Fusion Model for Traffic Sign Detection, Tracking and Recognition

Shubham Dhingra[1], G.Saranya[2], Shalini Diwakar[3], Manish Kumar[4]

[2]*Assistant Professor,*
[1, 2, 3, 4]*Department of Information Technology, SRM Institute of Science and Technology, Chennai, India*
*Corresponding Author Email: [1]*shubham.dhingra27@gmail.com,* [2]*saranya.g@ktr.srmuniv.ac.in*
[3]*shalini44diwakar@yahoo.in,* [4]*manish.khanted007@gmail.com*

Abstract

A video-input traffic sign recognition is an advanced application which is a part of Intelligent Transport System(ITS) that provides information to the vehicles in order to make them safe and coordinated on road. The approach is to take a video as an input, divide it into series of frames and implement detection and recognition under mobile conditions. There are three major components: 1) detection; 2) recognition; 3) classification. We implement our technology on real data sets to obtain results in real-time manner.

Keywords: traffic sign, detection, tracking, recognition.

I. Introduction

Traffic signs are one of the most crucial parts of our daily life. Without them, our everyday travel would be as unrealistic as the phrase a castle in the air. As the proposed system is explained further in this paper, we hope to diversify into explaining the project's advanced features: 1) to improve detection performance by applying the spatial distribution prior to traffic signs; 2) enhance tracking staging and accuracy under mobile conditions; 3) scale-based intra-frame fusion is used to evaluate the data sets in comparable fashion.

Various methods are introduced stating that they have been able to make a difference in the vehicle displacement considered to be as Intelligent Transport System(ITS) [1], [2]. Detection and tracking have enormous components in acquainting the drivers with the information required to have a safe travel. In this paper, with recognition we provide them with accurate assistance for their navigation making their driving experience a great and much safer one and hence the approach can be implemented on various smart systems such as autonomous driving [3], advanced driver assistance systems(ADAS) [4], etc.

The recognition has been a major problem for early years but as we are advancing into a new intellectual and technological era, it is sufficing to say that this can be solved using machine learning methods which has tried to be taken care of with utmost precision. The researches in this field have acted as an incentive for this framework to be developed. A few previous work shows detection, tracking and recognition of real world images [5]. Our strategy uses Viola-Jones algorithm for the identification of the traffic signs as by certain defined parameters. KNN or the k-Nearest Neighbour algorithm is used for tracking the signs as it is used for both classification and predictive techniques which helps us evaluate the outputs at an ease.

The online and offline detectors to be used present an integrated picture of the whole traffic signs image profile thus helping us create an almost foolproof and secure environment for the transportation cause. The Kalman Filter algorithm has been cast-off for detecting signs and integrate output by fusion of the individual frames. The complete method that retrieves the traffic sign detections along with the recognition process is the core of this paper.

2. Related Work

Basically traffic sign acknowledgment is a multi-class classification issue that has turned into a genuine test for PC vision and machine learning methods. Traffic sign acknowledgment has been tended to by an extensive number of classification procedures, from straightforward layout coordinating, to advanced machine learning methods. Presently, the most generally utilised approach in traffic sign acknowledgment is the outstanding Support Vector Machine (SVM) [6], [7]. The framework was required to characterise hopeful blobs into a shape class before acknowledgment. As an arrangement, just the pixels that were a piece of the sign were utilised to build the element vector. A few extra techniques were displayed to enhance the precision of a programmed traffic sign location and acknowledgment framework and, in the meantime, to lessen the quantity of help vectors required and in this manner, diminish the memory and time prerequisites to test new examples.

Since the Traffic Sign Recognition (TSR) structure is basic and has magnificent application potential, a gigantic measure of researches have conveyed. There are two parts involved: detection and classification.

There are an extensive variety of strategies that can be used for the acknowledgment of signs. Since colour and condition of all the signs are similar, many shading division methods are grasped.

All of the procedures generally change over the RGB space [7]-[13] to spaces of other colours in order to diminish affected ability of light, after that usage shading thresholding or shading move up to evacuate areas of interest.

Other than these shape and colour strategies, data science methods have been progressively utilised. The methods used regard location to be a classifier, utilising SVM and course classifier. Viola-Jones detector is a decent decision for real- time TSR. Inspired by this, countless depend on the course of helped classifiers. These strategies more often than not centre around the outline of more illustrative highlights, for example, Aggregated Channel Features, and separated dipoles. But, if the initial data set isn't sufficiently extensive containing different onset transforms, it might fall flat in case of inconspicuous targeted traffic signs and foundations. However, a sufficiently huge data set is difficult to gather for real life application, hence it's constraint is self-evident.

3. Our Approach

A video is input, the offline trained detectors are utilised to distinguish between candidate sign and the noise first, then the outcome is visible which is the estimation of cumulative frames while moving. While preparing the consequent casings, the movement is refreshed and used to anticipate the tracking outcomes. In the interim the online test accumulation calculation will analyse the accuracy of the anticipated outcomes by the movement demonstration. On the off chance that confident, the following outcomes at that point will be utilised as online tests to prepare an online identifier by fusing the results. In a case it does not happen, the identifier will be used to distinguish between signs while the movement is consistent. Output of tracking and recognition is based on the movement of input. Alongside detection and tracking part, the tracked outcomes are combined to get a final traffic sign recognised output.

Here, our approach has been introduced in the following five phases: off-line detection, tracking, collecting samples, fusion detector, and recognition.

A. Off-Line Detectors

When a new video is input in the algorithm, a pre-trained detector helps in detecting the candidate traffic signs. Detector remains consistent throughout the framework. This can guarantee a steady character of the entire structure. Colour is an essential part while detecting an object [14]. Let us consider, traffic signs have consistent shape and colour, the detector is designed based on the Aggregated Channel Features [15].

B. Tracking

• A video based system provides more character information than the system based on images. The tracking results from Kalman Filter are fused with the results of detection for higher accuracy of localisation.

• Because of two following reasons, we use Kalman Filter: it is easy to implement and requires less space hence suitable for implementing on real life applications.
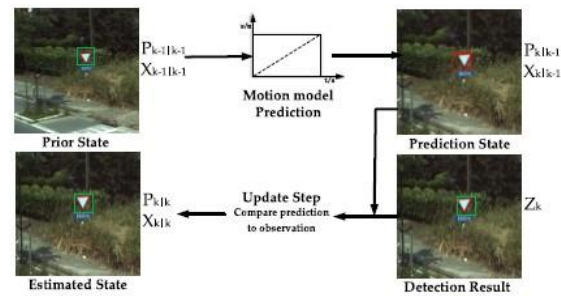
C. Collecting Samples

On-line learning is done for tracking of traffic sign under mobile conditions. There are two sides of traffic signs. First is the front side which is considered to be the positive and the back side is considered to be the negative side. To separate positive from negative, Region Of Interests (ROI's) are marked with separate variables. This is done to avoid noise while training the algorithm. In this, we fuse detected and tracked results to check if the candidate sign is genuine or not.

There are following three steps in collecting samples:

• The consistency in the shape. Traffic signs are normally composed of similar shapes, that is a critical prompt in identification.

• Second, difference in anticipated appearance against the mean of already settled target signs.

• Separation between areas of the applicant and already settled target.

D. Incremental Detector

By gathering sample online, we can obtain positive instances and negative instances by past outlines close to analysed edge. At that point a detector which has more merits based on the present circumstance will be prepared. The detector which has been prepared will be acquired for identifying traffic sign satisfactorily, and refresh the KF display.
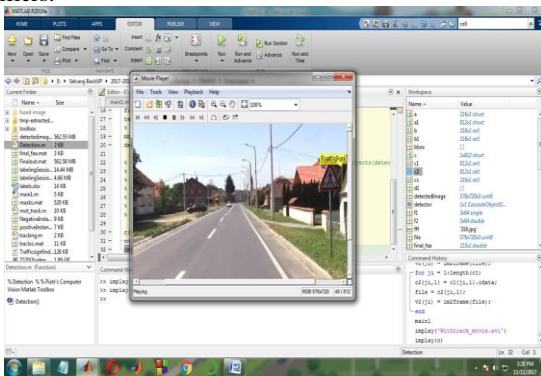
• Training: Incremental SVM is an essential part for learning online, and their performance is better while taking care of two-way classifiers issues for location assignment. For locating signs, an online SVM is created by fusing samples. Initially, SVM is instated by the offline samples. New coming sample is the estimation of KF regardless of whether it is true or not.

• Detection: Now the Kalman Filter have anticipated a state, the most encouraging traffic signs can be discovered considering the framework of motion model. For the online identification, the online model is utilised for restricting traffic signs in the vicinity of forecasted position, hence recognised outcome will be the finishing recognition result in the casing . Furthermore, this outcome is superior perception for refreshing KF parameters.

4. Recognition

SVM is used to recognise the tracked sign. We intertwine the aftereffects of numerous frames that have a place with the same physical sign together to improve accuracy. Considering the solid instinct that signs with bigger scale contains, we use Gaussian-based weighing capacity for melding the classification comes about at different levels with each other to get an outcome joining every one of the frames.
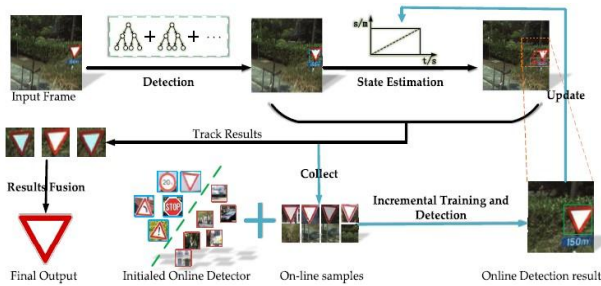
The initial part of our project is detection of traffic sign. For this, we import toolbox in Matlab which provides computer vision and video processing tools. We use Viola Jones for face detection, but we can custom train it for detection of other applications(whose aspect ratio doesn't vary by much, for example; traffic signs).

It constructs a "strong" classifier by combining several "weak" classifiers.

$$f(x) = Sgn\left(\sum_{i=1}^{N} \alpha i h i\right)$$

$$f_i(x) = \begin{cases} -s_i \\ s_j \end{cases} \quad \begin{array}{l} if \ f_i < \theta_i \\ otherwise \end{array}$$

5. Experimental Approach

Here, we will use the data set taken from various sources and implement evaluation techniques to analyse the accuracy of the proposed system.

A. Data Set

The implementation of the framework is carried out on the videos collected from various sources by mounting cameras on vehicles for continuous frames of traffic signs. There are plenty of data sets available to the public such as German TSR Benchmark (GTSRB) [16] and MASTIF data set [17], etc.

We use the MASTIF data of 2009, 2010 and 2011. Our algorithm is trained by the data of 2009 and data sets of 2010 and 2011 are used for the testing of accuracy of the algorithm.

B. Implementation

We use Matlab for the implementation of our project. Matlab was intended to disentangle the usage of numerical linear algebraic problems but it has since developed into something significantly greater and it is utilised to execute numerical calculations for an extensive variety of utilisation.

There are following modules involved:

1. Module 1

The threshold value θ_i and the polarity $s_i \in \pm 1$ are determined in the training, as well as the coefficients α_i.
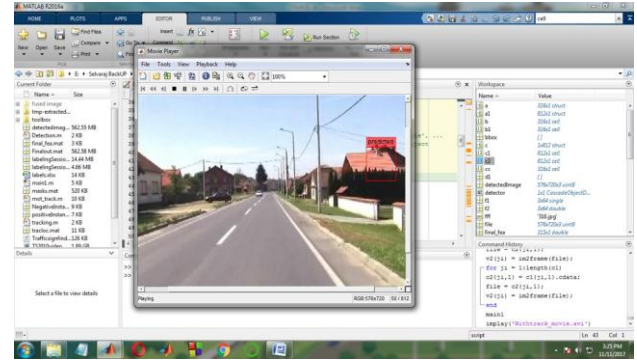
2. Module 2

The next part of the project is tracking of the detected sign. Once a road sign is detected, it is tracked from the next frame until it disappears. Kalman Filter is used for tracking , which is used to predict an object's future location.

The Kalman filter model assumes the true state at time k is evolved from the state at $(k - 1)$ according to

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k \text{ where,}$$

- F_k, the state-transition model;
- H_k, the observation model;
- Q_k, the covariance of the process noise;
- R_k, the covariance of the observation noise; and

sometimes

- B_k, the control-input model, for each time-step, k is described.

3. Module 3

To associate multiple physical objects with their corresponding tracks, k-nn algorithm is used to track and recognise the respective traffic sign.

- SVM is used to calculate constant b. Let the test set

6. Conclusion

We have introduced an efficient method of traffic sign detection, tracking and recognition by pre-training the algorithm and on-line updating it, which helps the predictor to predict the traffic sign using Kalman Filter. There are three key features. First is reducing negative instances using spatial distribution. Second is tracking using motion model and the be T, k-nn.

$T_{s\ v}$ set for SVM and K as the taken number for third is recognition of traffic sign using fusion model.

7. Future Enhancement
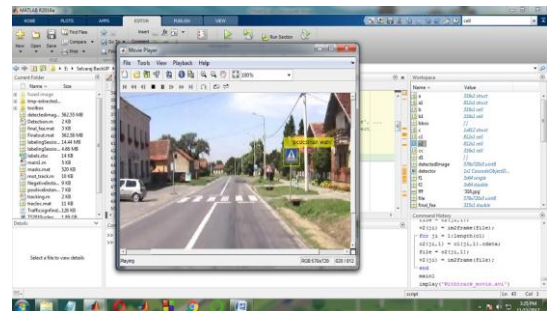
Future enhancement of this framework can be done by studying features and highlights of traffic signs for faster detection and recognition.

$$f(x) = sgn[\Sigma y_i a_i K(x_i . x) + b]$$

• If $|f(x)|$>threshold, make f(x) output directly, else transmit T_s $_v$, X and K to improve k_{-nn} algorithm classification.

8. Limitations

The proposed TSR framework imposes solid spatial appropriation for recognition. The dissemination helps earlier data to implement this framework for enhanced execution. In any case, it clampdowns algorithm for recognising genuine traffic signs which might not be the trained appropriation because of a few postures that vehicles exhibit while moving such as pivoting, coming down the hill, etc. Further confinement of this framework is that it needs horizon to be disappearing in the captured frame, that is normal for many situations.

9. Summary

Here, the evaluation of the whole system is carried out. For detection, prior training of the algorithm is done. While tracking, results of fusion are combined for better quality of state. Then we implement fusion strategy by combining several instances for offline training and then training the algorithm online. Hence, comparison of the detected and classified output is done to exhibit merits of the model. To sum up, this paper demonstrate the usefulness and effectiveness of the proposed framework.

Reference

[1] M. Da Lio et al., "Artificial co-drivers as a universal enabling technology for future intelligent vehicles and transportation systems," IEEE Trans. Intell. Transp. Syst., vol. 16, no. 1, pp. 244–263, Feb. 2015.

[2] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," IEEE Trans. Intell. Transp. Syst., vol. 12, no. 4, pp. 1624–1639, Dec. 2011.

[3] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W. V. Seelen, "An image processing system for driver assistance," Image Vis. Comput., vol. 18, no. 5, pp. 367–376, 2000.

[4] R. Timofte, V. A. Prisacariu, L. J. V. Gool, and I. Reid, "Combining traffic sign detection with 3D tracking towards better driver assistance," in Emerging Topics in Computer Vision and its Applications, E. C. H. Chen, Ed. Singapore: World Scientific, 2011.

[5] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-sign detection and classification in the wild," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 2110–2118.

[6] F. Zaklouta and B. Stanciulescu, "Real-time traffic sign recognition in three stages," Robot. Auto. Syst., vol. 62, no. 1, pp. 16–24, 2014.

[7] 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 Trans. Intell. Transp. Syst., vol. 8, no. 2, pp. 264–278, Jun. 2007.

[8] R. Timofte, K. Zimmermann, and L. J. V. Gool, "Multi- view traffic sign detection, recognition, and 3D localisation," in Proc. IEEE Workshop Appl. Comput. Vis., Dec. 2009, pp. 1–8.

[9] X. W. Gao, K. Hong, P. Passmore, L. Podladchikova, and D. Shaposhnikov, "Colour vision model-based approach for segmentation of traffic signs," J. Image Video Process. (EURASIP), vol. 2008, no. 1, pp. 1–7, 2008.

[10] X. W. Gao, L. Podladchikova, D. Shaposhnikov, K. Hong, and N. Shevtsova, "Recognition of traffic signs based on their colour and shape features extracted using human vision models," J. Vis. Commun. Image Represent., vol. 17, no. 4, pp. 675–685, 2006.

[11] S. Lafuente-Arroyo, S. Salcedo-Sanz, S. Maldonado- Bascón, J. A. Portilla-Figueras, and R. J. López-Sastre, "A decision support system for the automatic management of keep-clear signs based on support vector machines and geographic information systems," Expert Syst. Appl., vol. 37, no. 1, pp. 767–773, 2010.

[12] P. Gil-Jiménez, S. Maldonado-Bascón, H. Gómez- Moreno, S. Lafuente-Arroyo, and F. López-Ferreras, "Traffic sign shape classification and localization based on the normalized FFT of the signature of blobs and 2D homographies," Signal Process., vol. 88, no. 12, pp.2943–2955, 2008.

[13] V. A. Prisacariu, R. Timofte, K. Zimmermann, I. Reid, and L. V. Gool, "Integrating object detection with 3D tracking towards a better driver assistance system," in Proc. Int. Conf. Pattern Recognit., 2010, pp. 3344–3347.

[14] F. S. Khan, R. M. Anwer, J. van de Weijer, A. D. Bagdanov, M. Vanrell, and A. M. Lopez, "Color attributes for object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 3306–3313.

[15] P. Dollár, R. Appel, S. Belongie, and P. Perona, "Fast feature pyramids for object detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 8, pp. 1532–1545, Aug. 2014.

[16] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The german traffic sign detection benchmark," in Proc. Int. Joint Conf. Neural Netw., Aug. 2013, pp. 1–8.

[17] S. Segvi´c, K. Brki´c, Z. Kalafati´c, and A. Pinz, "Exploiting temporal and spatial constraints in traffic sign detection from a moving vehicle," Mach. Vis. Appl., vol. 25, no. 3, pp. 649–665, 2014.