IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 2, February 2023

DOI: 10.17148/IJARCCE.2023.12255

Deep Learning For Traffic Sign Detection and Recognition

Prof. NAYANA S[1], RANJITHA R[2], SAMBRAMA LOKESH U[3]

Assistant Professor Of The Department, Computer Science, East West Institute of Technology, Bangalore, India[1]

Student, Computer Science, East West Institute of Technology, Bangalore, India[2-5]

Abstract: The automatic detection and recognition of traffic signs is crucial for managing the inventory of traffic signs. With the least amount of human effort, it offers an accurate and timely manner of monitoring traffic-sign inventory. In the realm of computer vision, the recognition and detection of traffic signals is a well-researched topic. In-depth driver-assistance and autonomous systems successfully use the vast majority of current technologies to understand traffic signs. It is yet unknown how well the remaining traffic signs will perform when used to replace manual labor-intensive traffic-sign inventory management because this only covers a small part of all traffic signs (about 50 categories out of several hundred). This study focuses on the difficulty of locating and categorising a large variety of categories for traffic signs that may be used to inventory management software. We employ a convolutional neural network (CNN) method dubbed the mask R-CNN to address the whole pipeline of detection and identification in autonomous end-to-end learning. We offer a number of suggestions that improve overall performance and are evaluated on the ability to recognise traffic signs. Using this technique, 200 traffic-sign types represented in our novel dataset are found. The results are reported for very challenging traffic-sign categories that haven't yet been included in prior studies. We provide a comprehensive analysis of the deep learning method for identifying traffic signs with significant intra-category appearance variation and show below 3% mistake rates.

Keywords: Traffic sign recognition, traffic sign detection, image processing, convolutional neural network

I. INTRODUCTION

Since the early 2000s, ConvNets have been utilised successfully for the identification, segmentation, and recognition of objects and regions in images. For each of these tasks, biological picture segmentation was necessary, especially for connectomics, as well as the recognition of faces, text, pedestrians, and human bodies in photographs taken in the real world. For each of these activities, relatively labelled data was available. ConvNets has experienced substantial recent practical success with face recognition. It is important to be able to categorise photos down to the pixel level since technology like self-driving automobiles and mobile robots will exploit this capability.

Businesses like Mobileye and NVIDIA will leverage these ConvNet-based methods in their next vision systems for autos. Other applications that are becoming more common include speech recognition and natural language understanding. Despite these successes, ConvNets were largely disregarded by the mainstream computer-vision and machine-learning sectors until the ImageNet competition in 2012. Using deep convolutional networks to a data set of over one million web-sourced photographs with 1,000 possible classifications resulted in astounding results, with error rates that were almost half that of the best competing techniques.

This success was made feasible by the efficient use of GPUs, ReLUs, a novel regularisation technique called dropout, and techniques to generate more training cases by distorted the existing ones. ConvNets are currently the dominating method for nearly all recognition and detection tasks and approach human performance on multiple tasks as a result of this achievement, which has spurred a revolution in computer vision. and ConvNets recurrent net modules are used in a recent astonishing demonstration for the production of picture captions. There are ten to twenty layers of ReLUs in current ConvNet topologies, hundreds of millions of weights, and billions of connections between units.

Two years ago, it might have taken weeks to train such massive networks; now, training takes only a few hours due of improvements in algorithm parallelization, software, and hardware. The efficiency of ConvNet-based vision systemsbased goods and services for comprehending images. ConvNets can be easily implemented in chips or field-programmable gate arrays as effective hardware. ConvNet chips are being developed by a number of firms, including NVIDIA, Mobileye, Intel, Qualcomm, and Samsung, to enable real-time vision applications in smartphones, cameras, robotics, and self-driving cars.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⋇ Impact Factor 7.918 ⋇ Vol. 12, Issue 2, February 2023

DOI: 10.17148/IJARCCE.2023.12255

II. LITERATURE REVIEW

A. Transportation inventory management systems must keep an accurate record of the quantity, kind, location, and condition of low-cost, high-quantity roadside assets like traffic signs. Although constantly updated databases like Google Street View contain street-level photos of all traffic signs, their potential for developing inventory databases has not yet been properly investigated. The main advantage of these databases is that, after traffic signs are identified, their geographic coordinates may also be computed and shown on the same interface. This study introduces a new system for compiling inventories of traffic signs using Google Street View pictures. Using the use of computer vision, traffic signs are identified and categorised into four groups: regulatory, warning, stop, and yield signs.from the Street View API of Google. The most likely location of each traffic sign is determined using the discriminative classification scores from all photos that contain a sign, and is then displayed on Google Maps using a dynamic heat map. Moreover, a data card is created with details on the position and description of each recognised traffic sign. The inventory of traffic signs may finally be managed more effectively thanks to a number of data mining interfaces that are introduced.

B. A roadway asset management system must include the detection, identification, and placing of road signs. In this study, a system based on stereo vision is created to automatically catalogue road signs. The system synchronises and integrates data streams from many sensors of high-resolution cameras, differential GPS, and other sources in real time. receivers, a measuring device for distance, and an inertial measurement unit. The positions of the moving vehicle and the orientation of the cameras are determined by algorithms that are generated based on data sets from the numerous positioning sensors. The main conclusions of the study concern feature extraction and analysis that are applied for automated sign detection and recognition in Right-of-Way (ROW) images, the implementation of a tracking algorithm of the candidate sign region among the image frames so that the same signs are not counted more than once in an image sequence, and the use of stereo vision technique to compute the world coordinates of the road sign from the stereopaired ROW images. Onboard the vehicle, certain methods are used to carry out all data collection and processing in real time. This programme is an efficient and advanced replacement for conventional inventory procedures.

C. One of the most crucial components of a highway asset management system is the regular evaluation and updating of the status of traffic signs and mile markers. The majority of current techniques involve manually gathering and analysing data, which must be done for millions of kilometres of roadways and must be done repeatedly. The use of video feeds obtained from cameras installed on cars has significantly improved the data collection process, but the processing has mostly remained a manual and labor-intensive procedure. Due to the interclass heterogeneity of traffic signs, anticipated variations in illumination, occlusion, sign position, and orientation, automating the analysis from the gathered videos is especially difficult. To clarify This work offers three computer vision algorithms for traffic sign recognition and classification in the presence of cluttered backgrounds and static and dynamic occlusions, and analyses their performance. The task focuses in particular on (1) extracting two-dimensional (2D) candidate windows from already collected video streams that may contain traffic signs-without making any assumptions about their locations-, (2) detecting signs in these 2D candidate windows, and (3) categorising them into warning, regulatory, stop, and yield sign categories based on their shape and colour. Around 11,000 annotated photos of U.S. traffic signs with a wide variety of posture, scale, backdrop, illumination, and occlusion variation are introduced as a new complete benchmark data set for validation. According to experimental findings, the average accuracy is For the methods of (1) Haar-like features with Cascade classifiers, (2) histograms of oriented gradients (HOG) with numerous one-versus-all support vector machine (SVM) classifiers, and (3) HOG+C with the SVM classifiers, a variation of the second method with histograms of colours concatenated to HOG, respectively, the results were 76.20%, 89.31%, and 94.83%. The experimental findings show the potential of combining SVM discriminative classifiers with joint representation of texture and colour in HOG+C as a practical method for compiling current and comprehensive inventory of traffic signs for U.S. highways.

III. PROPOSED SYSTEM

The two phases of the system that we suggest in this paper are detection and recognition. Simply said, an indication from the environment is found during the detection phase. When a car is travelling at a given speed, the camera records the scene, and our algorithm determines whether or not a sign is visible in that frame. The colour and shape of the traffic sign are used to identify it. The suggested algorithm categorises the detected sign during the recognition stage. An ensemble of convolutional neural networks is used to do this.

Detection

Before the detecting procedure begins, the image captured by the car's camera is preprocessed at this step. An HSV image is created by transforming the received RGB image during general preprocessing procedures. HSV (Hue Saturation

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 2, February 2023
DOI: 10.17148/IJARCCE.2023.12255

Value) colour space is recommended over RGB (Red, Green, and Blue) colour scheme for detection. When compared to an RGB image, HSV is closer to what the human eye actually sees. A HSV image has a wider spectrum of colours than an RGB image, which specifies colours in terms of the three main colours. A HSV image is also less sensitive to variations in ambient light. The HSV image is equalised, which changes the histogram of the image, to change the contrast in the picture. The next stages are taken after obtaining the HSV image. would be to recognise items by their colour, then determine their shape and confirm that the object is a traffic sign.

Color Based Detection

The colour of a sign is the first and most crucial aspect that we observe. We understand that the board at the side of the road is actually a traffic sign once we notice the colour red. We presented our method of detection using the same reasoning. The suggested system learns from the acquired frames to look for a sign based on the colour red. A sign is tested to see if it is there if a piece of the image falls below the necessary red threshold. The red part's outlines are discovered when the red threshold is confirmed.

Shape Based Detection

The contours found in the preceding step are used to compute the number of edges. The DouglasPeucker algorithm [4] is used for this. In this essay, we discuss the shapes of the triangle and the circle. The Douglas-Peucker algorithm first determines the number of edges, and then it determines the area of the contour. A triangle is deemed to be large enough if it has three edges and satisfies the minimal requirement for area. The contour is also regarded as a suitably large circle if the number of edges discovered is higher than or equal to 6 and the size of the contour satisfies the minimal requirement. Finding the bounding box is crucial after the shapes have been located. The Region of Interest (ROI) is separated from the surrounding environment by the bounding box. The outside triangle or circle of the detected contour is typically touched by the bounding box. A triangular sign has two triangles: the inner triangle, which is enclosed by the outer triangle but does not contact the bounding box, and the outer triangle, which touches the enclosing box.

Sign Validation

The sign needs to be verified after the bounding box has been located. The image that has had a threshold filter applied to it is first inverted, turning the inner triangle from black to white and the outer triangle from white to black. It is not a sign if the white triangle hits the boundary, and it is a traffic sign if it does not touch either the outer triangle or the boundary box. This is changed to the next step and enlarged to 48x48. A green outline is drawn around the discovered sign to show its location. The detected sign is marked with a green border, as can be seen in Fig. 2(b). OpenCV was utilised to carry out to perform steps above.

Recognition

The symbol must also be classified after it is discovered. A convolutional neural network can be created using TensorFlow, Google's open source machine learning framework. The initial step in our implementation of the recognition phase was to take the identified sign from the previous phase and run some straightforward image preprocessing on it. The most crucial aspect of this phase is the training and testing of the CNN. For training and testing, we made use of the Belgian Traffic Signs data collection and the German Traffic Sign Benchmark. In that it comprises neurons, which in turn have weights and biases, a convolutional neural network is extremely similar to the brain. Each neuron takes an input and then carries out an action. The next neuron receives the output as its input. The first layer of a convolutional neural network is always the input layer, while the last layer is always the output layer. A hidden layer is anything extra that lies between those two. A feed-forward network with six convolutional layers is suggested in this research. In order to avoid overfitting, this also features completely connected hidden layers with dropout layers in between.

The sequential stack offered by Keras, an open source high level neural network library that can operate on top of TensorFlow, is used by the model put out in this research. Rectified Linear Unit (ReLu) activation is present throughout all layers of the proposed CNN [5]. The most common activation function is thought to be ReLu activation. regarding neural networks. A fully connected layer receives the output of the sixth convolutional layer and applies a flatten function to At that moment, flatten the output. The final layer, which employs SoftMax activation, receives the output that has been flattened. To increase processing performance, a max pooling layer is also added after every two convolutional layers. We employ a group of CNNs rather than simply one. The final outcome is determined by the sum of our three CNNs. Compared to just one CNN, this results in a considerably more accurate conclusion. It is necessary to mention the model's loss, optimizer, and metrics. The loss of the model is estimated as a value between 0 and 1 rather than using percentages because it is configured to categorical cross entropy. Stochastic Gradient Descent with Nesterov momentum is used by the optimizer. flatten The accuracy metric is employed. Epochs with a backward pass are utilised to enhance the machine's training. Epochs raise the prediction's level of precision.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 2, February 2023
DOI: 10.17148/IJARCCE.2023.12255

IV. CONCLUSION

This article proposes a deep learning-based method for recognising traffic signs that focuses mostly on circular traffic signs. This approach can recognise and identify traffic signs by using picture preprocessing, traffic sign detection, recognition, and classification. According to test results, this method's accuracy is 98.2%.

REFERENCES

[1]. Department. Transp., 2011 Traffic Signs Image Database, London, U.K.
[2]. "Real-Time Detection and Recognition of Road Traffic Signals," IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 13, NO. 4, DECEMBER 2012. Jack Greenhalgh and Majid Mirmehdi.
[3]. A. Shustanov and P. Yakimov, "CNN Design for Real-Time Traffic Sign Recognition," ITNT-2017, third international conference on information technology and nanotechnology, 25–27 April 2017, Samara, Russia.
[4]. May 2017, Bangalore, India, pp. 476-480. [10] "Traffic Sign Detection- A Novel Method and Recognition Using Convolution Neural Network," IEEE Region 10 Humanitarian Technologies Conference, December 2017, Dhaka, Bangladesh, by P. Dhar, Md. Z. Abedin, T. Biswas, and A. Datta.
[5]. "Traffic sign identification and categorization using colour feature and neural network," International Conference on Intelligent Control Power and Instrumentation, 2016. Sheikh M. A. A., Kole A., and Maity T
[6]. "Traffic sign identification and recognition using fully convolutional network directed proposals," Neurocomputing, 2016:S092523121630741X by Zhu Y, Zhang C, Zhou D, et al