e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:12/December-2022 Impact Factor- 6.752 www.irjmets.com

A FUSION MODEL WITH YOLOV3 AND SHUFFLENETV2 NETWORKS FOR TRAFFIC SIGN RECOGNITION

Wu Caipeng*1, Wu Xinghuan*2, Jia Yueyun*3

*1School Of Management, Shijiazhuang Tiedao University, Shijiazhuang, Hebei, China.

*2Hebei Dongyu Construction Project Management Co., Ltd, Shijiazhuang, Hbei, China.

*3School Of Continuing Education, Shijiazhuang University Of Applied Technology, Shijiazhuang, Hebei, China.

DOI : https://www.doi.org/10.56726/IRJMETS31957

ABSTRACT

The rapid and accurate recognition of traffic signs helps to improve the performance of advanced driver assistance systems and provides important safety guarantees for unmanned driving. Aiming at the existing object detection algorithms with the low accuracy of traffic sign recognition, weak generalization ability, and difficult detection for small targets, which cannot be well applied to practical applications, a lightweight deep network model for fast and accurate recognition of traffic signs is established by introducing the lightweight deep learning network ShuffleNetV2 and improving the regression-based object detection network model YOLOv3. Five data augmentation techniques were used to amplify the public dataset and the network parameters were trained using the transfer learning strategy, indicating that the new network was less computationally intensive and could achieve an accuracy of 96%.

Keywords: Traffic Sign Recognition; Deep Learning; Convolutional Neural Networks; Transfer Learning; Object Detection.

I. INTRODUCTION

With the rapid development of computer vision, real-time communication technology, and high-precision sensor equipment, intelligent transportation systems have increasingly become a research hotspot in academia and industry [1-3]. Sensors are typically deployed on vehicles and work together with visual feature recognition tools. Commonly used sensor equipment now includes radar equipment (including laser and millimeter wave radar, etc.) and image acquisition equipment (including monocular and binocular cameras and GPS positioning systems). This allows the car to perceive the traffic environment around the vehicle and collect the status of the vehicle when driving on the actual road. At the same time, the on-board computer could be used to intelligently calculate and analyze the collected data in real-time, therefore instructions can be issued in advance, and assist the driver in making judgments or operate intelligently to avoid driving risks, which helps to reduce the occurrence of accidents. With the combination of 5G communication and computer vision, traffic signs can be effectively identified and reported to the driver in advance, thereby reducing traffic accidents caused by human visual differences or fatigue driving. The Advanced Driving Assistance System (ADAS) is a product of this demand [4-6]. The role of traffic sign recognition in ADAS is to realize the adaptive cruise of the vehicle, and notify the road ahead of the speed limit, turning, warning, indication, and other information in advance. In recent years, driverless technology has been an important research object for many enterprises and universities. The technical fields supporting the functions of driverless cars are wide, among which vision-based traffic sign recognition is an important research direction. The driverless car does not need the driver's control during the whole process of driving, and will automatically adjust the driving state of the vehicle according to the actual situation on the road. Driverless systems intelligently connect pedestrians, vehicles, and roads. An important prerequisite for realizing intelligent interconnection is the recognition of traffic signs. The information is quickly transmitted to the driverless platform for calculation and analysis, and the next operation instructions are given to ensure that the vehicle drives safely and smoothly in accordance with traffic rules on the road. Another application is in-vehicle intelligent navigation systems. The car navigation system can intelligently give the optimal driving route between the departure point and the destination. The combination of traffic sign detection and recognition with the in-vehicle navigation map reduces the adverse effects of temporary changes in road conditions that cannot be predicted by the navigation system in advance,

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:12/December-2022 Impact Factor- 6.752 www.irjmets.com

thereby reducing driving risks. At present, the function of the smartphone navigation system has been relatively perfect, but the real-time update of the navigation software relies on the support of wireless networks. In some special sections, such as culverts, and tunnels, the information cannot be updated in time, however, the traffic sign recognition function in the car navigation will not be affected by the signal strength and can continue to work.

Traffic signs are important recognition objects for traffic images. There are many types of traffic signs, and the factors that affect the recognition of traffic signs are complicated. On the one hand, the on-board camera collects traffic sign images under real-time road conditions, so it will be affected by factors such as rain and snow, lighting, and aging and damage of signs. In addition, the image acquisition of traffic signs will be affected by blurring due to movement, distortion caused by continuous changes in the acquisition angle, and incomplete picture capture. In addition, obstacles such as trees can also affect the quality of signal acquisition. The above are some difficult problems affecting the recognition of traffic signs, but the value brought by accurate identification is very meaningful, which can reduce the incidence of traffic accidents, promote unmanned driving and intelligent navigation technology and assist drivers. Besides, traffic management departments can also benefit from it and achieve scientific management and maintenance of traffic signs.

There are two methods for traffic sign feature extraction, one is based on color and shape features [7-9], and the object detection method based on color features is mainly to detect the region of interest (ROI). Due to the influence of uneven illumination and discoloration of traffic signs, it will cause seriously missed detection when using color features-based object detection methods. The shape-based object detection method first extracts the shape features of the objects in the picture and then applies the shape feature operators to different features, and finally realizes the object detection through the shape-matching algorithm. Another approach for traffic sign recognition is machine learning [10, 11]. Compared with color- and shape-based object detection methods, object detection methods in the driving environment based on machine learning have greater adaptability, wider application range, and better robustness. In recent years, traffic sign recognition based on deep learning methods [12-14] has been developed rapidly. So far, the object detection model based on deep learning is mainly divided into two categories: one is based on regional recommendation, and its representative network models include the R-CNN series network models [15-19]; the other type is the regression-based object detection network model, which includes the YOLO series network models [11, 20-22] and SSD network models[23-26].

The focus of traffic sign detection should be on accuracy and real-time detection because only object detection models with high accuracy and strong real-time performance can be widely used. This paper aims to design a high-performance object detection network with YOLOv3 as the main framework. The main idea is to replace the backbone network of YOLOv3 with the more efficient ShuffleNetv2 to improve performance. Compared with the original YOLOv3 model, the prediction speed can be increased by 10ms~20ms. The model size is less than one-eighth of the original, and the accuracy of recognition is roughly equivalent to YOLOv3.

II. METHODOLOGY

The adopted recognition model takes YOLOv3 [27] as the main framework, and replaces the feature extraction network of YOLOv3 with ShuffleNetv2 [28]. The reason for using YOlOv3 as the detection network is that the detection part of YOlOv3 adopts the feature pyramid networks, which detects the output feature map on three different scales and then fuses the features of these three scales to obtain detection results for objects of large, medium and small scales. Considering the different sizes of traffic sign images obtained by vehicle cameras, the use of structures such as YOLOv3 can make the classification effect better, and the recognition rate of small targets is expected to improve. In addition, YOLOv3's accuracy is slightly better than SSD, almost equal to Faster R-CNN. YOLOv3 is at least twice as fast as SSD and Faster R-CNN. The architecture of the Yolov3 network is shown in Figure 1. Yolov3's backbone network uses the DarkNet-53 network, whose structure references the residual network. The multi-scale feature map is used to detect the target, and 9 sizes of prior boxes are clustered on the feature map of 3 sizes so that the accuracy is improved while ensuring real-time detection.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:12/December-2022 Impact Factor- 6.752 www.irjmets.com

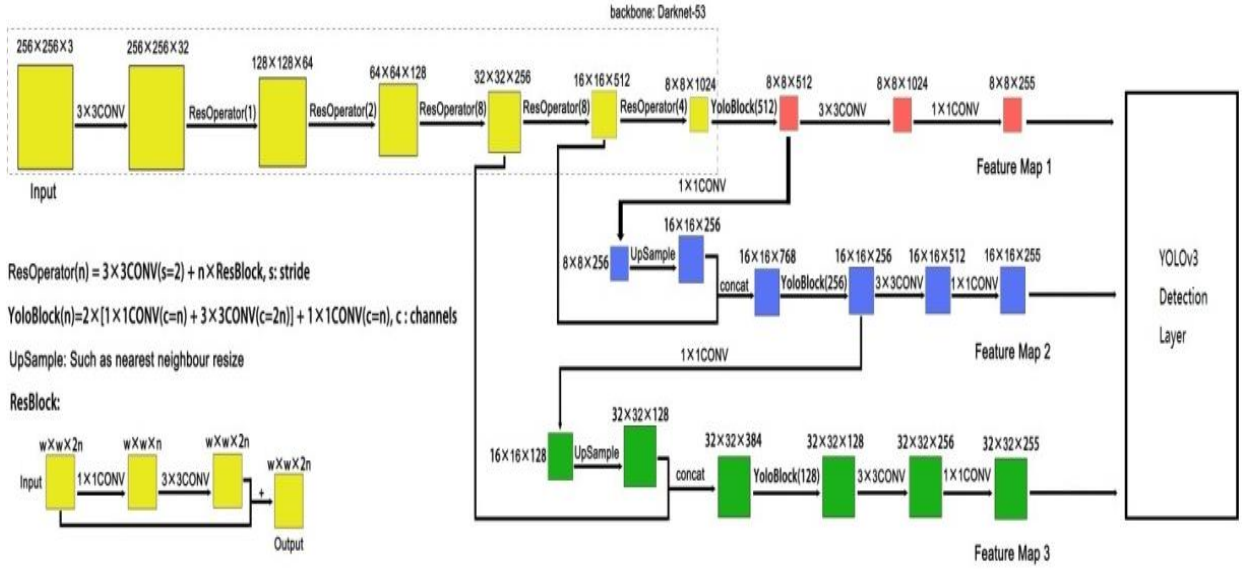


Figure 1: Architecture of the Yolov3 network.

The backbone network of the new model is ShuffleNetV2, which was proposed by Ma et al. in 2018 and has a good balance between speed and accuracy. At the same level of complexity, ShuffleNetv2 is more accurate than ShuffleNetV1 and MobileNetv2. Because the network structure of ShuffleNetv2 is relatively simple, it is suitable for mobile terminal devices and has good application prospects in the field of advanced driver assistance systems and unmanned driving. The network structure and parameters of ShuffleNetV2 as the main component are shown in Table 1.

Table 1. Network structure and parameters of ShuffleNetV2

Layer	Output size	KSize	Stride	Repeat	Output channels			
					0.5×	1×	1.5×	2×
Image	224×224				3	3	3	3
Conv1	112×112	3×3	2	1	24	24	24	24
MaxPool	56×56	3×3	2					
Stage2	28×28		2	1	48	116	176	244
	28×28		1	3				
Stage3	14×14		2	1	96	232	352	488
	14×14		1	7				
Stage4	7×7		2	1	192	464	704	976
	7×7		1	3				
Conv5	7×7	1×1	1	1	1024	1024	1024	2048
GlobalPool	1×1	7×7						
FC					1000	1000	1000	1000
FLOPs					41M	146M	299M	591M
# of Weights					1.4M	2.3M	3.5M	7.4M

The starting point for this replacement is that a common measure of model complexity today is FLOPs, which is an indirect metric because it is not exactly equivalent to computational speed. Two models of the same FLOPs may have different speeds. This inconsistency is mainly due to two reasons, the first of which is that other factors that affect speed, such as memory access cost (MAC), can be bottlenecks for GPUs, so it can't be ignored. In addition, the degree of parallelism of the model also affects the speed, and the model with high parallelism is relatively faster. Therefore, grouped convolution should not be used excessively, and 1x1 convolution should be used to balance the channel size of input and output to minimize memory access. In addition, it is recommended to avoid network fragmentation to improve parallelism; and reduce element-level operations. The model structure of the newly designed traffic sign recognition network is shown in Figure 2.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:12/December-2022 Impact Factor- 6.752 www.irjmets.com

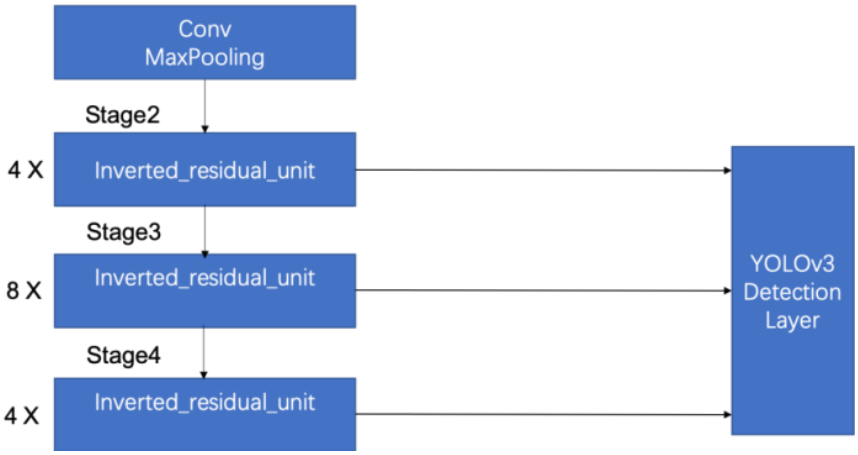
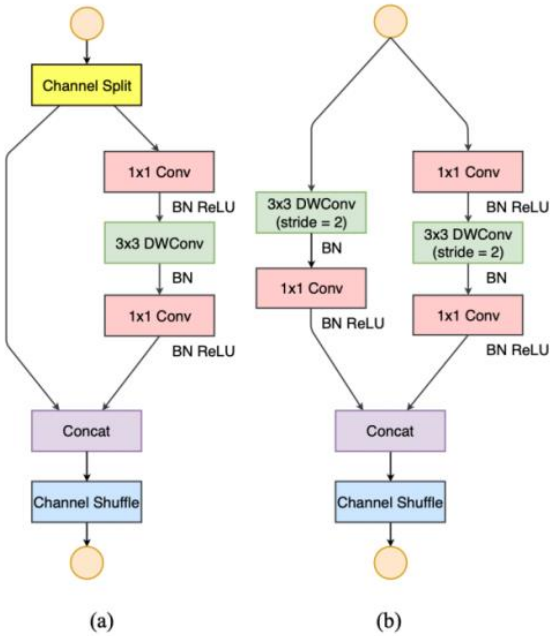


Figure 2: The architecture of the new CNN model.

The components of the main structural (Inverted_residual_unit) are shown in Figure 3:



(a) ShuffleNet-V2 basic unit; (b) ShuffleNet-V2 unit for spatial down sampling (2×)

Figure 3: The components of Inverted_residual_unit.

The loss function is used to describe the degree of difference between the predicted value of the model and the true value, and the value of the loss function determines the detection effect of the network model. The loss function of the YOLOv3 network model includes three types of losses: box loss, confidence loss, and classes loss. The box loss is divided into two items, namely the loss of the center point coordinates x, y and the loss of the width and height of the box w, h, both of which are calculated using the mean square error loss function.

The loss function for the center point coordinates is as follows:

$$L_{(x,y)} = \sum_{i=0}^{M^2} \sum_{j=0}^{N} I_{ij}^{obj} \left[(x_i^j - \hat{x}_i^j)^2 + (y_i^j - \hat{y}_i^j)^2 \right] \quad (1)$$

where x, and y are the center point coordinate values of the real box, respectively, and $\hat{x}$ and $\hat{y}$ are the center point coordinate values of the prediction box, respectively.

The loss function of the width and height of the border is as follows:

$$L_{(w,h)} = \sum_{i=0}^{M^2} \sum_{j=0}^{N} I_{ij}^{obj} \left[(w_i^j - \hat{w}_i^j)^2 + (h_i^j - \hat{h}_i^j)^2 \right] \quad (2)$$

where w, and h are the width and height of the real box, respectively, and $\hat{w}$, $\hat{h}$ are the width and height of the prediction box, respectively.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:12/December-2022 Impact Factor- 6.752 www.irjmets.com

The confidence loss is also divided into two items, namely the confidence loss with the corresponding object in the prediction box and the box confidence loss without the corresponding target in the prediction box, and the confidence loss without the target in the prediction box. The confidence loss with the target in the prediction box is shown in Equation (3):

$$L_{conf\_obj} = -\sum_{i=0}^{M^2}\sum_{j=0}^{N} I_{ij}^{obj}\left[\hat{C}_i^j \log\left(C_i^j\right) + (1 - C_i^j)\log\left(1 - C_i^j\right)\right] \quad (3)$$

where C_i^j and $\hat{C}_i^j$ are the C and $\hat{C}$ values of the j-th predicted box of the i-th element on the feature map, respectively.

The confidence loss without the target in the prediction box is shown in Equation (4):

$$L_{conf\_noobj} = -\lambda_{noobj}\sum_{i=0}^{M^2}\sum_{j=0}^{N} I_{ij}^{noobj}\left[\hat{C}_i^j \log\left(C_i^j\right) + (1 - \hat{C}_i^j)\log\left(1 - C_i^j\right)\right] \quad (4)$$

Only if the j-th prior box of the i-th element of the feature map has a target, the prediction box corresponding to the prior box will be used to calculate the classes loss, which is calculated using the binary cross entropy loss function, shown in Equation (5):

$$L_{cls} = -\sum_{i=0}^{M^2} I_{ij}^{obj}\sum_{c\in classes}[\hat{p}_i(c)\log(p_i(c)) + (1 - \hat{p}_i(c))\log(1 - p_i(c))] \quad (5)$$

where $\hat{\mathbf{p}}_\mathbf{i}(\mathbf{c})$ is the probability that the target predicted by the prediction box belongs to a certain class, $\mathbf{p_i(c)}$ is its true value, and if it belongs to a certain class, $\mathbf{p_i(c) = 1}$, otherwise $\mathbf{p_i(c) = 0}$.

The total loss function is shown in Equation (6).

$$Loss = L_{(x,y)} + L_{(w,h)} + L_{conf\_obj} + L_{conf_{noobj}} + L_{cls} \quad (6)$$

The dataset uses the German Traffic Sign Detection Benchmark (GTSDB)[29]. The dataset contains the following traffic signs: speed limit, restriction ends, no overtaking, priority at next intersection, priority road, give way, stop, no traffic both ways, no trucks, no entry, danger, bend left, bend right, bend, uneven road, slippery road, road narrows, construction, traffic signal, pedestrian crossing, school crossing, cycles crossing, snow, animals, go right, go left, go straight, go right or straight, go left or straight, keep right, keep left, and roundabout. The dataset has a total of 900 original images. Given insufficient training data, model parameters may be overfitting. Therefore, the data augmentation method is adopted to increase the amount of data to improve the robustness of the model and avoid overfitting. Techniques such as adding random noise, translation, changing lighting, rotation, and mirroring are used, as shown in Figure 4.

(a) Original image (b) Mirrored image (c) Image after lighting changes

(d) Image after rotation (e) Image after translation (f) Image after random noise

Figure 4: Data augmentation method.

An additional 4500 images were obtained using the above data augmentation method. The resulting dataset contains 5400 images, of which 3780 are used for training and 1620 photos are used for verification. A transfer learning strategy was used for training. The hardware platform used includes an Intel Core i7-12700 processor and 32GB DDR5 4400MHz memory. The programming language is Python 3.6, the deep learning framework is TensorFlow 2.0, and the neural network library is Keras 2.3. Table 2 shows the training option settings.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:12/December-2022 Impact Factor- 6.752 www.irjmets.com

Table 2. Training options

training options	epochs	mini batch size	warm-up period	L2 regularization factor	penalty threshold	learning rate
value	80	8	1000 iterations	0.0005	0.5.	0.001

III. RESULTS AND DISCUSSION

Typical traffic sign recognition results are shown in Figure 5.

Figure 5: Typical traffic sign recognition results.

It can be seen from Figure 5 that although the traffic signs belong to small targets in the image, the network structure used in this paper can still identify these small targets, and the recognized probability of belonging is above 0.8. This can be attributed to the multi-scale feature map adopted by Yolov3.

The change curve of the loss function with epochs is shown in Figure 6. It can be seen from Figure 6 that the loss function decreases rapidly and has basically stabilized when training to 20 epochs, indicating that replacing the feature extraction network of YOLOv3 with ShuffleNetv2 reduces the number of parameters, thereby greatly reducing the amount of computation, which is critical for the on-board computing system. The change curve of accuracy with epochs is shown in Figure 7.

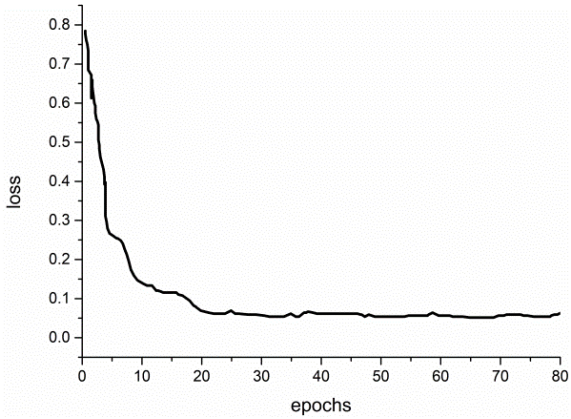


Figure 6: Change curve of the loss function with epochs.

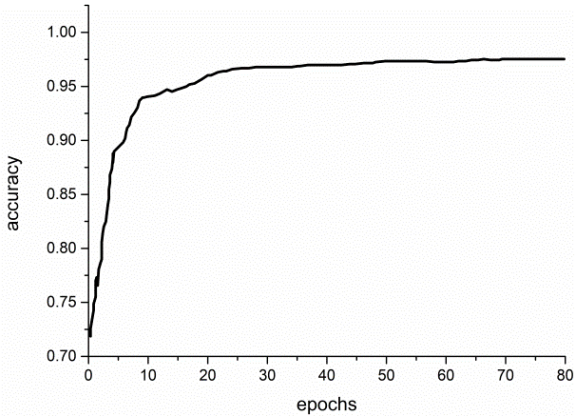


Figure 7: Change curve of accuracy with epochs.

As can be seen from Figure 7, the accuracy rate increases rapidly after training starts. After 20 epochs of training, the accuracy rate stabilized at more than 96%. The recognition time for a single image is about 20 ms. It shows that after using data augmentation technology, the overfitting phenomenon is effectively avoided, and after using transfer learning, the high accuracy of traffic sign recognition is obtained with a low computational cost.

IV. CONCLUSION

By replacing the feature extraction network of YOLOv3 with the lightweight deep learning network ShuffleNetV2, a novel convolutional neural network for traffic sign recognition is obtained. By using data augmentation technology and a transfer learning strategy, it is found that the network has fast training speed and high accuracy, which provides a new choice for real-time and accurate recognition of traffic signs.

V. REFERENCES

[1] Juan Guerrero-Ibáñez, Sherali Zeadally, Juan Contreras-Castillo, “Sensor Technologies for Intelligent Transportation Systems,” Sensors, Vol. 18, Issue. 4, pp. 1212, 2018.

[2] Maozhu Jin, Qian Zhang, Hua Wang, Yuan Yuan, "Research on Intelligent Transportation System Based on Internet of Things," International Journal of Heavy Vehicle Systems, Vol. 27, Issue. 3, pp. 247-257, 2020.

[3] AHMED S, KAMAL U, HASAN M K, "DFR-TSD: A Deep Learning Based Framework for Robust Traffic Sign Detection Under Challenging Weather Conditions," IEEE Transactions on Intelligent Transportation Systems, Vol. 23, Issue. 6, pp. 5150-5162, 2022.

[4] Jiali Yin, Bohao Chen, Kuohua Robert Lai, Ying Li, "Automatic Dangerous Driving Intensity Analysis for Advanced Driver Assistance Systems From Multimodal Driving Signals," IEEE Sensors Journal, Vol. 18, Issue. 12, pp. 4785-4794, 2018.

[5] Farag Wael, "A Lightweight Vehicle Detection and Tracking Technique for Advanced Driving Assistance Systems," Journal of Intelligent & Fuzzy Systems, Vol. 39, Issue. 3, pp. 2693-2710, 2020.

[6] Michael Gerstmair, Martin Gschwandtner, Rainer Findenig, Oliver Lang, A. Melzer, M. Hueme, "Miniaturized Advanced Driver Assistance Systems: A Low-Cost Educational Platform for Advanced Driver Assistance Systems and Autonomous Driving," IEEE Signal Processing Magazine, Vol. 38, Issue. 3, pp. 105-114, 2021.

[7] Selcan Kaplan Berkaya, Huseyin Gunduz, Ozgur Ozsen, Cuneyt Akinlar, Serkan Gunal, "On Circular Traffic Sign Detection and Recognition," Expert Systems with Applications, Vol. 48, Issue. C, pp. 67-75, 2016.

[8] Mohammed Boumediene, Christophe Cudel, Michel Basset, Abdelaziz Ouamri, "Triangular Traffic Signs Detection Based on RSLD Algorithm," Machine Vision and Applications, Vol. 24, Issue. 8, pp. 1721-1732, 2013.

[9] Manisha Vashisht, Brijesh Kumar, "Effective Implementation of Machine Learning Algorithms Using 3D Colour Texture Feature for Traffic Sign Detection for Smart Cities," Expert Systems, Vol. 39, Issue. 5, pp. e12781, 2022.

[10] Zhanwen Liu, Mingyuan Qi, Chao Shen, Yong Fang, Xiangmo Zhao, "Cascade Saccade Machine Learning Network with Hierarchical Classes for Traffic Sign Detection," Sustainable Cities and Society, Vol. 67, pp. 102700, 2021.

[11] Jie Hu,Zhanbin Wang,Minjie Chang,Lihao Xie,Wencai Xu,Nan Chen, "PSG-Yolov5: A Paradigm for Traffic Sign Detection and Recognition Algorithm Based on Deep Learning," Symmetry-Basel, Vol. 14, Issue. 11, pp. 2262, 2022.

[12] Kwangyong Lim, Yongwon Hong, Yeongwoo Choi, H. Byun, "Real-Time Traffic Sign Recognition Based on A General Purpose GPU and Deep-Learning," Plos One, Vol. 12, Issue. 3, pp. e0173317, 2017.

[13] Di Zang, Zhihua Wei, Maomao Bao, Jiujun Cheng, Dongdong Zhang, Keshuang Tang, Xin Li, "Deep Learning-Based Traffic Sign Recognition for Unmanned Autonomous Vehicles," Proceedings of the Institution of Mechanical Engineers Part I-Journal of Systems and Control Engineering, Vol. 232, Issue. 5, pp. 497-505, 2018.

[14] Jameel Ahmed Khan, Donghoon Yeo, Hyunchul Shin, "New Dark Area Sensitive Tone Mapping for Deep Learning Based Traffic Sign Recognition," Sensors, Vol. 18, Issue. 11, pp. 3776, 2018.

[15] Faming Shao, Xinqing Wang, Fanjie Meng, Jingwei Zhu, Dong Wang, Juying Dai, "Improved Faster R-CNN Traffic Sign Detection Based on a Second Region of Interest and Highly Possible Regions Proposal Network," Sensors, Vol. 19, Issue. 10, pp. 2288, 2019.

[16] Jianming Zhang, Zhipeng Xie, Juan Sun, Xin Zou, Jin Wang, "A Cascaded R-CNN With Multiscale Attention and Imbalanced Samples for Traffic Sign Detection," IEEE Access, Vol. 8, pp. 29742-29754, 2020.

[17] Jinghao Cao, Junju Zhang, Xin Jin, "A Traffic-Sign Detection Algorithm Based on Improved Sparse R-cnn," IEEE Access, Vol. 9, pp. 122774-122788, 20121.

[18] Xiang Gao, Long Chen, Kuan Wang, Xiaoxia Xiong, Hai Wang, Yicheng Li, "Improved Traffic Sign Detection Algorithm Based on Faster R-CNN," Applied Sciences-Basel, Vol. 12, Issue. 18, pp. 8948, 2022.

[19] Tianjiao Liang, Hong Bao, Weiguo Pan, Feng Pan, "Traffic Sign Detection via Improved Sparse R-CNN for Autonomous Vehicles," Journal of Advanced Transportation, Vol. 2022, pp. 3825532, 2022.

[20] Yuchen Liu, Gang Shi,Yanxiang Li, Ziyu Zhao, "M-YOLO: Traffic Sign Detection Algorithm Applicable to Complex Scenarios," Symmetry-Basel, Vol. 14, Issue. 5, pp. 952, 2022.

[21] Xinyue Ren, Weiwei Zhang, Minghui Wu, Chuanchang Li, Xiaolan Wang, "Meta-YOLO: Meta-Learning for Few-Shot Traffic Sign Detection via Decoupling Dependencies," Applied Sciences-Basel, Vol. 12, Issue. 11, pp. 5543, 2022.

[22] Jing Yu, Xiaojun Ye, Qiang Tu, "Traffic Sign Detection and Recognition in Multiimages Using a Fusion Model With YOLO and VGG Network," IEEE Transactions on Intelligent Transportation Systems, Vol. 23, Issue. 9, pp. 16632-16642, 2022.

[23] Zhongyi Lin, Jeffrey M. Ota, John Douglas Owens, Pınar Muyan-Özçelik, "Benchmarking Deep Learning Frameworks and Investigating FPGA Deployment for Traffic Sign Classification and Detection," IEEE Transactions on Intelligent Vehicles, Vol. 4, Issue. 3, pp. 385-395, 2019.

[24] Miguel Lopez-Montiel, Ulises Orozco-Rosas, Moisés Sánchez-Adame, Kenia Picos, Oscar Humberto Montiel Ross, "Evaluation Method of Deep Learning-Based Embedded Systems for Traffic Sign Detection," IEEE Access, Vol. 9, pp. 101217-101238, 2021.

[25] Yanzhao Zhu, Weiqi Yan, "Traffic Sign Recognition Based on Deep Learning," Multimedia Tools and Applications, Vol. 81, Issue. 13, pp. 17779-17791, 2022.

[26] Jianjun Wu, Shaowen Liao, "Traffic Sign Detection Based on SSD Combined with Receptive Field Module and Path Aggregation Network," Computational Intelligence and Neuroscience, Vol. 2022, pp. 4285436, 2022.

[27] Redmon Joseph, Farhadi, Ali, "YOLOv3: An Incremental Improvement," arXiv, Vol. 1804, pp. 02767, 2018.

[28] Ningning Ma, Xiangyu Zhang, HaiTao Zheng, Jian Sun, "ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design," arXiv, Vol. 1807, pp. 11164, 2015.

[29] Sebastian Houben, Johannes Stallkamp, Marc Schlipsing, Jan Salmen, Christian Igel, "Detection of Traffic Signs in Real-World Images: the German Traffic Sign Detection Benchmark," The 2013 International Joint Conference on Neural Networks (IJCNN), pp. 1-8. 2013.