Comparison of Several Features of Building Detection in Remote Sensing Image

Juncheng Wu[1, a], Jingwen Xu[1,b*], Junfang Zhao[2], Ning Li[1] and Surong Xiang[1]

[1] College of Resources, Sichuan Agricultural University, Chengdu 611130, China

[2] Chinese Academy of Meteorological Sciences, CMA, Beijing 100081, P.R. China

[a] email: wu7575@126.com, *[b] Corresponding author: x.j.w@163.com

Keywords: remote sensing; building detection; Haar; LBP; machine learning; deep learning; AdaBoost; evaluation

Abstract. AdaBoost algorithm based on Haar features and Local Binary Patterns (LBP) features are widely used in machine learning method, and deep learning algorithm is a hot research field in recent years. Remote sensing image recognition meets the needs of human beings in the social development of Information Technology Intelligent. Combined machine learning method with the three features, a certain number of buildings in the remote sensing image were trained and performance test, and then the test results were evaluated, respectively. The results show that the features performance of the deep learning algorithm is better than the other two kinds of features, and the accuracy can meet the requirements of the application of remote sensing image interpretation.

Introduction

As a new technology developing in the early 1960s, the automatic interpretation of remote sensing images has been one of the goals of the development of 3S technology. Automatic detection and recognition by computer can save a lot of labor force, material resources and time, so as to accelerate the speed of human access and utilization of geographic information. The building is an important place for human life and production, automatic access to building regional location can be quickly used for government decision-making and construction planning guidance. But building detection has been a difficult problem, because the shape of the building is not the same and the outline of the terrain easily by the terrain, the shadow, vegetation cover, and the features of the building is not easy to find. In the last few decades, with the rapid development of computer technology, it is possible to detect and identify buildings accurately.

Based on object-oriented building detection, remote sensing image segmentation is divided into many objects, and a variety of features of the object are detected [1]. Ensemble learning and deep learning are the most widely used machine learning algorithms, and have achieved some results in various fields [2]. For machine learning algorithm, AdaBoost is the most typical algorithm of ensemble learning, and the Haar and LBP features are widely used in AdaBoost [3-4]. PCANet make the simplified training process simpler, and can adapt to different tasks, different data types of deep learning network model [5]. In this study, three kinds of building image detection algorithms are applied to the building detection of remote sensing images, and the performance of each learning method is tested and evaluated using a building image samples library. We can provide a reference for building image detection, which can be used to achieve precision classification as soon as possible.

Dataset

At present, there is no standard database for remote sensing image detection. So we establish an image database to perform research. Cropped images from UAV remote sensing image, Danling County, Sichuan Province, China, by the 1000 building images and 2000 non building images. The positive sample is a 42×48 pixel RGB image, including various types of independent buildings. The negative sample is a RGB image with a resolution of 42×48 pixels, which includes all types of building elements. We randomly selected 100, 200, 500 images of building from the 1000 buildings as the training group,

respectively, the remaining 500 images as the test group. The same from 2000 non building images were randomly selected from 200, 400, 1000 as the training group and the positive sample to participate in training, and the remaining 1000 as a test group. Then we exchange the training group and the test group to perform another experiment, from which we could get two sets of data. In order to compare the detection performance of two kinds of machine learning methods under the same conditions, the same training sample and the same training method and test sample are provided.

Various Features and Models

AdaBoost Algorithm

AdaBoost is the abbreviation for "Boosting Adaptive", which was proposed by Yoav Freund and Robert Schapire in 1995[6]. AdaBoost is an adaptive iterative algorithm, which is composed of several weak classifiers. Setting up a training sample set $T=\{(x_1,y_1), (x_2,y_2)\ldots (x_n,y_n)\}$, the same weight is given to each of the same weight of 1/N, and then use the algorithm on the training set L round. When the error rate is small to a predetermined value, it will enter the next round. Each training sample is given a certain value, if the sample is accurately classified the weight is reduced, the specific provisions of the wrong will be improved. The prediction function is used to determine the classification problem by using the weighted average method.

Haar Features

Haar feature is a series of rectangular features, which is composed of edge features, linear features, central features and diagonal features. The feature template has a white and black two rectangles, and defines the features values of the template for a white rectangle and minus the black rectangle pixels. In the determination of the featuress of the featuress of the number of Haar feature is determined by the size of the training sample image matrix, feature templates in the sub window arbitrary placement, a form known as a feature to identify all the sub window is the basis for the weak classification training [7].

LBP Features

Local Binary Patterns are in a region, with the window center pixel as the threshold, the gray value of the adjacent pixels and the comparison, the value of the center pixel value is 1 and the other is 0. In the field, the value of the points in the field is combined into binary number and converted to decimal, and the final value is the LBP value of the center point, which is used to reflect the image texture information [8].

PCANet Models

The effect of image classification is largely determined by the shape, texture, color and other featuress of the classified objects. The featuress of Haar and LBP mentioned above are all the featuress of a large number of manual extraction. Deep learning is a feature that automatically learns from the region of interest. It uses a more abstract expression of things. PCANet model is a simplified deep learning network, which contains a cascade of principal component analysis (PCA), binary hashing, and block-wise histograms of three parts [5]. First of all, we are given N input training images of size $m \times n$, around each pixel, we take a $k_1 \times k_2$ patch, then subtract patch mean from each patch and get the features vector XX^T, Take the maximum eigenvalues as the filter for next step:

$$W_l^1 = q_1(XX^T) \in R^{k_1 k_2}, l = 1,2 \ldots, L_1 \tag{1}$$

The second stage is also the PCA, the need to be the first step in the input of the image and filter to do the second phase of the training image. The picture is divided into I_i ,then put the $L_1$of the image after the convolution of the sub block results are obtained:

$$Y = [Y_1^l, Y_2^l, \ldots, Y^{L_1}] \in R^{k_1 k_2 \times L_1 N_a} \tag{2}$$

The feature vector of the YY^T is the largest eigenvalue of the L_2, which is used as a filter:

$$W_\ell^2 - q_\ell(YY^T) \in R^{k_1 k_2}, \ell - 1,2 \ldots, L_2 \tag{3}$$

Output stage: hashing and histogram:

$$f_i = [\mathrm{Bhist}(T_i^1), \ldots, \mathrm{Bhist}(T_i^{L_1})]^T \in R^{(2^{L_2})L_1 B} \tag{4}$$

Results and Comparative Studies

Performance of Haar Features

Haar features and LBP features are the most common features, which are directly integrated in the OpenCV library. The Haar feature selection is the right upper feature and 45 degree rotation feature, Table 1 shows that the AdaBoost algorithm is based on the Haar feature detection results in different samples. The detection rate reached the best when the number of positive samples was 500, but still could not reach more than half of the detection rates.

Sample numbers	100	200	500
1st hit rates	0.33	0.38	0.40
2nd hit rates	0.30	0.40	0.41
1st error rates	0.27	0.23	0.19
2nd error rates	0.25	0.23	0.22
Average hit rates	0.315	0.39	0.405
Average error rates	0.26	0.23	0.205

Table.1 Detection performance of Haar features in different samples

Performance of LBP Features

The LBP feature has its own rotation invariant features, and we set the default parameters for training by OpenCV. Table 2 shows that the performance of LBP is still in the best performance of the 500 samples after different numbers of samples. Compared with the feature of LBP, the LBP features are more fast and accurate than the Haar.

Sample numbers	100	200	500
1st hit rates	0.54	0.64	0.67
2nd hit rates	0.49	0.58	0.66
1st error rates	0.23	0.18	0.18
2nd error rates	0.23	0.16	0.14
Average hit rates	0.515	0.61	0.665
Average error rates	0.23	0.17	0.16

Table.2 Detection performance of LBP features in different samples

Performance of PCANet Model Extract Features

The feature of PCANet model is extracted from the sample, which is more abstract for image feature extraction. As can be seen from table 3, in the case of only 100 of the sample, the detection accuracy is still more than the best results of the Haar features and LBP features mentioned above. When the number of samples is 500, the accuracy rate can reach 91%, and the false positive rate is only 9.5%. Under the limitation of experimental conditions, feature extraction of the deep learning algorithm is satisfactory. The training speed is faster than the LBP feature classifier, and the detection speed is the same.

Sample numbers	100	200	500
1st hit rates	0.83	0.86	0.92
2nd hit rates	0.81	0.87	0.90
1st error rates	0.09	0.12	0.10
2nd error rates	0.13	0.14	0.09
Average hit rates	0.82	0.865	0.91
Average error rates	0.11	0.13	0.095

Table.3 Detection performance of deep learning algorithm in different samples

Conclusion and Discussion

In our research, we have finished the testing and comparisons about three features, including Haar features, LBP features and the features extracted by PCANet, and during the test we chose the same samples. The difference of performance which resulted from the problem of the sample could be

reduced by the training, testing and exchanging the different numbers of samples. It suggested that the features extracted by PCANet have achieved a relatively good detection performance, and it is the most suitable for building image detection in remote sensing images. Although Haar features and LBP features, which once have an advantage in the field of face recognition, have no advantage in the field of the detecting of buildings in remote sensing images, owing to failing to reflect the features of building images. At present, deep learning has become a hot spot in the field of machine learning due to its excellent performance and generalization ability. How to strengthen the deep learning structure, improve the algorithm and make the remote sensing automatic interpretation more suitable are the key focus of this study.

Acknowledgements

This study is financially supported by the National Non-profit Research Foundation for Meteorology (GYHY201506016), Open Research Fund Program of the Meteorological Center for Huaihe watershed (HRM200905).

References

[1] Wei Y, Zhao Z, Song J. Urban building extraction from high-resolution satellite panchromatic image using clustering and edge detection[C]//Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International. IEEE, 2004, 3: 2008-2010.

[2] Valentini G, Masulli F. Ensembles of learning machines [M]//Neural Nets. Springer Berlin Heidelberg, 2002: 3-20.

[3] Wilson P I, Fernandez J. Facial feature detection using Haar classifiers [J]. Journal of Computing Sciences in Colleges, 2006, 21(4): 127-133.

[4] Guo Z, Zhang L, Zhang D. Rotation invariant texture classification using LBP variance (LBPV) with global matching [J]. Pattern recognition, 2010, 43(3): 706-719.

[5] Chan T H, Jia K, Gao S, et al. PCANet: A simple deep learning baseline for image classification? [J]. arXiv preprint arXiv: 1404.3606, 2014.

[6] Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting [J]. Journal of computer and system sciences, 1997, 55(1): 119-139.

[7] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]//Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2001, 1: I-511-I-518 vol. 1.

[8] Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002, 24(7): 971-987.