DOI: 10.46300/9106.2022.16.20 Volume 16, 2022

Finely Crafted Features for Traffic Sign Recognition

Wei Li, Haiyu Song, Pengjie Wang
School of Computer Science and Engineering, Dalian Minzu University,
Dalian, 16600
China

Received: July 9, 2021. Revised: December 11, 2021. Accepted: January 5, 2022. Published: January 7, 2022.

Abstract-Traffic sign recognition (TSR) is the basic technology of the Advanced Driving Assistance System (ADAS) and intelligent automobile, whileas high-qualified feature vector plays a key role in TSR. Therefore, the feature extraction of TSR has become an active research in the fields of computer vision and intelligent automobiles. Although deep learning features have made a breakthrough in image classification, it is difficult to apply to TSR because of its large scale of training dataset and high space-time complexity of model training. Considering visual characteristics of traffic signs and external factors such as weather, light, and blur in real scenes, an efficient method to extract high-qualified image features is proposed. As a result, the lower-dimension feature can accurately depict the visual feature of TSR due to powerful descriptive and discriminative ability. In addition, benefiting from a simple feature extraction method and lower time cost, our method is suitable to recognize traffic signs online in real-world applications scenarios. Extensive quantitative experimental results demonstrate the effectiveness and efficiency of our method.

Keywords-Traffic sign recognition, Advanced Driving Assistance System, Feature extraction; computer science

I. Introduction

Traffic signs Recognition (TSR) is a important topic for both industry and academia, and it has been a hot area of research over the past decades. As one of the vital fields in the intelligent transportation system (ITS), driverless system and automatic driving system, a robust TSR system can assist drivers by providing helpful information that makes the driving more convenient and safe. Moreover, the TSR system could deal with complicated problems that obstruct recognition effectiveness. The problems include variations in illumination (fog, rain, light levels, shadow, and twilight), signs occlusion, and motion blur. Owing to the problems, traffic signs recognition is a challenging research task. Nevertheless, for improving driving safety, the development of such an automatic traffic sign recognition system has made great progress.

The researches of TSR have made much research progress in recent years. Traditional traffic sign recognition approaches include a wide variety of algorithms, such as template matching [1,2], handcrafted features with classifier [3,4], and deep learning method [5,6]. In the template matching approaches, the traffic sign region is compared to a set of traffic signs template exemplars (models) labeled with discrete class in order to find out the most similar traffic sign class. Since the strict requirements for the traffic sign templates, the approaches based on template matching are difficult to get high performance. The handcrafted features include color-based features, histogram of oriented gradient (HOG), scale-invariant feature transform (SIFT), etc and the multi-feature fusion is widely applied in traffic sign recognition. Feature selection and fusion are the difficulty of the method base on feature extraction. In recent years, deep learning, especially Convolution Neural Networks (CNN) [7], has been successfully applied in pattern recognition. Benefiting from the abundant German Traffic Sign Recognition Benchmark (GSTRB) dataset [8], the methods based on CNNs have achieved the higher classification accuracy. The CNN with learned color and spatial transformation achieve a satisfying accuracy of 99.59%, but the designed structure is complex. Therefore the TSR methods based on deep learning are time-consuming or need to learn a large amount of parameters.

Since the color, the shape, and the size are unique attribute information of traffic signs and color is the most differentiated feature of traffic signs, we introduce a finely crafted feature based on feature fusion of color-histogram-based feature and HOG feature. We also improve the color-histogram-based feature and HOG feature by reducing the dimensions within categories using PCA, which greatly enhance the execution speed of the algorithm to 16.91 millisecond with higher performance (Accuracy of our method is 99.99%) . To recognize traffic signs in real traffic scene, our research is based on the public GTSRB dataset [8], which contains a large number of traffic sign images in different environments and conditions.

Our contribution in this paper can be summarized as follows.

First, since the shortcoming of the TSR method based on deep learning is calculating complexity and too many training parameters, we introduce the finely crafted feature based on the color-histogram-based features and HOG features. The dimension of the extracted features is reduced by PCA, which makes the running time of traffic sign recognition significantly reduced.

Second, we further introduce a feature dimension reduction method based on PCA within categories. Our method not only improves the algorithm efficiency but also improves its accuracy (test time perimage is 16.91ms and the accuracy of our method is 99.99%).

II. Related work

The first research on TSR can be traced back to 1987 when Akatsuka and Imai [11,12] worked on earliest TSR system. The TSR systems usually consist of two stages, one is traffic sign detection, and the other is traffic sign classification. In the first stage, the regions of interest (ROI) are extracted from the traffic signs. In the second stage, the ROI is identified as a certain traffic sign through classification methods. The classification methods that were used for the ROI including color-based techniques, as in [13-16]. Some methods used Histogram of Oriented Gradient (HOG) features as the feature extracted from traffic signs [17-18]. Takaki proposed a Scale-Invariant Feature Transform (SIFT) [19] that can be applied to traffic sign's ROI. The classifying process generally includes two parts: first is feature extraction, that is, extracting the feature information from the ROI of traffic signs; second is classifying traffic signs into different classes, that is, judging the class of the ROI area of traffic signs. In recent years, traffic sign classification has also made much progress and the main TSR methods are based on models, machine learning.

A.Traffic Sign Detection

In the detection stage, the image is segmented based on the visual characteristics of traffic signs, such as color and shape. Traffic sign detection methods can be divided into three categories, the first is color-based methods, the second is shape-based methods, and the third is machine learning methods[20].

1) Traffic Sign Detection based on color features

Traffic signs have bright primary colors contrast strongly with background environments, which represent basic information of the signs. Many methods segment the traffic signs in a specific color space. Typically, images taken by cameras are based on RGB color space, whereas, the RGB color space is sensitive to the illumination variations. Therefore, some researches[21] applied a color ratio between the intensity components of RGB, while some other researches [22] used only one RGB component as a reference to detect the sign colors [23]. The Hue Saturation Intensity (HSI) [24] and Hue Saturation Value (HSV) [25] color spaces have been frequently used to reduce the dependency on illumination variation.

2) Traffic Sign Detection based on shape features

The method based on fast-radial symmetry transform gave a regular polygon detector for detecting traffic signs, which has close similarity with Hough transform. The approach in [27] detect rectangular and triangular signs by the Harris corner detector method and inquiry corners in six pre-defined control areas of a region. The shape of the sign is then judged through the configuration of the control regions which the corners are found there. The method in [28] used a distance transform (DT) based on template matching method, and edges in the original image are found in which the DT image is built. Radically, the DT image is an image with every pixel representing the distance to the closest edge, and the basic theory is to match a template against the DT image to obtain the shape of an image.

3) Traffic Sign Detection methods base on Machine Learning

DOI: 10.46300/9106.2022.16.20 Volume 16, 2022

Manual features, such as HOG [29], SIFT [30], enhancement information of typical color or geometric shape, fail in many difficult environments. However, CNNs are considered to be robust and powerful in many applications and some relevant work based on CNNs has been done. A novel framework was proposed with two deep learning components including fully convolutional network (FCN) and deep convolutional neural network (CNN) which guiding traffic sign proposals and object classification respectively in [31]. The proposed approach compared itself with R-CNN [32] in the experiments. A multi-class fully convolutional network [33] was proposed to detect and classify traffic signs simultaneously. Besides the network, another detection network was proposed for treating all traffic signs as one category in the detection step. In [33], the performance of the original Fast R-CNN [34] for traffic sign recognition is analyzed in detail. In [35], the original Faster R-CNN [36]was applied to perform the task, and the results show that this approach is valuable, however, the accuracy and speed are unable to meet practical needs.

B. Traffic Sign Classification

To ensure a prominent classification, there are mode-based methods and machine-learning-based methods.

1) Traffic Sign Classification based on models

A group of works are based their identification process on traffic signs models. In fact, the traffic sign region is compared to a set of traffic signs template models labeled with discrete categories in order to find out the most similar traffic sign category. To perform traffic signs matching, some comparison metrics are applied as the normalized correlation between the templates and the traffic sign regions [23,38].

2) Traffic Sign Classification based on machine learning

The machine-learning-based traffic sign classification methods are based on a training stage in which different artificial technics can be applied, such as Neural Network (NN) [39,40] and Support Vector Machine (SVM) [41]. They regard traffic signs as a global entity whose characteristics are learned and require some prior knowledge about the traffic sign structure.

The traffic sign classification methods based on machine learning using neural networks with their different extended networks have been widespread used. Actually, some researchers applied CNNs [39] while others applied the radial-based neural networks [40]. The SVM classifier are also widely used to identify the corresponding traffic sign categories[41]. Furthermore, the Adaboost algorithm was applied to recognize traffic signs with group classifiers [23,42].

With the development of deep learning, the field of traffic sign classification is one of the earlier areas applying deep learning technology [43-45]. Many of the top methods in the IJCNN2011 traffic sign recognition competition had adopted Convolution Neural Network methods, including LeCun with their team of Turing Award 2018 for deep learning. Paper [40] proposed Multi-scale CNN to identify traffic signs and achieved great success. However, most of the deep learning models proposed or adopted in this field are often combined with other feature or complex classification models.

The method in [9] used the CNN to train step-by-step after the traditional feature extraction with the first stage feeding into the classifier along with the second stage features, thus, it produced more accurate results than any other mechanism by increasing the network's depth or disregarding the color information[10].

III. Method

Recently, deep learning features have shown significant performance in image classification and recognition. Deep features always are more general and stable than manual features, but they do not outperform customized finely crafted manual features for specific applications. In addition, in most cases, deep feature extraction requires expensive space-time overhead, which limits their application to TSR. Different from the traditional feature of color histogram concatenating three-channel histograms, in our method, each bin combine information from three channels. To obtain a compact and discriminative feature, we reduce dimension using PCA. As a result, we propose a comprehensive feature vector for TSR, combining color histogram and HOG feature with dimension reduction, and perform traffic sign recognition using an SVM classifier.

A. Color-histogram-based method

In our method, we use RGB color space to extract color feature. The values of color histogram are required to be counted to reflect the number of features of different color grayscale in the image. The traditional procedure is as follows.

As the RGB image is a three-dimensional gray image consist of a red channel, a green channel, and a blue channel, the image is divided into three dimensions when it is processed. Commonly, the quantity distribution of color in each dimension is calculated to form a histogram, then the histogram is connected in series to form the feature vector of the image (see Figure 1 (a)), the dimensions of the color-histogram feature vector is 48.

The traditional color histogram feature is defined by taking the ratio of the number of pixels conforming to the pixel value in a certain interval to the total number of pixels in each channel as the bins of the current channel. The final color histogram is generated by concatenating the histograms of each channel and the dimension of the feature is the total number of bins of three channels. The improved color histogram adopted in our model is generated as a global histogram by calculating the ratio of the number of pixels in a certain pixel value range of three channels to the total number of pixels and the dimension of the feature is Nr*Ng*Nb (Nr, Ng and Nb are the numbers of pixel intervals of three channels in RGB color space respectively). Therefore, compared with the traditional color histogram feature, the dimension of the improved color histogram feature will be much higher, the description ability of the improved feature for the image is more differentiated. Consequently, the improved color-histogram-based feature can better reflect the characteristics of the image.

B. HOG-based method

The methods based on shape features mainly use Histogram of Oriented Gradient (HOG) features. HOG is a feature descriptor used for object recognition in computer vision and HOG features are constructed by calculating and statistics of the gradient histogram of the local region in an image. Generally, the shape features of local targets can be well described by gradient or edge oriented density data distribution. The procedure is as follows.

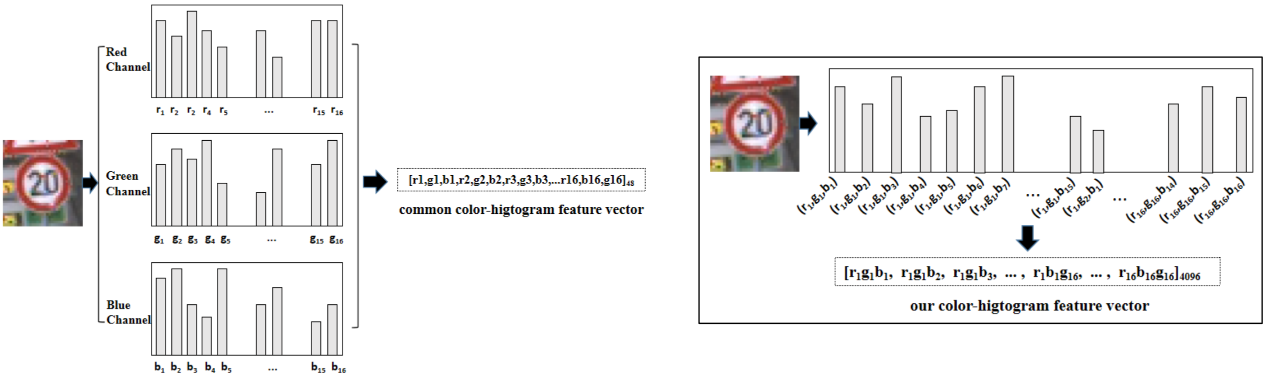
Graying the images and normalizing them by Gamma correction method, then each image is divided into different cells; Making a statistic of the gradient histogram of each cell, then we get the descriptor; Combining several cells into one block with overlapping ones to form the block descriptors; We obtain the HOG features of the images by connecting the blocks in series.

C.The dimension reduction method of features based on Principal Component Analysis (PCA)

To describe color features more differentiated, we make the color-histogram of three color channels cross connection with each other and the step size of the histograms is unequal length (see Figure 1 (b)). Thus, the dimensions of the color-histogram feature vector are too high (the dimension of our color-histogram feature vector is 4096). The high dimensionality increases computational complexity and time-consuming. We are inspired by the characteristics that image features have correlations between images. Therefore, if we can remove the related content of an image feature, the computational complexity would be greatly reduced with low loss of information entropy. Representing images as different features, such as color histogram, HOG, the feature vectors have a certain correlation, redundant features and noise error factors with a high dimension. We apply dimension reduction method to compress the multi-dimensional data corresponding to image features, retain the most important information dimension and remove those unimportant information. Due to the reduction of dimension, we can show excellent memory management. The final features are selected by deleting unnecessary feature lists, so as to improve the performance of the model. Furthermore, fewer dimensions make the computing efficiency higher, the training model faster, the model performance higher, and reduce the complexity of the model and the risk of over fitting greatly.

Since Principal Component Analysis (PCA) method is widely applied in the fields of data mining and machine learning, which has a strict mathematical foundation for dimension reduction, we introduced PCA in reducing the dimension of an image color-histogram and HOG feature to reduce the coupling relation between features. PCA transforms the original feature matrix into a set of linear independence representations between different dimensions by linear transformation. The main features are extracted, the dimension of the vector is reduced, and the running time is improved largely with PCA.

PCA is calculated as follows. Supposing A is the feature matrix as M*N, each row of the matrix represents a data record and the number of the records is M; Each column of the matrix represents one class of color-histogram feature information and the number of feature classes is N. The dimension reduction procedure is as follows. First, the mean value is subtracted from each value for all N columns and we get the covariance matrix A subtracting the mean. Second, we calculate and get the eigenvalues of the covariance matrix and its corresponding feature vectors. Third, these feature vectors are sorted by their values and the first K columns of the feature vectors form into a new matrix B, then the matrix of N column is reduced into the matrix of K column (K<N). Last, we get the new matrix P =B*A after dimension reduction by PCA technique.



(a) (b)

Fig. 1 The extraction procedure of common color-histogram feature (a) and our color-histogram feature (b)

The covariance matrix formula is shown as formula 1:

$$X = \begin{bmatrix} X1 \\ X2 \\ \vdots \\ Xn \end{bmatrix} \tag{1}$$

$$\sum = \begin{bmatrix} E[(X1-\mu1)(X1-\mu1)] & E[(X1-\mu1)(X2-\mu2)] & \cdots & E[(X1-\mu1)(Xn-\mu n)] \\ E[(X2-\mu2)(X1-\mu1)] & E[(X2-\mu2)(X2-\mu2)] & \cdots & E[(X2-\mu2)(Xn-\mu n)] \\ E[(X3-\mu3)(X1-\mu1)] & E[(X3-\mu3)(X2-\mu2)] & \cdots & E[(X2-\mu2)(Xn-\mu n)] \\ \vdots & \vdots & \vdots & \vdots \\ E[(Xn-\mu n)(X1-\mu1)] & E[(Xn-\mu n)(X2-\mu2)] & \cdots & E[(Xn-\mu n)(Xn-\mu n)] \end{bmatrix}$$

D.The dimension reduction method of Color-histogram and HOG features based on PCA

GTSRB dataset consists of standard traffic sign images. Although the signs are different from each other, signs within one target category have some common properties. For example, Prohibitory signs have circular red borders, Danger signs have triangular red borders and Mandatory signs have blue backgrounds and white arrows. HOG feature is a widely used image feature in computer vision, however, it lacks color description of an image. Considering the characteristic of traffic sign images in color and shape, we combine the color-histogram and HOG to get higher accuracy. Consequently, the combined features have more description ability with more comprehensive information. We also make the color-histogram of three color channels cross-connection with each other. Thus the dimension of the color-histogram is too high to calculate, so we apply PCA to reduce the dimension of the color-histogram-features (see Figure 2).

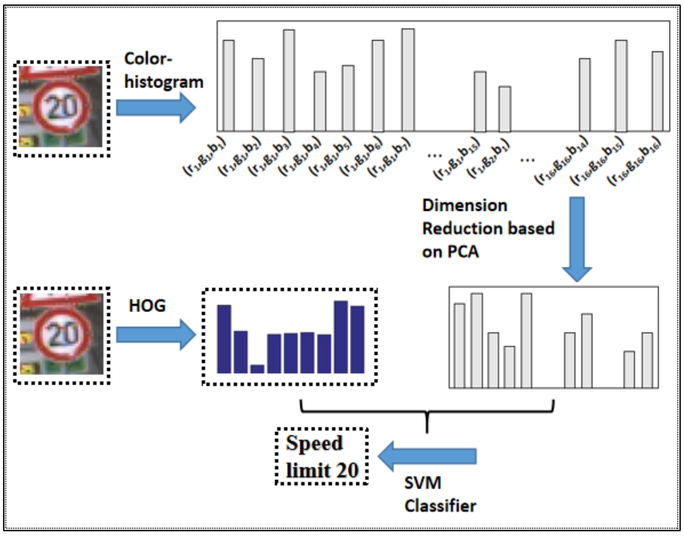


Fig. 2 The architecture of our method

As the high discrimination of traffic signs, the color and shape features of one category should be significantly different from those of other categories, but they are similar to their own category. Therefore, we extract the principal feature components of each category and merge them into a matrix dimension reduction features.

The dimension reduction procedure of our method is as follows. Supposing A is the feature matrix of a certain category as M*N. First, the mean value is subtracted from each value for all columns and the values were squared to become positive after subtract the mean. Second, after the values of each column are divided by the variance of the column, we get the normalized values of Gauss, which can determine the similarity of features, that is, the values are inversely proportional to the similarity. Last, we get a new dimension reduction matrix by choosing the first few columns features with small overall similarity in each category and give these similar features some weights.

The Gauss normalized formula of the matrix column is as formula 2:

$$X=\begin{bmatrix} X1 \\ X2 \\ \vdots \\ Xn \end{bmatrix} \qquad Y=\begin{bmatrix} (X1-\overline{X})(X1-\overline{X})/Var(X) \\ (X2-\overline{X})(X2-\overline{X})/Var(X) \\ (X3-\overline{X})(X3-\overline{X})/Var(X) \\ \vdots \\ (Xn-\overline{X})(Xn-\overline{X})/Var(X) \end{bmatrix} \qquad (2)$$

E. Classification methods

We applied the Support Vector Machine (SVM) as the classifier after feature extraction of traffic sign images. SVM is a method to obtain the optimal classification surface in the case of binary pattern recognition and linear separability.

In order to solve the problem of dimension disaster, when using the SVM classifier to solve problems, it is necessary to replace the dot product calculation with appropriate kernel functions according to the characteristics of problems. It is normally that transforming the dot product operation of high dimensional feature space into the kernel function operation of low dimensional original space implicitly. The kernel function not only satisfies the Mercer condition in theory but also reflects the distribution characteristics of sample data in practical application. Therefore, the selection of kernel function is the core problem in SVM theory.

In practice, the most commonly used kernel functions include the Polynomial kernel, the Radial basis function (RBF) kernel , Multilayer Perceptron (MLP) kernel, the Fourier series kernel, B-spline kernel, and Sigmoid kernel. Since RBF kernel and Polynomial kernel are better than Sigmoid kernel in classification effect, and the parameters' number of RBF is less than Polynomial kernel, we choose RBF as the kernel of our SVM classifier. RBF is a universal kernel and reflects the complexity of the model, so it can be applied in any sample distribution by selecting the appropriate kernel parameters.

The RBF is adopted as formula (3):

$$K(x,z)=\exp\left(-\frac{(x-z)^2}{2\sigma^2}\right)=\exp(-\gamma\cdot(x-z)^2) \quad (3)$$

Formula 3 is the most widely used kernel function in the base of the support vector. Where x the sample of the feature, $\sigma$is variance, and γ is the only hyperparameter which defines the generalization ability of SVM classifier. Therefore, we adopt the RBF as the kernel function of SVM classifier and the classification decision function is as formula (4).

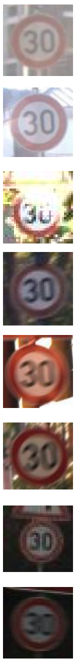
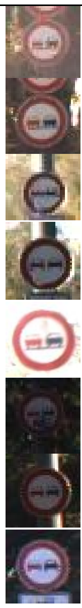
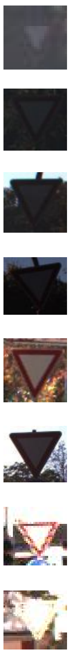
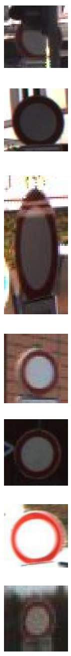
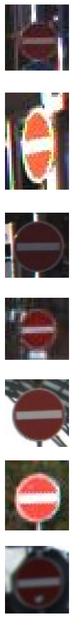
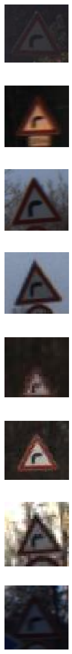
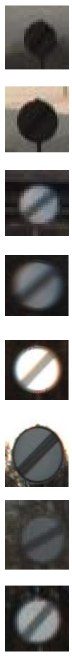
$$f(x) = sign(\sum_{i=1}^{N_s} a_i^* y_i \exp(-\frac{\| x - z \|^2}{2\sigma^2}) + b^*) \quad (4)$$

IV. Experiment

A. Datasets

GTSRB is the traffic sign recognition dataset constructed by the traffic sign detection and recognition competition jointly organized by the German academic and industrial circles at IJCNN2011 of the international conference in 2011[8]. GTSRB is a large, lifelike dataset of more than 50,000 traffic sign images in 43 categories. The dataset contains a large number of images under various conditions, such as low resolution, different light intensity, partial occlusion, tilt, motion blur shown in Table I.. It can reflect the real traffic scene and is difficult to recognize [8].

TABLE I. Samples in GTSRB traffic sign recognition dataset

Category number	0	1	2	9	11	13	15	17	20	32	42
Samples in GTSRB dataset											

The images of the GTSRB dataset (see in Table I) are detected and saved as separated traffic signs from the real world. The definition of the images in the dataset is a bit different and the clarity of most images is so poor that many images can not be recognized by experts. All these cases reflect the real condition of the application in traffic sign recognition technology, so it is hard to recognize the traffic signs accurately.

After 2011, although GTSRB has become the most widely used traffic sign recognition dataset in the field of traffic sign recognition, for many reasons, the experimental datasets used in the published results in this field are still diverse. Some are small datasets or some are collecting datasets by researchers themselves, so it is difficult to compare performance between different methods directly. Considering the characteristics of the GTSRB dataset, such as large sample size, various complex scenes, difficulty of recognition, we experiment with our method on the GTSRB dataset and compare the performance with some state-of-the-art methods.

B. Evaluation metrics

We apply the precision, recall, F1-measure and accuracy to evaluate our traffic sign recognition method. The evaluation metrics is calculated as equation (5)-(8).

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

$$F1 = \frac{2 * Precision * Recall}{Precison + Recall} \tag{7}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{8}$$

Where TP is the number of correct traffic signs predicted, FP is the number of wrong ones predicted, TN is the number of correct negative case, and FN is the number of false negative case. F1-measure combines Precision and Recall, which indicates the integrated result.

C. Training details

Since there are 50000 images in the GTSRB dataset, it will take a great deal of time to train and test the data. To save running time, we choose two images as the training set in each track of the original training data set and one image as the test set randomly in each track. Thus, we get the better parameters with approximate accuracy, but the execution time can be greatly reduced.

Although there are many parameters in SVM, we only describe parameters c, v, g here in detail. SVM classifier requires to optimize parameters c and g when using RBF kernel. Parameter c is the penalty coefficient: if c is too small, it is easy to be over-fitting; if c is too large, the fitting results will be very different from the actual results. Parameter g is the coefficient of the RBF kernel and the value of g will affect the ability of SVM to find an optimal hyperplane that separates multiple types of data. Parameter v represents the number of cross-validated groups.

We introduced the cross-validation function SVMcg to determine optimal parameters. SVMcg function defines the range of parameters c and g in $[2^{-5}, 2^{+5}]$. Thus, the best parameter c and parameter g can be obtained quickly and effectively when training.

D. Experimental results

Our experiment is implemented in Corei7 CPU with 24G RAM.

Firstly, the extracted RGB color-histogram feature and HOG feature are directly trained and tested without dimension reduction. The test results are as following:

TABLE II. EXPERIMENTAL RESULTS BASED ON COLOR-HISTOGRAM AND HOG FEATURES

Image feature	Precision	Recall	F1-measure	Accuracy	Test time
Color-histogram-based	69.21%	58.08%	65.37%	64.73%	602.75ms
HOG	87.28%	70.56%	73.99%	94.88%	611.62ms

Obviously, the test time in table II is time-consuming and the extraction method without dimension reduction can not satisfy the purpose of assisting drivers. The accuracy of the method based on improved color-histogram-based or HOG features can not meet the requirement of reality in TSR also, therefore, the extraction method of image features is required to be improved.

TABLE III. EXPERIMENTAL RESULTS AFTER DIMENSION REDUCTION BASED ON PCA

Image feature	Precision	Recall	F1-measure	Accuracy	Test time
PCA on Color-histogram	68.23%	57.92%	62.65%	62.21%	101.88 ms
PCA on HOG	95.45%	89.87%	92.58%	94.57%	179.12 ms

PCA on HOG + color-histogram (our method)	99.85%	92.78%	96.19%	99.99%	16.91 ms

The precision, recall, F1-measure and accuracy in table III of the image features extraction method improved by PCA dimension reduction technology is similar to the data in table 2, however, the test time is greatly reduced. The test time of the color-histogram-based feature is 101.8 milliseconds and the test time of the HOG-based feature is 179.12 milliseconds respectively. The accuracy of 94.57% is still a gap from the actual application of TSR.

From table III, we can find that the test time is reduced and the accuracy is improved greatly with the dimension reduction method base on PCA within categories. Especially, the evaluation metrics of the dimension reduction method applying in the concating feature of HOG and color-histogram-based feature can reach the requirement of the actual application.

TABLE IV. COMPARISON OF OUR METHODS WITH OTHER EXISTING METHODS [37]

method	Precision	Recall	F1-measure
LDA	96.09%	89.47%	92.12%
Multi-Scale CNNs	98.72%	91.23%	94.35%
Random Forest（HOG）	94.56%	85.69%	88.56%
Our method	99.85%	92.78%	96.19%

According to the experimental results, it can be seen that the PCA within categories on the HOG feature dimension reduction method applying in the GTSRB traffic sign dataset is superior to other methods. Whether in terms of test time or evaluation metrics, our method is a little better than the state-of-art methods (see Table IV).

V. CONCLUSIONS

In this paper, we introduce PCA to improve the finely crafted features based on the color-histogram and HOG features of traffic sign images, and improve the features based on PCA within categories to obtain excellent performance further. The accuracy is better than the state-of-art methods in the field of TSR and the test time has been greatly reduced in our method especially. Our main contributions are as follows:

1) We applied the color information of traffic signs fully and used the improved color-histogram-based feature to increase the discrimination between images.

2) We adopted PCA algorithm to reduce the dimension of the improved color-histogram-based feature so as to increase the running speed of our method.

3) In order to enhance the expression ability of features, we applied the HOG feature of images with concatenating the improved color-histogram-based feature after dimensionality reduction.

Nevertheless, our method is only verified on the GTSRB dataset, and lack of comparison of recognition effects of traffic signs of other countries. We cannot declare that our method is suitable for all traffic sign recognition tasks.

In future research, we would experiment our method on traffic sign recognition datasets of different countries, so as to further verify the effectiveness of our method. Furthermore, we would aim to experiment other feature descriptors and classifiers as well as comparing the performance of our method with the most recent methods.

References

[1] A. M. Moussa, M. I. Habib and R. Y. Rizk, “FRoTeMa:

Fast and Robust Template Matching," International Journal of Advanced Computer ence and Applications, 2015, 6(10),pp. 195-200.

[2] L. D. Stefano, S. Mattoccia and F. Tombari, "ZNCC-based template matching using bounded partial correlation," Pattern Recognition Letters 2005, 26, PP. 2129–2134.

[3] S. Ardianto, C. J. Chen and H. M. Hang, "Real-time traffic sign recognition using color segmentation and SVM," 2017 International Conference on Systems, Signals and Image Processing (IWSSIP), **2017**, pp. 1–5.

[4] H. N. Thanh, "Morphological Classification for Traffic Sign Recognition," Electrical and Electronic Engineering **2014**, 4, PP. 36–44.

[5] Y. Du, R. Zhang, A. Zargari and T. C. Thai, et al., "Classification of Tumor Epithelium and Stroma by Exploiting Image Features Learned by Deep Convolutional Neural Networks," Annals of Biomedical Engineering 2018, 46, pp. 1988–1999.

[6] X. B. Quan and W. X. Xiong, "Real-Time Embedded Traffic Sign Recognition Using Efficient Convolutional Neural Network," IEEE Access **2019**, PP, 1–1.

[7] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems 25, 2012, Vol. 25, pp. 1097–1105.

[8] J. Stallkamp, M. Schlipsing, J. Salmen and C. Igel, "The German Traffic Sign Recognition Benchmark: A multi-class classification competition," The 2011 International Joint Conference on Neural Networks, 2011, pp. 1453–1460.

[9] P. Sermanet and Y. LeCun, "Traffic sign recognition with multi-scale Convolutional Networks," The 2011 International Joint Conference on Neural Networks, 2011, pp. 2809–2813.

[10] S. Hussain, M. Abualkibash and S. Tout, "A Survey of Traffic Sign Recognition Systems Based on Convolutional Neural Networks," 2018 IEEE International Conference on Electro/Information Technology (EIT), 2018, pp. 570–573.

[11] B. Hoferlin and K. Zimmermann, "Towards reliable traffic sign recognition," 2009 IEEE Intelligent Vehicles Symposium, 2009, pp. 324–329.

[12] H. Akatsuka and S. Imai, "Road Signposts Recognition System. SAE International Congress and Exposition," 1987, Vol. 96, pp. 936–943.

[13] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez-Moreno and F. Lopez-Ferreras, "Road-Sign Detection and Recognition Based on Support Vector Machines," IEEE Transactions on Intelligent Transportation Systems **2007**, 8, pp. 264–278.

[14] R. Timofte, K. Zimmermann and L. V. Gool, "Multi-view traffic sign detection, recognition, and 3D localisation," machine vision applications **2014**, 25, pp. 633–647.

[15] J. Miura, T. Kanda, S. Nakatani and Y. Shirai, "An Active Vision System for On-Line Traffic Sign Recognition," IEICE Transactions on Information and Systems **2002**, 85, pp. 1784–1792.

[16] A. de la Escalera, J. Armingol and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles," Image and Vision Computing **2003**, 21, pp. 247–258.

[17] F. Zaklouta and B. Stanciulescu, "Real-Time Traffic Sign Recognition Using Tree Classifiers," IEEE Transactions on Intelligent Transportation Systems **2012**, 13, pp. 1507–1514.

[18] A. Ruta, Y. Li and X. Liu, "Robust Class Similarity Measure for Traffic Sign Recognition," IEEE Transactions on Intelligent Transportation Systems **2010**, 11, pp. 846–855.

[19] M. Takaki and H. Fujiyoshi, "Traffic Sign Recognition Using SIFT Features" Ieej Transactions on Electronics, Information and Systems **2009**, 129, pp. 824–831.

[20] P. Pandey and R. K. Kulkarni, "Traffic Sign Detection for Advanced Driver Assistance System," 2018 International Conference On Advances in Communication and Computing Technology (ICACCT) **2018**, pp. 182–185.

[21] S.K. Berkaya, H. Gunduz, O. Ozsen, C. Akinla rand S. Gunal, "On circular traffic sign detection and recognition," Expert Systems With Applications **2016**, 48, pp. 67–75.

[22] A. de la Escalera, L.Moreno, M. Salichs and J. Armingol, "Road traffic sign detection and classification," IEEE Transactions on Industrial Electronics **1997**, 44, pp. 848–859.

[23] N. B. Romdhane, H. Mliki and M. Hammami, "An improved traffic signs recognition and tracking method for driver assistance system," 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), 2016, pp. 1–6.

[24] Y. Han, K. Virupakshappa and E. Oruklu, "Robust traffic sign recognition with feature extraction and k-NN classification methods," 2015 IEEE International Conference on Electro/Information Technology (EIT),

2015, pp. 484–488.

[25] F. Mariut, C. Fosalau, M. Avila and D.Petrisor, “Detection and recognition of traffic signs using Gabor filters,” 2011 34th International Conference on Telecommunications and Signal Processing (TSP), 2011, pp. 554–558.

[26] G. Loy and N. Barnes, “Fast shape-based road sign detection for a driver assistance system,” 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), 2004, Vol. 1, pp. 70–75.

[27] C. Paulo and P. Correia, “Automatic Detection and Classification of Traffic Signs,” Eighth International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS ’07), 2007, pp. 11–11.

[28] D. Gavrila, “Traffic Sign Recognition Revisited,” Mustererkennung 1999, 21. DAGM-Symposium, 1999, pp. 86–93.

[29] Deepika, S. Vashisth and S. Saurav, “Histogram of Oriented Gradients Based Reduced Feature for Traffic Sign Recognition,” 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 2206–2212.

[30] B. Hoferlin and K. Zimmermann, “Towards reliable traffic sign recognition,” 2009 IEEE Intelligent Vehicles Symposium, 2009, pp. 324–329. doi:10.1109/IVS.2009.5164298.

[31] Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai and W. Liu, “Traffic sign detection and recognition using fully convolutional network guided proposals,” *Neurocomputing* **2016**, *214*,pp. 758–766.

[32] R. Girshick, J. Donahue, T. Darrell and J. Malik, “Region-Based Convolutional Networks for Accurate Object Detection and Segmentation,” IEEE Transactions on Pattern Analysis and Machine Intelligence **2016**, *38*, pp. 142–158.

[33] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li and S. Hu, “Traffic-Sign Detection and Classification in the Wild,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2110–2118.

[34] R. Girshick, “Fast R-CNN,” 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 1440–1448.

[35] E. Peng, F. Chen and X. Song, “Traffic Sign Detection with Convolutional Neural Networks,” International Conference on Cognitive Systems and Signal Processing **2016**, pp. 214–224.

[36] S. Ren, K. He, R. Girshick and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence **2017**, *39*, pp. 1137–1149.

[37] J. Li and Z. Wang, “Real-Time Traffic Sign Recognition Based on Efficient CNNs in the Wild,” IEEE Transactions on Intelligent Transportation Systems **2019**, *20*, pp. 975–984.

[38] A. U. Peker, O. Tosun, H. L. Akin and T. Acarman, “Fusion of map matching and traffic sign recognition,” Intelligent Vehicles Symposium Proceedings, 2014 IEEE, 2014, pp. 867–872.

[39] R. Qian, B. Zhang, Y. Yue, Z. Wang and F. Coenen, “Robust chinese traffic sign detection and recognition with deep convolutional neural network,” 2015 11th International Conference on Natural Computation (ICNC), 2015, pp. 791–796.

[40] M. M. Lau, K. H. Lim and A. A. Gopalai, “Malaysia traffic sign recognition with convolutional neural network,” 2015 IEEE International Conference on Digital Signal Processing (DSP), 2015, pp. 1006–1010.

[41] C. Liu, F. Chang and Z. Chen, “High performance traffic sign recognition based on sparse representation and SVM classification,” Natural Computation (ICNC), 2014 10th International Conference on, 2014, pp. 108–112.

[42] L. Chen, Q. Li, M. Li and Q. Mao, “Traffic sign detection and recognition for intelligent vehicle,” 2011 IEEE Intelligent Vehicles Symposium (IV), 2011, pp. 908–913.

[43] D. A. Alghmgham, G. Latif, J. Alghazo, et al., “Autonomous Traffic Sign (ATSR) Detection and Recognition using Deep CNN,” Procedia Computer Science, 2019, 163,pp. 266-274.

[44] L. Shangzheng, “A Traffic Sign Image Recognition and Classification Approach Based on Convolutional Neural Network,” 2019 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), 2019, pp. 408-411, doi: 10.1109/ICMTMA.2019.00096.

[45] D. Xie , E. Nuakoh, P. Chatterjee, A. Ghattan, K. Edoh and K. Roy, “Traffic Sign Recognition for Self-driving Cars with Deep Learning,” In: Advanced Machine Learning Technologies and Applications. AMLTA 2020. Advances in Intelligent Systems and Computing, vol 1141.pp. 207-217, Springer, Singapore.

https://doi.org/10.1007/978-981-15-3383-9_19

Wei Li received the B.S. and M.S. degrees in computer software and theory from Jilin University, in 2002 and 2005, respectively. She received the Ph.D. degree in the Transportation College, Jilin University, China. She is also a Lecture with the School of Computer Science and Engineering, Dalian Minzu University, China. Her current research interests include image understanding, computer vision, and machine learning.

Haiyu Song received the B.S.,M.S.,and Ph.D. degrees in computer software and theory from Jilin University, in 1996, 2003, and 2012, respectively. He is currently a associate professor with the School of Computer Science and Engineering, Dalian Minzu University, China. His current research interests include image understanding, computer vision, and machine learning.

Pengjie Wang received the B.S. and M.S. degrees in computer software and theory from Jilin University, in 2001, 2004, and the Ph.D. degree in computer application from Zhejiang University, in 2011. He is currently a professor with the School of Computer Science and Engineering, Dalian Minzu University, China. His current research interests include computer vision, virtual reality, and machine learning.

Author Contributions: Please, indicate the role and the contribution of each author:

Wei Li was responsible for investigating relevant literature, planning the system, and completing the simulation and optimization work.

Haiyu Song was responsible for the construction of the experimental model of the fourth part.

Pengjie Wang was responsible for completing the experimental part.