

Traffic Sign Classification Using Transfer Learning: An Investigation of Feature-Combining Model

Lim Wee Sheng¹, Ahmad Fakhri Ab. Nasir^{2,3}, Mohd Azraai Mohd Razman¹, Anwar P.P. Abdul Majeed^{1,3}, Nur Shazwani Kamarudin², and M. Zulfahmi Toh²

¹Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pekan Pahang, Malaysia.

²Faculty of Computing, Universiti Malaysia Pahang, 26600 Pekan Pahang, Malaysia.

³Centre for Software Development & Integrated Computing, Universiti Malaysia Pahang, 26600 Malaysia.

ABSTRACT – The traffic sign classification system is a technology to help drivers to recognise the traffic sign hence reducing the accident. Many types of learning models have been applied to this technology recently. However, the deployment of learning models is unknown and shown to be non-trivial towards image classification and object detection. The implementation of Transfer Learning (TL) has been demonstrated to be a powerful tool in the extraction of essential features as well as can save lots of training time. Besides, the feature-combining model exhibited great performance in the TL method in many applications. Nonetheless, the utilisation of such methods towards traffic sign classification applications are not yet being evaluated. The present study aims to exploit and investigate the effectiveness of transfer learning feature-combining models, particularly to classify traffic signs. The images were gathered from GTSRB dataset which consists of 10 different types of traffic signs i.e. warning, stop, repair, not enter, traffic light, turn right, speed limit (80km/s), speed limit (50km/s), speed limit (60km/s), and turn left sign board. A total of 7000 images were then split to 70:30 for train and test ratio using a stratified method. The VGG16 and VGG19 TL-features models were used to combine with two classifiers, Random Forest (RF) and Neural Network. In summary, six different pipelines were trained and tested. From the results obtained, the best pipeline was VGG16+VGG19 with RF classifier, which was able to yield an average classification accuracy of 0.9838. The findings showed that the feature-combining model successfully classifies the traffic signs much better than the single TL-feature model. The investigation would be useful for traffic signs classification applications i.e. for ADAS systems.

ARTICLE HISTORY

Received: 2nd June 2021

Revised: 8th July 2021

Accepted: 30th July 2021

KEYWORDS

Traffic Sign Classification

Transfer Learning

VGG Model

Random Forest

Neural Network

INTRODUCTION

Artificial Intelligence (AI) has become a more important technology in our daily life where it is able to replace manpower. Most of the high technology applications are integrated with AI systems such as autonomous vehicles, medical assistance, security vision detection, and other machines. AI is not only used to reduce manpower but it can also help to increase human safety [1]. As road accidents keep on increasing every year, automotive engineers have started to develop the Advanced Driver Assistant System (ADAS) to reduce the accident by helping the driver control the vehicle [2]. In 2004, the World Health Organization (WHO) report of the Global Burden of Disease Project said that 1.27 million people die from road accidents every year. Crashes in road traffic cost most nations 3% of their gross domestic product [3].

To solve this issue, ADAS technology was started to be developed to increase safety on the road. In 2020, there are already a bunch of ADAS technologies where it can help users to detect the obstacle, recognize the traffic sign and auto parking. For traffic sign classification, when the traffic sign is classified by the system, it will give the warning signal to the driver. In a dangerous situation, the ADAS system will control the car immediately to prevent a serious car accident from occurring [4]. For example, when the ADAS system recognizes the bumble sign it will give the information to the driver and in a certain distance the vehicle will automatically reduce the speed [5]. The ADAS system recognizes the bumble sign and automatically slows down the vehicle speed if the driver is not aware the bump is ahead.

Another useful application is traffic sign recognition in road maintenance. Nowadays, to verify the existence and status of the traffic sign board, a human operator needs to watch the video to verify it. This job is very challenging and tedious work because the traffic sign board will appear from time to time, hence the operator needs to be focused and check it one by one. In future, the intelligent vehicles will be depending on the traffic sign board classification for analysis and give the warning signal to the driver [5]. Recognition of traffic signage requires a classification process to determine the meaning of the traffic signage derived using different classification algorithms. Sometimes the system classified the image of the traffic sign board will fail due to the images may include noise, occlusion, undesired background blur and the image pixels [6].

RELATED WORK

From the quick literature survey on traffic sign classification domain, many researchers used the convolution neural network (CNN) to execute both feature extraction and classification. From the study conducted in [7], the researcher implemented the traffic sign recognition by using the CNN, the it was trained by using the German Traffic Sign Recognition Benchmark (GTSRB) dataset with 43 different classes of traffic sign. Four difference models that were tested by adjusting the hyperparameter of the batch size and epochs to test the accuracy of each model. From the obtained results, the fourth model contained the highest accuracy with 95%, where it used the batch size of 64 and 110 of the epochs. The research in [8] uses the GTSRB dataset which contains the highest accuracy (VGG = 99.3% and ResNet = 99.6%) compared with the BTSC dataset (VGG = 98.3% and ResNet = 98.8%). In addition, the GTSRB dataset was used to test with five different deep learning architectures and the ResNet achieved the highest accuracy with 99.6%. The reported result showed that the proposed method can be effectively implemented for real time applications and provide an acceptable accuracy outperforming human performance.

Another research was done by [9], which also uses a GTSRB dataset containing 43 differences of classes. The research was using five machine learning algorithms namely XGBOOST, KNN, Random Forest, SVM and CNN to test the classification accuracy. The results showed that CNN method gave the highest classification accuracy with 97.33% while SVM method had the lowest classification accuracy with 49.08%. Thus, the results show that CNN is the best fit with the highest accuracy for our proposed system when compared to other supervised learning models. Other than that, the researcher in [10] also employed the CNN using the GTSRB dataset. The best obtained classification accuracy was 97.8% using 32 batch size and the 15 epochs parameters. Last but not least, the study in [11] used SVM and Random Forest classifiers. The model tested on the GTRSB dataset and the result showed that the combination of the SVM and Random Forest classifier granted the highest classification accuracy with 98.76% while Random Forest classifier achieved 97.94% and SVM classifier achieved 98.51%.

It is obviously shown that from the literature the implementation of machine learning models towards traffic sign classification is non-trivial. To the best of the authors' knowledge, the capability of a feature-combining Transfer Learning (TL) model for multi-classes traffic sign classification is not yet investigated. In this paper, we describe a traffic sign classification by using several TL models such as VGG16, VGG19 as well as a hybridisation of both models for feature extraction and combined with neural network and random forest classifier. Generally, TL is used in huge processing tasks to reduce the amount of computational power. TL can be more accurate, faster with lesser data training and successfully implemented in several studies for instance in [12-14].

METHODOLOGY

Figure 1 shows the overall process of the research. The dataset used for this research is gathered from the German Traffic Sign Recognition Benchmark (GTSRB). Sample of dataset is shown in Figure 2.

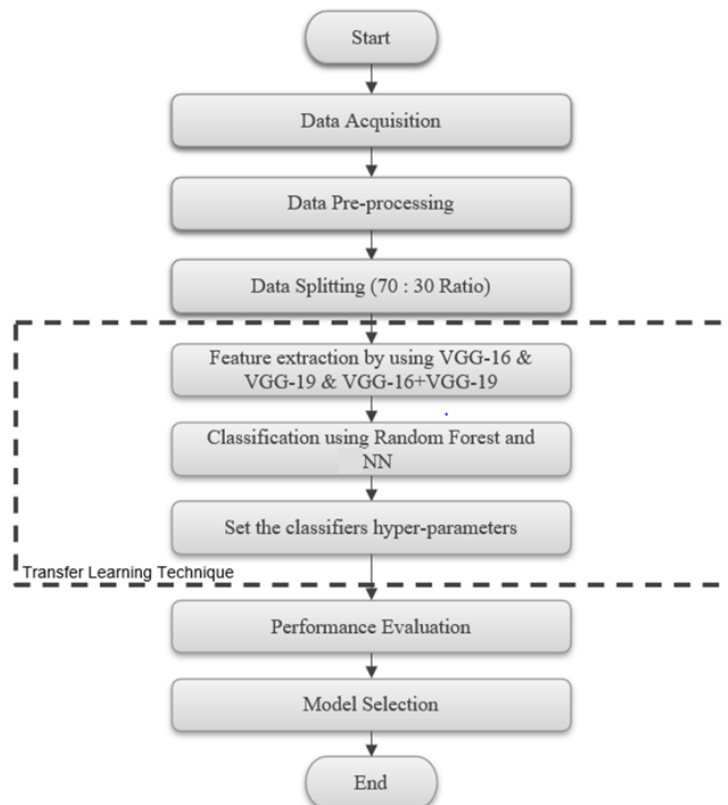


Figure 1. Research flowchart.

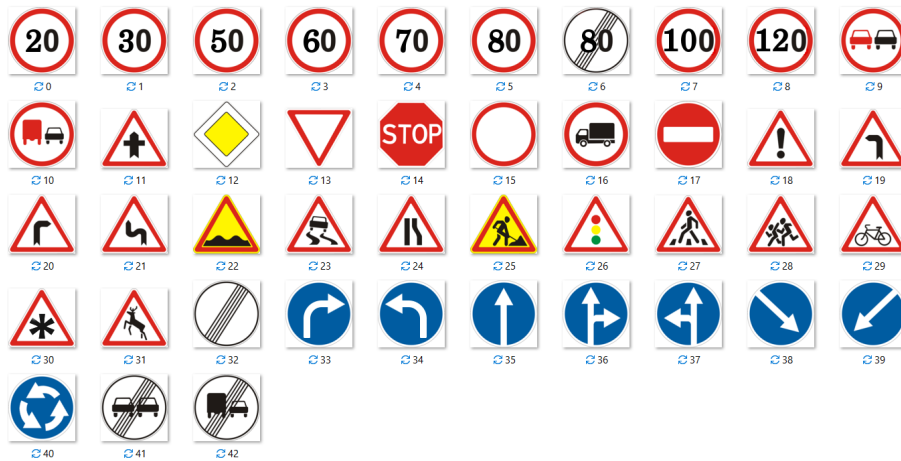


Figure 2. Sample of images used in this study.

The GTSRB dataset contains 50,000 images with 43 classes. The size of the images are between 15x15 pixels to 250x250 pixels. In this study, 10 classes of the traffic sign board were chosen i.e. warning, stop, repair, not enter, traffic light, turn right, speed limit (80km/s), speed limit (50km/s), speed limit (60km/s) and turn left sign board. The total number of images were 7000, 700 images per class.

In the data pre-processing, all the images required to be resized, append the feature, and reset the channel. The size of the image needs to be fixed with the dimension of the 244x244 with the height and width because the input dimension of the used TL models is 244x244. The GTSRB dataset there has a different size of the dataset hence it should be resized. In addition, images need to be appended to the feature and label it to tell the system which images are under which classes. Next, the image channel needs to be set with three channels because input image to the TL models requires it to be with an RGB channel. The dataset was then split into 2 sets for performance evaluation, test and train with 70:30 ratio.

Two TL models were selected for this study namely VGG-16 and VGG-19. VGG-16 contains 16 weight layers, which are 13 convolution layers and three fully connected layers. Meanwhile, VGG-19 contains a network with 19 layers, which is 16 convolution layers and three fully connected layers. Both models worked on convolution layers of 3x3 filter with a phase instead of having a large number of hyper-parameters and using the same padding and max pooling layer of 2x2 stride two filter. The input image dimension size must be (224, 224, 3), 224 is the weight and the height of the image and using 3 channels of RGB. Figure 3 and Figure 4 shows the basic structure of model VGG-16 and VGG-19 respectively. The fully connected layers were discarded (frozen) since in the present study this model was used for feature extraction. Those features were then coupled with Neural Network and Random Forest as classifiers for classification purposes.

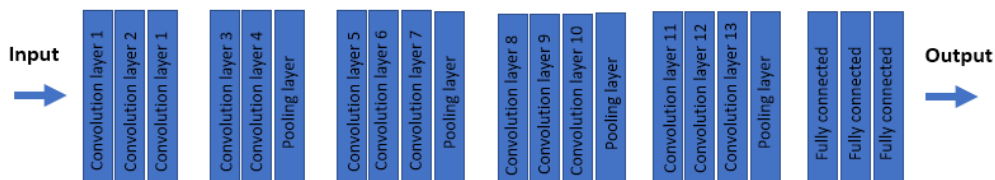


Figure 3. Basic structure of VGG-16.

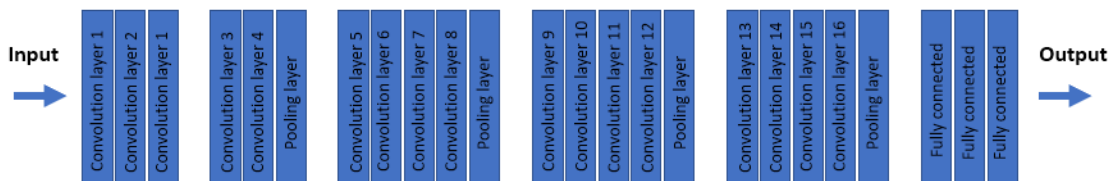


Figure 4. Basic structure of VGG-19.

In present study, there are two classifiers were use which is Random Forest (RF) and Neural Network (NN). The hyper-parameters setting of these two classifiers is picturised in Table 1.

Table 1. List of hyperparameters values setting in RF and NN classifier.

Classifier	Hyperparameters	Values
RF	Number of trees	100
	Limit depth of individual trees	10
	Do not split subset smaller than	4
NN	Learning rate	0.01
	Batch size	32
	Activation function	ReLU

EXPERIMENTAL RESULTS

The classification accuracy (CA) of the different pipelines is shown in Table 2. From the obtained result, no overfitting occurred for all pipelines since the CA for testing data and CA for training data is almost similar. Both feature combining models (VGG16 + VGG19) for NN and RF achieved the higher CA as compared to a single feature model (either VGG16 or VGG19). The result proved that the feature combination model achieved highest accuracy since it has more relevant features (strong features) that gained from both VGG16 and VGG19. The best pipeline of the selected model would be the VGG16 + VGG19 with Random Forest classifier. During the prediction, the pipeline has the minimum number of misclassifications with 18 images (0.38%) in the training dataset and 34 images (1.62%) testing dataset.

Table 2. The classification accuracy on each pipeline

Pipeline	Classification accuracy		
	Testing accuracy	Training accuracy	Average
VGG16 with NN classifier	84.81%	86.08%	84.45%
VGG19 with NN classifier	90.62%	90.98%	90.80%
VGG16 + VGG19 with NN classifier	93.90%	95.31%	94.61%
VGG16 with Random Forest classifier	97.81%	99.51%	98.66%
VGG19 with Random Forest classifier	98.00%	99.86%	98.93%
VGG16 + VGG19 with Random Forest classifier	98.38%	99.63%	99.01%

The recall performance of the best pipeline is illustrated in Table 3. From the tabulated results, only two classes failed to achieve perfect recall which is speed limit (80km/s) and speed limit (60km/s). Both classes were misclassified since the number '6' and '8' are almost similar. Other classes were perfectly classified on this pipeline. The overall performance metric is tabulated in Table 4.

Table 3. The recall performance of VGG16 + VGG 19 with RF classifier pipeline

Class	Recall		
	Training	Testing	Average
Warning	1.00	0.98	0.990
Stop	1.00	1.00	1.000
Repair	1.00	1.00	1.000
Not Enter	1.00	1.00	1.000
Traffic Light	1.00	1.00	1.000
Turn Right	1.00	1.00	1.000
Speed Limit (80km/s)	0.98	0.99	0.985
Speed Limit (50km/s)	1.00	0.94	0.970
Speed Limit (60km/s)	0.99	0.94	0.965
Turn Left	1.00	1.00	1.000
Average	0.954	0.94	0.947

Table 4. Overall classification report for VGG16 + VGG 19 with RF classifier pipeline

Dataset	Multiclass	Performance Metric				Sample Size
		Precision	Recall	F1 Score	Accuracy	
Test	Macro Average	0.98	0.98	0.98	0.98	2100
	Weight Average	0.98	0.98	0.98		2100
Train	Macro Average	1.00	1.00	1.00	1.00	4900
	Weight Average	1.00	1.00	1.00		4900

CONCLUSION

The study evaluated six different TL coupled with NN and RF pipelines in the classification of traffic signs. It was shown from the preliminary investigation carried out that the feature combining model, VGG16+VGG19 with NN and RF pipeline is the best and could attain a CA of 93.90% and 98.38% for the testing dataset as well as 95.31% and 99.63% for the training dataset respectively. The outcome of the study is non-trivial, mainly towards the realisation of a larger traffic signs classifications implementation. Future studies shall attempt on the evaluation of other TL pipelines particularly on the feature combining TL model, classifiers as well as classifier hyperparameters optimisation. Besides, in realisation of largest traffic signs classes classification, the further studies shall add massive traffic signs images from many more classes with different countries set up to further generalise the pipelines.

REFERENCES

- [1] M. Hengstler, E. Enkel, and S. Duelli, "Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices", *Technological Forecasting and Social Change*, vol. 105, pp. 105-120, 2016, doi: 10.1016/j.techfore.2015.12.014.
- [2] J. Shuttleworth, 2019, Accessed 16 March 2021, "SAEJ3016 Levels of Driving Automation", SAE International <<https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic>>.
- [3] World Health Organization, 2021, Accessed 16 March 2021, "Road traffic injuries", WHO <<https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>>.
- [4] S. Khvoynitskaya, 2020, Accessed 16 March 2021, "3 types of autonomous vehicle sensors in self-driving cars", itransition <<https://www.itransition.com/blog/autonomous-vehicle-sensors>>.
- [5] J-H Shi and H-Y Lin "A Vision System for Traffic Sign Detection and Recognition", *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*, pp. 1596-1601, 2017, doi: 10.1109/ISIE.2017.8001485.
- [6] K. Q. W. Luo, 2015. "Traffic-signage detection and recognition based on k-means clustering and support vector machine". Master's thesis, Universiti Teknologi Malaysia.
- [7] S. Narejo, S. Talpur, M. Memon, and A. Rahoo, "An Automated System For Traffic Sign Recognition Using Convolution Neural Network", *3C Technology*, pp. 119-135, Special Issue November 2020, doi: 10.17993/3ctecno.2020.specialissue6.119-135.
- [8] R. Ayachi, Y. E. Said, and M. Atri, "To Perform Road Signs Recognition for Autonomous Vehicles Using Cascaded Deep Learning Pipeline Applied", *Artificial Intelligence Advances*, vol. 1, pp. 1-10, 2019, doi: 10.30564/aia.v1i1.569.
- [9] R. Jaju, 2019. "Multiclass Classification of Road Traffic Signs Using Machine Learning Algorithms". MSc Research Project, National College of Ireland.
- [10] V. Ushaa, V. Kayalvizhib, S. Bharathic, and T. Johnpeterd, "Traffic Sign Classification Using Deep Learning", *Turkish Journal of Computer and Mathematics Education*, vol. 12 No. 9, pp. 250-253, 2021
- [11] Y. Ma and L. Huang, "Hierarchical Traffic Sign Recognition Based on Multi-feature and Multi-classifier Fusion", *First International Conference on Information Science and Electronic Technology (ISET 2015)*, pp. 56-59, 2015, doi: 10.2991/iset-15.2015.15.
- [12] J. L. Mahendra Kumar, M. Rashid, R. Muazu Musa, M. A. Mohd Razman, N. Sulaiman, R. Jailani, and A. P.P. Abdul Majeed, A. "An Evaluation of Different Fast Fourier Transform - Transfer Learning Pipelines for the Classification of Wink-based EEG Signals", *MEKATRONIKA*, vol. 2, no. 1, pp. 1-7, 2020, doi: 10.15282/mekatronika.v2i1.4881.
- [13] J. A. Mat Jizat, A. P.P. Abdul Majeed, A. F. Ab. Nasir, Z. Taha, and E. Yuen, "Evaluation of the machine learning classifier in wafer defects classification", *ICT Express*, 2021, doi: 10.1016/j.ict.2021.04.007.
- [14] F. N. Mohd Noor, W. H. Mohd Isa, I. Mohd Khairuddin, M. A. Mohd Razman, J. A. Mat Jizat, A. F. Ab Nasir, R. Muazu Musa, A. PP Abdul Majeed, "The Diagnosis of Diabetic Retinopathy: A Transfer Learning with Support Vector Machine Approach", *Advances in Robotics, Automation and Data Analytics: Selected Papers from ICITES 2020*, vol. 1350, pp. 391-398, 2021.