Traffic Sign Recognition With R-CNN

Kalaga Sahitya
Department of Computer Science and Engineering

Karaka Jyoshna
Department of Computer Science and Engineering

Boddeda. Haritha Lakshmi
Assistant Professor
Department of Computer Science and Engineering
Vignan's Institute of Engineering For Women

K Satya Priya
Department of Computer Science and Engineering

Karanam Pooja
Department of Computer Science and Engineering

Abstract— **One of the important fields of Artificial Intelligence is Computer Vision. Computer Vision is the science of computers and software systems that can recognize and understand images and scenes. Traffic sign recognition system (TSRS) is a significant portion of intelligent transportation system (ITS). Being able to identify traffic signs accurately and effectively can improve the driving safety. This paper brings forward a traffic sign recognition technique on the strength of deep learning, which mainly aims at the recognition of circularsigns. The detected road traffic signs are classified based on deep learning. In this article, a traffic sign detection and identification method on account of the image processing is proposed, which is combined with convolutional neural network (CNN) to sort traffic signs. On account of its high recognition rate, CNN can be used to realize various computer vision tasks. TensorFlow is used to implement CNN. In the German data sets, we are able to identify the circular symbol with more than 98.2% accuracy.**

Keywords—Keras, OpenCV, Convolutional Neural Networks(CNN), Deep Learning, Traffic sign recognition.

I. Introduction

Traffic sign recognition system is a crucial research direction in computer vision and a significant section of Advanced Driver Assistance System (ADAS). It can be grouped into two technologies, traffic-sign detection and traffic-sign recognition. The correctness of detection will directly lead to the final identification results. Traffic signs contain necessary messages about vehicle safety and they show the latest traffic conditions, define road rights, forbid and allow some behaviors and driving routes, cue dangerous messages and so on. They can also help drivers identify the condition of the road, so as to determine the driving routes.

Traffic signs have some constant characteristics that can be used for detection and classification, among them, color and shape are important attributes that can help drivers obtain road information. The colors used in traffic signs in each country are almost similar, usually consisting of simple colors (red, blue, yellow, etc.) and fixed shapes (circles, triangles, rectangles, etc.). the image of traffic signs is often affected by some external factors, such as weather conditions. Therefore, traffic-sign recognition is a challenging subject and also a valuable subject in traffic engineering research. In this a variety of traffic-sign identification technologies have been developed. It's CNN based on transfer of learning method is put forward. Deep CNN is trained with big data set, and then effective regional convolutional neural network (RCNN) detection is obtained through a spot of standard traffic training examples.

In this project, a multi-resolution feature combination network fabric is come up with, which is able to study many useful features from diminutive - sized objects, also the traffic sign detection framework is divided into spatial sequence classification and regression tasks to obtain more information and improve the detection performance. For purpose of realize the real-time of CNN detection with recognition of traffic signs.

This text mainly realizes traffic-sign detection and identification through three parts: pre-processing, detection and classification, an image containing the area of interest is output, and the area of traffic signs is located. In the recognition and classification stage, the extracted 3 and segmented traffic sign area is used as input, and the convolutional neural network in deep learning is used to identify and classify the detected information.

II. Objective

In the world of Artificial Intelligence and advancement in technologies, many researchers and big companies like Tesla, Uber, Google, Mercedes-Benz, Toyota, Ford, Audi, etc. are working on autonomous vehicles and self-driving cars. So, for achieving accuracy in this technology, the vehicles should be able to interpret traffic signs and make decisions accordingly.

There are several different types of traffic signs like speed limits, no entry, traffic signals, turn left or right, children crossing, no passing of heavy vehicles, etc. Traffic signs classification is the process of identifying which class a traffic sign belongs to.

We will build a deep neural network model that can classify traffic signs present in the image into different categories. With this model, we are able to read and understand traffic signs which are a very important task for all autonomous vehicles.

Also we will build a custom graphical user interface which will allow anyone to upload an image and classify it into any of the traffic signs using our own saved neural network model.

III. Literature Survey

The last decade shows a growth evolution in the development of intelligent transportation systems (ITS) and especially Advanced Driver Assistance System (ADAS) and Self-Driving Cars (SDC). In these systems, traffic signs detection and recognition is one of the difficult tasks that confront researchers and developers. This issue is addressed as a problem of detecting, recognizing, and classifying objects (traffic signs) using computer vision and still be a challenge until now.

The work presented in this project focuses on traffic signs recognition without the consideration of the detection step. Traffic signs recognition is divided in two parts. In the first, several methods have been proposed, including edge detection, scale invariance feature (SIFT), speeded-up robust feature (SURF), Histogram of gradient (HOG) and others. Bag of Words exploiting SURF and k-means classifier was used. Typically, the output of this step is the input of the classification algorithms for the recognition of the road signs. Many algorithms have been used such as K-Nearest Neighbor (KNN) classifier, Support Vector Machine (SVM) and neural network for traffic signs classification. Many proposed the evaluation of three methods namely, Artificial Neural Network (ANN), Support Vector Machine (SVM) and Ensemble subSpace KNN using BoW where every road sign is encoded with 200 features. The Multi-layer Perceptron Neural network provides better results.

Currently, Convolutional networks are gradually replaced traditional computer vision algorithms for different applications such as object classification and pattern recognition. It is used for the extraction and the learning of depth description of the traffic signs. This solution overcomes the step of descriptors extraction which is very sensitive to different factors. This network takes 2D image and processes it with convolution operations. It has the ability to learn a representative description of image.

IV. TRAFFIC SIGNS DATASET

A rich dataset is needed in object recognition based on neural network in order to train the system and evaluate its results. For the purpose of traffic signs classification, we used the German Traffic Sign Recognition Benchmark (GTSRB) which contains 43 classes divided into 3 categories as represented in table I.

TABLE I
THE DATASET DISTRIBUTION

Category	Task	Number of images	Shape
Training data	Used to train the network	34799	4 dimensions tensor to determine the image index in the dataset, the pixel's row-column and the information it carries (Red Green Blue value)
Validation data	Allows to supervise the network performances while training it (a reduced version of testing data)	4410	
Testing data	Used to evaluate the final network	12630	

V. THE NEURAL NETWORK ARCHITECTURE

Using a fully connected neural network to make an image classification requires a large number of layers and neurons in the network, which increases the number of parameters leading the network to over-fitting (memorizing the training data only). The input image may also lose its pixels correlation properties since all neurons (carrying pixels values) are connected to each other.

Convolutional neural networks have emerged to solve these problems through their kernel filters to extract main features of the input image and then inject them into a fully connected network to define the class.

In this text, CNN is mainly used for classification. The most important part of the training and testing of CNN, in this article, the GTSRB data set is used to train and test CNN. CNN is a multi-layered network, it is very similar to brain. Each layer of CNN is composed of multiple neurons. Each neuron receives an input to perform a task, and some operations and outputs pass as input to the next neuron. Convolutional layer is the core part of convolutional neural network, and its main function is feature selection. Compared with traditional machine learning, CNN has good classification performance.

In this paper, CNN is used to classify the detected signs and a light-weight CNN classifier is designed. The light-weight CNN consists of two convolutional layers, two pooling layers and two full connection layers. In this text, the kernel size of the convolution layer is set as 5x5, the quantity of convolution kernel is set as 32, and the step size is set as 1.The quantity of hidden layer nodes in the first convolution layer is 16, and in the second convolution layer is 32. The size of feature graphs is 32x32 and 16x16 respectively. The kernel size of the pooling layer is 2x2, the hidden nodes of the full connection layer are 512 and 128, and the quantity of hidden nodes of the last output layer is 43.The initial value of the learning rate can be selected as a larger value to improve thetraining speed, or a smaller value to quicken the rate of convergence. The initial value of the learning rate in the text is set as 0.0001. In order to prevent the phenomenon of over- fitting in the network, the hidden layer of the full connection is Dropout (regularization) processing. During the training process, data of some nodes are randomly discarded to prevent over-fitting. Dropout set the node data to 0 to discard some eigenvalues. This process of feature extraction and classification by CNN is displayed in the Fig. 1.

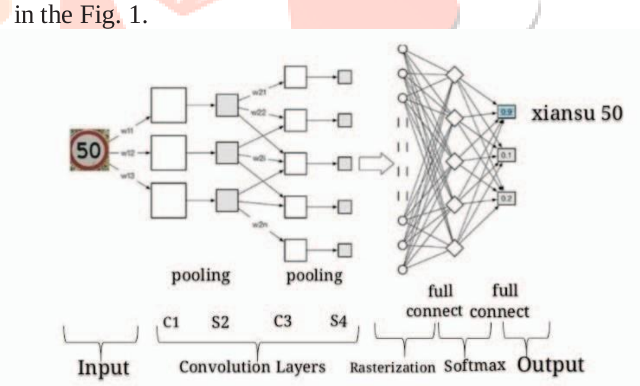

Fig.1 CNN feature extraction structure diagram

The model presented in this paper makes use of an open source advanced string stack provided by Keras, a neural network library that can run on TensorFlow. All the proposed CNN layers use ReLU as the activation function to learn complex features. ReLU is considered as the most commonly used activation function in NN, which can make the learning of network model faster and conducive to real-time learning. At the same time, in order to prevent over-fitting, Dropout layer is added to the CNN to suppress over-fitting. Dropout can increase the randomness of the network. GTSRB and self- designed traffic sign data sets were used for training and testing throughout the network. In neural network training, the dropout method is established to prevent from over-fitting by shutting down some neurons during the training phase.

VI. Implementation results

To built and train the network, the TensorFlow deep learning library is used. Training and testing were implemented using the dataset described in section IV and the developed method succeeds in classifying the 43 traffic signs classes.

The implementation results of the network and its improvement operations show the impact of each changed element. As represented the entire dataset is divided into two parts, 90% of which are training sets and the remaining 10% are test sets. When training the model, the accuracy of training increases gradually, and the loss rate decreases gradually with the increase of training times. The curve of accuracy and loss rate of training is shown in the fig.2.

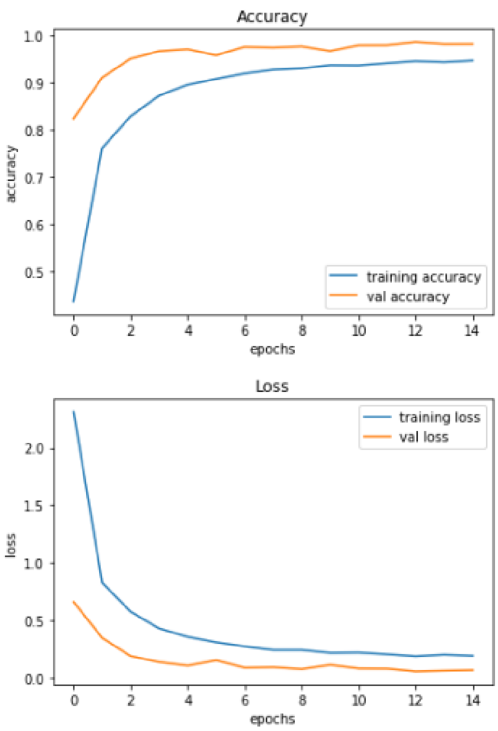

Fig.2 Accuracy and loss rate

The obtained results show the effectiveness of the developed method since a validation accuracy of 97,1% is achieved.

VII. Conclusion

We have successfully built an end-to-end neural network system that can automatically view an image and generate a reasonable description in plain English. It is based on a convolution neural network that encodes an image into a compact representation, followed by a recurrent neural network that generates a corresponding sentence.

The model is trained to maximize the likelihood of the sentence given the image. Experiments on several images show the robustness in terms of qualitative results (the generated sentences are very reasonable) and quantitative evaluation, using either ranking metrics or RELU, a metric used in machine translation to evaluate the quality of generated sentences.

It is clear that, as the size of the available datasets for image description increases, so will the performance of the proposed model. Furthermore, it will be interesting to see how one can use unsupervised data, both from images alone and text alone, to improve image description approaches.

VIII. Future enhancement

- A vaster dataset can help in improving the accuracy of the dataset and also will allow the more classes of the traffic signs to be classified.

- The integration with a real time camera system will allow our model to classify images of the traffic signs in real time and also can be implemented with a video system which will constantly classify the traffic signs and store the data and location of the certain signs for future reference.

Furthermore this can be used in more advanced artificial intelligence (AI) systems for cars and various vehicles to detect and act according to the different traffic signs.

IX. References

[1] RaimondaStaniute and Dmitrij Šešok: A Systematic Literature Review on Image Captioning (2019) Sauletekio.

[2] Marc Tanti, Albert Gatt and Kenneth P Camilleri: Where to put Image in an Image Caption Generator (2018) University of Malta.

[3] Marc Tanti, Albert Gatt and Kenneth P Camilleri: What is the Role of Recurrent Neural Networks in an Image Caption Generator? (2018) University of Malta.

[4] Jason Brownlee: Caption Generation with Inject and Merge Encode-Decoder Models (2017).

[5] Kaustubh: ResNet, AlexNet, VGGNet, Inception: Understanding Various Architectures of Convolutional Networks (2019) .

[6] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-JingZhu: BLEU: A method for Automatic Evaluation of Machine Translation (2002) IBM, New York, USA.

[7] Rafaella Bernardi, RuketCakici, Desmond Elliot: Automatic Description Generation from Images: A Survey of Models, Datasets, and Evaluation Measures (2017).