

# TSD-CPI: Traffic Sign Detection Technique Based on Centroid Position Identification in Text Mining



R. Karthika, S. Murugan

**Abstract:** Detecting and Identifying traffic sign is a complicated issue due to the changing variability in cloud conditions. Hence, it is necessary to identify and detect of traffic signs during journey. The traffic text sign identification fails due to noise, blur, distortion and occlusion. In order to identify the text, a technique should be adapted that recognizes the text with improved accuracy. In existing algorithms such as Histogram of Oriented Gradients (HOG) and Support Vector Machine (SVM) were not detecting the Centroid position. In this paper, the text Centroid of position sign is detected using text color, font and size. During journey, if the text is blurred, this Traffic Sign Detection Technique based on Centroid Position Identification (TSD-CPI) K-means algorithm for clustering is possible to use. As a result, it detects the text that with improved accuracy. Ultimately, it reduces the processing time. The experimental result reveals that using WEKA-3.8 with the proposed technique shows improvement over the existing algorithms in terms of precision and Recall which enhance the accuracy in text mining.

**Keywords:** Histogram, Gradients, Support Vector Machine, Centroid and K-Means.

## I. INTRODUCTION

Traffic sign detection is a prominent topic in text mining and Intelligent Transportation System (ITS). It is used for intimation about the present road condition and other important information [1]. They are simple attractive text-based colors and it is easy to read the information they carry. However, accidents occur when drivers don't pay attention. Hence, an automatic driver classification is necessary to identify and classify the symbols. In precedent few decades, the research group and IT industry have attained progress in detecting the signs. Image based traffic sign may be low, blur and occlusion, because of low pixel resolution. Image traffic signs are also poor due to environmental conditions such as weather, sign color fading, etc. Detection of Traffic sign and identification is still a problem. To overcome such issues, an efficient technique is proposed on the basis of Centroid position identification.

Revised Manuscript Received on December 30, 2019.

\* Correspondence Author

**R. Karthika**, Research Scholar, Department of Computer Science, Memorial College (Autonomous), Puthanampatti-621007 (Tamil Nadu) India. E-mail: r.karthi1711@gmail.com

**S. Murugan**, Associate Professor, Department of Computer Science, Memorial College (Autonomous), Puthanampatti-621007 (Tamil Nadu) India. E-mail: murugan\_nmc@hotmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The proposed Traffic Sign Detection based on Centroid Position Identification (TSD-CPI) classifies the traffic sign text using color TSD-CPI based segmentation method to identify traffic sign regions. Ultimately, this reduces the searching delay and improves the accuracy.

This work consists of two parts. One is the segmentation of traffic sign and the other is the classification using K-means clustering algorithm. The texts are extracted by segmentation method from various regions. Eventually, traffic texts are identified, and outliers get eliminated by clustering algorithm. In the existing algorithms such as HoG and SVM were not suitable for detecting and recognizing Centroid position of the text. Figure 1 show various traffic sign images with Centroid positions [2]



**Figure 1: Sample images for Traffic Signs with centroid positions**

The method proposed here is evaluation based on traffic sign dataset. The result of the experiment demonstrates the proposed technique has higher detection rate than other algorithms. In addition, the proposed algorithm reduces the processing time and enhances the accuracy. The content in the other parts are as follows: Section 2 places of interest the connected installation. Section 3 explains the technique TSD-CPI that is proposed and computes the processing time that enhances the accuracy. Section 4 displays the result and discussion. Section 5 draws the conclusion.

## II. REVIEW OF LITERATURE

Jose et al., [1] invented Road recognition depends on illuminate invariance. The author used Hue-Saturation-Intensity (HSI) algorithm and utilized the parameters such as shape, size.

As a result, they reduced the false detection and image region. Yuwan et al., [2] designed an incremental framework traffic detection tracking and recognition algorithm. The contractual information and special distribution were used and utilized parameters such as shape, colours, and pictogram. Hence the false positive detection was reduced. Jack et al., [3] developed automatic detection of traffic sign. In this paper, the author used Optical Character Recognition (OCR). They have used parameters such as size, shape and colours were used. It reduced the storage information from scene and total number of false positives. David et al., [4] examined with colour-based method shape - arc - estimation and SVM. The author used Support Vector Machine (SVM) for better performance. They used attributes such as pixel and RGB colours. They have improved the improved level of accuracy. Creusen et al., [5] developed the exploitation of colour in the Hog-based traffic sign detection. The author improved detection of HoG based on color segmentation. They used SVM for precise vision of traffic sign. They used the attributes such as color, pixel and grayscale image. The utilization of background information was increased and reduced the memory computation. Claw Bahlmann et al., [6] proposed a traffic sign detection system, to Track, and recognize using Color, Shape, and Motion Information. They proposed an algorithm called AdaBoost. They improved the accuracy in terms of detection and recognition per second. Yingying Zhu et al., [7] developed segmentation for detecting networks for text-based detection. The attributes such as height width and size were used. The work enhanced text detection by increasing the efficiency and accuracy. Also, it reduced the area of search and removed outside texts. Ellahyani et al., [8] explored detection system and recognition with combination features and forests randomly. The forests classification was for detecting circular, triangular and rectangular shapes of the images that were segmented. This method identified the information within the detected signs. They used HoG as main criteria for detecting traffic sign and used the parameters such as colour, shape, and pictogram. Hence this method has reduced the search space based on colour segmentation which yields 96% classification rate. Nadra Ben Romdhanem et al., [9] progressed the gratitude and tracking technique for driver support structure. The method illuminated colour-based segmentation for generating the candidate regions. HoG features were extracted for encoding, to generate the feature vector. The optical flow ensures continuous capture of the recognized traffic sign for executing time based on SVM and HOG acceleration. Finally, the method showed good rate of recognition even in threatening weather conditions and changing lights. Miguel Angel et al., [10] examined the fast detection of traffic signs under various lights. Hough transform was used to detect edges of the image. Then, it was classified using a neural network. They used the techniques such as Connected-Components (CCC) and Kalman filter with the help of image pixel and colors. They improved accuracy and recognized the speed limit. Jin Zhao et al., [11] proposed the system for detecting real-time signs using Speeded up Robust Features (SURF) and Features on Field Programmable Gate Array (FPGA). This work was developed by video-based traffic sign detection of driver-assistance systems. The author implemented the work

by using colour and shape attributes. As a result, it reduced the computational complication, equivalent construction and statistics stream systems. Carlos Filipe Paulo et al., [12] explored pictogram contours for recognition of the signs of traffic. This was detected based on colours, shape, information and special cases. Traffic signs were recognized based on the pictograms. They used image processing algorithm for implementation. Hasan Fleyeh [13] by a fuzzy approach, detected road and traffic sign color and segmented them. It was implemented based on HSV of all pixel in the color break. This same rule was used to extract colors of road signs. Safat et al., [14] designed the matching shapes and segmented the colors based on the system of detection of traffic signs. The automatic traffic sign detection system by an imaging sensor and assisted the driver to properly operate the vehicle was proposed. The result showed as higher accuracy of traffic sign detection rate lower the computational time.

### III. TSD-CPI: METHOD

The segmentation of the traffic sign texts detection is proposed using k-means clustering. It is a group of objects where pixels of a group can be defined by the same relationship. It is a classification technique which is unsupervised, as it automatically classifies the objects based on the text input data. K here describes clustering algorithm to segregate the texts. This segmentation is used to detect signs from the text-based images. The methodology for the aforesaid technique for traffic sign uncovering is illustrated with Figure 2. Segmentation of the traffic sign images are considered as input, where this preprocessing stage of the image is converted to color. If there is any noise that will be removed using filtering technique. After preprocessing, centric position of the image is captured using clustering. The segmentation and traffic image are detected. Morphological filtering is performed to enhance the text of the detected portion.

#### Algorithm Begin

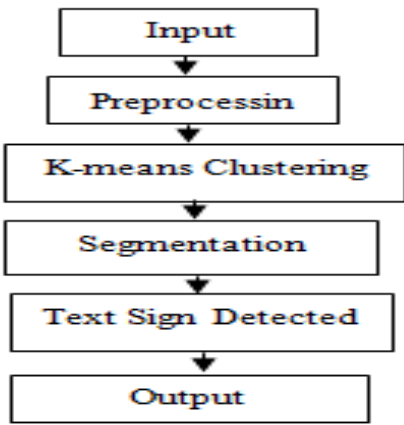
1. Let  $X_1, \dots, X_n$  are data points in the input image, let  $k$  be the number of clusters it is given by the user
2. Choose  $C_1, \dots, C_k$  cluster centers
3. Find the distance between each pixel and each cluster center
4. The distance function is given by  

$$J = |X_i - C_j| \text{ for } i=1, \dots, N \text{ and for } j=1, \dots, k, \text{ where } |X_i - C_j|$$
5. Distribute the data points  $x$  among the  $k$  clusters using the relation  

$$X \in C_j \text{ if } |X - C_j| < |X - C_i| \text{ for } i=1, 2, \dots, k, i \neq j, \text{ where } \text{denotes the set of data points whose cluster centre is } C_j$$
6. Repeat from Step 1 to Step 5 till convergence is met.

End

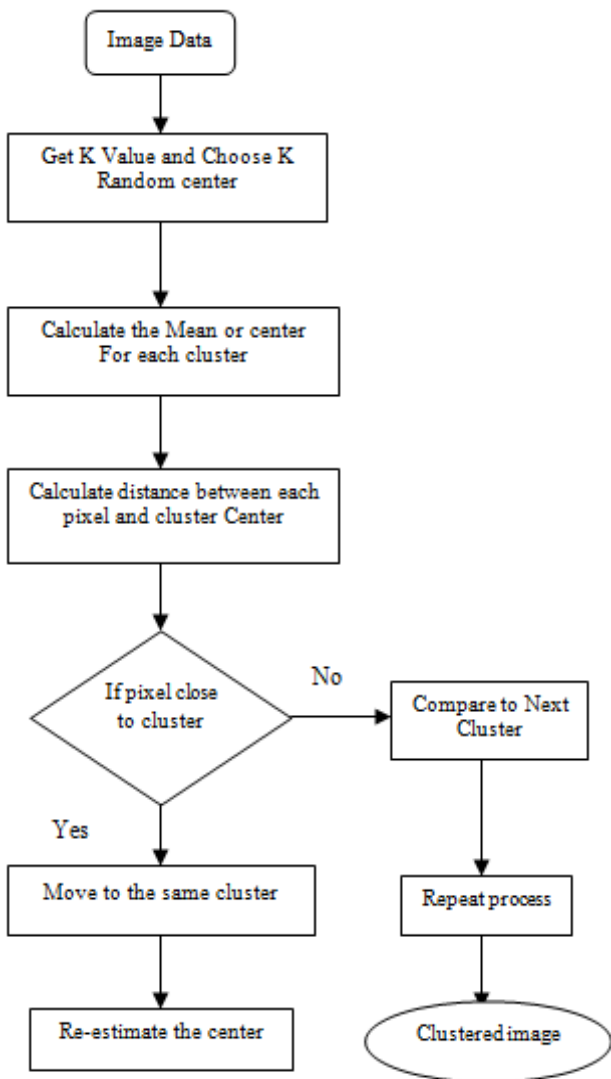




**Figure 2: The Methodology of the Proposed TSD-CPI Technique**

**3.1 K-Means Clustering**

K number of objects is selected at random in this algorithm, which earlier represented cluster mean or centre. Based on entity with this group mean’s distance, group is assigned an object, which computes a new mean for all group. This method iterates until the measure task congregates as depicted in Figure 3.



**Figure 3: Flowchart for the Proposed TSD-CPI Technique**

**3.2 Segmentation**

This is the partitioning process of a digital image into manifold sections (sets of pixels, also known as super pixels). It aims to modify the image to make it more meaningful and easily understandable. It can be used for locating objects and boundaries.

The aim of the segmentation of traffic sign is to:

To Identify Region of Interest i.e. locate text sign, blur and other distortion

To measure text color, font and size

Let the feature vectors derived from l clustered data be  $X = (X_i | i=1, 2, \dots, l)$ . The generalized architecture instigates k cluster centroids  $C = (C_j | j=1, 2, \dots, k)$  by randomly selecting k feature vectors from X. Later, the feature vectors are grouped into k clusters. The residue is calculated using Equation (1).  $D = |X_i - C_j|$  (1)

The following step is to group the centroids based on the group members, and then include them in their new group. This procedure comes to a stop after the convergence of all the cluster centroids.

**3.3 Traffic Sign Filtering (TSF)**

This step assesses the shapes and structure more sensibly and the channels of filter are organized from the central operation. The image is then sorted as segment and illumination images. The centre pixel represents the origin of the sorting out part. The sorting out part judges the fitment of the of the image, when it passes over the zone of the information picture. In case of issues in fitment, the geometric characteristics of the image are hidden from the neighborhood. Deterioration and extension are the major filters, where image visibility decreases during break down and thickens the articles in the image. Every single piece of the pixel in the neighborhood information is used by extension and the most diminished part is used by the breakdown.

**IV. SIMULATION RESULTS**

The proposed TSD-CPI has been simulated in WEKA tool 3.8. The Centroid position and identification of traffic sign on SVM, HoG and the proposed method TSD-CPI are shown in Table 1.

**Table 1: Centroid Position and Identification of Traffic Sign**

| Detected as a Sign | Conversion Result |
|--------------------|-------------------|
|                    |                   |
|                    |                   |
|                    |                   |
|                    |                   |
|                    |                   |

Performance Metrics

Performance metrics are used to evaluate the presentation of the suggested architecture. The TSD-CPI is contrasted with SVM and HoG based on precision and recall as shown in Table 2.

**Table 2: Traffic Sign Dataset Evaluation for the Proposed and Existing Algorithms**

| Sign Name           | SVM based |        | HoG based |        | Proposed TSD-CPI |        |
|---------------------|-----------|--------|-----------|--------|------------------|--------|
|                     | Precision | Recall | Precision | Recall | Precision        | Recall |
| Pedestrian Crossing | 87.9      | 87.2   | 91.03     | 93.45  | 96.03            | 98.52  |
| Right Side (Turn)   | 93.8      | 95.33  | 96.53     | 97.77  | 98.82            | 99.17  |
| No Stopping         | 66.8      | 71.7   | 77.27     | 81.46  | 97.14            | 99.2   |
| Speed Limit         | 76.12     | 77.27  | 80.56     | 90.48  | 98.45            | 99.76  |
| School Zone         | 47.76     | 79.68  | 71.5      | 89.4   | 95.7             | 97.89  |

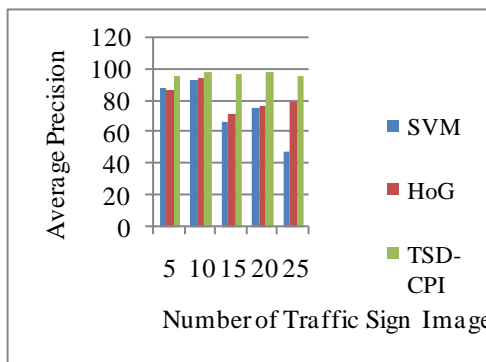
Precision

It is the portion of occurrences that are relevant among all of the retrieved instances. Precision is calculated as tracks:

$$\text{Precision} = \frac{\text{No. of Correctly Detcted Signs}}{\text{No. of Detected Signs}}$$

**V. EXPERIMENTAL RESULTS AND ANALYSIS**

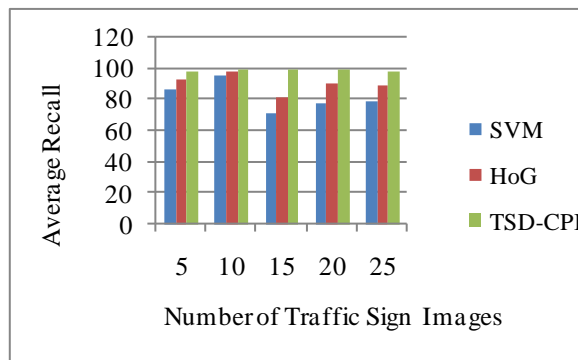
The average precision comparison of the proposed technique TSD-CPI with SVM and HoG under various number of traffic sign images as shown in Figure 4.



**Figure 4: Result of Average Precision for Various Traffic Sign Images**

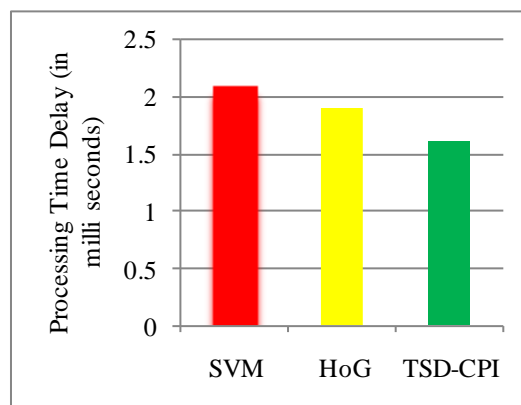
Precision detection rate is high in TSD-CPI compared to SVM and HoG even as the number of sign images is increased. Whereas, the processing time delay is less in TSD-CPI due to the centroid position identification which improves the accuracy. Likewise, recall can be calculated as follows. The performance of the TSD-CPI average recall under various number of traffic sign images as shown in Figure 5.

$$\text{Recall} = \frac{\text{No. of Correctly Detcted Signs}}{\text{No. of True Signs}}$$



**Figure 5: Result of Average Recall Various Traffic Sign Images**

Figure 6 depicts the proposed TSD-CPI outperforming existing algorithms with respect to processing time delay. The X axis represents techniques which have been compared and Y axis represents the processing time delay in milliseconds.



**Figure 6: Processing Time Delay among SVM, HoG and TSD-CPI**

Figure 7 depicts the proposed TSD-CPI outperforming existing algorithms with respect to accuracy. The X axis represents techniques which have been compared and Y axis represents the accuracy. Accuracy can be calculated as follows:

$$\text{Accuracy} = \frac{\text{No. of Correctly Detcted Signs}}{\text{Total No. of Traffic Signs}}$$

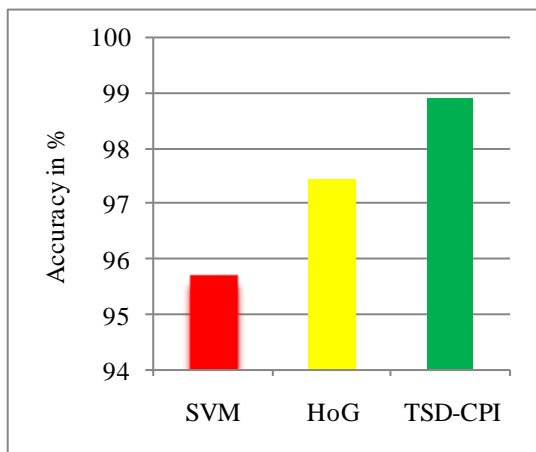


Figure 7: Accuracy among SVM, HoG and TSD-CPI

## VI. CONCLUSION

TSD-CPI is proposed for detection and gratitude of the traffic signs Centroid position. The central involvements of this paper are TSD-CPI based on k-means clustering algorithm, for detecting signs of traffic, thereby reducing the search area and ensuring accuracy. Experimental results demonstrate that the suggested TSD-CPI accomplishes in elevated accuracy with minimum processing time delay. In future, it can be extended to generate and develop real-time system for signs of traffic based on TSD-CPI.

## REFERENCES

1. José M. Álvarez and Antonio M. Lopez, "Road Detection Based on Illuminant Invariance" IEEE Transactions on Intelligent Transportation Systems, Vol. 12, No. 1, 2011, pp. 184-193.
2. Yuan Yuan, Zhitong Xiong, and Qi Wang, "An Incremental Framework for Video-Based Traffic Sign Detection, Tracking, and Recognition" IEEE Transactions On Intelligent Transportation Systems, Vol. 18, No. 7, 2017, pp.1918-1929.
3. Jack Greenhalgh and Majid Mirmehdi "Recognizing Text-Based Traffic Signs" IEEE Transactions On Intelligent Transportation Systems, pp.1-10.
4. David Soendoro, Iping Supriana, "Traffic Sign Recognition with Color-based Method, Shape-arc Estimation and SVM ", International Conference On Electrical Engineering And Informatics, 2011,pp. 1-6.
5. I.M. Creusen, R.G.J. Wijnhoven, E. Herbschleb, "Color Exploitation In Hog-Based Traffic Sign Detection", IEEE 17th International Conference on Image Processing , pp.1-4.
6. Claw Bahlmann, Ying Zhu, Visvanathan Ramesh, Martin Pellkofert , Thorstea Koehled, "A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information", IEEE Xplore, pp.255-260.
7. Yingying Zhu , Minghui Liao , Mingkun Yang , and Wenyu Liu , "Senior Member Cascaded Segmentation-Detection Networks for Text-Based Traffic Sign Detection", IEEE Transactions On Intelligent Transportation Systems, Vol. 19, No. 1,2018,pp.209-219.
8. Ayoub Ellahyani Mohamed EL Ansari, Ilyas EL Jaafari " Traffic Sign Detection and Recognition using Features Combination and Random Forests", International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, 2016, pp.1-8.
9. Dhanush V, Mahendra A R, Kumudavalli MV,Samanta D , Application of deep learning technique for automatic data exchange with Air-Gapped Systems and its Security Concerns, Proc. of IEEE International Conference on Computing Methodologies and Communication , 18-19, July 2017 , Erode.
10. R Kumar, Rishabh K,Samanta D , M Paul, CM Vijaya Kumar, A Combining approach using DFT and FIR filter to enhance Impulse response, Proc. of IEEE International Conference on Computing Methodologies and Communication , 18-19, July 2017 , Erode.
11. G Ghosh, Samanta D, M Paul, N Kumar Janghel, Hiding Based Message Communication techniques depends on Divide and Conquer Approach, Proc. of IEEE International Conference on Computing Methodologies and Communication, 18-19, July 2017, Erode.

12. R. K. Singh, T. Begum, L. Borah, Samanta D, Text Encryption: Character Jumbling, Proc. of IEEE International Conference on Inventive Systems and Control @IEEE, 19-20 January 2017, Coimbatore.
13. Nadra Ben Romdhane, Hazar Mliki, Mohamed Hammamim,"An Improved Traffic Signs Recognition and Tracking Method for Driver Assistance System" IEEE Xplore , pp.1-6
14. Miguel Angel Garcia-Garrido, Miguel Angel Sotelo and Ernesto Martin-Gorostiza, " Fast Traffic Sign Detection and Recognition Under Changing Lighting Conditions", IEEE Intelligent Transportation Systems Conference Toronto, Canada, 17-20, 2006,pp.811-816.
15. Jin Zhao, Sichao Zhu, and Xinming Huang, "Real-Time Traffic Sign Detection Using SURF Features on FPGA",IEEE Xplore,pp.1-6.
16. Carlos Filipe Paulo, Paulo Lobato Correia, "Traffic Sign Recognition Based on Pictogram Contours", IEEE Xplore,pp.67-70.
17. Hasan Fleyeh,"Color Detection and Segmentation For Road and Traffic Signs", IEEE Xplore, pp. 809-814
18. Safat B. Wali, Mahammad A. Hannan and Shahrum Abdullah, Aini Hussain, Salina A. Samad, "Shape Matching and Color Segmentation Based Traffic Sign Detection System", pp. 36-40.

## AUTHORS PROFILE



**R. Karthika**, Received her Master of Philosophy degree in computer Science from Bharathidasan University,Tiruchirappalli,Tamilnadu,India in 2014.She has been pursuing Doctor of Philosophy in the field of Text Mining. Currently working as a Guest Lecturer of Computer Science at Government Arts College, Kulithalai-639 120.Karur-Dt. Her Researches interest Data Mining, Image Processing.



**Dr. S. Murugan** is working as an Associate Professor of PG & Research Department of Computer Science at Nehru Memorial College, Puthanampatti-621 007.Tiruchirappalli-Dt. He has Published many articles in various reputed journals.His research interest include Data Mining, Cryptography, Image Processing.