Intermediary report - traffic sign detection and recognition

1st Mihai Cosmin Cristea
1310A Group
RMA02_01

2nd Andrei Razvan Rautu
1310A Group
RMA02_01

***Abstract*—The traffic sign detection and recognition is necessary for the safety and proper navigation of drivers. Intelligent driver assistance systems have great potential in emerging technologies. This paper presents an efficient algorithm which detects the traffic sign from video based on colour and shape information. Then the auto associative neural networks are performed to recognise the traffic signs. This paper concentrates on red colour traffic signs. The algorithm is robust and able to detect the traffic signs with complex and randomly changing natural scenes. The algorithm is trained to find the best architecture for recognition. The experimental results shows the accuracy of detection and recognition of traffic signs.**

Keywords-Traffic sign recognition, Colour Segmentation, Auto associative neural networks(AANN).

I. Introduction

Every year nearly 1.24 million people die in road crashes. There is always a possibility that the driver may not notice the traffic sign and result in road disasters. Also many accidents may occur due to the mood fluctuations of the driver. Thus incorporation of ATSDR system in the vehicle ensures the safety and comfort of drivers by warning them about the presence of traffic signs in different ways. Traffic signs are designed with distinguishable features such as contrast color, shape, and pictogram, so that it can be easily noticed. Many factors influence the visibility and makes it unfavorable for recognition by the system. Prominent among them are [1],

(a) Lighting condition: It varies due to the change in weather conditions (example: rain, twilight, dust, etc.) and also the observed variations throughout the day, especially in the night when the beam light of vehicle falls on traffic sign.

(b) Occlusion: Traffic signs on urban road are more prone to partial occlusion which alter the shape of symbol.

(c) Rotation: Circular shift and distortion in camera projection produces in-plane and out-of-plane rotation respectively.

(d) Multi-class classification: Large number of classes makes classification more challenging.

(e) Cluttered signs and background: The horizontal or vertical arrangement of multiple traffic symbols and presence of unwanted background such as traffic light, static or moving vehicle, car indicator, etc., are to be considered.

(f) Poor image quality: It may be caused due to vibration of vehicle or poor camera quality.

(g) Distance between camera and traffic sign: Traffic sign varies with different scales depending on the distance between camera and traffic sign.

The objective of this work is to show the possibility of video processing algorithm capable of segmenting the traffic signs from images captured and recognise those using auto associative neural networks which is fast, reliable and able to notify drivers about the traffic signs.

This work contains:

2. State-of-the-Art
3. Related Work
4. Method description
5. Preliminary results
6. Preliminary conclusions

II. State-of-the-Art

Traffic-sign recognition algorithms are divided, in most cases, into two stages: 1) detection and 2) recognition.

In many works, the first block of the detection system consists of a segmentation stage by thresholding with a given color space to extract the sign color from the image. Direct thresholding over the red green blue (RGB) space is seldom used because it is very sensitive to lighting changes. Thus, in [4], a color ratio between the intensity of the specific RGB color components and the sum of intensity of RGB is used to detect strong colors in the image. A different relation between the RGB components is employed in [2], where one component is taken as a reference. Although other spaces are used, for example, the YUV system is considered in [3] to detect blue rectangular signs, the most frequently employed space is the hue saturation intensity (HSI) system because color information, which is encoded with hue and saturation components, presents low variations for objects of interest with a similar color. In [4], proper thresholds on hue and saturation bands are fixed to extract the red and blue colors. In [5], a nonlinear transformation over hue and saturation is employed to enhance the desired colors in the image (red and blue) using two lookup tables for every color for which we are looking. A similarity measurement between the hue component and the previously stored hue values of particular colors in road signs is calculated in [6], and this measurement is fed into a perceptual analyzer that is implemented by a neuronal network.

III. Related Work

Many benchmarks in the domain of driver assistance have been widely accepted once they were made available. In most cases, the contribution was not limited to publishing novel data, but included the definition of appropriate evaluation methodologies.

Traffic sign detection is a currently well-studied and broad field of research. The survey by Møgelmose et al. [7] provides detailed analysis on the most recent developments.

Most approaches make use of two prominent features of traffic signs: color and shape. Due to diverse natural lighting conditions the treatment of color is difficult and many heuristics have been applied [8], [9]. Regarding shape, one can say that two paradigms are currently persued: model-based and Viola-Jones-like methods.

The model-based approaches rely on robust edge detection and aim at connecting them to regular polygons or circles, usually via a Hough-like voting scheme or template matching. The Viola-Jones-like detectors compute a number of fast and robust features and try to identify trained patterns by use of different possibly weak classifiers.

Beyond our benchmark, there are publicly available datasets that are important to mention here: the Summer Swedish Traffic Signs dataset provides the remarkable amount of 20,000 images from video sequences of which 20% have been annotated by the authors. The MASTIF project has started to assemble data packages with traffic sign sequences every year since 2009 containing 1,000 to 6,000 images. The stereopolis database comprises 847 images with 273 road signs from 10 different classes.

However, these datasets consist of continuous video sequences, mostly recorded on a single tour and day. Accordingly, the same traffic sign instance will repeat several times in the dataset. More important, lighting conditions and driving scenario (rural, urban, highway) have little variance. We try to adress this issue by providing single images and presenting most traffic sign instances only once.

IV. Method description

Exploratory data analysis EDA

First we will familiarize ourselves with the data itself. The German sign data consists of many signs as shown below. We further expect each sign to be present only at relevant locations, therefore there is a difference in number of signs one would expect to see. In this data set the most common sign was the 20 kmph sign. We will not change the relative number of these signs because the incidence rates in the data reflects prior probability of observing a new sign. Leaving relative ratio of images unchanged biases the model towards predicting more frequent sign when the model is unsure between two signs.

Data augmentation and preprocessing

A big limitation of deep neural networks is that they may have millions of parameters, tuning which requires a vast data set. This however is not always possible. In such cases, data augmentation helps us generate additional training examples. We will generate additional data samples by applying affine transformation to the image. Affine transformations refer to transformations that do not alter the parallelism of lines, i.e. can be represented as a linear operation on the matrix. We will specifically use rotation, shearing and translation to simulate the effect of viewing the sign from different angles and different distances. Figure below presents original image and augmented images generated from it.

I applied the following preprocessing for image data,

1. I first applied histogram equalization so the effect of brightness is removed. I used openCV'2 cv2

Model architecture

Figure below presents the model architecture we will use. This architecture was converged upon after trying several different architectures. The first module in the model above is comprised of 3 1X1 filters. These filters have the effect of changing color maps. In most applications, changing color map can result in significant improvements in performance. However, it is not clear what the best color map is for different applications, therefore using 3 1X1 filters results in a case were the network itself ends up choosing the best color map.

The next 3 modules are composed of 32, 64 and 128 (respectively) 3X3 filters followed by maxpooling and dropouts. The output from each of the convolution module is fed into a feedforward layer. Rationale being that the fully connected layer has access to outputs from low level and higher level filters and has the ability to choose the features that works the best. The feedfoward layers are composed of 2 hidden layers with 1024 neurons in each layer. Additional dropout layers are applied after each of the fully connected layers.

Training

We will first start with large augmentation so the model learns overall features of traffic sign, and we will gradually reduce the augmentation to fine tune the model. The training is carried out in the following steps,

Generate 10 new images per image in the training set using data augmentation Split data into training and validation sets such that the validation set is 25After first 10 epochs, lower the augmentation by a factor of 0.9 per epoch.

Model performance

Once all the parameters were identified, the model took 1/2 an hour of training on Nvidia's Titan X.

2. I first scaled images between -.5 and .5, by dividing by 255. and subtracting .5. As there was limited data, and the number of images in each class were different, I generated additional data by jittering the images. For jittering, I rotated the images by random number generated between +/- 40 degress, shifted them by +/- 10 pixels along vertical and horizontal, and a final shearing transformation. After this transformation, the image get transformed as follows.

V. Preliminary results

	ClassId	SignName	Occurance
0	2	Speed limit (50km/h)	2250
1	1	Speed limit (30km/h)	2220
2	13	Yield	2160
3	12	Priority road	2100
4	38	Keep right	2070
5	10	No passing for vechiles over 3.5 metric tons	2010
6	4	Speed limit (70km/h)	1980
7	5	Speed limit (80km/h)	1860
8	25	Road work	1500
9	9	No passing	1470

```
data_pd_sorted.tail(10)
```

	ClassId	SignName	Occurance
33	39	Keep left	300
34	29	Bicycles crossing	270
35	24	Road narrows on the right	270
36	41	End of no passing	240
37	42	End of no passing by vechiles over 3.5 metric ...	240
38	32	End of all speed and passing limits	240
39	27	Pedestrians	240
40	37	Go straight or left	210
41	19	Dangerous curve to the left	210
42	0	Speed limit (20km/h)	210

VI. Preliminary conclusion

Traffic sign recognition is a challenging task and attracting many computer vision communities in recent years. Traffic hazards are mainly caused when the driver fails to notice the traffic sign. This paper proposes a novel automatic detection and recognition of traffic sign using multiple thresholding and subspace approach. The proposed system works effectively in both daylight and night illuminations. It is able to detect the shapes using NCC based correlation measure with recall and precision rate of 90 % and 97 % respectively.

In our future work, we intend to detect all variety of traffic symbols including white, green, yellow, and blue. Further, this could be combined with tracking algorithm to track the multiple traffic signs in different video sequences.

References

[1] Gudigar A, Chokkadi S, Raghavendra U (2016) A review on automatic detection and recognition of traffic sign. Multimed Tools Appl 75(1):333–364

[2] H. Kamada, S. Naoi and T. Gotoh, ”A compact navigation system using image processing and fuzzy control”, Proc. Southeastcon, vol. 1, pp. 337-342, 1990-Apr.

[3] A. de la Escalera, L. E. Moreno, M. A. Salichs and J. M. Armingol, ”Road traffic sign detection and classification”, IEEE Trans. Ind. Electron., vol. 44, no. 6, pp. 848-859, Dec. 1997.

[4] J. Miura, T. Kanda and Y. Shirai, ”An active vision system for real-time traffic sign recognition”, Proc. IEEE Intell. Transp. Syst., pp. 52-57, 2000-Oct.

[5] P. Arnoul, M. Viala, J. Guerin and M. Mergy, ”Traffic signs localisation for highways inventory from a video camera on board a moving collection van”, Proc. IEEE Intell. Veh. Symp., pp. 141-146, 1996-Sep.

[6] A. de la Escalera, J. M. Armingol, J. M. Pastor and F. J. Rodrguez, ”Visual sign information extraction and identification by deformable models for intelligent vehicles”, IEEE Trans. Intell. Transp. Syst., vol. 15, no. 2, pp. 57-68, Jun. 2004.

[7] A. Møgelmose, M. Trivedi, and T. Moeslund, ”Vision based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey, ” IEEE Transactions on Intelligent Transportation Systems, Special Issue on Machine Learning for Traffic Sign Recognition, 2012.

[8] S. Houben, ”A single target voting scheme for traffic sign detection, ” in Proceedings of the IEEE Intelligent Vehicles Symposium, 2011, pp. 124-129.

[9] A. Arlicot, B. Soheilian, and N. Paparoditis, ”Circular road sign extraction from street level images using colour, shape and texture database maps, ” in Proceedings of the Workshop on City Models, Roads and Traffic, 2009.

[10] R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[11] Duda RO, Hart PE, Stork DG (2000) Pattern classification, 2nd edn. Wiley-Interscience

[12] Cai D, He X, Zhou K, Han J Bao H (2007) Locality sensitive discriminant analysis. In: Proceedings of the 20th international joint conference on artifical intelligence. Hyderabad, India, pp 708–713

[13] Tenenbaum JB, Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science 290(5500):2319–2323