International Journal of Computer Science and Information Technology
ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 4, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v4n1.18
Journal homepage: https://wepub.org/index.php/IJCSIT/index

A Tobacco Brand Recognition Method Based on HOG and SVM

Shujie Liu *

Jinan Cigarette Factory, China Tobacco Shandong Industrial Co., Ltd., Jinan, 250101, China
*Corresponding Author: alotove.liu@hotmail.com

ABSTRACT

The appearance inspection system cannot automatically switch the brand, and a person needs to log into the system and manually switch the inspection brand in the process of production and brand change. To address this problem, a deep learning-based automatic cigarette brand recognition method is developed and validated. Experiments show that: the system can automatically detect the production brand, effectively save the time to change the brand, and has important reference significance for the application in the field of brand identification and automatic switching detection of the brand.

KEYWORDS

Deep learning; Hog; Svm

1. INTRODUCTION

With the development of information technology detection technology, the intelligence level of tobacco generation is getting higher and higher. However, cigarette production work is highly planned in China. There is a possibility of dynamic adjustment of the production grades of a certain machine, which brings challenges to the appearance inspection activities. Due to the differences in the appearance of the small boxes of each brand, when changing the production brand, the metrologist is required to manually change the type of inspection, which slows down the speed of brand change. This paper focuses on a deep learning based brand recognition method to achieve automatic switching of production brand numbers and improve the speed of brand change.

2. SVM CLASSIFICATION BASED ON HOG FEATURE EXTRACTION

HOG feature descriptor in an image, the appearance and shape of the local target (appearance and shape) local gradient distribution is well described, even if we do not know the corresponding gradient and the location of the edges (essence: the statistical information of the gradient, the gradient is mainly present in the edge edge or corner corner of the place). A feature descriptor is a way to simplify the representation of an image by extracting useful information from the image and discarding irrelevant information.The HOG feature descriptor converts a 3-channel color image into a feature vector of a certain length. In HOG feature descriptor, the distribution of gradient direction, that is, the histogram of gradient direction is considered as a feature. The gradient (x and y derivatives) of an image is very useful because the gradient magnitude around edges and corners (regions with sudden changes in intensity) is large and edges and corners contain more information about the shape of an object than flat regions.The main purpose of the HOG algorithm is to perform a gradient computation on an image, counting the gradient direction and gradient magnitude of the image. He extracts edge and gradient features that capture the local shape well, and because the image is Gamma-corrected and normalized using the cell approach, it is very invariant to geometric and optical changes, and transformations or rotations have very little effect on small enough regions.

The basic idea is to divide the image into many small connected regions, i.e., cell units Cell, and then vote statistics on the gradient magnitude and direction of Cell to form a histogram based on gradient properties. The histogram is normalized over a larger range of the image (aka interval or Block). The normalized block descriptor is called HOG descriptor (feature descriptor). The HOG descriptors of all the blocks in the detection window are combined to form the final feature vector.HOG+SVM is a frequently used classification and detection algorithm that possesses simple principles and low computational resource requirements. Firstly HOG is used to extract the image features and SVM classifier takes these features as input and performs category prediction based on the features.

2.1. HOG Features

HOG (Histogram of oriented gradient) [1] feature is an image description method proposed by French researcher Dalal at CVPR 2005 to enable human target detection, which constitutes a feature by calculating and counting the histogram of gradient orientations in local regions of an image. Its advantages are that it can maintain good invariance to geometric and optical deformations and is highly robust to changes in the environment.

The main idea of the feature is that the appearance and characteristics of local targets in an image can be well described by the gradient or the orientation density of the edges (essentially: statistical information about the gradient, which is mainly present at the edges). In practice, the image is divided into small cells and the histogram of the gradient direction (or edge direction) is computed for each cell. In order to have better invariance to light and shadows, the histogram needs to be contrast normalised, which can be achieved by forming cell cells into larger blocks and normalising all cell cells within the block. We refer to the normalised block descriptors as HOG descriptors. The HOG descriptors of all blocks in the detection window are combined to form the final feature vector. It is usually necessary to convert the image to grey scale during HOG feature extraction, and the result of feature extraction is shown in Fig. 1.

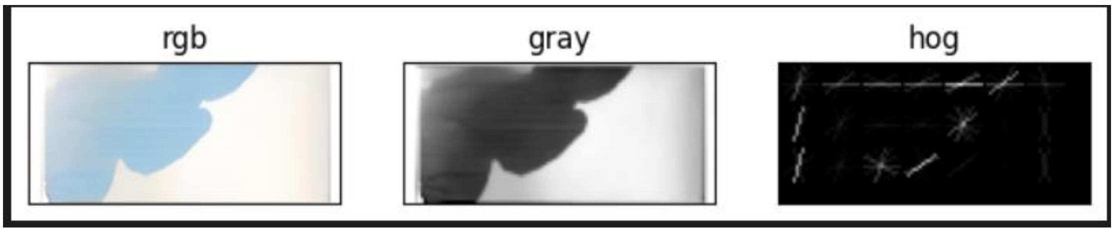

Figure 1. HOG feature extraction

2.2. SVM Classifier

Support Vector Machine [2] is a binary classification model, first proposed by Vladimir N. Vapnik and Alexey Ya. Chervonenkis in 1963. It maps the feature vectors of instances to some points in space, and the purpose of SVM is to draw a line that ‘best’ distinguishes between these two types of points, to the point that if new points are added later, this line will make a good classification. In order to quickly identify image features, PCA [3] dimensionality reduction is also required [4], the main principle of which is to map the initial n-dimensional data matrix from a high-dimensional space to a low-dimensional space using linear transformations [5]. In this paper, SVM algorithm in sklearn library is used, and the test results of various kernels of SVM algorithm are shown in Figure 2.

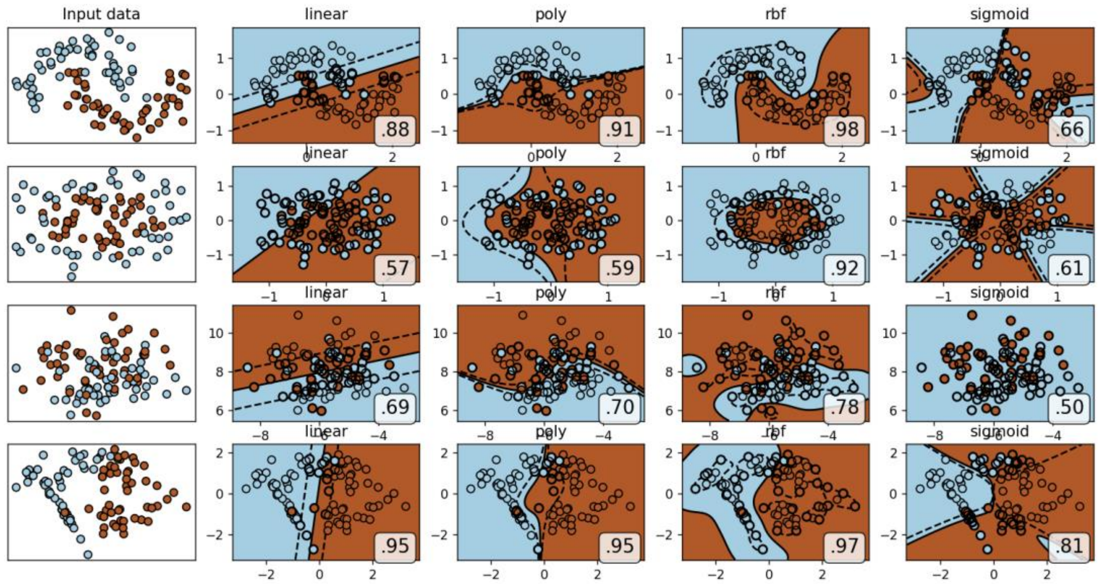

Figure 2. SVM test results

The main purpose of the Linear Kernel is to formally unify the "problem in mapped space" and the "problem in mapped space" (meaning: sometimes we can write code or formulas by simply writing a template or a generic expression and substituting it into a different kernel). When we write code or formulas, we just need to write a template or a generic expression, and then substitute different kernels, and that's it. Here, it is formally unified, no need to find a linear and a nonlinear) Linear kernel, mainly used in the case of linearly differentiable, we can see that the dimension of the feature space to the input space is the same. Finding the optimal linear classifier in the original space has the advantage of less parameters and faster speed. For linearly differentiable data, its classification is ideal, the disadvantage is that it can only solve linearly differentiable problems.

$$\kappa(x_1, x_2) = \langle x_1, x_2 \rangle \tag{1}$$

The Polynomial Kernel polynomial kernel function enables the mapping of a low-dimensional input space to a high-dimensional feature space. The Polynomial Kernel is suitable for orthogonally normalized (vectors are orthogonal and modulo 1) data, and is a global kernel function that allows data points that are far apart to have an effect on the value of the kernel function. The larger the parameter d, the higher the dimension of the mapping and the more computationally intensive it will be. The advantage is that the nonlinear problem can be solved and the summarized prediction can be achieved by setting Q subjectively. However, the polynomial kernel function has many parameters, when the order of the polynomial d is relatively high, due to the learning complexity will also be too high, easy to appear "overfitting" phenomenon, the kernel matrix will tend to infinity or infinitesimal element value, the computational complexity will be unable to calculate the Avenue.

$$k(x_i, x_j) = \left(x_i^t x_j\right)^d, d \geq 1 \tag{2}$$

$$K(x_i, x_j) = \left(x_i \cdot x_j + 1\right)^d \tag{3}$$

Radial Basis Function Radial basis kernel function is a commonly used kernel function in SVM. Radial Basis Function is a function that uses vectors as independent variables and can output a scalar based on vector distance operations. Gaussian radial basis function is a localized kernel function, which can map a sample to a higher dimensional space, the kernel function is one of the most widely used, no matter large samples or small samples have a better performance, and its relative to the polynomial kernel function with fewer parameters, so most of the cases in the case of do not know what kind of kernel function to use when the preference for the use of Gaussian kernel function. The advantages are that it can be mapped to the wireless dimension and has only one parameter, which makes it easy to choose compared to the polynomial kernel. The disadvantages are also the same, its poor interpretability, slower computation and easy overfitting.

$$k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \sigma > 0 \quad (4)$$

The Sigmoid kernel function is derived from neural networks and is widely used in deep learning and machine learning. When using the Sigmoid function as the kernel function, the Support Vector Machine implements a multi-layer perceptron neural network, applying the SVM method, the number of hidden layer nodes (which determines the structure of the neural network), and the weights of the hidden layer nodes on the input nodes are all determined automatically in the process of design (training). Moreover, the theoretical foundation of Support Vector Machines determines that it ultimately finds the global optimum rather than the local optimum, and also ensures its good generalization ability to unknown samples without over-learning linearity.

$$k(x_i, x_j) = \tanh(\beta x_i^t x_j + \theta), \beta > 0 \quad (5)$$

2.3. Training Results and Analysis

After extracting features by HOG the feature image is fed into SVM for fitting training, after many tests, when extracting features by HOG algorithm, set the number of pixels per cell pixels_per_cell=(18, 18), how many cells are in each block cell_per_block=(10, 10), use L2 norm for Normalisation block_norm=‘L2’, the best results were achieved with the number of orientation boxes orientations=9, the test result Accuracy was 75.0% and the confusion matrix is shown in Figure 3

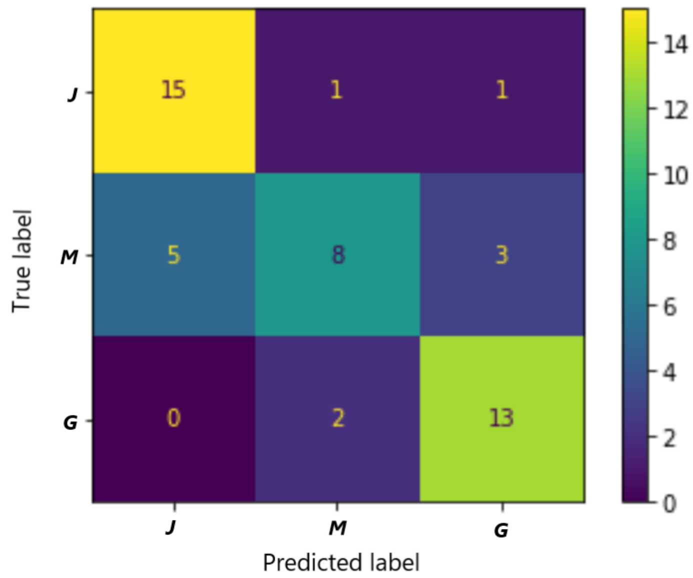

Figure 3. Confusion matrix of SVM classification results

From the figure, it can be seen that J category has the best effect, M category has the worst effect, and G category is in the middle position, and in the case of misidentification, M misidentifies J the most, and there is no G misidentifies J.

The confusion matrix is calculated to get the accuracy and recall of each category as shown in Table 1, and it can be seen that G is the highest and M is the lowest in the accuracy, while the recall is the highest and M is the lowest in J.

Table 1. Accuracy and recall

	Precision	Recall
J	0.75	0.88
M	0.73	0.50
G	0.76	0.87

3. SUMMARY

It can be learnt from the above experimental results that the accuracy rate can reach more than 70% in both target detection. The above method can quickly determine the tobacco brand and reduce manual intervention. More accurate algorithms will be studied in the follow-up work to improve the switching accuracy.

REFERENCES

[1] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection[C]//IEEE Computer Society Conference on Computer Vision & Pattern Recognition. IEEE, 2005. DOI:10.1109/CVPR.2005.177.

[2] Cortes C, Vapnik V. Support-Vector Networks [J]. Machine Learning, 1995, 20(3):273-297. DOI:10.1023/A:1022627411411.

[3] Jolliffe I T. Principal Component Analysis [J]. Journal of Marketing Research, 2002, 87(4):513. DOI:10.2307/3172953.

[4] HU Mingying, TAO Sheng. A image scaling method of PCA saliency detection combined with correlation analysis [J]. Journal of Jimei University (Natural Science Edition), 2022, 27(4):379-384.

[5] CUI Weiqing, DANG Changchun, ZHANG Wang, et al. A HOG feature template matching algorithm [J]. Mechanical Management and Development, 2018, 33(11):252-253.