Volume 3, Issue IX, September 2024, No. 44, pp. 573-584
Submitted 22/08/2024; Final peer review 15/9/2024
Online Publication 18/9/2024
Available Online at http://www.ijortacs.com

THE USE OF MACHINE LEARNING MODEL FOR AUTOMATIC DETECTION OF ROAD TRAFFIC SIGNS

*[1]Oji Nkechi Blessing
Department of Computer Science, Federal University of Technology Owerri (FUTO), Imo State, Nigeria
Corresponding Author Email: nkechi.oji@futo.edu.ng
Tel: +2348035057652

Abstract

This study introduces an improved method for traffic sign detection and recognition, specifically designed for complex urban environments and varying lighting conditions. The proposed system is structured into three main phases: segmentation, shape classification, and recognition. In the segmentation phase, images are enhanced and segmented using the Hue, Saturation, and Value (HSV) color space, effectively isolating Regions of Interest (ROIs) despite the presence of similarly colored objects and lighting variations. The shape classification phase utilizes a Random Forest classifier combined with Distance to Borders (DtBs) features to accurately classify traffic signs into triangular, circular, and rectangular shapes. In the recognition phase, Random Forest and Support Vector Machine (SVM) classifiers are employed along with a range of feature extraction techniques, such as Histogram of Oriented Gradients (HOG), Gabor filters, Local Binary Patterns (LBP), and Local Self-Similarity (LSS). Experimental results on a publicly available dataset show high recall and accuracy, with an Area under the Curve (AUC) of 94.50%. These findings demonstrate that the proposed approach offers a reliable and efficient method for traffic sign detection and recognition, with potential applications in autonomous vehicles and real-time traffic monitoring systems. Future research will focus on integrating additional data sources and exploring advanced deep learning techniques to further enhance system performance.

Keywords: Traffic Sign; HSV; HOG; Gabor filters; LBP; SVM, AUC, Distance to Borders

1. INTRODUCTION

The goal of constructing and operating a road transport system is to provide sustainable and socioeconomically efficient mobility. People and businesses rely on transportation that is environmentally friendly, economical, and safe. The integration of automation and information technology into traditional transport vehicles and infrastructure is known as the Intelligent Transport System (ITS). It encompasses traffic control, monitoring, and vehicle routing (Rauber et al., 2021). ITS is broadly categorized into two main types: a. Information Communication Technologies (ICT). b. Intelligent Detection and Recognition Systems. While Intelligent Detection and Recognition Systems use artificial intelligence and image analysis to navigate and monitor traffic, ICT, based on distributed applications, manages data circulation within the network of the intelligent transport system (Okfalisa et al., 2017).

An ITS collects information from traffic signs to operate autonomous vehicles. Typically, drivers must remain highly alert to recognize traffic signs when necessary. However, a driver's ability to identify traffic signs, particularly in adverse weather conditions, is heavily influenced by their physical and mental state. Factors such as stress, fatigue, illness, or intoxication can impair a driver's ability to process visual information. Consequently, ITS is increasingly developing predictive strategies to optimize transportation efficiency and safety, with traffic sign detection and recognition playing a crucial role (Anand et al., 2018).

A functional digital infrastructure is essential for the effective operation of ITS. It is necessary to collect traffic data and maintain operational databases, such as the National Road Database (NVDB). Once this digital infrastructure is in place, ITS can be used for road traffic management and other services (Divakarla et al., 2019). Advanced Driver Assistance Systems (ADAS) are one of the fastest-growing sectors in the automotive electronics industry. ADAS technologies rely on active sensor systems, vehicle data networks, vision systems, and more. These sensors extract various types of data from driving situations. However, one of the key challenges ADAS faces is understanding the environment and guiding vehicles in real-world scenarios (Herrmann et al., 2018).

The purpose of traffic signs is to direct, warn, and regulate traffic. They provide drivers with vital information. However, in real-world conditions, traffic signs are not always easily visible. Poor visibility at night or during adverse weather, as well as distractions from the headlights of passing vehicles, can make traffic signs harder to recognize. These conditions may lead to traffic accidents and severe injuries (Kellett et al., 2019). Therefore, a vision-based road sign detection and identification system is ideal for capturing a driver's attention and helping them avoid traffic hazards. In addition to enhancing ADAS, these systems play critical roles in autonomous driving, urban scene understanding, and sign maintenance monitoring, among other practical applications. By providing drivers with timely information on road hazards and updates on traffic signs, such systems can significantly improve safety (Inagaki et al., 2019).

The traffic sign recognition system proposed in this work consists of two key modules: the detection module and the classification module. After receiving images from the camera, the detection module identifies all potential areas that may contain traffic signs. The classification module then determines the type of sign present in each identified area. Traffic signs are characterized by their visual features, such as color, shape, and pictogram, which provide important information. Therefore, the detection and identification modules rely primarily on the color and shape features of traffic signs.

2. TRAFFIC SIGN DETECTION

Finding the positions and sizes of traffic signs in photos of nature scenes is the aim of traffic sign detection. The two primary indicators for identifying traffic signs are their distinct colours and forms. Thus, colour-based and shape-based detection techniques may be separated into two groups. Colour-based techniques are often quick and independent of scale, translation, and rotation. Since colour is easily influenced by lighting, the primary challenge for color-based algorithms is to be lighting-condition invariant. These techniques often have the same structure:

the picture is converted to a colour space, and then a thresholder is applied. Although this thresholding is quite sensitive to changes in light, some writers choose to conduct it directly in RGB (Red, Green, Blue) space. Simple formulae linking the red, green, and blue components are used to get around this. To segment the desired hue, for instance, Escalera et al. (1997) exploited various relations between the R, G, and B components. The difference between R and G and the difference between R and B channels are used in (Benallal and Meunier, 2003) to produce two reliable characteristics in the identification of traffic signs. Ruta et al. (2010) extracted red, blue, and yellow blobs using colour enhancement. In the RGB colour space, this transform highlights the pixels where the specified colour channel predominates over the other two.

Other colour schemes like YUV and Hue Saturation Intensity (HIS) are also utilised in addition to RGB space. For instance, (Miura et al., 2000) considers the YUV system to identify blue rectangular signs. A segmentation technique in both the La-b and HSI colour spaces is employed in (Lillo-Castellano et al., 2015) to extract candidate blobs for chromatic signals. White signs are simultaneously identified using an achromatic decomposition. Subsequently, a post-processing stage is executed to eliminate uninteresting areas, merge fractured signs, and divide signs situated at identical posts.

2.1 TRAFFIC SIGN RECOGNITION

Classifying the observed traffic signs into their respective sub-classes is the aim of traffic sign recognition. In the context of the recognition issue, machine learning algorithms frequently make use of certain attributes. To identify signs, Maldonado et al. (2007) used several one-vs-all SVMs using Gaussian kernels for each colour and shape categorisation. In (Salti et al., 2015), candidate areas supplied by the interest region detectors are classified using SVMs and HOG characteristics. The durability of local elements, which usually occur in outside data, notably dramatic lighting and size changes, allows it to endure large appearance alterations. In order to get high detection accuracy, Zaklouta (2012) employs random forest-based classification in conjunction with various sized HOG features. In order to lower the computational complexity of traffic sign detection, Tang (2013) suggests an effective approach of traffic sign identification that leverages complimentary characteristics. The traffic sign classification is then implemented using the SVM. Tang (2013) employed complimentary characteristics, namely HOG (Dalal and Triggs, 2005) and LBP (Ojala et al., 1996). Convolutional Neural Networks (CNNs) are yet another technique for classifying traffic signs. (Stallkamp et al., 2011) provides evidence that CNN performs better than humans in classifying traffic signs.

A CNN and a Multi-Layer Perception (MLP) that has been trained on HOG features were utilised in (Ciresan et al., 2011). A CNN with multi-scale features employing a layer-skipping link is provided in (Sermanet and LeCun, 2011).The authors of (Jin et al., 2014) propose a hinge loss stochastic gradient descent technique for convolutional neural network training. The accuracy rates produced by the procedure are high. CNNs do, however, come with a hefty computational cost when training the data. Any study on traffic sign recognition (TSR) often yields data of varying quality depending on the research group doing the investigation. Due in large part to the absence of a common collection of road picture data, it is exceedingly challenging to determine which strategy produces superior overall results. Since it is typically unclear from the various research if pictures with low lighting have been utilised in the tests, it is impossible to tell, for example, how effectively the systems adapt to changes in image illumination. Since compiling a set of road scene photographs takes a lot of effort, another drawback of not having a standardised database of road images is that some studies are based on a tiny group of images. Working with limited data sets is a challenge in terms of assessing the dependability of the outcomes.

2.2 Data Collection

Although search engines have generated general image datasets like ImageNet and Microsoft COCO by retrieving images from the Internet using keywords, very few Internet users upload real-world images of traffic signs as they might be seen on the street, and even when they do, the traffic signs are incidental, none of the signs in these images will be named. Here, such a method is not appropriate. Moreover, pictures devoid of traffic signs must to be incorporated into the benchmark as well, to replicate a real-world application environment and assess the detector's ability to discriminate between actual traffic signs and other items that resemble them. We found that extracting data from Tencent Street Views would be the best method for gathering good photographs. The gathered photos were then manually annotated. Following international standards, Nigerian traffic signs fall into three categories: prohibited (mostly white with a red circle surrounding it and possibly a diagonal bar), mandatory (mostly blue circles with white information) and warning (mostly yellow triangles with a black boundary and information). There are other signals that seem like traffic signs but aren't; some of these are seen in Figure 1. These signals belong to a different class within a given category. We noted the bounding box, boundary vertices, and class label for each traffic sign throughout the annotation process. We employ the polygon mode and the ellipse mode to get the pixel mask for the sign.

Figure 1: Signs like traffic-signs, but with other meaning

3 PROPOSED SYSTEM DESIGN

Three steps make up the suggested system: segmentation, form detection, and recognition. Our initial step is to extract ROIs from the photos by segmenting them. In the second, we use the ROIs to identify the desired forms. During the last phase, we identify the data included in the identified traffic signs. The suggested method's algorithm scheme is shown in Figure 2. We describe in full each stage of the suggested procedure in this section.

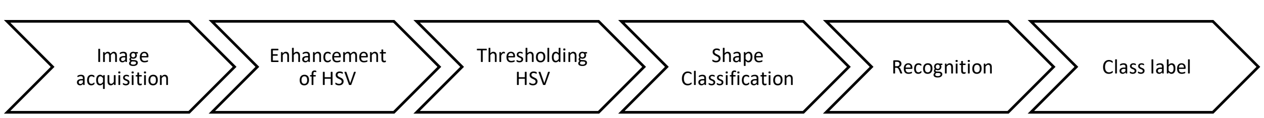

Figure 2: Algorithm scheme

Corresponding Author Tel: +234 8035057652

3.1 Segmentation

A number of factors, including sunlight, shadows, object orientation with respect to the sun, and weather, might affect colour segmentation algorithms. Scenes with dense metropolitan areas often see these factors changing. on addition, a lot of other items on the roadway have the same colours as the red and blue traffic signals. As a result, categorisation is not done and the colour information is solely used to create ROIs.

We employ the HSV colour space in our system to mitigate the challenges caused by variations in light and potential degradation of the signs. We apply thresholding as well as enhancing methods. We start by optimising the input image in the HSV colour space. Next, we use predefined thresholds to segment the picture. Images of traffic signs were used to empirically determine these criteria. After that, the binary picture is post-processed to remove unnecessary ROIs and lower the total number of ROIs needed to give the shape classification step.

1) **Enhancement:** HSV colour space is a good option for enhancing colour images and has been validated by several trials. HSV components only have a weak association with one another, meaning that changing one will only marginally alter the other. Regretfully, a little variation in HSV can cause significant colour distortion in some circumstances. This study solely enhances the value component of the supplied image, leaving the hue and saturation components unaltered. Two processes are involved in this enhancement process: contrast enhancement and luminance enhancement.
2) **Thresholding:** We use thresholding to divide the image into regions of interest (ROIs) following the enhancing procedure. Each component of a picture is categorised according on its hue, saturation, and value. The threshold values shown in Table 1 are used to determine if a pixel's colour is red or blue. The resultant hue (H) falls between [0, 360], whereas the saturation (S) and intensity (I) fall between [0, 255]. In addition, we make reference to the achromatic decomposition that divides white colour.

Table 1: Thresholds used for road sign detection

	Red	Bleu
Hue	$0 < H < 12$ Or $300 < H < 360$	$190 < H < 260$
Saturation	$25 < S < 250$	$70 < S < 250$
Value	$30 < V < 200$	$56 < V < 128$

Following the segmentation step, we are left with a binary picture (see Figure 3(b)), where the pixels of interest are predominantly white. Next, we removed noise and blobs deemed uninteresting based on the size and aspect ratio of the blobs. Based on road image data, the boundaries for both criteria such as size and aspect ratio were empirically determined (see Figure 3(b)).

(a) (b)

Figure 3: Segmentation results.

3.2 Shape Classification

The blobs that were collected from the segmentation step are now categorised in this stage based on their shapes. Only triangles, circles, and rectangles are considered. As a result, feature vectors for the inputs of a random forest classifier are Distance to Borders (DtBs). The lengths between a blob's outer edge and its bounding box are measured in DtBs. These characteristics are frequently used to categorise forms, and several studies on traffic sign identification demonstrate how well they function. These measurements are displayed for a triangle in Figure 4.

The ROIs are classified into suitable forms using a random forest classifier following the computation of these attributes. A Random Forest is a collection of classification trees, where each tree casts one vote to determine which class is more frequently present in the input data. It gives bagging an extra degree of unpredictability. Random forests alter the construction of the classification or regression trees in addition to use distinct bootstrap samples of the data for each tree. Every node in a standard tree is divided using the variable that yields the best split. Every node in a random forest is divided using the best of a subset of randomly selected predictors at that node. In comparison to many other classifiers, this apparently counterintuitive approach performs quite well and is resistant to overfitting. Because random forests can be more accurate and noise-resistant than single classifiers, they are gaining popularity.

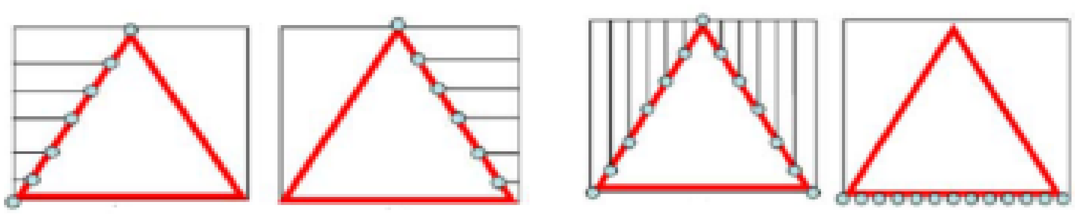

Figure 4: Distance to Borders for a triangular shape

3.3 Recognition

The recognition procedure begins once the candidate blobs are categorised into a form class. The primary goals of this stage are to employ a method with a high degree of accuracy while also minimising the amount of memory and algorithmic complexity. In this study, we compare the state-of-the-art SVM classifier with the Random Forests classifier. SVMs perform worse than random forests in terms of accuracy rate and execution time. Motivated by the current ones, we attempt to offer new features through various combinations for the feature extraction. In this study, HOG, Gabor filters, LBP, and LSS are employed. Along with the classifier features, these features' performance and execution time

a. **Features extraction**: Four different types of features such as HOG, Gabor, LBP, and LSS were employed in this work. The HOG feature is employed initially. The fundamental notion behind HOG features is that, even in the absence of exact information of the matching gradient or edge placements, the look and shape of a local item may frequently be very effectively described by the distribution of the local intensity gradients or edge orientations. The histogram may be computed rapidly due to the method's simplicity and speed. The Gabor feature is the second one. Numerous issues with pattern recognition and signal processing have been addressed using Gabor filters. They can investigate the image's local spectrum properties. A spatial band-pass filter that is selective to both orientation and spatial frequency is called a 2D Gabor filter.
b. **Classifiers:** In this work, two classifiers were used: SVMs and random forests. In section 4, the classification results are illustrated. Because random forests can be more accurate and noise-resistant than single classifiers, they are gaining more and more attention. The fact that Random Forests only require two parameters which are the number of variables in the random subset at each node and the total number of trees in the forest and are often not very sensitive to their values is another benefit of using them. The basic principle of random forests is that an indefinite number of simple trees are used, and the mode of each tree's predictions represents the final predicted class for a test item. On the other hand, SVMs are employed to broaden our analysis of TSR classifiers.

The algorithm tries to distinguish between good and bad instances. The fundamental idea of SVM is to use a nonlinear transform to convert the input vectors into a higher dimensional space. From there, an optical hyper-plane that divides the data may be located. The greatest generalisation ability should be had by this hyper-plane. A linear function is often unable to separate the data. In these situations, using a kernel function becomes crucial. The purpose of SVM is to address binary classification issues. On the other hand, classification is achieved by combining binary classification problems for the road sign inventory problem, which is a multiple classification problem.

4 RESULTS

The outcomes of the suggested method are shown in this section. The attributes and classifier evaluations are provided to support the selection of the suggested system. A 2.7 GHz Intel i5 CPU was used for all of the tests, which were run on the public dataset.

Figure 5: Sample of an Environment before Detection of Traffic Sign Detection

Figure 6: The scene when traffic sign has been detected

The precision-recall curve is used to evaluate the detection stage, and the recall and precision values are calculated as follows: The suggested method's precision-recall curves using the COCO dataset. Table 2 lists the optimal trade-off between the recall and accuracy levels as well as the detection module's Area under Curve (AUC). It is evident that the approach produces the greatest results, with an accuracy of 95.12% and a recall of 93.41%. 94.50% is the accuracy-recall curve's AUC.

Table 2: The best trade-off between the recall and precision values as well as the AUC obtained by the detection method on STS data set in %.

	Recall (%)	**Precision (%)**	**AUC (%)**
On the COCO Dataset	93.41	95.12	94.50

An example of test photos and the accompanying segmentation phase results is presented in Figure 6. The original test photos, shown in the top row with the labels (a) and (b), seem to be scenes from a road setting, maybe showing street signs and road markings. The results of segmentation are displayed in the centre row, designated (c), where pertinent characteristics from the photographs have been selected or emphasised against a contrasting backdrop. It seems likely that the purpose of this segmentation is to identify certain items or regions of interest, such road boundaries or signage. In order to facilitate additional image processing or identification tasks, the bottom row, which is also a component of (c), further refines the segmentation findings by isolating crucial regions such as traffic signs or other significant road-related characteristics in a binary format (black and white). The ultimate outcome of the suggested system's traffic sign detection is shown in Figure 7.

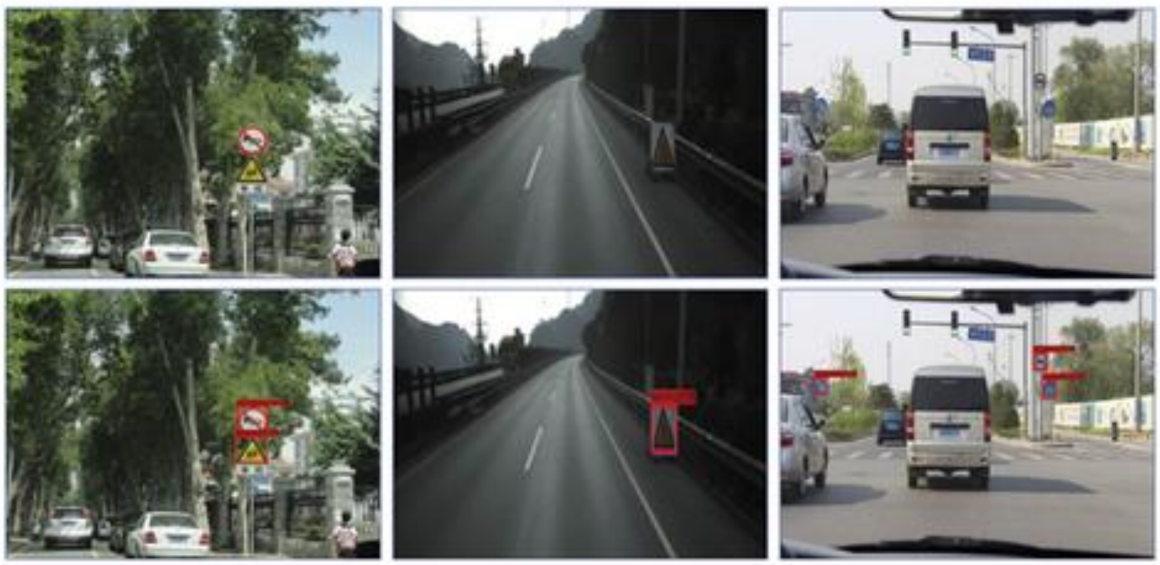

Figure 7: Example of test images (a) and (b), (c) segmentation phase's results.

Figure 8: Final detection results by the proposed detection method.

The final detection results obtained by the suggested detection approach are shown in Figure 8. The situation in the image on the left is one in which the system has successfully detected an item since it is surrounded by a red bounding box, most likely a traffic sign. A close-up of a traffic sign is shown in the image on the right, which is also designated with a red bounding box. This confirms that the sign is correctly identified and located inside the road context. These bounding boxes are used to illustrate how well the detection algorithm recognises and isolates pertinent characteristics in real-world situations.

5 CONCLUSION

In order to attain high accuracy in real-world circumstances, this work proposes a complete approach to traffic sign detection and identification by merging several methodologies. The suggested system efficiently recognises and classifies traffic signs from intricate urban landscapes by utilising color-based segmentation, shape classification, and sophisticated feature extraction techniques. Using the HSV colour space, the traffic sign recognition phase was able to achieve strong segmentation in spite of various lighting conditions and the presence of objects in the surrounding environment with similar colours. ROIs were extracted more easily with the use of enhancement and thresholding approaches. These ROIs were then further improved to remove unnecessary blobs and noise. At the form classification step, traffic signs were divided into triangle, circular, and rectangular shapes using DtBs characteristics and a Random Forest classifier. This strategy worked well, producing excellent detection accuracy and demonstrating resilience to the wide range of environmental factors seen in metropolitan settings. Random Forests and SVM classifiers were used in conjunction with a variety of feature extraction techniques, such as HOG, Gabor filters, LBP, and LSS, during the recognition phase. Random Forests is the recommended option for traffic sign identification in this study because, according to a comparison analysis, it performed better in terms of accuracy and execution time. With an AUC of 94.50%, the experimental findings demonstrated the system's high recall and accuracy rates when assessed on a public dataset. The system's potential use in autonomous cars and real-world traffic monitoring systems is highlighted by its ability to reliably identify and recognise traffic signs in a variety of difficult settings. To sum up, the suggested approach provides a dependable and effective means of identifying and detecting traffic signs. To improve the system's resilience and condition adaptability, subsequent study may investigate the incorporation of other data sources and the application of sophisticated deep learning methods.

REFERENCES

Anand, S., Gupta, S., Darbari, V., & Kohli, S. (2018). Crack-pot: Autonomous road crack and pothole detection. In 2018 Digital Image Computing: Techniques and Applications (DICTA) (pp. 1–6). https://doi.org/10.1109/DICTA.2018.8615819

Benallal, M., & Meunier, J. (2003). Real-time color segmentation of road signs. In Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on (Vol. 3, pp. 1823–1826). IEEE.

Ciresan, D., Meier, U., Masci, J., &Schmidhuber, J. (2011). A committee of neural networks for traffic sign classification. In Neural Networks (IJCNN), The 2011 International Joint Conference on (pp. 1918–1921). IEEE.

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886–893). IEEE.

Divakarla, K., Emadi, A., Razavi, S., Habibi, S., & Yan, F. (2019). A review of autonomous vehicle technology landscape. International Journal of Electric and Hybrid Vehicles (IJEHV), 11, 320–345.

Escalera, D. L., Moreno, L. E., Salichs, M. A., &Armingol, J. M. (1997). Road traffic sign detection and classification. IEEE Transactions on Industrial Electronics, 44(6), 848–859.

Herrmann, A., Brenner, W., & Stadler, R. (2018). Autonomous Driving: How the Driverless Revolution Will Change the World.

Inagaki, T., & Sheridan, T. (2019). A critique of the SAE conditional driving automation definition, and analyses of options for improvement. Cognition, Technology & Work, 21, 569–578.

Jin, J., Fu, K., & Zhang, C. (2014). Traffic sign recognition with hinge loss trained convolutional neural networks. IEEE Transactions on Intelligent Transportation Systems, 15(5), 1991–2000.

Kellett, J., Barreto, R., Hengel, A., & Vogiatzis, N. (2019). How might autonomous vehicles impact the city? The case of commuting to central Adelaide. Urban Policy and Research, 37, 442–457.

Lillo-Castellano, J., Mora-Jimenez, I., Figuera-Pozuelo, C., & Rojo-Alvarez, J. (2015). Traffic sign segmentation and classification using statistical learning methods. Neurocomputing, 153, 286–299.

Maldonado-Bascon, S., Lafuente-Arroyo, S., Gil-Jimenez, P., Gomez-Moreno, H., & Lopez-Ferreras, F. (2007). Road-sign detection and recognition based on support vector machines. IEEE Transactions on Intelligent Transportation Systems, 8(2), 264–278.

Miura, J., Kanda, T., & Shirai, Y. (2000). An active vision system for real-time traffic sign recognition. In Intelligent Transportation Systems, 2000. Proceedings. 2000 IEEE (pp. 52–57). IEEE.

Ojala, T., Pietikainen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. Pattern Recognition, 29(1), 51–59.

Okfalisa, I., Mustakim, M., & Reza, N. (2017). Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification. In 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE) (pp. 294–298). https://doi.org/10.1109/ICITISEE.2017.8285514

Rauber, T., da Silva-Loca, A., de Boldt, F., Rodrigues, A., & Varejao, F. (2021). An experimental methodology to evaluate machine learning methods for fault diagnosis based on vibration signals. Expert Systems with Applications, 167, 114022. https://doi.org/10.1016/j.eswa.2020.114022

Ruta, A., Li, Y., & Liu, X. (2010). Real-time traffic sign recognition from video by class-specific discriminative features. Pattern Recognition, 43(1), 416–430.

Salti, S., Petrelli, A., Tombari, F., Fioraio, N., & Di Stefano, L. (2015). Traffic sign detection via interest region extraction. Pattern Recognition, 48(4), 1039–1049.

Sermanet, P., & LeCun, Y. (2011). Traffic sign recognition with multiscale convolutional networks. In Neural Networks (IJCNN), The 2011 International Joint Conference on (pp. 2809–2813). IEEE.

Stallkamp, J., Schlipsing, M., Salmen, J., & Igel, C. (2011). The German traffic sign recognition benchmark: A multi-class classification competition. In Neural Networks (IJCNN), The 2011 International Joint Conference on (pp. 1453–1460). IEEE.

Tang, S., & Huang, L.-L. (2013). Traffic sign recognition using complementary features. In Pattern Recognition (ACPR), 2013 2nd IAPR Asian Conference on (pp. 210–214). IEEE.

Zaklouta, F., &Stanciulescu, B. (2012). Real-time traffic-sign recognition using tree classifiers. IEEE Transactions on Intelligent Transportation Systems, 13(4), 1507–1514.