# Multi-stage Deep Learning Technique for Improving Traffic Sign Recognition


Pubudu Sanjeewani
*Centre for Intelligent Systems*
*Central Queensland University*
Brisbane, Australia
p.thihagodagamage@cqu.edu.au

Brijesh Verma
*Centre for Intelligent Systems*
*Central Queensland University*
Brisbane, Australia
b.verma@cqu.edu.au

Joseph Affum
*Transport Safety*
*Australian Road Research Board (ARRB)*
Brisbane, Australia
joseph.affum@arrb.com.au



***Abstract*—The automatic detection and interpretation of all roadside signs with text information is a very challenging and difficult problem. The main challenge is to distinguish signs with important information which are similar in shapes, sizes, and colors. In this paper, we propose a multi-stage deep learning-based technique combined with Optical Character Recognition (OCR) for automatic detection and recognition of text in speed limit traffic signs. The proposed technique consists of novel concept in which deep learning detection and OCR interpretation of attributes is introduced. The technique has been evaluated on real world dataset. A comparative analysis of results using the proposed technique has been performed and it has shown a significant improvement in speed sign recognition accuracy and misclassifications.**



***Keywords*—*deep learning, speed limit sign recognition, fully convolutional neural network (FCN), optical character recognition***


## I. INTRODUCTION

Detecting and recognizing traffic signs automatically and information written on them is a challenging problem and an essential part of road environment with a number of important application areas [1]-[2], including advanced driver assistance systems, road surveying, and autonomous vehicles. Many traffic signs carry information such as speed limits, road conditions and warnings such as roadworks, school zones, railway stops, etc. which are very useful for road surveys, autonomous vehicles, and road users.

Researchers have been investigating various challenges in traffic sign and information recognition and have achieved some good results in identifying individual signs. Even though there exists much research in detecting and recognizing both symbol-based traffic signs [3]-[4] and recognizing text in real scenes [5]-[7], a smaller number of research has been specifically conducted on recognizing text on traffic information signs [8]-[9]. Mainly, the research is categorized as region-based method and Connected Component (CC)-based methods [5]-[7]. Local features such as texture are used in region-based text detection methods [4] to locate text regions while information such as intensity, colour distribution, and edges are used by CC-based methods to segment text characters individually. Most existing state-of-the-art methods includes both detection and recognition [10]-[11]. Meanwhile, some researchers used advanced deep learning-based techniques. Tabernik et al. [12] proposed a system based on Convolutional Neural Network (CNN) and mask R-CNN for detecting and recognizing traffic-signs in wide range with better performance for many traffic-sign categories, including several complicated ones with large intra-class variability. Lee et al. [13] proposed traffic sign detection and boundary estimation using CNN and obtained robust results on a low power mobile platform. Avramovic et al. [14] proposed a neural-network-based traffic technique for detecting and recognizing signs in high-definition images using region focusing and parallelization. Bangquan et al. [15] proposed a real-time embedded traffic sign recognition using efficient CNN and achieved 98.6% accuracy on the German Traffic Sign Recognition Benchmark (GTSRB). Some researchers used OCRs for text recognition tasks in various applications. Čakić, et al. [16] used OCR for number recognition for food tracking and tracing. But ended up in failure when presented with images that contain shadows. Solanki et al. [17] used OCR to detect and recognize traffic signs and obtained promising results for test images taken in different time intervals and different weather conditions. Zhu et al. [18] proposed a cascaded segmentation-detection networks for text-based traffic sign detection using FCN to segment candidate traffic sign areas providing candidate Region of Interests (RoIs) followed by a fast neural network to detect texts on the extracted RoIs and achieved good accuracy.

The main purpose of our research is to develop a novel technique which can improve accuracy of speed limit detection in our complex AusRAP (Australian Road Assessment Program) attribute detection system. The goal of AusRAP attribute detection system is to detect and identify all road attributes to assess road safety. The major problem we are currently facing is the misclassification of attributes with text which have similar colours, shapes, and sizes. To reduce misclassifications and improve accuracy, we propose a novel technique based on deep learning segmentation and OCR interpretation for speed sign recognition. The main original contributions of this paper are as follows:

1. A multi-stage deep learning-based technique for improving speed sign recognition accuracy in our AusRAP safety attribute detection system is proposed.
2. Novel concept of introducing OCR combined with deep learning to detect text on speed signs is proposed and investigated.
3. A large number of experiments have been conducted on real-world video data which has been obtained from our industry partner. The dataset consists of all state roads in Queensland.

4. A comparative analysis of experimental results is presented.

This research paper is organized as follows. Section II describes the proposed technique. Section III presents experiments while results and discussion are presented in section IV. A comparative analysis is presented in section V. Finally, section VI summarizes the research with conclusion and future work.

## II. Proposed Technique

An overview of the proposed technique is shown in Fig. 1. The proposed technique consists of multistage detection and recognition technique. In detection stage, deep learner model is trained on all AusRAP attributes. In this stage, all speed limit signs are identified as a single attribute. In recognition stage, RoIs are obtained from input images based on detected pixels in detection stage. RoIs are fed to OCR which outputs actual speed limits. The recognized speed limits are used to replace speed classes in predicted frames and final output frames with recognized attributes are produced. Each attribute contains different color pixels for visualization purposes.

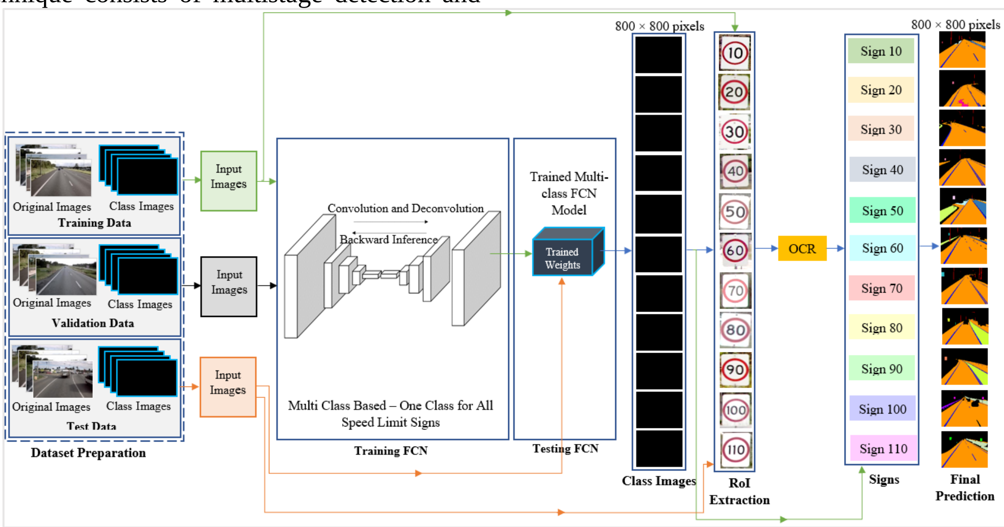


Fig. 1. Architecture of the proposed technique.

As illustrated in Fig. 1, the proposed technique has several steps: data preparation, FCN training, FCN testing, RoI extraction and OCR classifier.

### A. Dataset Preparation

The training, validation and test images used in this research was created by using Queensland Road videos which were provided by our industry partners, the Department of Transport and Main Roads (DTMR), Queensland and Australian Road Research Board (ARRB). Three custom datasets for training, validation and testing have been created in the data preparation stage. Each of these datasets comprises of original resized images and annotated resized images converted to class images. Images were annotated using Adobe Photoshop CC 2019 software and annotated images were converted to class images using a MATLAB program. All images were resized to 800 × 800 pixels using a MATLAB program.

Altogether, there are 61 AusRAP attributes to be identified. The attributes include different types of road safety attributes such as, road markings, road, defects on road, traffic signs, roadside objects such as barriers, deep drainage ditches, rigid structures, aggressive vertical faces, boulders, unprotected barrier ends, pedestrian fencing, trees, poles, guideposts, flexiposts, sidewalks, medians, dedicated bicycle paths, pedestrians, bicycles, motorcycles, and, etc. The traffic signs identified in this research include, chevron alignment marker, railway passive sign, railway stop signage, pedestrian crossing sign, merge lane sign, intersection legs-3 sign, intersection legs-4 sign, curvature sign, roundabout sign, one way sign, minor roadworks sign, major roadworks sign, signed shared roadway, school zone warning sign, slip lane sign, and all speed limit signs. Among these 61 attributes, 11 attributes were speed limit signs. All speed limit sign boards have speed limit text written in black color around a red color circle in white background. The shapes of all the speed limit sign boards are all rectangular.

All 61 attributes in the dataset were annotated using separate 61 RGB colors carefully selected by a MATLAB program avoiding colors present in the background. Each attribute was assigned a separate class and all the background information were included in a class called "unlabeled" which means, altogether there were 62 classes. Annotated images have been included in test dataset to compare the Ground Truth (GT) with the prediction results and calculate the prediction accuracy for comparative analysis. In this research, the training dataset consisted of 549 images while test dataset consisted of 106 images. For each attribute, minimum 20 objects were included in the training dataset. For each experiment, same dataset was used for testing and to obtain the final prediction.

### B. Training FCN

The deep learning model is based on FCN architecture. It consists of 800 × 800 × 3 inputs and same number of outputs. It uses the pre-trained model: imagenet-vgg-verydeep-19 beta 16, number of convolutional layers having similar padding, rectified linear activation function, max pooling for down sampling and Adam optimizer using the backpropagation algorithm.

In training stage, FCN was trained using one class for all speed limit signs. The network parameters such as number of iterations, layers, etc. have been optimized experimentally to achieve highest accuracy. Altogether, there were 61 attributes in the system. But the number of classes were reduced to 52 with the use of one class for all speed limit signs. FCN models were automatically generated incrementally and at each 5,000 iterations. Validation accuracy was calculated for each model using validation dataset. Out of the 32 models auto generated at the end of 160,000 iterations, the model having the highest accuracy and lowest misclassifications was selected as the best model for OCR stage.

### *C. Testing FCN*

In testing/prediction stage, the best trained model from training stage is evaluated on test data. The testing stage consists of many stages: determination of RoIs where the speed limit sign is anticipated to be located, detection of all possible candidates within these regions, applying contextual constraints to reduce candidates, interpretation by OCR and production of final prediction results combining all the 61 attributes. The prediction images generated by best FCN model were black color class images with class numbers as prediction. From these class images, the regions detected as having the speed limit class were extracted and used to extract the same regions in their corresponding original images. An appropriate threshold value was applied before extracting RoI images to reduce large number of detected candidate regions with noise, unlikely candidates, and smaller dots like objects identified as speed limit signs. These images were saved as RoI images. When the speed limit sign has been located, the next stage of the algorithm is to recognize text within the located region. All the RoI images extracted in this way were then fed to an OCR for interpretation. In OCR, pre-processing techniques were applied to RoI images to remove noise and isolate the speed limit text from the background. Then, the speed limit was read and the signs in the prediction class images were updated with the recognized speed limit sign. If OCR detects the RoI image not as a speed limit sign, that region was updated as unlabeled class. The process within OCR is described in detail in section D below. Fig. 2 below shows some of the sample RoIs which were input to OCR.

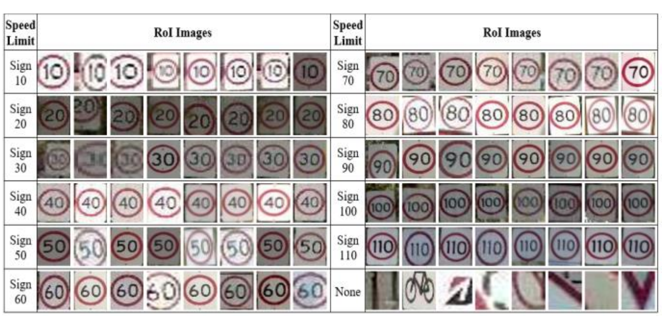


Fig. 2. RoI images as input to OCR.

### *D. OCR Classifier*

In OCR stage, RoIs extracted from input frames based on trained FCN's decision are fed to OCR which outputs recognized speed sign. In this research, even though the RoI image appears as an easy input for OCR to classify, it is a challenging task in real scenario because of the non-uniform background around text which cause poor text segmentation. In this research, the following steps were followed when implementing OCR.

Step 1: Remove edges of the RoI images as they contain part of the background.

Step 2: Increase the size of the regions a bit to incorporate additional background pixels around the text characters.

(Eventually, this supports improving the internal heuristics used in determining the polarity of the text on the background. Therefore, RoI images were resized by a specifying scale factor.)

Step 3: Convert RoI images to binary images by replacing all pixels in the input having luminance greater than specified level with 1 (white) and all other pixels with the value 0 (black).

Step 4: Remove the circle around the speed limits.

(For better results by OCR, RoI images were pre-processed, and OCR received particular regions in the RoI image that it should process. In this research, these regions are those regions that just contain the numbers. A connected component analysis was used to isolate the text in RoI images. Almost in all RoI images, there was a red color circle around the text. By applying this analysis, the circle around the speed limit was removed.)

Step 5: Dilate RoI images to grow the text a bit to get a bigger RoI around them.

Step 6: Determine the bounding boxes around the text.

Step 7: Erode the binary image and make the numbers thinner to remove artifacts.

(After eliminating the background variation, the numbers were apparent in the binary image. However, there can be few artifacts surrounding the numbers that might obstruct accurate detection by OCR. Additional pre-processing using morphological reconstruction was used to remove these artifacts and introduce a cleaner image for OCR. To remove these artifacts, the binary image was eroded, and the numbers were thinned.)

Step 8: Remove all the connected components of area lesser than 150 pixels from the binary image.

(There are small, connected regions within the RoI image which are not likely to contain any text. Therefore, all the connected components (objects) that have fewer than 150 pixels area were removed from the binary image, producing another binary image.)

Step 9: Disable the automatic layout analysis using the "TextLayout" parameter to help improving the results.

(If the text is scattered and irregularly located in the background, the heuristics used for document layout analysis within OCR becomes unsuccessful in determining the text blocks within the image, and, consequently, the text recognition becomes faulty. Therefore, the automatic layout analysis is disabled using the "TextLayout" parameter to help improving the results. After applying these pre-processing steps, the numbers are now well segmented from the background.)

Step 10: Constraint OCR to only choose the best equivalents from the set "0123456789".

(Now the processing of RoI image is finished and only characters are visible. However, still there is a possibility to falsely recognize several non-numeric characters in the image as digits. This may take place when a non-digit character closely looks similar to one of the digits. There can be "noise" in the results due to alphabet letters. Therefore, another approach to improve the results was using priori knowledge about the text within the image. In this research, the text we are interested in contains only numeric digits. Therefore, the results are further improved by limit OCR to only choose the best equivalents from the set "0123456789".)

Step 11: Remove any trailing characters, such as white space or new lines using the deblank function.

Step 12: Display the recognized text on the original image.

Images in Fig. 3 show how RoI images are processed in OCR to get the final prediction for some sample speed limit signs.

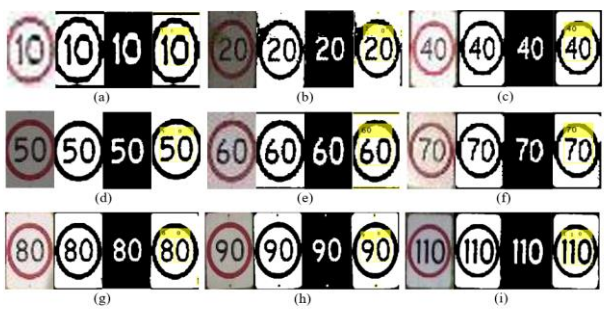

Fig. 3. Prediction results of OCR for sign 10 (a), sign 20 (b), sign 40 (c), sign 50 (d), sign 60 (e), sign 70 (f), sign 80 (g), sign 90 (h), and sign 110 (i).

## III. Experimental Results

Many experiments were performed on High-Performance Computing (HPC) facility for the proposed technique. FCNs have been implemented using Python (version 3.5), TensorFlow in Anaconda environment. As a primary metric, we report classification accuracy attribute wise and pixel wise as a percentage which is commonly used for visual object detection. Pixel wise accuracy in this context is the percentage of correctly classified pixels over the total number of pixels of an object. Attribute wise accuracy is the percentage of correctly classified objects over the total number of objects for an attribute. An object is identified as a prediction based on the majority prediction for the regions belonging to an attribute. Misclassifications were recorded as a percentage of total number misclassified images for an attribute over total number of images in the test dataset. Mainly, two types of experiments (proposed technique with OCR and without OCR) were conducted to evaluate the performance of the proposed technique.

*1) Proposed technique without OCR:* A multi-class model with all separate classes including 11 speed signs (total 61 classes) and accuracy for all the 61 attributes were recorded at intervals of 5,000 iterations. The model having comparatively a higher accuracy as well as lower misclassifications were selected as the best model and this model was saved. Accuracy and misclassifications obtained for speed signs at the best model were recorded as shown in Table I.

*2) Proposed technique with OCR:* A multi-class model was trained for all the 61 attributes by including all the 11 speed limit signs in one class and all the other attributes in separate classes. The model having comparatively a higher accuracy as well as lower misclassifications was selected as the best model and was saved. The best model was obtained at 115,000 iterations whose attribute wise accuracy and pixel wise accuracy of speed sign classes were recorded as 97.5% and 96% respectively. RoI images from corresponding original images were extracted and these RoI images were passed to an OCR to recognize text in speed limit signs. Results obtained were recorded as shown in Table II.

TABLE I. Best Classification Accuracy and Misclassifications Obtianed by Proposed Technique Without OCR

| **Speed Limit** | **Accuracy (%)** | | **Misclassifications (%)** |
|---|---|---|---|
| | ***Attribute wise*** | ***Pixel wise*** | |
| Sign 10 | 50.00 | 48.00 | 0.94 |
| Sign 20 | 33.34 | 13.00 | 1.89 |
| Sign 30 | 0.00 | 0.00 | 0.94 |
| Sign 40 | 33.34 | 28.00 | 0.94 |
| Sign 50 | 0.00 | 0.00 | 0.94 |
| Sign 60 | 66.69 | 36.00 | 2.83 |
| Sign 70 | 14.29 | 12.00 | 3.77 |
| Sign 80 | 100.00 | 76.00 | 0.00 |
| Sign 90 | 50.00 | 50.97 | 0.94 |
| Sign 100 | 75.00 | 46.00 | 0.94 |
| Sign 110 | 50.00 | 36.00 | 0.94 |

## IV. Result Analysis and Discussion

The results have been visualized and analyzed as shown in figures below. Fig. 4 illustrate the best prediction results obtained for some speed limit signs from multi class model (Prediction 1) and multi class-based model with one class for all speed limit signs with OCR (Prediction 2) at 115, 000 iterations. Each prediction result is presented with its original image and GT image for easy comparison. As can be seen from Fig. 4, prediction images from multi class-based experiment with one class for all speed limit signs with OCR almost look similar to the corresponding GT images. Prediction images obtained by multi class experiment without OCR have much noise and many misclassifications.

TABLE II. BEST CLASSIFICATION ACCURACY AND MISCLASSIFICATIONS OBTIANED BY THE PROPOSED TECHNIQUE WITH OCR (ONE CLASS FOR ALL SPEED LIMIT SIGNS)

| Speed Limit | Attribute Wise Accuracy (%) | Misclassifications (%) |
|---|---|---|
| Sign 10 | 50.00 | 0.00 |
| Sign 20 | 100.00 | 0.00 |
| Sign 30 | 100.00 | 0.00 |
| Sign 40 | 100.00 | 0.00 |
| Sign 50 | 50.00 | 0.00 |
| Sign 60 | 83.33 | 0.00 |
| Sign 70 | 42.86 | 0.00 |
| Sign 80 | 50.00 | 0.00 |
| Sign 90 | 100.00 | 0.00 |
| Sign 100 | 50.00 | 0.00 |
| Sign 110 | 100.00 | 0.00 |

When observing Table I, it is seen that sign 10, sign 30, sign 40, sign 60, sign 90, sign 100, and sign 110 obtained 100% attribute wise accuracy with 0% misclassifications for sign 10. Among them, sign 30, sign 40, sign 90, sign 100, and sign 110 have very high pixel wise accuracy equal or above 90% and 75% pixel wise accuracy for sign 60. Sign 70 also have a comparatively higher accuracy (71.44% attribute wise and 77% pixel wise). Among them, sign 20 has the lowest attribute wise accuracy which is 33.34%. As shown in Table II, sign 20, sign 30, sign 40, sign 90, and sign 110 have 100% attribute wise accuracy for multi class-based experiment with one class for all the speed limit signs with OCR. Sign 60 also has a considerably higher attribute wise accuracy which is 83.33%. Sign 10, sign 50, sign 80, and sign 100 got 50% attribute wise accuracy. All the speed limit signs have 0% misclassifications as well.

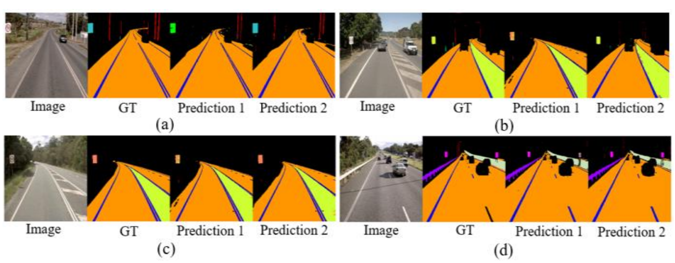


Fig. 4. Prediction results of sign 70 (a), sign 80 (b), sign 90 (c) and sign 100 (d) obtained by the proposed technique with OCR at 115, 000 iterations.

## V. COMPARATIVE ANALYSIS

We have analyzed and compared the results obtained by the proposed technique. Table III shows a comparison of attribute wise accuracy obtained by the proposed technique and DeepLabV3+ [20] without OCR and the proposed technique with OCR. Table IV below shows a comparison of misclassifications obtained by the proposed technique with and without OCR.

When observing Table III, it is seen that attribute wise accuracy of sign 20, sign 30, sign 40, sign 50, sign 60, sign 70, sign 90, and sign 110 have significantly increased with proposed multi-stage deep learning technique with OCR (column 4) when compared with both FCN8 and DeepLabV3+ without OCR. At the same time, accuracy of sign 80, and sign 100 have decreased when compared with both FCN8 and DeepLabV3+ without OCR. The accuracy of sign 10 has significantly improved in the proposed technique with OCR comparative to DeepLabV3+ and remains the same comparative to FCN8. The attribute wise accuracy with proposed technique without OCR is much lower than with OCR (column 2). As shown in Table IV, misclassifications also have decreased to 0% by the proposed technique with OCR. All misclassifications were non-speed sign numbers, so they were not misclassified as any other speed limit sign. This is the main advantage of proposed technique over other existing techniques. Therefore, it is appropriate to conclude that the proposed technique is well suited for improving speed sign accuracy. Fig. 5 shows some sample results obtained from pixels detected as speed limit signs by the segmentation network without (Prediction 1) and with (Prediction 2) OCR. Prediction 1 images have very high misclassification ratio for each segmented speed sign making the decision-making task for the target class very challenging. Sample images in Prediction 2 are shown to have perfect segmentation results eliminating the pixelwise misclassification issue from segmentation network.

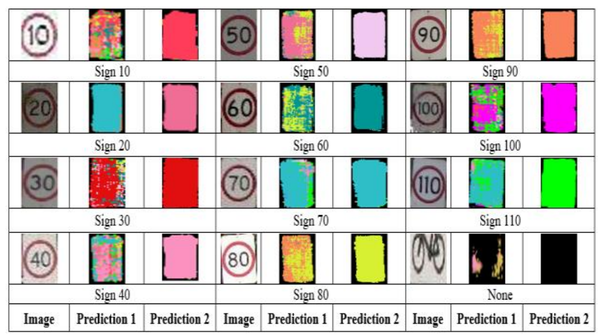


Fig. 5. Sample prediction results output by the proposed technique.

TABLE III. COMPARISON OF CLASSIFICATION ACCURACY (%)

| Speed Limit | Attribute Wise Accuracy (%) Without OCR | | Proposed Technique With OCR |
|---|---|---|---|
| | *FCN8* | *DeepLabV3+ [20]* | *FCN8* |
| Sign 10 | 50.00 | 0.00 | **50.00** |
| Sign 20 | 33.34 | 0.00 | **100.00** |
| Sign 30 | 0.00 | 0.00 | **100.00** |
| Sign 40 | 33.34 | 66.69 | **100.00** |
| Sign 50 | 0.00 | 25.00 | **50.00** |
| Sign 60 | 66.69 | 80.00 | **83.33** |
| Sign 70 | 14.29 | 28.58 | **42.86** |
| Sign 80 | 100.00 | 75.00 | **50.00** |
| Sign 90 | 50.00 | 50.00 | **100.00** |
| Sign 100 | 75.00 | 75.00 | **50.00** |
| Sign 110 | 50.00 | 100.00 | **100.00** |

TABLE IV. COMPARISON OF MISCLASSIFICATIONS

| Speed Limit | Misclassifications (%) | |
|---|---|---|
| | ***Proposed technique without OCR*** | ***Proposed technique with OCR*** |
| Sign 10 | 0.94 | **0.00** |
| Sign 20 | 1.89 | **0.00** |
| Sign 30 | 0.94 | **0.00** |
| Sign 40 | 0.94 | **0.00** |
| Sign 50 | 0.94 | **0.00** |
| Sign 60 | 2.83 | **0.00** |
| Sign 70 | 3.77 | **0.00** |
| Sign 80 | 0.00 | **0.00** |
| Sign 90 | 0.94 | **0.00** |
| Sign 100 | 0.94 | **0.00** |
| Sign 110 | 0.94 | **0.00** |

## VI. CONCLUSION AND FUTURE WORK

The proposed multi-stage deep learning technique for improving speed sign recognition accuracy has been presented and discussed in this research. The proposed technique has been implemented and evaluated on large real-world dataset. Many experiments were conducted to compare the performance of the proposed technique with and without OCR and find the best classification accuracy and minimum misclassifications. The evidence from experiments demonstrated that the proposed technique can produce high accuracy and accurately identify speed limit signs with zero misclassification rates. The overall performance of the proposed technique for many speed limit signs was promising including 100% classification accuracy for most of the speed limit signs. Recognition accuracy was vastly improved by the proposed technique with OCR component. In our future work, we will evaluate our technique on all state roads in Queensland and also extend it to recognize text in other traffic sign boards such as minor and major roadworks signs, school zone warning, one way sign and warning signs.

## ACKNOWLEDGMENT

This research was supported under Australian Research Council's Linkage Projects funding scheme (project number LP170101255).

## REFERENCES

[1] J. Cao, C. Song, S. Peng, F. Xiao and S. Song, "Improved traffic sign detection and recognition algorithm for intelligent vehicles," *Sensors*, vol. 19, 2019.

[2] D. Zang *et al.*, "Deep learning–based traffic sign recognition for unmanned autonomous vehicles," *Proc. Institution of Mech. Engineers, Part I: J. Syst. and Control Eng.*, vol. 232, no. 5, pp. 497-505, 2018.

[3] J. Greenhalgh and M. Mirmehdi, "Real-time detection and recognition of road traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1498–1506, Dec. 2012.

[4] J.-J. Lee, P.-H. Lee, S.-W. Lee, A. Yuille, and C. Koch, "AdaBoost for text detection in natural scene," in *Int. Conf. Document Anal. and Recognit. (ICDAR)*, Beijing, China, 2011, pp. 429–434.

[5] J. Zhang and R. Kasturi, "Character energy and link energy-based text extraction in scene images," in *Asian Conf. Comput. Vision* (*ACCV)*, no. 2, 2010, pp. 308–320.

[6] L. Neumann and J. Matas, "A method for text localization and recognition in real-world images," in *Asian Conf. Comput. Vision* (*ACCV)*, 2010, pp. 9–11.

[7] B. Epshtein, E. Ofek, and Y.Wexler, "Detecting text in natural scenes with stroke width transform," in *Proc. CVPR*, 2010, pp. 2963–2970.

[8] S. Xavier and Reshmi, "Automatic detection and recognition of text in traffic sign boards based on word recognizer," *Int J. for Innovative Res. in Sci. & Technol.*, vol. 3 , no. 4, Sep. 2016.

[9] J. Greenhalgh and M. Mirmehdi, "Recognizing text-based traffic signs," *IEEE Trans. Intell. Transp. Sys.*, vol. 16, no. 3, pp. 1360-1369, June 2015.

[10] A. Reina, R. Sastre, S. Arroyo and P. Jiménez, "Adaptive traffic road sign panels text extraction," in *Proc. WSEAS ICSPRA*, 2006, pp. 295–300.

[11] A. González, L. Bergasa and J. Yebes, "Text detection and recognition on traffic panels from street-level imagery using visual appearance," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 228–238, Feb. 2014.

[12] D. Tabernik and D. Skočaj, "Deep learning for large-scale traffic-sign detection and recognition," *IEEE Trans. Intell. Transp. Sys.*, vol. 21, no. 4, pp. 1427-1440, Apr.2020.

[13] H. S. Lee and K. Kim, "Simultaneous traffic sign detection and boundary estimation using convolutional neural network," *IEEE Trans. Intell. Transp. Sys.*, vol. 19, no. 5, pp. 16652-1663, May 2018.

[14] A. Avramović, D. Sluga, D. Tabernik, D. Skočaj, V. Stojnić and N. Ilc, "Neural-network-based traffic sign detection and recognition in high-definition images using region focusing and parallelization," *IEEE Access*, vol. 8, pp. 189855-189868, 2020.

[15] X. Bangquan and W. Xiao Xiong, "Real-time embedded traffic sign recognition using efficient convolutional neural network," *IEEE Access*, vol. 7, pp. 53330-53346, 2019.

[16] S. Čakić, T. Popović, S. Šandi, S. Krčo and A. Gazivoda, "The use of tesseract ocr number recognition for food tracking and tracing," in *24th Int. Conf. Inf. Technol. (IT)*, Zabljak, Montenegro, 2020, pp. 1-4.

[17] D. Solanki and Gireesh Dixit, "Traffic sign detection and recognition using feature based and OCR method," *IJRSET*, vol. 2, no. 2, pp. 32-40 , 2015.

[18] Y. Zhu, M. Liao, M. Yang and W. Liu, "Cascaded segmentation-detection networks for text-based traffic sign detection," *IEEE Trans. Intell. Transp. Sys.*, vol. 19, no. 1, pp. 209-219, Jan. 2018.

[19] P. Sanjeewani and B.Verma, "Single class detection-based deep learning approach for identification of road safety attributes," *Neural Comput. and Appl.*, pp. 1-12, Feb. 2021.

[20] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Eur. Conf. Comput. Vision (ECCV)*, 2018.