

Image Classification Algorithm Based on Depth Neural Network and Multi-Level Feature Learning

Qian Sun¹, Li Xin¹, Faliang Chang², Zengshun Zhao^{1,2,3}

¹College of Electronics Communications and Physics, Shandong University of Science and Technology, Qingdao, 266590, China

²School of Control Science and Engineering, Shandong University, Jinan, 250061, China

³Department of Electrical & Computer Engineering, University of Florida, Gainesville, FL 32611, USA

Keywords: Deep Neural Network; Multi-Layer Feature Learning; Minimum Two-Order Norm; Image Classification; Image Features

Abstract: Image classification being widely applied to computer vision is an image processing method to distinguish the different category targets according to the different features reflected by the image information. BOW-SVM is a relatively typical image classification method with higher precision, however, it's unsatisfactory in operation performance. To improve the performance and precision more efficiently, a high-efficiency image classification method based on HOG-PCA is proposed. First of all, it is to make the feature whitening by extracting the Histogram of Oriented Gradients (HOG) features, secondly, make the random down-sampling for the scale unification, afterwards, adopt the principal component analysis (PCA) for feature mapping and finally make the nearest neighbor classification through the minimum two-order norm determination. In the experiment, the proposed method is realized and tested on the PASCAL 2012 data set through C++ on the basis of OPENCV and Darwin to compare the precision and operation performance of this method and BOW-SVM method; according to the experiment, the proposed has higher precision and better operation performance.

1. Introduction

Image classification is to make quantitative analysis on images with computer, classify the image or image area even each pixel point in image into one of several categories according to the different features reflected from the image information to achieve the classification goal. Image classification is widely applied to: target recognition of aeroplane and guided missile in military affairs; product quality inspection in industrial production; analysis on satellite cloud photograph and infrared picture in science; leukocyte recognition and ultrasonic liver image recognition in medical science and scene recognition in life. However, influenced by human and natural factors, there remain some questions in image classification. By now, image classification is still the research difficulty and hotspot in the image processing field [1]. An image classification method based on HOG-PCA is proposed. Firstly, HOG is adopted to optimize adaptability of light, scale and orientation to extract the image features; secondly, to whiten the features and reduce the abnormal points; finally, to make use of the computing performance superiorities of feature face. The biggest advantage of our method lies in the simplicity and favorable computing performance; the image categories are of large sufficiency and especially suitable to classification of natural images, its basic framework is as shown in Fig.1:

We test the proposed method in data set PASCAL 2012[9] to compare this method and the BOW-SVM method; proved by this experiment, the proposed method has obvious advantages in both precision and operation time.

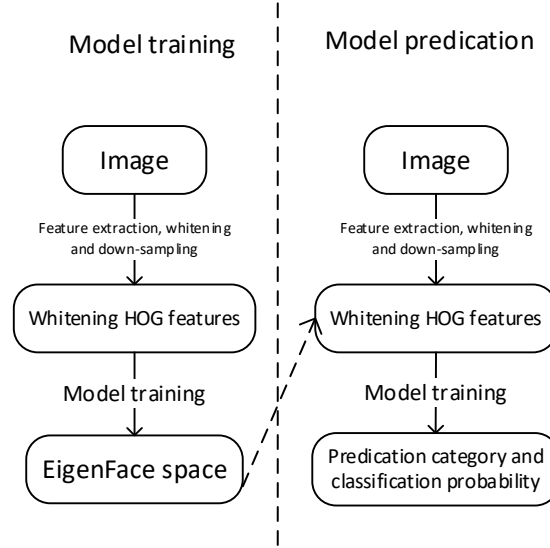


Fig.1. Image classification method based on HOG-PCA

2. Image Classification Algorithm Based on HOG-PCA

Image classification is generally composed of three steps: extraction and processing of image features, establishment of classification mode and classification evaluation[10]. First of all, the HOG is adopted to extract the image features for the whitening processing and down-sampling; secondly, to build the PCA mapping model for the HOG features of the training sample; finally, adopt the least square procedure of model for classification.

2.1 HOG features extraction

The core concept of Histogram of Oriented Gradients (HOG) is that the local external form of the inspected object can be described by the light intensity gradient or the distribution of edge direction. The HOG is calculated in five procedures:

(1) Image gamma and color standardization

In evaluating different color spaces such as gray scale, RGB and LAB, to unify the evaluation criterion, they are generally transformed into the gray level image. What's more the gray level value is transformed from 0 to 1.

(2) Computing direction of one-dimensional gradient filtering

Gradient computing method is crucial to the correct description of image features and the simple method is always most effective. During the HOG calculation, the Sobel operator is adopted to calculate the image difference of the one-order x or y axis to obtain the favorable performance. The Sobel operator with Gaussian Blur and difference will make better anti-noise performance of the result.

(3) Weighted voting of space and direction cells

This is the most important step, the cumulative weight on direction of each cell shall be set; the weight algorithm on the calculation direction shall be:

$$x(i) = \cos(\theta) \quad (1)$$

$$y(i) = \sin(\theta) \quad (2)$$

$$\theta = \theta + \pi / (N_{direction} + 1) \quad (3)$$

In which: i is the direction number, θ is the angle, the initial value is 0, $x(i)$ is the weight of x axis difference on i direction, $y(i)$ is the weight of y axis difference on i direction.

Followed by the calculation amplitude and direction, amplitude is mean square value of the image differences on x axis and y axis. $N_{direction}$ is the total direction number, generally set as 9. The maximum weighted value on each direction is taken as the direction value.

(4) Building block and standardization

Features in cells are aggregated into blocks according to the following calculation method:

$$B(x) = (C(x) - B(\text{size}) + 1) / B(\text{step}) \quad (4)$$

In which: $B(x)$ is the total value of block x axis, $C(x)$ is the total value of cell x axis, $B(\text{size})$ is the block size, $B(\text{step})$ is the step size of block variation. Calculation of total block number on y axis is similar.

Standardization consists of two steps: (1) feature shearing, namely shear to propose values that are too small and too big according to the range of minimum feature value and maximum feature value, unify the feature values of the residue features into the range of 0~1; (2) feature standardization.

(5) Building the HOG of the image

Dimensionality reduction can reduce the feature number and lower the computation complexity, set the output dimensionality as required. Aggregation of energy and direction: Aggregate the feature values on different directions and blocks to form the HOG of images. Compared with other descriptors, the descriptors obtained by HOG have maintained the invariance of geometric or optical transformation. Here is the example for the HOG features.



Fig.2. Image and its HOG features

In Fig.2, the left is the image and the right is the colorful HOG features of the image.

2.2 HOG feature whitening and down-sampling

As there remain the height space coupling in HOG features, the HOG features shall be whitened for two purposes: (1) to reduce the height correlation between features, (2) to make the variances of training features the same. Feature whitening can be sketched as below[11]:

Assume a group of feature X in dimension d and line n ($d \times n$) will be changed like below to reduce the correlations between features

$$Y = WX \quad (5)$$

In which, Y is the feature after transformation, W is the whitening matrix; to make W eliminate the feature correlation, YY^T must be angle corner matrix, which means:

$$YY^T = (n-1)I \quad (6)$$

There are many W s satisfying the requirements of formula 6, so that it is limited as $W = W^T$, we can obtain W through the formula 7.

$$W = \sqrt{n-1}(XX^T)^{-\frac{1}{2}} \quad (7)$$

As the images to be classified are of inconsistent scales, random down-sampling or up-sampling is necessary. The down-sampling is necessary if there are more whitening features than the required features.

$$Z = \text{subsample}(Y, m) \quad (8)$$

In which, Y is the whitened feature, there are m features after Z sampling, function *subsample* random sampling method is adopted to extract features.

2.3 PCA mapping

The principal component analysis [12] reduces the data dimension by restraining the coefficients with fewer information quantity through the K-L transformation; however, when calculating the consumption of covariance matrix, PCA will transfer the multi-dimensional image information into several principal components that contain most of the image information to improve the classification efficiency [13].

Assume there are N sample features $\{x_1, x_2, \dots, x_N\}$, each feature is an n -dimensional space and part of a category of $\{X_1, X_2, \dots, X_c\}$. Consider a linear transformation, being mapped from n -dimensional space to m -dimensional space and satisfying $m < n$; the new feature vector $y_k \in R^m$ defines the linear transformation as below:

$$y_n = W^T x_k \quad k = 1, 2, \dots, N \quad (9)$$

Here, W^T is a matrix with orthogonal array. Assume the total sample scatter matrix is defined as below:

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (10)$$

In which, N is the sample size, $\mu \in R^n$ is the mean value of all samples. The linear transformation can be adopted then.

$$\begin{aligned} W_{opt} &= \arg \max_W |W^T S_T W| \\ &= [w_1, w_2, \dots, w_m] \end{aligned} \quad (11)$$

Here, $\{w_i | i = 1, 2, \dots, m\}$ is the m maximum feature vectors corresponding to the n -dimensional feature space scatter set S^T and the mapping feature of the model[14].

2.4 Classification discrimination

Calculate the two-order norm distance between each sample and the training sample for the category of the minimum distance and conclude it into the category of this training sample.

(1) Project the image features to be recognized into the PCA space.

(2) Calculate the distance between the projected features and the training samples in PCA space.

$$d(i) = \sqrt{(t_1(i) - p_1)^2 + (t_2(i) - p_2)^2 + \dots + (t_n(i) - p_n)^2} \quad (12)$$

In which, $d(i)$ is the distance from the No. i training sample, $t_1(i)$ is the first projection feature value of the No. i training sample, p_1 is the first projection feature value of a test sample. $t_2(i)$, p_2 , $t_n(i)$ and p_n have the same meaning.

(3) Classification with the minimum distance is taken as its classification in the end.

(4) Reliability calculation method of classification. The distance between the two of them is the reliability.

3. Conclusion

In this paper, the general principles, procedures and methods of image classification are analyzed; also, a simple and high-efficiency image classification method is proposed. Due to the inherent features of HOG, this method has taken factors such as image rotation, light and scale into consideration at the same time of classifying the images rapidly. In this experiment, we adopt the C++ to implement the proposed method, the efficiency of which is higher than the common Matlabsimulation method; compared with the classical BOW-SVM method, the proposed method has obvious advantages in operation performance, total accuracy, average precision ratio and average recall ratio. Image classification is widely applied; its classification precision relies on the further analysis and research of other more complicated influencing factor such as scene, color and shelter on images. It has always been the focus and difficulty in image classification research as

well as one of the important development fields for image classification.

Acknowledgement

National Natural Science Foundation of China (Grant No. 61403281);
Natural Science Foundation of Shandong Province, China (ZR2014FM002);
China Postdoctoral Science Foundation (2015T80717);

References

- [1] Fernandes, S.L., Gurupur, V.P., Sunder, N.R., Arunkumar, N., Kadry, S.A novel nonintrusive decision support approach for heart rate measurement. Pattern Recognition Letters, 2017.
- [2] Arunkumar, N., Ramkumar, K., Venkatraman, V., Abdulhay, E., Fernandes, S. L., Kadry, S., & Segal, S. Classification of focal and non focal EEG using entropies. Pattern Recognition Letters, 2017.
- [3] Arunkumar, N., Kumar, K.R., Venkataraman, V. Automatic detection of epileptic seizures using new entropy measures. Journal of Medical Imaging and Health Informatics, 2016.
- [4] Arunkumar, N., Ram Kumar, K., Venkataraman, V. Automatic detection of epileptic seizures using permutation entropy, Tsallis entropy and Kolmogorov complexity. Journal of Medical Imaging and Health Informatics, 2016.