WILEY | Hindawi

Research Article

The Implications of Weather and Reflectivity Variations on Automatic Traffic Sign Recognition Performance

Mudasser Seraj ⓘ**, Andres Rosales-Castellanos** ⓘ**, Amr Shalkamy** ⓘ**, Karim El-Basyouny** ⓘ**, and Tony Z. Qiu** ⓘ

Department of Civil and Environmental Engineering, University of Alberta, Edmonton T6G 2R3, Canada

Correspondence should be addressed to Mudasser Seraj; seraj@ualberta.ca

Automatic recognition of traffic signs in complex, real-world environments has become a pressing research concern with rapid improvements of smart technologies. Hence, this study leveraged an industry-grade object detection and classification algorithm (You-Only-Look-Once, YOLO) to develop an automatic traffic sign recognition system that can identify widely used regulatory and warning signs in diverse driving conditions. Sign recognition performance was assessed in terms of weather and reflectivity to identify the limitations of the developed system in real-world conditions. Furthermore, we produced several editions of our sign recognition system by gradually increasing the number of training images in order to account for the significance of training resources in recognition performance. Analysis considering variable weather conditions, including fair (clear and sunny) and inclement (cloudy and snowy), demonstrated a lower susceptibility of sign recognition in the highly trained system. Analysis considering variable reflectivity conditions, including sheeting type, lighting conditions, and sign age, showed that older engineering-grade sheeting signs were more likely to go unnoticed by the developed system at night. In summary, this study incorporated automatic object detection technology to develop a novel sign recognition system to determine its real-world applicability, opportunities, and limitations for future integration with advanced driver assistance technologies.

1. Introduction

Traffic signs are used to regulate, warn, and guide traffic on roadways and facilitate coordinated road usage [1], and their placement, orientation, and visibility are crucial for road operation and safety. Given their importance, automatic recognition of roadway signs by smart transportation technology is a current research interest with numerous potential applications [2]. Automatic sign recognition could facilitate the interpretation of information received from the detected signs by assisted and autonomous driving systems. For instance, traffic operation and maintenance authorities could develop traffic sign inventories and conveniently identify traffic sign maintenance needs with the help of such automated systems [3, 4]. In addition, rapid and accurate recognition of traffic signs is important for improving traffic safety, the primary goal of Intelligent Transportation Systems and Vision Zero initiatives [5–7].

While significant effort to develop a robust sign recognition system has been made by both academics and industry practitioners [8–15], the sensitivity of the developed systems to recognize signs in varying real-world conditions is still hypothetical. Our descriptive research, then, concentrates on the question: how does the diversity of weather and reflectivity conditions of the physical world influence the recognition performance of a TSR system, given that the system is developed by incremental training resources? To address this crucial research question, our study leveraged a state-of-the-practise object detection algorithm to develop a robust TSR system that can successfully recognize a wide range of traffic signs. The objective was to measure the recognition efficiency of the developed TSR system in various weather and reflectivity conditions with the aim of assessing the implications of those variations on the recognition performance. The scope of this research is limited to identifying the changes of performance pattern due to

variations in real-world weather and reflectivity conditions resulting from an industry-level object detection system. Hence, comparing the detection performance in comparison to other state-of-the-art TSR system in standard lighting, weather and reflectivity conditions are beyond the scope of this study. Additionally, our study emphasised on detection sensitivity to TSR system training resources (i.e., images of traffic signs from favourable environmental conditions) which restrained us from analyzing asymmetrical impact due to training resources with varying lighting and environmental features.

The findings presented here will contribute to the overall body of knowledge in several ways. The key contributions of this study are listed below:

Besides intrinsic features of traffic signs (i.e., size, shape, and color), which are mostly focused on earlier studies, this study paid attention to external factors and their influence on sign identification efficiency.

Traffic sign images collected from favourable lighting and weather conditions are used as training resources in the development phase of the TSR system and expected to establish a more proficient system with gradually increased resources. The identified pattern of progression will facilitate future research by providing a benchmark for the correlation between training datasets and the expected efficiency of a TSR system that is developed with the same system architecture.

As mentioned before, the TSR system was only trained with images from favourable environmental condition which implies the proficiency of the developed system in identifying the signs in unfavourable condition without being trained for such conditions.

2. Literature Review

Researchers and practitioners have attempted to establish automatic traffic sign detection systems over the last few decades that have increased in complexity as technology has advanced. The most conventional form of research has focused on systems that attempt to extract signs from their environments for identification based on colour and shape. In terms of colour-based recognition, a valuable set of investigations used RGB space [16–18]. A clustering method in a colour space was developed by Tominaga [19] for sign detection. Ohlander et al. [20] also used a recursive region splitting method to achieve colour segmentation and applied Hue, Saturation, Intensity (HSI), and sign area. In terms of shape-based recognition, studies used the basic sign shapes of circles, triangles, or rectangles [9, 11, 21–26]. This method does overcome brightness issues found with colour-based methods. However, a significant difficulty for shape-based detection is the rotation angle and distortions of signs. All road signs in this method require a nonzero angle between the optical axis of each camera and the normal vector to the sign surface and should be as high as 30°, depending on the distance between the sign and the cameras. The complexity of automatic shape-based detection increases for signs with

acute viewing angles as well as signs with torn corners and occluded parts.

Methods based on statistical machine learning and artificial intelligence (AI) have become prevalent, providing researchers with new tools to develop more efficient and reliable TSR systems using classification techniques, such as artificial neural networks (ANNs) [27, 28] k-nearest neighbour (KNN) [13, 29], support vector machine (SVM) [8, 18], and random forest [30, 31]. In 2005, Gil-Jimenez et al. [32] explored shape classification algorithms using SVMs for TSR. They found that using statistical classification methods such as SVM, colour, and shape, traffic signs can be roughly recognized. In 2006, Gao [33] did further research on shape and colour using human vision models, testing on 98 British traffic signs in various viewing conditions. The results were that the recognition rate increased to as high as 95%, especially for immobile traffic signs. In another study [34], unified visual saliency with Histograms of Oriented Gradients (HOG) featured learning for TSR. Here, the authors combined SVM with the HOG feather learning method and were able to achieve a high recognition rate. In 2010, Prisacariu et al. [35] proposed a real-time system that introduced region-based 3D tracking to single view detection, followed by adaptive boosting cascades and SVM to improve accuracy. Huang et al. [36] developed a method for detecting and recognizing speed-limit signs using only gray-level information. Hechri and Mtibaa [37] proposed and tested road sign detection using shape-filtering methods, with a multilayer, perception neural network classification model. Other studies have refined the details of real-time TSR. Sheng et al. [38] proposed treatment for recognizing signs of different colours, such as red/yellow/blue versus grayscale, and a probabilistic neural network to achieve final recognition. Li et al. [39] proposed a fuzzy shape recognizer to improve the robustness of traffic sign detection. Although it highlighted some success in sign detection and recognition, the authors emphasised the need to investigate and improve the overall performance in different weather and light conditions. However, while these studies were focused on developing TSR systems with higher accuracy through advanced techniques of machine learning and artificial intelligence, they were often silent on the implications of real-world variables on sign recognition.

A few examples of recent studies do couple the application of statistical machine learning and artificial intelligence with the influence of exogenous factors on sign detection and recognition. For instance, Sajjad et al. [40] developed a deep learning-based sign detection system as a part of an autonomous driving demonstration. Although the developed system performed well in a controlled environment, the detection and navigation accuracy are yet to be tested in real-world scenarios with complex challenges. Wang et al. [41] also developed an AI-based lightweight sign detection system that outperformed the Microsoft COCO benchmark [42]. However, the developed system was tested solely on images extracted from diverse real-world settings and not continuous detection from real-world driving condition videos that would include the added complexity from continuous sign tracking. The analysis performed by Muhammad et al. [43] on

multiple state-of-the-art approaches for sign detection from Swedish Traffic Sign Dataset [44] also suffered from similar limitations. Tabernik and Skocaj developed a convolutional neural network- (CNN-) based system capable of recognizing 200 categories of signs. While the literature demonstrated excellent performance on the tested dataset, the pair did not compare the developed TSR system's performance variations for different lighting and weather conditions. Garcia-Garrido et al. [45] tested a sign detection, classification, and tracking system under different weather and light conditions employing an algorithm designed to use a camera mounted on a vehicle's windscreen. While the authors stated that the developed TSR system was tested in different weather and reflectivity conditions, additional information regarding recognition performance in varying environmental scenarios was absent in the paper. Phu and Lwin Oo [46] introduced an RGB colour-based thresholding technique for sign detection and recognition, using an adaptive neuro-fuzzy inference system (ANFIS) to recognize different features points. Although they concluded that the system yields good results in sunny, cloudy, and rainy weather conditions, supporting analysis for this claim was not provided in the paper. Lim et al. [47] proposed a system for real-time recognition of speed limit signs in different illumination conditions using modified census transform (MCT) and support vector machine (SVM). High detection and recognition rates were obtained. However, the influence of incremental training on TSR system performance remained unaddressed. Hassaballah et al. [48] presented their conceptually similar research on detecting objects (i.e., vehicles) in adverse weather conditions. They restored visibility by improving raw image quality before object detection and tracking. The user's perspective of the testing dataset was static for all the different weather conditions collected from images. In real-world driving scenarios, the adverse weather conditions introduce added challenges of uninterrupted detection and tracking with continuously changing environments and backgrounds.

From our review, it is evident that numerous studies have been proposed to detect and recognize road traffic signs. While plenty of proposed methods rely on computer vision and artificial intelligence tools, challenges still exist in the field, including the effects of weather, reflectivity variations, limited classes of sign recognition, and human-machine interactivity. More importantly, the majority of established TSR systems have not been tested for real-world environmental challenges. To address these research gaps in sign recognition systems, we introduce a YOLO-based TSR system for detection and recognition of key regulatory and warning traffic signs in real time. Our study measured its performance in a variety of weather and reflectivity conditions and considered their influence on recognition ability. We contribute to the existing TSR system foundations by pairing it with an industry-level detection system and evaluating its performance in complex real-world scenarios. Furthermore, we developed several editions of the TSR system by training the system, introducing increasing numbers of traffic sign images to its training process. We went on to measure each edition's ability to provide direction on the necessary resources required to attain a certain level of success from TSR systems.

3. Data Collection

The data collection process played a pivotal role in our research progress as the diversity of collected data mandated the individuality of this research. In the beginning, a subset of available on-road traffic signs, coupled with signs used in earlier studies, was selected for investigation. Our initial selection criteria covered the majority of basic regulatory signs while expanding to include common warning signs. Ultimately, eight types of signs from two sign classes (i.e., regulatory and warning) were used, as listed in Table 1. Signs with different specifications (i.e., age and sheeting types) were collected from across Canada. Altogether, a total of twenty-eight signs were used for data collection. Fifteen of these signs were made of ASTM D4956 Type XI (usually known as and will be referred as diamond grade) sheeting with high reflectivity, and the remaining thirteen signs were made of Type I (usually known as and will be referred as engineering-grade) sheeting with low reflectivity. Out of twenty-eight signs, fifteen signs were new (age ≤1-year), while thirteen signs were three years or older. Although the collected sign inventory was diverse with respect to these features, we were unable to acquire each sign type in the different ages and sheeting types, proving to be an analysis limitation concerning reflectivity in sign recognition. Therefore, the analysis of reflectivity was performed by grouping signs into sign classes and comparing the performance of each sign class.

Once procured, the signs were installed on an access-controlled roadway segment (with a length of 750 meters in each direction) specifically built for research purposes at the University of Alberta, Edmonton, Alberta, Canada (Figure 1). A total of ten signposts were evenly spaced at designated spots on the test track, five for each direction of traffic. Each signpost was equipped with an adjustable holding mechanism to facilitate the placement and removal of traffic signs. On each round of data collection, ten signs were installed on the signposts. After installation, a vehicle equipped with a video recording camera drove along the roadway segment at 40 km/h recording the installed traffic signs. Although only ten signs were installed on the test track at any given time, all the available signs for this study were used by changing and/or moving them after each recording, thus generating different combinations for every round. For each combination of traffic signs, three recordings were made on each round to ensure the availability of superior quality videos for any given scenario. To maintain a standard recording environment, the elevations of the signs on their posts, as well as the position and angle of video camera inside the car, were kept consistent.

All videos recorded in the afternoon underwent weather impact analysis to filter for the influence of light. In total, 795 videos were recorded covering different weather and reflectivity conditions. Ninety seven (12%) of those were unusable due to exogenous factors (e.g., placement errors and obscured visuals). From the remaining videos, 15%, all with favourable conditions (i.e., daytime and sunny weather), were set aside for the image extraction required for TSR system development. The remaining videos were labelled and batched according to the weather and reflectivity conditions during recording.

TABLE 1: Selected traffic signs for this study.

Sign class	Sign type	Sign image
Regulatory signs	Speed limit	MAXIMUM 30 / MAXIMUM 80 / MAXIMUM 70
	Stop	STOP
	Yield	
Warning signs	Speed limit change ahead	
	Signal ahead	
	Stop ahead	
	Yield ahead	
	Curve ahead	

4. Sign Recognition System Development and Testing

In order to overcome the inherent limitations of self-created TSR systems, the search for an established system architecture based on artificial intelligence led us to YOLO. YOLO is an industry-grade object detection system that is extremely fast and accurate [49]. We chose version 3 (v3) for our study as it includes a new object classifier network that has outperformed earlier versions in different types of object detection. We then trained that classifier network to detect the traffic signs listed in Table 1 and subsequently tested its sign recognition ability in varying weather and reflectivity conditions.

Since adequate training datasets are the most important input for a deep learning approach, initial efforts went towards the extraction of video frames containing traffic signs, with the intention to gradually increase the number of images in the dataset. Each extracted image was labelled using Labellmg software [50] to extract individual regions of an image, create a bounding box, and generate annotation in a YOLOv3 compatible format, with a total of 4,445 annotations generated from the extracted frames. Once the annotations were generated, a python script was written to find corresponding images since image labelling software only develops individual annotations. Without finding the image for a specific annotation, the training process could not be executed. After developing a set of images with annotations, the images were further augmented to simulate potential variations of light, weather conditions, and physical distortions.

Using a python script, shape-based modification included rotating the captured frames at specific angles (i.e., 0°, 15°, 30°, and 45°) to imitate possible physical distortion of traffic signs that might be experienced on the video. Colours were also modified using the python script, to replicate various light scenarios by changing hue, saturation, and contrast of extracted image frames. While it would be impractical to account for all possible shape and colour-based inconsistencies experienced in real world, the augmented dataset provided some primary criterion for the algorithm to learn. Examples of data acquisition and shape and colour-based augmentation of image frames are provided in Figure 2.

Once the preprocessing of training data was finished, the images were fed into the YOLOv3 object detection system with a Darknet-53 feature extractor to develop our TSR system. YOLOv3 used a variant of Darknet [51] that originally had a 53-layer convolutional neural network (CNN) trained on ImageNet [52]. For the task of recognition, 53 additional CNN layers were stacked onto it, making it a 106-CNN layer underlying architecture for YOLOv3. As a result, the residual skip connection and upsampling features of the TSR system were enhanced. YOLOv3 performed recognition on feature maps of three different sizes at three different places in the network. Object recognition from images in different scales is a unique feature of YOLOv3, making it ideal for this study since traffic signs are small in relation to the larger image. Its recognition uses a detection kernel shaped as $1 \times 1 \times [B \times (5 + C)]$. Here, B is the number of possible predicted bounding boxes on a cell of the feature map, 5 is the four bounding box attributes and one object confidence, and C is the number of sign classes.

The first detection was made by the 82nd layer. For the first 81 layers, the image was down sampled by the network, such that the 81st layer had a stride of 32. For instance, if we had an image of 416×416 pixels, the resultant feature map would be 13×13. If one detection is made here using the 1×1 detection kernel, this gives us a detection feature map of $13 \times 13 \times 48$. The default value of $B = 3$ for YOLOv3 as it predicts 3 bounding boxes for every cell, where each bounding box specialises in detecting a certain kind of object; C is the number of classes, 11 in our case reflecting the number of detection signs. Hence, the kernel size is $1 \times 1 \times 54$ for each region on the image.

The feature map from layer 79 was subjected to a few convolutional layers before being upsampled by two-times to dimensions of 26×26. The map was then depth
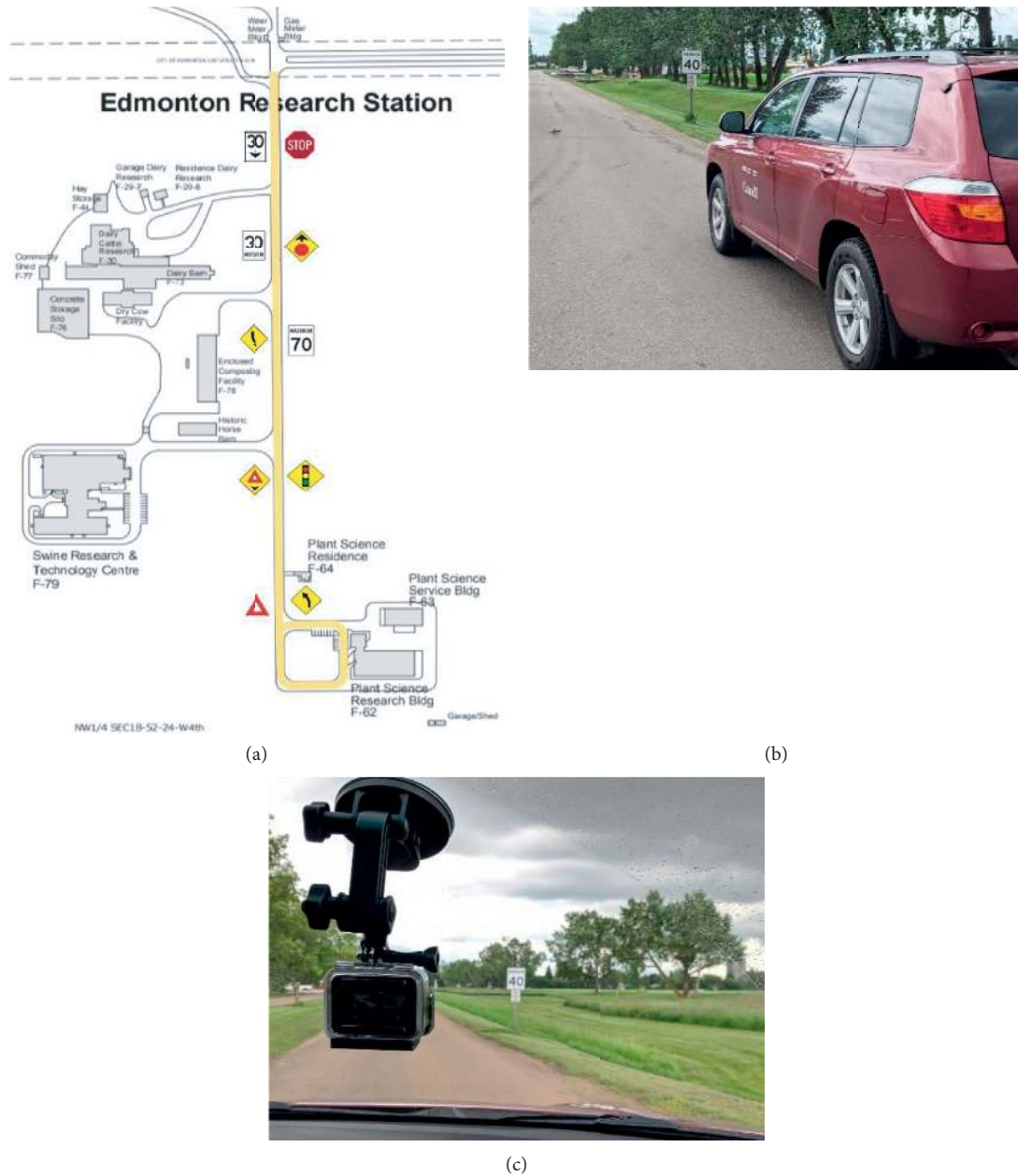
(a)

(b)

(c)

Figure 1: (a) Plan of data collection test track. (b) Vehicle. (c) Camera used for recording videos of traffic signs.

concatenated with the feature map from layer 61 and subjected to a few 1×1 kernel convolutional layers to fuse the features. The second detection was made on the 94th layer, yielding a detection feature map of $26 \times 26 \times 255$. A similar procedure was followed where the feature map from layer 91 was subjected to convolutional layers before being depth concatenated with a feature map from layer 36. Like before, a few 1×1 kernel convolutional layers followed to fuse the information. The final detection was made at 106th layer, yielding a feature map of $52 \times 52 \times 255$. YOLOv3 used nine anchor boxes, three for each image scale. Training YOLO on a custom dataset required using K-means clustering to generate these nine anchors. The anchors were arranged in descending order

of dimension, assigning the three biggest anchors for the first scale, the next three for the second scale, and the last three for the third scale.

Training the TSR system began by feeding the annotated training images into the YOLOv3 network. Training images used for all versions of the TSR system were collected from favourable lighting (i.e., daylight) and weather (i.e., sunny) conditions. Continuous monitoring of average loss for each epoch was made from the beginning of the training process. The objective of observing the average loss value was to stop the training after it reached a certain threshold or a point where the loss value became stationary, and it could be assumed that the network had converged. The training usually converged at a loss rate of 0.03 over a certain number
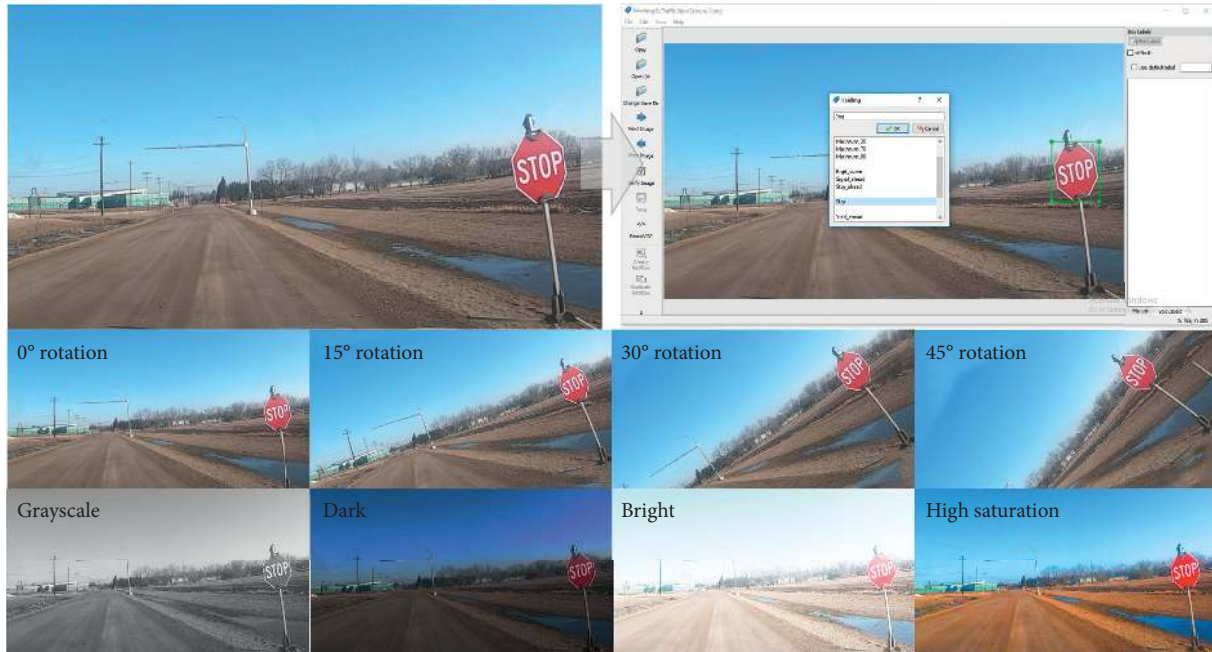
Figure 2: Different stages of data preprocessing.

of epochs (130 – 200 epochs). The training process involved iteratively updating the weights of the neural network based on the number of mistakes made on the training dataset. Weights represent the strength of the connection between neuron units, which helped to determine how much influence the input would have on the output. Every neuron of the network was given an input value and a bias value and was then multiplied by a weight value that was adjusted during each iteration of training.

Once the training task was completed by attaining a target average loss rate of 0.03, the most recent weight generated by the system was executed on Darknet to analyse the testing video dataset. These recorded videos contained the chosen traffic signs in different light, weather, and reflectivity scenarios. In order to obtain labelled videos, screenshots were captured during analysis by the developed TSR system. Additionally, a .csv file containing the information pertaining to the recognized traffic signs and image frames of labelled traffic signs from the analysed video was automatically stored in a secure server after each test video analysis was completed. An overview of the entire sign recognition system development and testing process is summarized in Figure 3.

In this study, five versions of the TSR system were developed by gradually increasing the number of training images, allowing us to explore the impact of training datasets on recognition performance. Each version, therefore, had a different training dataset but maintained the same parameters as outlined in Table 2. Training datasets for each version were prepared from extracted images recorded in bright, sunny conditions to maintain a standard training environment. In each version of the TSR system, 20% training images contained traffic signs and reminder of the training data were images without traffic signs. The number

of annotated and augmented images of traffic signs used for the training and validation of each system version is listed in Table 2.

5. Analysis of TSR System Performance and Progress

To perform analysis of system performance, a set of parameters was chosen from the literature. A combination of these parameters was used exclusively to evaluate the identification accuracy of different versions of the developed TSR system. The recognition performance of each sign type was evaluated using three parameters: precision, recall, and F1-score. To measure these parameter values, the recognition criterion was defined by measuring the number of True Positives, False Positives, and False Negatives on each dataset. The analysis was based on the following definitions of the parameters.

True Positive: TSR system successfully recognized (i.e., identified and classified) the presence of a test sign

False Positive: TSR system incorrectly recognized (i.e., identified and/or classified) the presence of a test sign

False Negative: TSR system failed to recognize (i.e., identify and recognize) the presence of a test sign

Using the above definitions, three key parameter values were calculated, and the following definitions were used to determine the parameter values:

Precision is the fraction of correct recognition instances out of total successful recognitions (equation (1)).

Recall is the fraction of correct recognition instances retrieved over total expected recognitions (equation (2)).
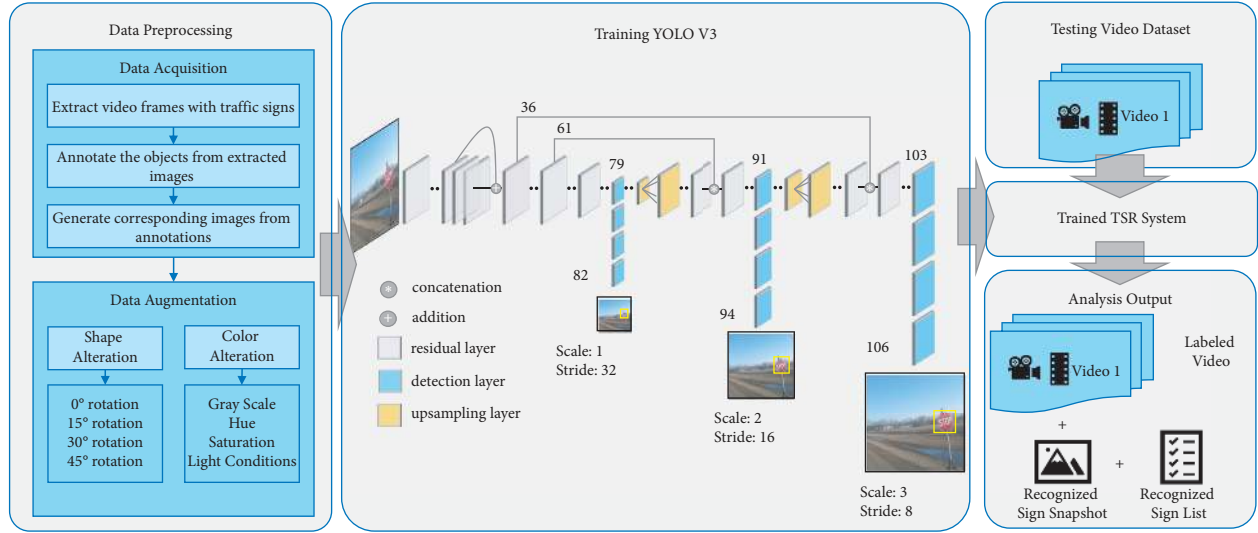
FIGURE 3: Flow chart of the developed TSR system.

TABLE 2: Information on training traffic sign image datasets for each version of the TSR system.

Groups	Number of images				
	Ver. 1	Ver. 2	Ver. 3	Ver. 4	Ver. 5
Max. 30 km/h	264	808	1594	1978	2232
Max. 70 km/h	120	502	750	1146	1408
Max. 80 km/h	291	963	1673	2257	2435
Stop	622	1388	2512	3732	4454
Yield	1272	3178	4408	5476	6046
Max. 30 km/h ahead	388	984	1578	2006	2192
Signal ahead	686	1724	2722	3232	3482
Stop ahead	932	2038	2962	4370	5014
Yield ahead	310	550	852	1192	1456
Right curve ahead	848	2196	4072	5138	5726
Left curve ahead	560	1078	1634	2110	2264
Annotated and unannotated image	5054	10231	18258	23215	26714
Training and validation annotated image	6293	15409	24757	32637	36709

$F1$-score is the weighted average of precision and recall. Since this takes both False Positives and Negatives into account, the $F1$-score provides a more useful interpretation of recognition performance (equation (3)).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}, \quad (1)$$

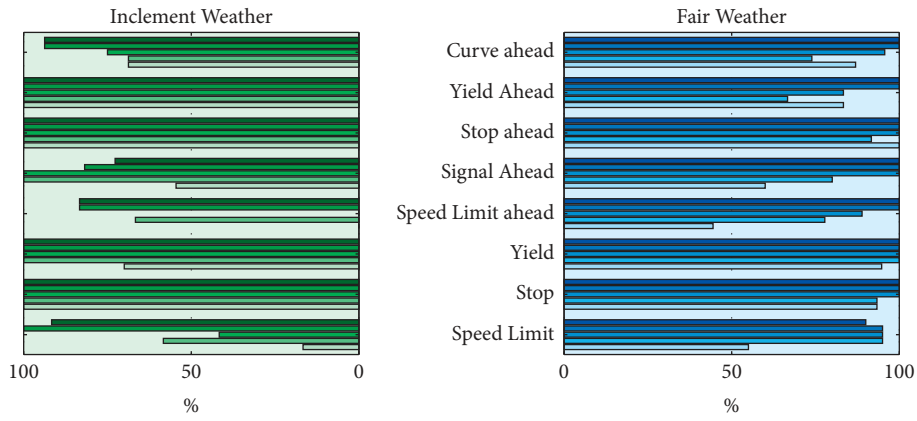$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}, \quad (2)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \quad (3)$$
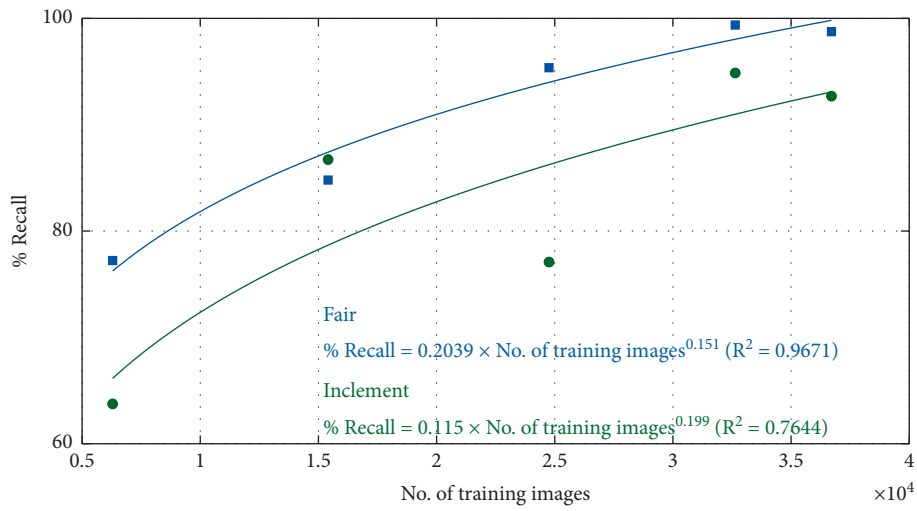
6. Implications of Weather Variations on TSR System Performance

The automatic sign recognition system's competence when introducing varying weather conditions was analysed by testing the videos through the established TSR system. Out of 370
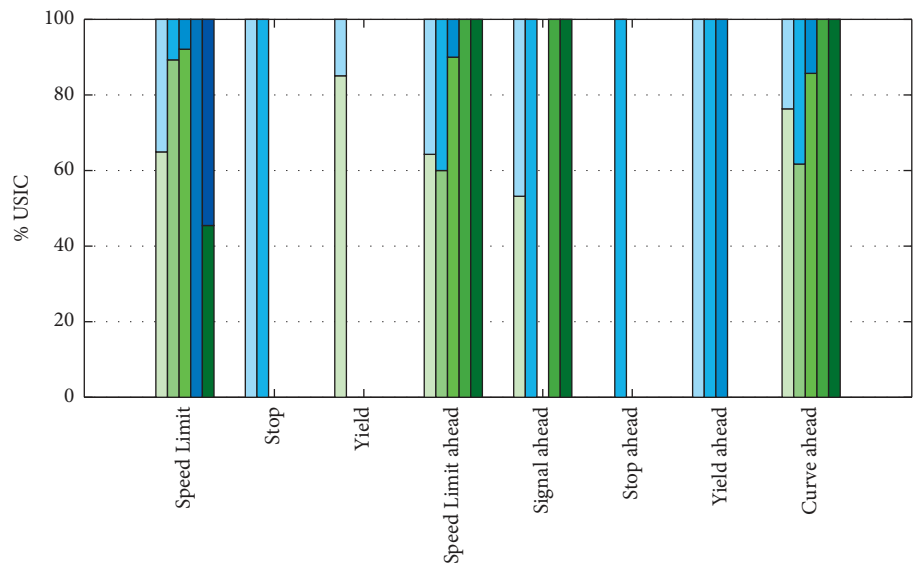
videos considered (based on lighting conditions), 193 videos were recorded in varying weather conditions, but consistent daylight was chosen to keep results clear of any impact from variations in light conditions. As already outlined, the comparison of recognition performance was between fair and inclement weather conditions. Fair weather was represented on 104 videos among the tested datasets. Figure 4(a) primarily compared the recall performance of tested sign types in incremental versions of the developed TSR system for two types of weather scenarios, inclement weather and fair weather. Each horizontal bar represents the recall value of a specific sign for a specific version of the TSR system. The incremental intensity of green/blue color represents the higher version of the TSR system. For instance, the bottom-most horizontal bar of Figure 4(a) illustrated the recall value of Speed Limit signs for version 1 of the TSR system, which was much higher in fair weather (57.43%) than in inclement weather (17.77%). Table 3 outlines the version-specific parameter values irrespective of sign types. Since the impact of weather conditions was not clear from this illustration, we carried out several significance tests on the evaluation parameters.

(a)

Fair
% Recall = $0.2039 \times$ No. of training images$^{0.151}$ ($R^2 = 0.9671$)

Inclement
% Recall = $0.115 \times$ No. of training images$^{0.199}$ ($R^2 = 0.7644$)

(b)

(c)

FIGURE 4: (a) Comparison of recall metric on studied weather conditions. (b) Progression of average recall values with an increasing number of training images. (c) Unidentified signs proportion of compared weather scenarios.

TABLE 3: Version-specific parameter values (i.e., mean and standard deviation) at different weather conditions.

Ver.	Precision		Recall		F1-score	
	Fair	Inclement	Fair	Inclement	Fair	Inclement
1	90.98%	80.14%	77.23%	63.75%	83.54%	71.01%
	(7.15%)	(5.47%)	(20.98%)	(38.49%)	(12.67%)	(8.87%)
2	87.52%	75.18%	84.79%	86.72%	86.13%	80.54%
	(10.68%)	(8.45%)	(11.80%)	(18.57%)	(13.21%)	(13.86%)
3	88.05%	67.70%	95.36%	77.08%	91.56%	72.09%
	(13.98%)	(14.92%)	(6.24%)	(37.47%)	(10.63%)	(16.55%)
4	81.49%	58.19%	99.38%	94.86%	89.55%	72.13%
	(16.16%)	(18.09%)	(1.77%)	(7.89%)	(5.19%)	(24.40%)
5	71.23%	40.02%	98.75%	92.68%	82.76%	55.90%
	(24.76%)	(23.48%)	(3.54%)	(10.01%)	(8.93%)	(11.81%)

Significance test (two-sample t-test) results found that the recall parameter in fair weather was significantly higher, at 95% confidence level, than in inclement weather conditions for versions 1 and 2 of the TSR system. Based on the chi-square contingency table test, the precision metric was found to be significantly different in versions 1 through 3, whereas an insignificant difference was observed on the F1-score for versions 3 through 5. Hence, it can be stated that the devised TSR system did not experience significant performance variations due to weather on versions that were trained with more than 20,000 training and validation images.

Additional analysis of the average recall metric of different versions revealed a pattern of recognition performance progression as the number of images increased. This pattern revealed that the system required a higher number of images to be trained when faced with scenarios such as imperfect conditions (i.e., inclement weather) to match the performance for perfect conditions. For instance, to attain an 80% recall on signs in inclement weather, the TSR system required training on more than 17,000 images. By contrast, approximately 9000 training images would be adequate to attain a similar performance for signs in fair weather conditions. Finally, the analysis on the percent of unidentified signs in imperfect conditions (%USIC) revealed that the majority were set within inclement weather conditions (average = 56.46% and standard deviation = 41.34%). With the developed versions of the TSR system, the proportions of unidentified signs were evenly distributed between fair and inclement weather (Figure 4(c)), and several sign types were completely identified with the higher numbered versions, irrespective of weather conditions.

Since different inclement weather types bring different identification challenges, the TSR system performance was further analysed based on both inclement weather types (i.e., cloudy and snowy) compared to 89 videos. Figure 5(a) showed the comparative performance of recall parameters by the developed TSR systems in cloudy and snowy weather conditions. Figures 5(b) and 5(c) illustrate sample identification snapshots of one sign in cloudy and snowy weather conditions, respectively. The average recall value in cloudy weather conditions (84.11%) was slightly higher than snowy weather conditions (81.93%), considering all versions of the TSR system. Most of the signs showed similar performance in both inclement weather conditions, except signs with white backgrounds (i.e., speed limit signs and speed limit ahead signs). On average, identification of speed limit signs was 9.43% higher in cloudy weather conditions. Similarly, speed limit ahead signs in cloudy weather conditions experienced 5.06% higher identification rates.

7. Implications of Sign Reflectivity Variations on TSR System Performance

The analysis of sign reflectivity was designed to comprehensively consider the reflectivity factors that influence recognition performance. Hence, the signs were initially compared based on two sheeting categories, namely, engineering grade and diamond grade. Each sheeting type was further divided into two lighting conditions, daytime and night-time, presuming lighting conditions would play a significant role in detectability. Finally, two more categories for sign age were also taken into account, i.e., signs less than one year old and signs three years or older, presuming that reduction in reflectivity accumulates over the year due to weathering, natural abrasion, and other factors. The goal in this portion of the analysis was to establish the effect of these factors on the TSR system's accuracy.

The classification of the tests for this section was not made sign-specific but was rather classed in a higher-order due to the lack of availability of some specific signs in one of the two sheeting types or different age groups. As a result, signs were divided into two classes, namely, warning signs (e.g., curve ahead and signal ahead) and regulatory signs (e.g., stop and speed limit). Figure 6 shows the distribution of data samples in the testing dataset that contained characteristics related to sheeting type, lighting conditions, and sign age. The two sheeting types selected for this study possess considerably different reflectivity features. Engineering grade or Type I sheeting typically meets the requirements of ASTM D4956 and contains some basic reflectivity properties. On the other hand, diamond grade or Type XI sheeting is designed to reflect close to 50% of the available light to the driver, enabling them to better recognize signs and at a greater distance [53]. Lighting conditions were considered for the second level of reflectivity factors. All the videos for daytime lighting conditions were collected between 10 am to 4 pm, and night-time data were collected from 1 hour after sunset till 10 pm. Finally, the signs were sorted based on two predominant age groups of available signs.
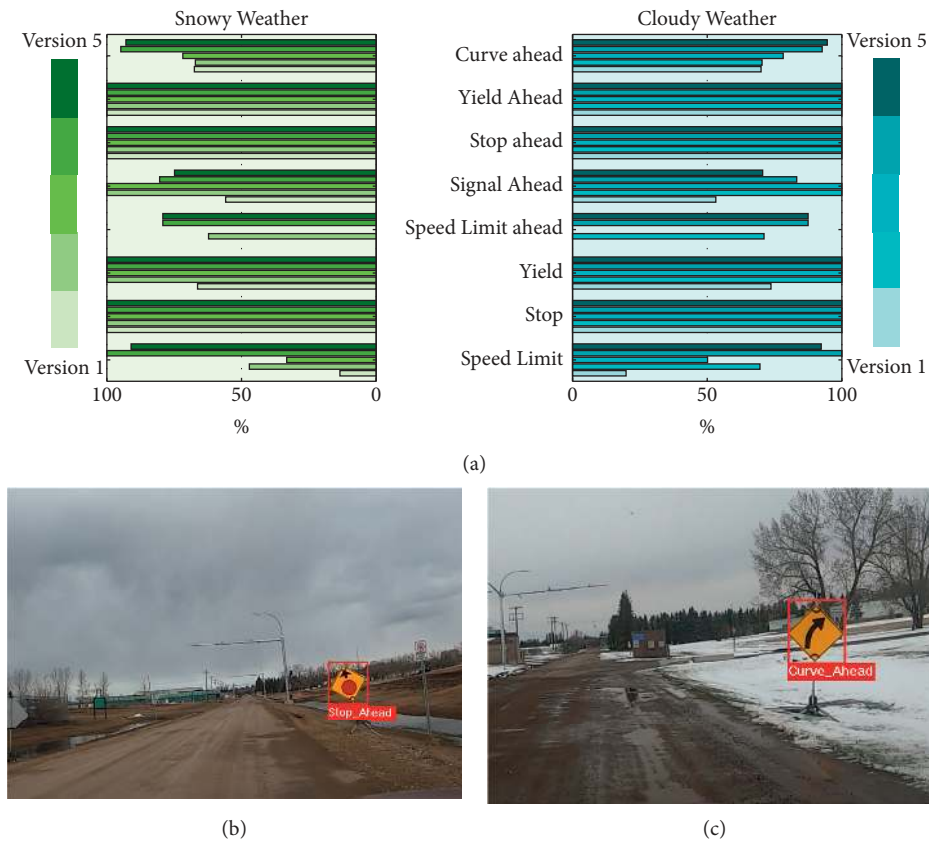
(a)

(b)

(c)

Figure 5: (a) Performance comparison in studied inclement weather scenarios and sample sign identification by the TSR system in (b) cloudy and (c) snowy weather.

Figure 6: Share of samples according to sign age, sheeting type, and lighting conditions.

The performed analysis is summarized in Figure 7, which shows enhanced performance for the newer and diamond grade signs in night conditions. Figure 7(a) shows a comparison of recall parameters for both sheeting types obtained from different versions of the developed TSR system. A significance test on the recall metric for this part of the analysis showed a significant difference at a 95% confidence level for all versions of the developed TSR system, irrespective of sign class. Further analysis was performed to associate the influence of lighting on sign recognition for the two distinct sheeting types. In general, diamond grade signs were found to be similarly identifiable as engineering-grade signs in daylight (Figure 7(b)).
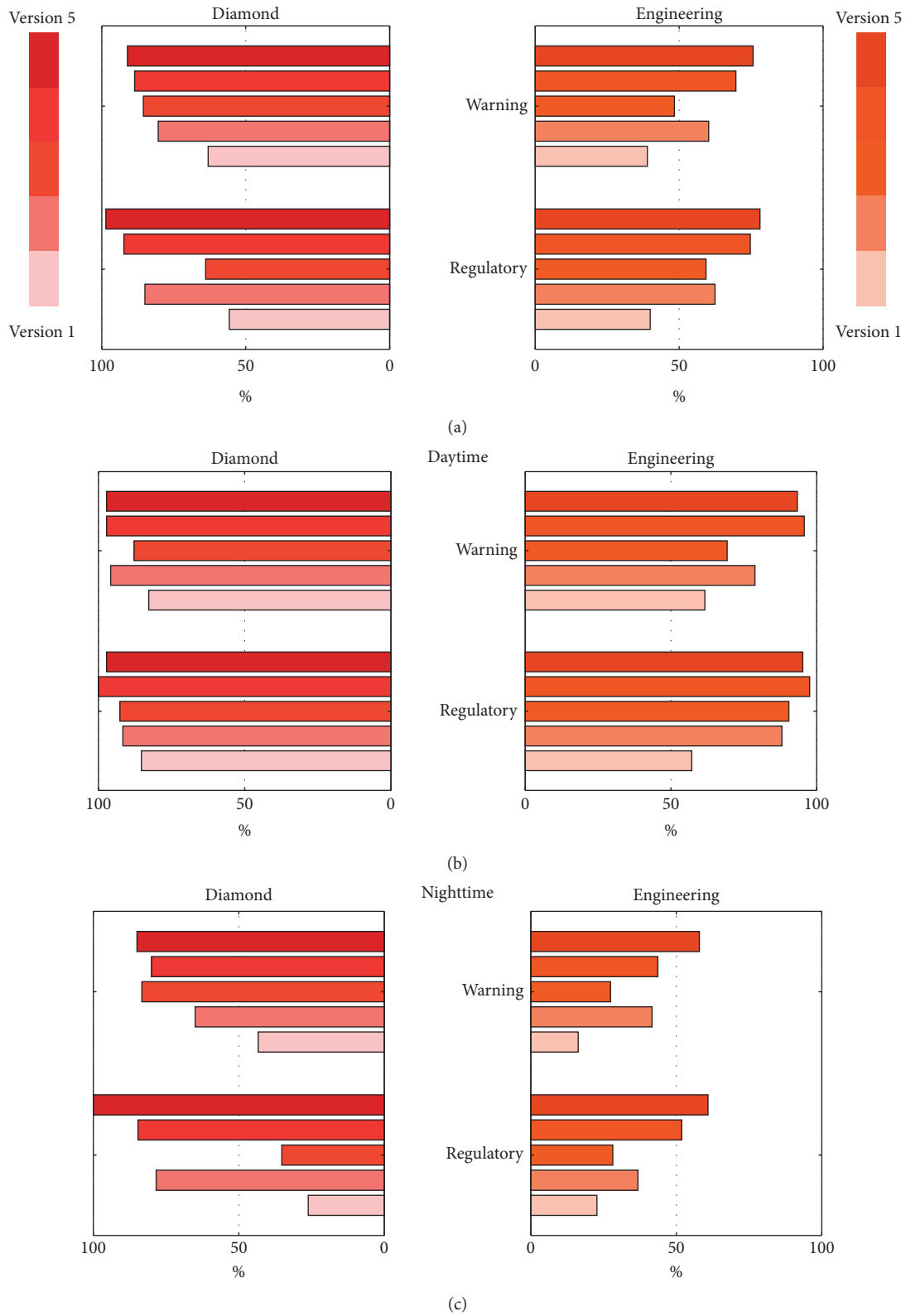
Figure 7: Recall parameter comparison between (a) sheeting types, (b) sheeting types with daytime lighting, and (c) sheeting types with night time lighting.
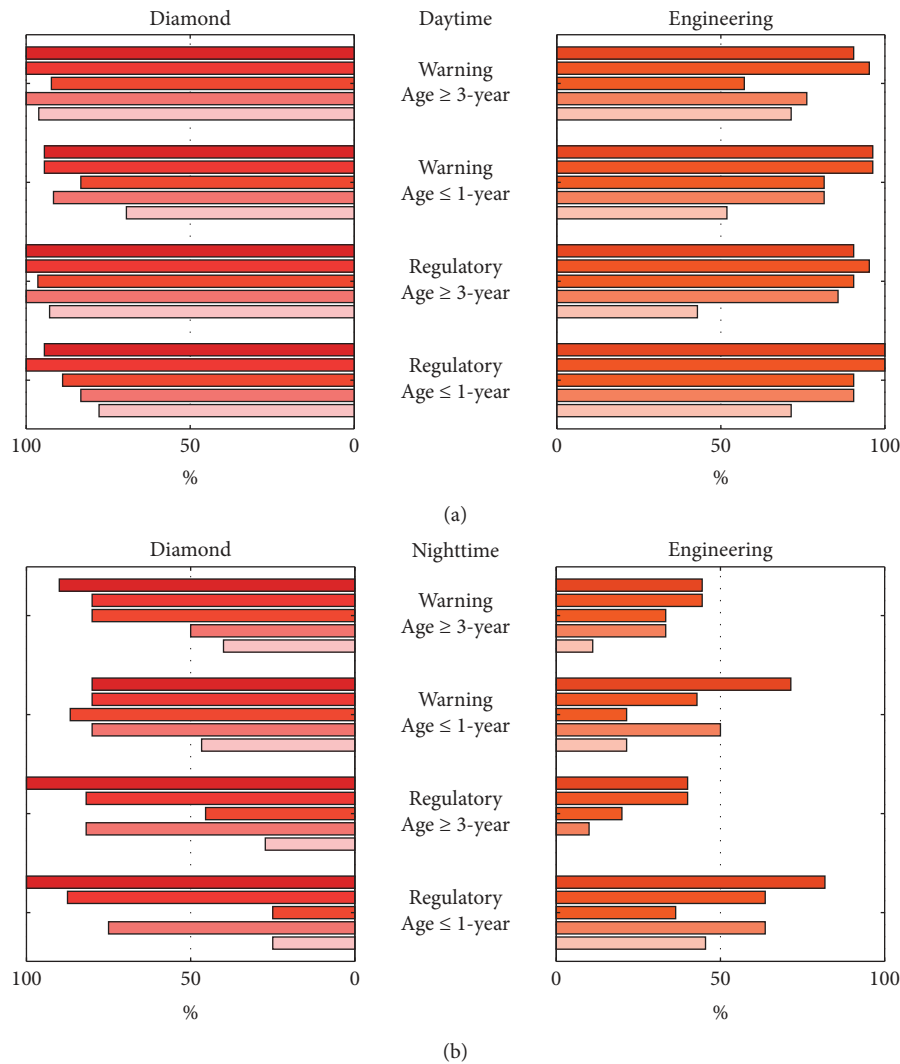
(a)

(b)

FIGURE 8: Recall parameter comparison between (a) sheeting types in daytime lighting and (b) sheeting types in night time lighting with different age groups.

Version 1 (Z value = 2.28) for regulatory signs and version 1 (Z value = 2.23), 2 (Z value = 2.01), and 3 (Z value = 2.08) for warning signs showed significantly higher levels of recognition for diamond grade signs during daytime. However, at night, diamond grade signs showed significantly enhanced recognizability for both sign classes in all versions (Figure 7(c)).

Figures 8 outlines daytime and night-time performance, respectively, via sign class-specific recall values in each version of the developed TSR system for the two sheeting type scenarios under scrutiny and further divided by sign age. At first glance, the daylight recall results seem to outperform night-time results, although both types of sheeting seem to produce a comparable outcome with slight favouring of diamond grade signs. Generally, the significance test results showed a significant difference between the two sheeting types, irrespective of sign class and age, for an earlier version of the TSR system at daylight condition (Figure 8(a)). More evolved forms of the TSR system varied less in responses to sign age. A striking difference in sign recognition performance was observed at night. For night-time samples (Figure 8(b)), the

diamond grade signs provided all versions with a significantly higher recall rate. The comparison between different ages of engineering-grade signs showed significantly higher levels of recognition of new signs (age ≤1 year) as compared to older signs (age ≥3 years) in both sign classes (i.e., regulatory and warning). On the contrary, sign recognition was less sensitive to age for both sign classes when using diamond grade sheeting. This part of the analysis proved beyond a reasonable doubt that sheeting types of signs could play a significant role in recognition by the TSR system, exclusively during the night. Furthermore, older engineering-grade signs were more likely to be missed by a well-developed TSR system at night, in comparison to new engineering-grade signs, diamond grade signs of both age groups, and sign images taken in daylight.

8. Conclusion and Future Research

This study has contributed to the development of a new, robust automatic TSR system through the unique integration of industry-ready technology with an experimental

investigation of the TSR system's limitations and opportunities. The primary objective of the research was to study the influence of weather and reflectivity variations on the TSR system's performance, particularly as the system was improved consistently through the gradual increase in the number of training images. The results from the recognition performance analysis found that the impact of reflectivity conditions was far more significant than that of weather variations. However, scenarios with high reflectivity, fair weather, and generous lighting proved to perform better than their counterparts. Comprehensive significance testing was conducted on the evaluation parameters to identify statistically significant differences between the compared scenarios. This study also revealed that, even if the TSR system was trained with sign images from amenable natural conditions, it could attain a relatively comparable level of recognition based on imperfect real-world conditions. Furthermore, this research provides a benchmark for the resources required to train a TSR system of similar architecture with a specific level of accuracy. Our study not only developed a real-time TSR system with the capability of recognizing several sign types situated in real-world scenarios with varying environmental conditions but it also specified the scope and restrictions of the established system.

While this study focused primarily on the performance variations caused by changes of two factors (i.e., weather and reflectivity), the implications of other factors such as recognition distances, deterioration due to damage, and other exogenous variations will be explored in future studies. We believe that this study will assist in the sustainable and consistent growth of TSR system development as a part of advanced driver assistance systems. In the broader field, interested industry partners can work towards overcoming the identified limitations of TSR systems, while researchers can identify additional applications of this technology to improve traffic operation and the safety of all road users.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] Federal Highway Administration, *Manual on Uniform Traffic Control Devices for Streets and Highways*, Federal Highway Administration, Washington, D.C., USA, 2012.

[2] U.S. Department of Transportation, *Preparing for the Future of Transportation: Automated Vehicle 3.0*, U.S. Department of Transportation, Washington, D.C., USA, 2018.

[3] L. Yang, K. Kwon, K. Moon, S. Lee, and S. Kwon, "Broken traffic sign recognition based on local histogram matching," in *Proceedings of the 2012 Computing, Communications and Applications Conference*, Hong Kong, China, January 2012.

[4] S. Maldonado-Bascón, S. Lafuente-Arroyo, P. Siegmann, H. Gómez-Moreno, and F. J. Acevedo-Rodríguez, "Traffic sign recognition system for inventory purposes," in *Proceedings of the 2008 IEEE Intelligent Vehicles Symposium*, pp. 590–595, Eindhoven, Netherlands, June 2008.

[5] B. Johansson, "Road sign recognition from a moving vehicle," Master thesis, Uppsala University, Uppasala, Swedan, 2002.

[6] B. Höferlin and K. Zimmermann, "Towards reliable traffic sign recognition," in *Proceedings of the 2009 IEEE Intelligent Vehicles Symposium*, pp. 324–329, Xi'an, China, June 2009.

[7] J. Abele, C. Kerlen, S. Krueger, H. Baum, T. Geißler, and S. Grawenhoff, "Exploratory study on the potential socio-economic impact of the introduction of intelligent safety systems in road vehicles," Final report, Publications Office of the European Union, Luxembourg, 2005.

[8] J. Greenhalgh and M. Mirmehdi, "Real-time detection and recognition of road traffic signs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1498–1506, 2012.

[9] W. Gang, K. S. Ishwar, and S. Susanta, "Traffic sign detection and recognition for safe driving," *Intelligent Engineering Systems Through Artificial Neural Networks*, vol. 9, pp. 821–826, 1999.

[10] H. Guan, W. Yan, Y. Yu, L. Zhong, and D. Li, "Robust traffic-sign detection and classification using mobile LiDAR data with digital images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 5, pp. 1715–1724, 2018.

[11] J. Miura, T. Kanda, S. Nakatani, and Y. Shirai, "An active vision system for on-line traffic sign recognition," *IEICE - Transactions on Info and Systems*, vol. 85, no. 11, pp. 1784–1792, 2002.

[12] 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 *Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN)*, Dallas, TX, USA, August 2013.

[13] M. Gu and Z. Cai, "Traffic sign recognition using dual tree-complex wavelet Transform and 2D independent component analysis," in *Proceedings of the 10th World Congress on Intelligent Control and Automation*, Beijing, China, July 2012.

[14] 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 Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 264–278, 2007.

[15] M. Soilán, B. Riveiro, J. Martínez-Sánchez, and P. Arias, "Traffic sign detection in MLS acquired point clouds for geometric and image-based semantic inventory," *ISPRS*

Journal of Photogrammetry and Remote Sensing, vol. 114, pp. 92–101, 2016.

[16] H. Kamada, S. Naoi, and T. Gotoh, "A compact navigation system using image processing and fuzzy control," in *IEEE Proceedings on Southeastcon*, New Orleans, LA, USA, April 1990.

[17] R. Janssen, W. Ritter, F. Stein, and S. Ott, "Hybrid approach for traffic sign recognition," in *Proceedings of the Intelligent Vehicles'93 Symposium*, Tokyo, Japan, July 1993.

[18] H. Gomez-Moreno, S. Maldonado-Bascon, P. Gil-Jimenez, and S. Lafuente-Arroyo, "Goal evaluation of segmentation algorithms for traffic sign recognition," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 917–930, 2010.

[19] S. Tominaga, "Color image segmentation using three perceptual attributes," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 628–630, Miami Beach, FL, USA, June 1986.

[20] R. Ohlander, K. Price, and D. R. Reddy, "Picture segmentation using A recursive region splitting method," *Computer Graphics and Image Processing*, vol. 8, no. 3, pp. 313–333, 1978.

[21] D. M. Gavrila, "Traffic sign recognition revisited," in *Proceedings of the Mustererkennung 1999: 21st DAGM Symposium*, Bonn, Germany, September 1999.

[22] C.-Y. Chiung-Yao Fang, S.-W. Sei-Wang Chen, and C.-S. Chiou-Shann Fuh, "Road-sign detection and tracking," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 5, pp. 1329–1341, 2003.

[23] X. Gao, D. Shaposhnikov, V. Gusakova, K. Hong, and L. Podladchikova, "The foveal system for traffic signs recognition (FOSTS)," in *Proceeding of the 14 International Conference on Neurocybernetics*, Rostov-on-Don, Russia, 2005.

[24] G. Bae, J. Ha, J. Y. Jeon, S. Y. Jo, and H. Jeong, "LED traffic sign detection using rectangular hough transform," in *Proceedings of the 2014 International Conference on Information Science & Applications (ICISA)*, Seoul, Republic of Korea, May 2014.

[25] M. A. Garcia-Garrido, M. Ocana, D. F. Llorca, M. A. Sotelo, E. Arroyo, and A. Llamazares, "Robust traffic signs detection by means of vision and V2I communications," in *Proceedings of the 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Washington, DC, USA, October 2011.

[26] R. Belaroussi and J. Tarel, "Angle vertex and bisector geometric model for triangular road sign detection," in *Proceedings of the 2009 Workshop on Applications of Computer Vision (WACV)*, Snowbird, UT, USA, December 2009.

[27] S. KumarSaha, D. Chakraborty, and M. Al-Amin Bhuiyan, "Neural network based road sign recognition," *International Journal of Computer Applications*, vol. 50, no. 10, pp. 35–41, 2012.

[28] R. Rajesh, K. Rajeev, K. Suchithra, V. P. Lekhesh, V. Gopakumar, and N. K. Ragesh, "Coherence vector of oriented Gradients for traffic sign recognition using neural networks," in *Proceedings of the 2011 International Joint Conference on Neural Networks*, San Jose, CA, USA, July-August 2011.

[29] F. Zaklouta and B. Stanciulescu, "Warning traffic sign recognition using a HOG-based K-d tree," in *Proceedings of the 2011 IEEE Intelligent Vehicles Symposium (IV)*, Baden-Baden, Germany, June 2011.

[30] A. Z. Kouzani, "Road-sign identification using ensemble learning," in *Proceedings of 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, June 2007.

[31] J. Greenhalgh and M. Mirmehdi, "Traffic sign recognition using MSER and random forests," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, August 2012.

[32] P. Gil-Jiménez, S. Lafuente-Arroyo, S. Maldonado-Bascón, and H. Gómez-Moreno, "Shape classification algorithm using support vector machines for traffic sign recognition," in *Proceedings of the 8th international conference on Artificial Neural Networks: computational Intelligence and Bioinspired Systems*, Barcelona, Spain, June 2005.

[33] 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," *Journal of Visual Communication and Image Representation*, vol. 17, no. 4, pp. 675–685, 2006.

[34] Y. Xie, L. Liu, C. Li, and Y. Qu, "Unifying visual saliency with HOG feature learning for traffic sign detection," in *Proceedings of the 2009 IEEE Intelligent Vehicles Symposium*, Xi'an, China, June 2009.

[35] V. A. Prisacariu, R. Timofte, K. Zimmermann, I. Reid, and L. Van Gool, "Integrating object detection with 3D tracking towards a better driver assistance system," in *Proceedings of the 2010 20th International Conference on Pattern Recognition*, Istanbul, Turkey, August 2010.

[36] Y. Huang, Y. Le, and F. Cheng, "A method of detecting and recognizing speed-limit signs," in *Proceedings of the 2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Piraeus-Athens, Greece, July 2012.

[37] A. Hechri and A. Mtibaa, "Lanes and road signs recognition for driver assistance system," *IJCSI international journal of computer science*, vol. 8, no. 6, 2011.

[38] Y. Sheng, K. Zhang, C. Ye, C. Liang, and J. Li, "Automatic detection and recognition of traffic signs in stereo images based on features and probabilistic neural networks," in *Proceedings of the SPIE Photonics Europe*, Strasbourg, France, April 2008.

[39] L. Li, J. Li, and J. Sun, "Robust traffic sign detection using fuzzy shape recognizer," in *Proceedings of the Sixth International Symposium on Multispectral Image Processing and Pattern Recognition*, M. Ding, B. Bhanu, F. M. Wahl, and J. Roberts, Eds., p. 74960Z, Yichang, China, November 2009.

[40] M. Sajjad, M. Irfan, K. Muhammad et al., "An efficient and scalable simulation model for autonomous vehicles with economical hardware," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1718–1732, 2021.

[41] Z. Wang, J. Wang, Y. Li, and S. Wang, "Traffic sign recognition with lightweight two-stage model in complex scenes," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–11, 2020.

[42] T.-Y. Lin, M. Maire, S. Belongie et al., "Microsoft COCO: common objects in context," in *Proceedings of the 13th European Conference on Computer Vision -- ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., pp. 740–755, Springer International Publishing, Zurich, Switzerland, September 2014.

[43] K. Muhammad, A. Ullah, J. Lloret, J. D. Ser, and V. H. C. de Albuquerque, "Deep learning for safe autonomous driving: current challenges and future directions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4316–4336, 2021.

[44] F. Larsson and M. Felsberg, "Using fourier descriptors and spatial models for traffic sign recognition," in *Proceedings of the 17th Scandinavian Conference on Image Analysis*, A. Heyden and F. Kahl, Eds., pp. 238–249, Springer, Ystad, Sweden, May 2011.

[45] M. A. Garcia-Garrido, M. A. Sotelo, and E. Martin-Gorostiza, "Fast traffic sign detection and recognition under changing lighting conditions," 2006.

[46] K. T. Phu and L. Lwin Oo, "Traffic sign recognition system using feature points," in *Proceedings of the 2018 12th International Conference on Research Challenges in Information Science (RCIS)*, pp. 1–6, Nantes, France, May 2018.

[47] K. Lim, T. Lee, C. Shin, S. Chung, Y. Choi, and H. Byun, "Real-time illumination-invariant speed-limit sign recognition based on a modified census transform and support vector machines," in *Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication-ICUIMC'14*, pp. 1–5, Siem Reap, Cambodia, January 2014.

[48] M. Hassaballah, M. A. Kenk, K. Muhammad, and S. Minaee, "Vehicle detection and tracking in adverse weather using a deep learning framework," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4230–4242, 2021.

[49] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018.

[50] Tzutalin. LabelImg. Git Code. https://github.com/tzutalin/labelImg. Accessed Jun. 6, 2018.

[51] J. D. Redmon, "Open source neural networks in C," 2018, https://pjreddie.com/darknet/.

[52] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," 2009.

[53] A. Halstuch and Y. Yitzhaky, "Properties of light reflected from road signs in active imaging," *Applied Optics*, vol. 47, no. 22, p. 4031, 2008.