Multi-stage Deep Learning Technique for Improving Traffic Sign Recognition

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

Brijesh Verma
Center 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

INTRODUCTION

- **Problem:** 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.

- **What is Proposed:** In this research, we offer a multi-stage deep learning-based approach combined with Optical Character Recognition (OCR) for automatic identification and recognition of text in speed limit traffic signs. The proposed technique consists of novel concepts in which deep learning detection and OCR interpretation of attributes is introduced. The technique has been evaluated on real world dataset. A comparison of the data obtained utilizing the proposed approach revealed a significant improvement in speed sign recognition accuracy and misclassifications.

- **Data Preparation:** The 549 training images and 106 test images extracted from roadside video data provided by the Department of Transport and Main Roads (DTMR), Queensland, Australia and ARRB have been used to analyze the proposed technique.

PROPOSED TECHNIQUE

The proposed technique consists of a multistage detection and recognition stage. In detection stage, deep learner model is trained on all Australian Road Assessment Program (AusRAP) attributes. All speed limit signs are classified as a single attribute at this stage. In recognition stage, Region of Interests (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.

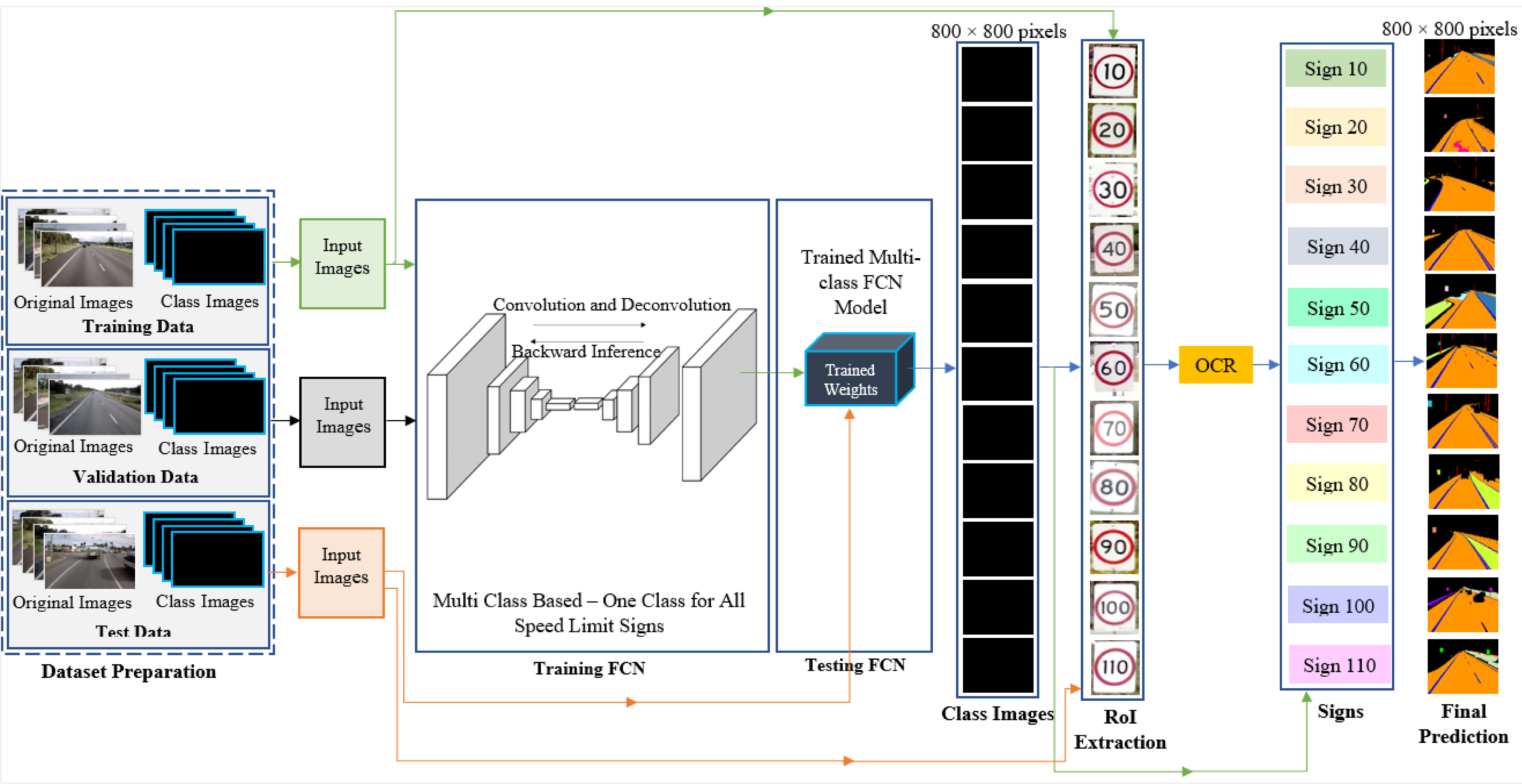

Fig. 1: Architecture of the proposed technique.

Fig. 2 below shows some of the sample RoIs which were input to OCR. Fig. 3 below shows some sample prediction results of OCR.

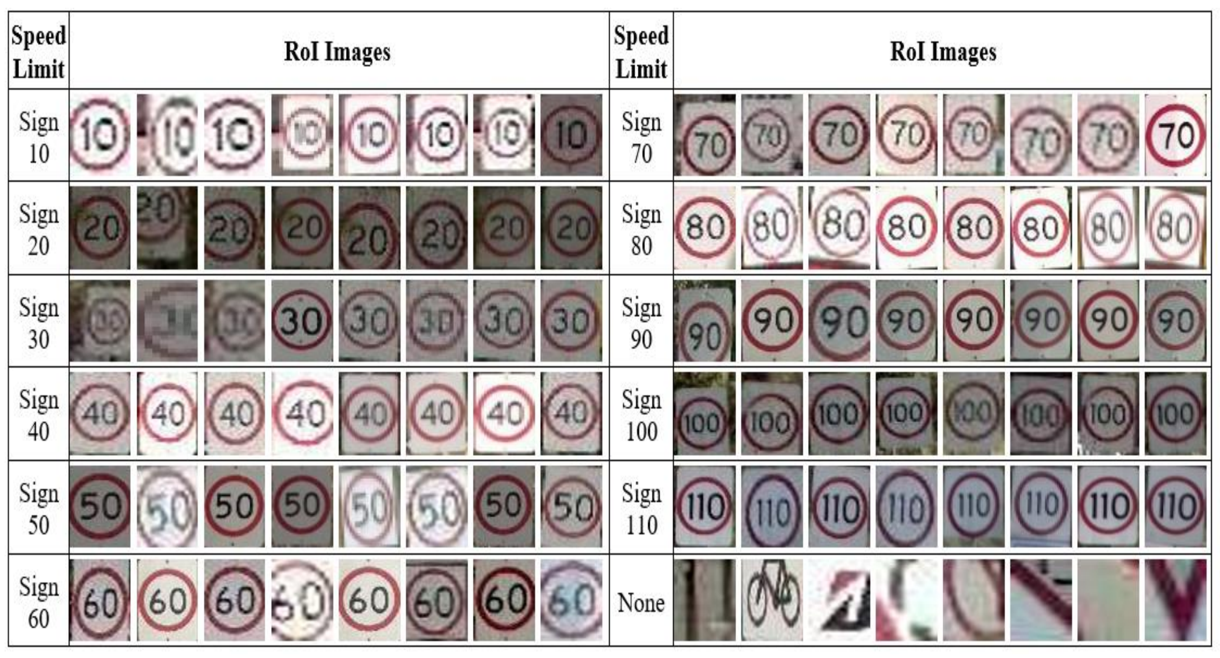

Fig. 2: ROI images input to OCR.

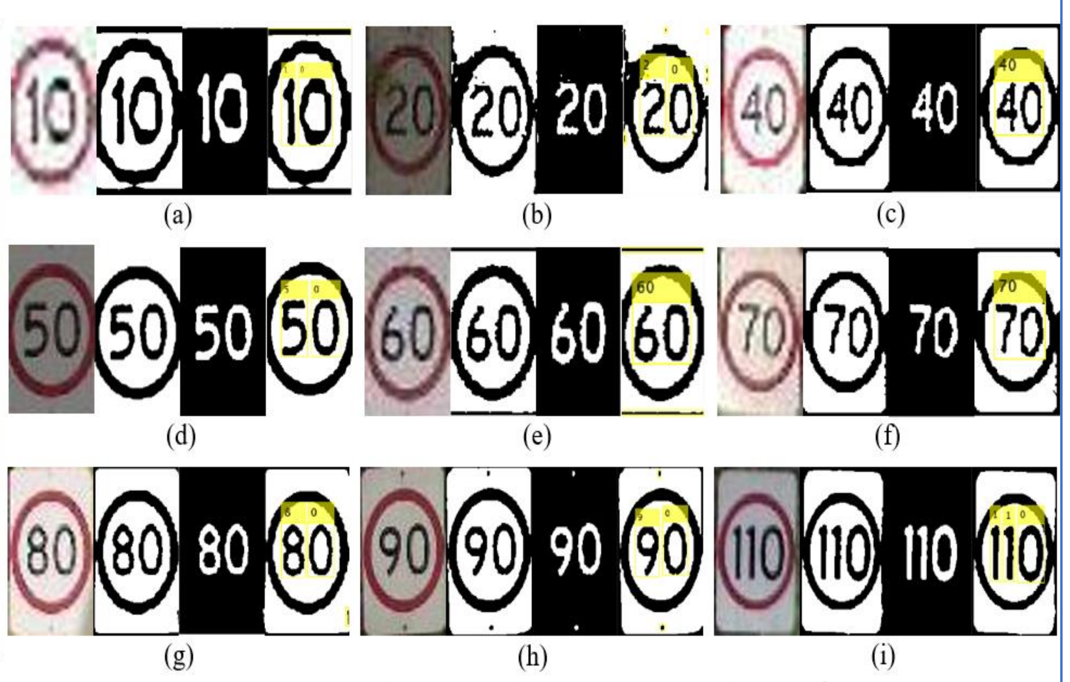

Fig. 3: Sample prediction results of OCR.

EXPERIMENTAL RESULTS

- We have implemented the proposed approach shown in Fig. 1 in Anaconda environment using Python (version 3.5) and TensorFlow. Several experiments were carried out on a High-Performance Computing (HPC) facility to validate the proposed technique.

- Two types of experiments (proposed technique with OCR and without OCR) were conducted to assess the effectiveness of the proposed approach.

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 was selected as the best model and this model was saved.

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. RoI images from corresponding original images were extracted and these RoI images were passed to an OCR to recognize text in speed limit signs.

- Fig. 4 shows the best prediction results obtained for some speed limit signs from a multi class model (Prediction 1) and a multi class-based model with one class for all speed limit signs with OCR (Prediction 2).

- Following the examples with results, we provide empirical evidence of the classification accuracy and misclassifications of the proposed approach.

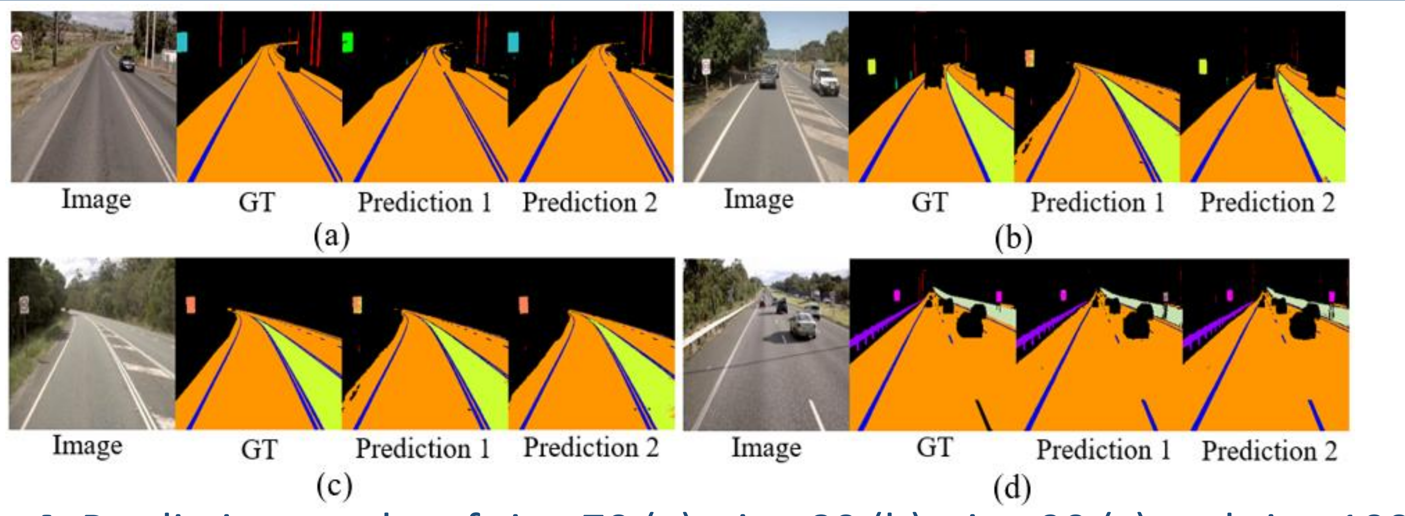

Fig. 4: Prediction results of sign 70 (a), sign 80 (b), sign 90 (c) and sign 100 (d) obtained by the proposed technique with OCR at the best model.

COMPARATIVE ANALYSIS

The results of pixels recognized as speed limit signs by the segmentation network without (Prediction 1) and with (Prediction 2) OCR are shown in Fig. 5.

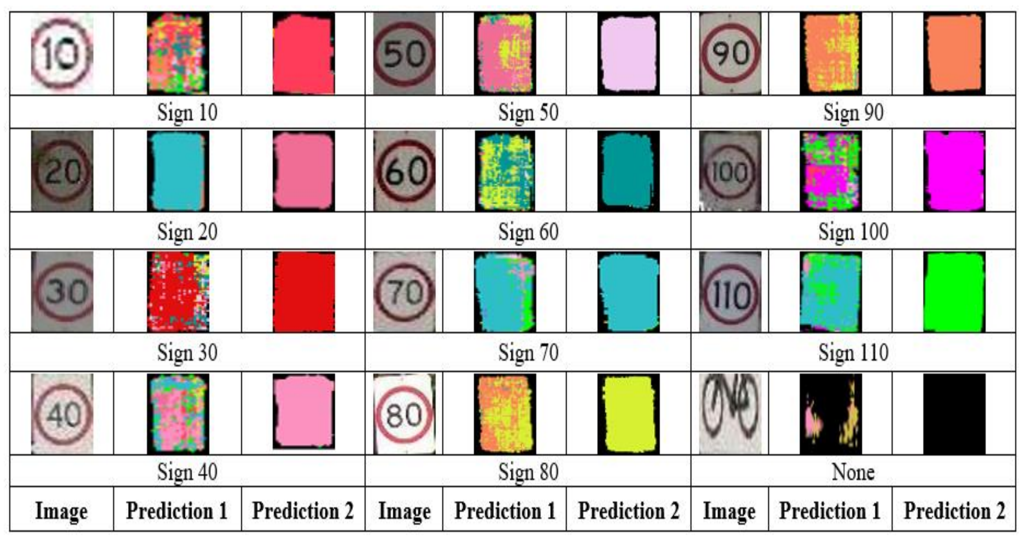

Fig. 5: Sample prediction results output by the proposed technique.

Table 1 compares the attribute-wise accuracy of the proposed approach, DeepLabV3+ [1], and the proposed technique with OCR. Table 2 shows a comparison of misclassifications acquired with and without OCR using the proposed approach.

Table 1: Comparison of Classification Accuracy (%)

Speed Limit	Attribute wise accuracy (%) without OCR		Proposed technique with OCR
	FCN8	DeepLabV3+ [1]	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 2: 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**

DISCUSSION AND CONCLUSION

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 performance of the proposed technique for many speed limit signs was promising overall including 100% classification accuracy for most of the speed limit signs. Recognition accuracy was vastly improved by the proposed technique with OCR component.

REFERENCES

[1] 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.