Figure 5. A single map layer of waters (up) and Unearthed Grounds (down)

Figure 6. A single map layer of field

2.3 Method of accuracy evaluation based on CNN model

The evaluation method based on convolutional neural network is closer to the current quality evaluation approach, with the neural network model as a visual interpretation expert. Although the existing hardware and software cannot offer a neural network that has equivalent classification accuracy the same level of an interpretation expert, but in the fast evaluation of huge volume data situation, the automatic interpretation can deliver a rapid and stable evaluation. At the same time, under the conservative parameter setting, the network can automatically give a score of the full image. Combined with the future algorithm and program encapsulation in this paper, there will be an end to end function of the initial quality evaluation from DOM-DLG data, with the overall scoring and the possible errors.

Based on the overall evaluation, the whole confusion matrix can be formed by a single score after a balance. On the basis of the assumption that the network recognition is correct, the consistency between the classification vector and the neural network recognition is expressed by the Kappa coefficients of the confusion matrix. The range of the value of the Kappa coefficients is generally in the range of 0 to 1, when the value is in the range of 0-0.2 they are considered to have a slightly consistency, in 0.21-0.4 is fair consistent ,with 0.41-0.6 has a moderate degree of consistency, with 0.61-0.8 it's substantial consistency, and the consistency of 0.81-1 is called almost perfectly identical.

Since the Kappa coefficient itself has the same attribute of evaluation, the final evaluation of the whole map can be scored by the scoring system consisting of all the images, and the total score of the percentile system consisting of Kappa coefficients. The analogy Kappa coefficient, 0 - 20 points as rank E, 21 - 40 points as rank D, 41 - 60 points as rank C, 61 - 80 points as rank B, 81 - 100 points as rank A.

3. EXPERIMENT ON AUTOMATIC ACCURACY EVALUATION OF LAND COVER CLASSIFICATION

The experimental part is combined with the previous remote sensing image. First, the full-scale pre-processing is used to get a single class map layer, then all the spots are input, and the final output confusion matrix is used to screen and mark the possible errors. The results of the experimental results include 1604 images of the segmentation map.

Finally, convolution neural network is applied to automatic classification accuracy evaluation of images with 9360 pixels long and 5970 pixels wide, which totally takes 16 seconds. Combined with conservative single estimation, the confusion matrix is given.

Fact \ Forecast	Meadow	Building	Farmland	Forest	Water	Excavated ground	Garden plot
Meadow	148	14	3	3	3	9	3
Building	2	175	0	1	0	1	0
Farmland	1	0	140	0	1	4	0
Forest	17	1	0	539	3	4	5
Water	18	3	1	17	413	3	11
Excavated ground	0	1	2	1	0	31	0
Garden plot	3	0	0	1	0	2	18

Table 4. The confusion matrix of the whole image

There are:

$$p_0 = \frac{148 + 175 + 140 + 539 + 413 + 31 + 18}{1604} = 0.9127$$

$$p_e = \frac{183 \times 189 + 179 \times 194 + 146 \times 146 + 569 \times 562 + 468 \times 420 + 35 \times 54 + 24 \times 37}{1604 \times 1604} = 0.237$$

So:

$$k = \frac{p_0 - p_e}{1 - p_e} = \frac{0.9127 - 0.237}{1 - 0.237} = 0.89$$

The Kappa coefficient of the confusion matrix of the image is 0.89.

4. CONCLUSION

In the quality evaluation of land-cover classification, classification accuracy evaluation is an important content. Under the background of major national projects such as National Geographic Census, it demands a rapid, accurate and effective quality evaluation for large amounts of data. By having novel approach based on convolutional neural network technology with deep structure, constructing the land cover classification accuracy evaluation method, is of great theoretical and practical significance on the construction stage of quality evaluation of related products of Surveying and mapping engineering and new automation and intelligent quality inspection system.

ACKNOWLEDGEMENTS

Dr.Zhang Jixian, Dr.Ma Wei and Mr.Lu Fujun are greatly appreciated for their constructive suggestions and discussions. We thank the anonymous reviewers for their constructive comments and helpful suggestions. This work is partly supported by National Science Foundation of China (grant No. 41671440).

REFERENCES

L. Zhang, L. Zhang, and V. Kumar, Deep learning for Remote Sensing Data, *IEEE Geoscience and Remote Sensing Magazine*, 4(June), 22–40 (2016).

J. Wan, D. Wang, S. C. H. Hoi, et al., Deep learning for content-based image retrieval: A comprehensive study, *The 22nd ACM international conference on Multimedia*, 157–166 (2014).

I. Arel, D. C. Rose, and T. P. Karnowski, Deep Machine Learning - A New Frontier in Artificial Intelligence Research, *IEEE Computational Intelligence Magazine*, 5(4), 13–18 (2010).

H. Deborah, N. Richard, and J. Y. Hardeberg, A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6), 3224–3234 (2015).

P. Dollar, C. Wojek, B. Schiele, et al., Pedestrian detection: An evaluation of the state of the art, *IEEE transactions on pattern analysis and machine intelligence*, 34(4), 743–761 (2012).

M. Fauvel, Y. Tarabalka, J. A. Benediktsson, et al., Advances in spectral-spatial classification of hyperspectral images, *Proceedings of the IEEE*, 101(3), 652–675 (2013).

https://doi.org/10.5194/isprs-archives-XLII-3-257-2018

M. Hussain, D. Chen, A. Cheng, et al., Change detection from remotely sensed images: From pixel-based to object-based approaches, *ISPRS Journal of Photogrammetry and Remote Sensing*, 80, 91–106 (2013).

D. Lunga, S. Prasad, M. M. Crawford, et al., Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning, *IEEE Signal Processing Magazine*, 31(1), 55–66 (2014).

Rwanga S S, Ndambuki J M. Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS. *International Journal of Geosciences*, 2017, 8(4):611-622.

Kramm T, Hoffmeister D, Curdt C, et al. Accuracy Assessment of Landform Classification Approaches on Different Spatial Scales for the Iranian Loess Plateau. *International Journal of Geo-Information*, 2017, 6(11).

Song H, Zhang X. Accuracy assessment of existing global land cover products by using authority data. *Boletin Tecnico/technical Bulletin*, 2017, 55(6):6-12.