IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 4, May 2022

Traffic Sign Recognition

Tejaswini Salunkhe[1], Sakshi Gaikwad[2], Sadaf Sayyed[3], Prof. Kshitija Potdar[4]

Students, Department of Computer Engineering[1,2,3]

Guide, Department of Computer Engineering[4]

Dhole Patil College of Engineering, Pune, Maharashtra, India

Abstract: *Traffic sign detection and recognition are significant in the development of intelligent vehicles. And enhanced traffic sign detection and recognition algorithm for intelligent vehicles is proposed to address problems such as how easily affected traditional traffic sign detection is by the environment, and poor real-time performance of deep learning-based methodologies for traffic sign recognition. Firstly, the HSV is used for spatial threshold segmentation, and traffic signs are effectively detected based on the shape features. Secondly, the model is considerably improved on the basis of the classical convolutional neural network model by using Gabor kernel as the initial convolutional kernel, adding the batch normalization processing after the pooling layer and selecting Adam method as the optimizer algorithm. Finally, the traffic sign categorization and recognition experiments are conducted based on the German Traffic Sign Recognition Benchmark. The favourable prediction and specific recognition of traffic signs are achieved through the continuous training and testing of the network model. Experimental results show that the specific recognition rate of traffic signs reaches 99.75%, and the average processing time per frame is 5.4 ms. Compared with other algorithms, the proposed algorithm has remarkable accuracy and real-time performance, strong stimulus generalization ability and high training efficiency. The specific recognition rate and average processing time are markedly improved. This improvement is of considerable importance to reduce the accident rate and enhance the road traffic safety situation, providing a strong technical guarantee for the steady development of intelligent vehicle driving assistance.*

Keywords: Driving Assistance; Intelligent Vehicles; Traffic Sign Detection; Convolutional Neural Network; Traffic Sign Recognition.

I. Introduction

Traffic sign recognition is highly important in driving assistance systems because it can help to understand the environment and prevent vehicle accidents. Thus, this function has a significant potential benefit for supporting drivers. In addition, traffic sign recognition is a part of autonomous driving vehicles, a current hot topic worldwide. Detecting and recognizing distant objects is currently a challenging problem and a requirement for vehicles driving at high speed. It is a contract fact that when using images with large resolution, detection and recognition can improve their performance because the objects in the images are detailed, but the images require a large computational cost. On the other hand, images with smaller resolution require a lower computational cost; however, the detail of an image is reduced, leading to the detection and recognition method performing more poorly. In our recognition method, observing that traffic signs in Korea usually appear in the high positions so that drivers can see them easily. We remove a lower part of an image. Consequently, we can reduce the computational cost to achieve real-time computation. In addition, we do not recognize traffic signs directly. However, we design a two-step approach that solves the problem of traffic sign recognition. We observe that traffic signs often have circle, rectangle, and triangle shapes. Therefore, we first detect if the shape appears in an image. This can also reduce computational cost because the method in this stage only deal with three patterns instead of many more. After that, the image regions (that will be processed more) become much smaller. We divide our recognition method into two steps: detecting traffic signs first and then recognizing the detailed information of the traffic signs. After determining that an image region (usually small image region) contains a traffic sign, we up-sample the image region before performing recognition. HOG was originally designed for human detection. The sizes of humans in images are often large enough to be detected. However, traffic signs in real images are usually small. Using HOG with a small finding patch increases false alarms significantly. Recently, deep learning methods have been popular in traffic sign detection and recognition. However, deep learning-based methods for traffic sign detection are far from real-time computation and are even supported with GPU.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 4, May 2022

II. Related Work

Generally, traffic sign recognition contains two parts: detection and categorization. The purpose of detection is to find the locations and sizes of the existing traffic signs in an image, and the task of categorization is to assign a class label to each detected traffic sign. Related works on these two parts are reviewed respectively in this section.

2.1 Traffic Sign Detection

In todays' rapidly growing world, it is difficult to recognize every block in the original image with the traffic sign recognition algorithm for real-time performance. Therefore, traffic sign detection algorithm should be adopted to extract the blocks that might be traffic signs from the original image in a rapid and precise way. For the similar reason, it is also impossible to detect every block in the original image with the traffic sign detection algorithm. There must be a step in the algorithm to extract the regions of interest (ROIs). Following the same set of colour and shape standards, traffic sign images from the same region carry similar features. Traffic signs usually contain four main colours red, yellow, blue, and white, and the typical shapes triangle, rectangle, circle, and octagon. Hence, the ROIs in the original image are separated based on the hue in the hue-saturation-value (HSV) colour space, and then denoised based on some shape features.

Mostly traffic signs include borders and background colour. The meaning of traffic signs mostly reside in the pictures within the borders. The borders and background colour makes the original image slightly complicated, which in return reduces the detection accuracy. To overcome this problem, the traffic sign detection algorithm should also remove the borders, convert the original image into a binary image, and eliminate noises, leaving only the important shape features in the ROIs. For combined processing, the traffic sign detection algorithm should map the ROIs to a standard size, and describe the shape features of the ROIs with distance to border (DtB). Hence, both hue and shape features could be occupied to depict the ROIs. The SVM is exceptionally good at obtaining the discriminant model of binary categorization problems quickly and specifically. As a result, the SVM has frequently been adopted as a classifier in traffic sign detection algorithms. Extracted from the camera data of intelligent vehicles, the traffic sign images may have abnormalities like occlusion, and colour fading. These abnormalities will make the hue and shape features of the images different from expected. A good traffic sign detection algorithm should be able to specifically detect all traffic sign images with abnormalities. Taking all normal images as simple samples and all deviant images as complex samples, the classifier of the traffic sign detection algorithm could be trained by the CL strategy. Under this strategy, the characteristic parameters of simple samples are extracted from simple curriculums, while the extra characteristic parameters of complex samples are extracted from complex curriculums. With the extracted parameters, the obtained computing model could distinguish.

As many samples as possible, and thus approximate the real model. Of course, it is a heavy work to classify original images into curriculums, owing to the sheer number of traffic sign samples. Thus, this paper decides to train the SVM by the SPL, and integrates the training method with the learning method into a novel algorithm called SPSVM. The performance of the SPSVM was tested in three steps: Firstly, the ROI extraction method was implemented to extract the candidate blocks from the training set and the test set; Next, every block was given a label about whether it is a traffic sign; Finally, the discriminant model was obtained by applying the SPSVM on the training set, and used to discriminate the training set and the test set, producing the accuracy of the SPSVM on the two datasets.

2.2 Traffic Sign Recognition

The images confirmed by the traffic sign detection algorithm as traffic sign images are the input data of the traffic sign recognition algorithm, which determines the semantics of these images, i.e. allocate each image into its class of traffic signs. Considering the various types of traffic signs, the recognition process is a multi-categorization problem. The DL algorithms boast relatively high categorization accuracy for multi-categorization problems. Among them, the CNN has been widely reported as the most specific traffic sign recognition algorithm. Therefore, this paper takes the CNN as the classifier of the traffic sign recognition algorithm. Our CNN consists of three convolutional layers and a fully connected layer. The first convolutional layer has 100 7×7 filers, each of which convolve a 7×7 neighbourhood in the input image. The second convolutional layer has 150 4×4 filters; The third convolutional layer has 250 4×4 filters; The fully connected layer contains 300 neurons, and outputs 43 features. The input data of the CNN are black and white traffic sign images scaled to the pixel size of 30×30. The output data are the probability of each input image belonging to a type of traffic signs. The input data of the traffic sign recognition algorithm contain both normal samples and deviant samples. The abnormality of deviant samples refers to the deviant shape features induced by complex situations, such as occlusion, jitter, and noise. The CL strategy could also be applied to train the CNN. Under this strategy, the CNN could extract the characteristic parameters of simple samples from simple curriculums, and the extra characteristic parameters of complex samples from complex curriculums. Therefore, this paper decides to train the CNN by the SPL, and integrates the training method with the learning method into a novel algorithm called SPCNN. The performance of the SPCNN was tested in three steps: Firstly, the images detected from training set and test set were all given labels about which type of traffic signs they belong to; Next, the traffic sign images detected from the training set were used to train the CNN; Finally, the trained CNN was adopted to recognize all the images detected from the training set and the test set, producing the accuracy of the SPCNN on the two datasets.

2.3 The SPL

The details of the SPL are given below. Let $D = \{(xi , yi)\}i$=1 n be the training set, where xi is the i-th sample, and yi is the label of the i-th sample. Suppose L(yi, g(xi, w)) is the loss of the prediction model brought by the i-th sample, whereg is the prediction function, and w is the parameter of the prediction model. Then, the objective function of the SPL optimization can be defined as: (( ) ( ) ( ) , [0,1] 1 min (, ;) , , , n w v n i i i i i E w v v L y f x w h v λ λ ∈ = = + ∑ (1) where, h is the self-paced regular function, whose attributes and properties are introduced by Meng et al. [28]. The following self-paced regular function is selected for this research: h v v (,) λ λ = − (2) For the CNN, the loss function can be expressed as: L y f x w p y y x ( ) i i , , log | ( ) = − = ( ) i (3) where, p(y=yi|x) is the value of CNN output corresponding to the class yi. For the SVM, the hinge loss function can be expressed as: ( ) , , max 0,1 ( ) { ( ) } T L y f x w i i = − + w x b y i i (4) where, b is the bias. Obviously, after each training, the loss of each sample can be obtained easily from the outputs of the CNN and the SVM. Eq. (1) can be solved by the expectation-maximization (EM) algorithm (Algorithm 1). In the EM algorithm, dataset D is the set of all samples; the array v is used to select the samples for each training; w is the parameter value of the MACHINE LEARNING model; L is the value of the loss function; λ is the threshold for sample groups; s is used to determine the increment of sample complexity in each training. During the SPL, the training is divided into multiple rounds. In each round, the training samples include all the simple samples and part of the complex samples. The sample composition reflects the knowledge acquisition in education process: reviewing simple knowledge, and probing deep into complex knowledge.

Algorithm 1: The EM algorithm

Input: Input dataset D, w_0, λ, and s
Output: w of $min_v\, E(w, v; \lambda)$.
repeat
  Calculate L of every sample by w;
  Sort the samples in ascending order of their loss values L;
  for i=1 **to** n
    if $L(y_i, f(x_i, w))<\lambda$ **then** v_i=1; // select this sample
    else
      v_i=0; // do not select this sample
    end
  end
  Update w by solving $\min \sum_{i=1}^{n} v_i L_i(y_i, f(x_i, w)) + \frac{1}{2}\|w\|_2$
  $\lambda=\lambda+s$
until w **is not changed or** v_i**=1,** i**=1, 2, ..., n.**

III. EXPERIMENTS

3.1 Datasets

Four datasets were selected for our experiments, including GTSDB, KITTI, BTSDB, and GTSRB. The former two were for training, and the latter two were for testing. The GTSRB contains 51,840 images containing traffic signs; the KITTI includes 400 1,242×375 images captured by two cameras (left and right); the GTSDB offers 900 1,360×800 images; the BTSDB provides 3,133 images shot by eight cameras. 3.2 Traffic sign detection The SVM is well known for its speed and effectiveness in binary categorization problems. Therefore, the SPSVM was applied to detect the ROIs extracted by the classifier, and judge whether the ROIs contain traffic signs. The SPSVM was compared with ChnFtrs, 3D TS, Bayesian classifier, and the original SVM, which does not use the SPL. The detection results were evaluated by four metrics,

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

IJARSCT
Impact Factor: 6.252

Volume 2, Issue 4, May 2022

Accuracy, False Rate, Area under Curve (AUC), and Average Precision (AP). Accuracy is the ratio of the number of correctly recognized traffic signs to the number of all traffic signs; False Rate is the ratio of the number of incorrectly recognized traffic signs to the number of all traffic signs; AUC is the area under the receiver operating characteristic (ROC) curve; AP is the area under the Precision-Recall (PR) curve. The AUC has been adopted in many experiments to evaluate the performance of the classifier. However, the AP outperforms the AUC when the number of samples is unevenly distributed across different categories. This metric was selected for our experiments for the following reasons: In our samples, the number of images with traffic sign has a huge difference from that of images without traffic sign. In our experiments, the images were not classified by the labels (i.e. mandatory, dangerous, and prohibited) of the datasets. Instead, every ROI extracted in the previous step was allocated one of the two categories: "Yes" (the ROI contains traffic sign) and "No" (the ROI does not contain traffic sign), making traffic sign detection a binary categorization problem rather than a multi-categorization problem. The results (Table 1) show that the five algorithms had similar performance in traffic sign detection on the training set. However, the SPSVM outperformed the other algorithms on the test set. Besides, the Accuracy values of the SPSVM on the two datasets differed by less than 1%. This means the SPL indeed enhances the stimulus generalization ability of the SVM. 3.3 Traffic sign recognition The SPCNN code was programmed on MATLAB, and compared with IDSIA DNN, IDSIA MCDNN, and Multi-Scale CNN. The detection results were evaluated by three metrics: Accuracy, mean Area under Curve (mAUC), and mean Average Precision (mAP). According to the results of traffic sign recognition (Table 2), the proposed SPCNN achieved better performance than the contrastive methods on the test set. Therefore, the SPL training overcomes the overfitting problem, which is common to the DL algorithms, and improves the stimulus generalization ability of the CNN.

Table 1. The results of traffic sign detection

	Training				Test			
	Accuracy	False Rate	AUC	AP	Accuracy	False Rate	AUC	AP
ChnFtrs	**98.53**	3.21	0.9721	0.9272	94.32	9.21	0.9215	0.8927
3D TS	97.72	4.54	0.9613	0.9193	95.54	10.13	0.9223	0.8902
Bayesian	95.36	5.18	0.9549	0.9247	94.72	10.35	0.9107	0.8824
SVM	97.39	4.23	0.9687	0.9231	94.21	10.87	0.9114	0.8892
SPSVM	98.15	**2.73**	**0.9837**	**0.9612**	**97.17**	**5.53**	**0.9822**	**0.9531**

Table 2. The results of traffic sign recognition

	Training			Test		
	Accuracy	AUC	AP	Accuracy	AUC	AP
IDSIA DNN	98.52	0.9415	0.9012	88.27	0.8443	0.8093
IDSIA MCDNN	**99.46**	0.9512	0.9194	87.16	0.8310	0.8024
Multi-Scale CNN	99.17	0.9426	0.9163	88.92	0.8469	0.8102
SPCNN	99.35	**0.9748**	**0.9284**	**95.78**	**0.9353**	**0.9017**

IV. Conclusion

In traffic sign detection and recognition, the recognition rates of many algorithms vary greatly from dataset to dataset. To solve the problem, this paper introduces the SPL to train the MACHINE LEARNING algorithms, and verifies the stimulus generalization ability of the trained model through numerical experiments. The experimental results show that our model could achieve similar Accuracies on training set and test set. This means the SPL could reduce the overfitting problem, and enhance the stimulus generalization ability of the MACHINE LEARNING algorithms. Thus, the SPL is suitable for training traffic sign recognition models with a high requirement on precision. In this research, the SPL is innovatively applied to the detection and recognition of traffic signs. However, the datasets are not sufficiently large, the traffic sign images are not highly diversified, and the complex situations are rather limited. As a result, the obtained model cannot be directly applied to actual unmanned driving control systems. In the future research, more largescale datasets from different countries will be learned, and the scale of the CNN will be expanded, aiming to create a highly specific recognition model for industrial use. The SPL has not been extensively studied, when compared with immensely popular learning methods like reinforcement learning, active learning , and ensemble machine learning . Many researches fail to consider the SPL, because this learning strategy, only capable of enhancing the stimulus generalization ability of the target algorithm, performs poorly on the training set. However, many experiments have proved that the SPL helps to develop highly universal models for the MACHINE LEARNING algorithms, making the learning results are felicitous to real-world scenarios. As a result, it is of great value to promote the SPL.

References

[1]. Gao, H., Liu, C., Yu, Y., Li, B. (2014). Traffic signs recognition based on PCA-SIFT. In Proceeding of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, pp. 5070-5076. https://doi.org/10.1109/WCICA.2014.7053576

[2]. Zaklouta, F., Stanciulescu, B., Hamdoun, O. (2011). Traffic sign categorization using KD trees and random forests. In The 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, pp. 2151-2155. https://doi.org/10.1109/IJCNN.2011.6033494

[3]. Feng, G., Ma, L., Tan, X. (2014). Ground traffic signs recognition based on Zernike moments and SVM. In 2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference, Chongqing, China, pp. 478-481. https://doi.org/10.1109/ITAIC.2014.7065096 .

[4]. T. Ojala and M. Pietikinen and D. Harwood, “Performance evaluation of texture measures with categorization based on Kullback discrimination of distributions”, International Conference on Pattern Recognition, vol. 1, pp. 582-585, 1994.

[5]. Belongie S. and Malik J. and Puzicha J., “Shape matching and object recognition using shape contexts”, Transactions on Pattern Analysis and Machine Intelligence, 24(24), pp. 509-522, 2002.

[6]. Viola and Jones, “Robust Real-time Object Detection”, IJCV, 2001.

[7]. Viola and Jones, "Rapid object detection using a boosted cascade of simple features", Computer Vision and Pattern Recognition, 2001.

[8]. J. Greenhalgh and M. Mirmehdi, “Real-time detection and recognition of road traffic signs,” IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1498–1506, Dec. 2012.

[9]. Y. Yang, H. Luo, H. Xu, and F. Wu, “Towards real-time traffic sign detection and categorization,” IEEE Trans. Intell. Transp. Syst., vol. 17, no. 7, pp. 2022–2031, Jul. 2016.

[10]. Y. Wu, Y. Liu, J. Li, H. Liu, and X. Hu, “Traffic sign detection based on convolutional neural networks,” in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Aug. 2013, pp. 1–7.