

DOI: 10.37943/12YZFG6952

Sharipa Temirgazyeva

Master's Student of the "Information Systems" Program
temirgazyevash@gmail.com, orcid.org/0000-0003-1562-4812
Al-Farabi Kazakh National University, Kazakhstan

Batyrkhan Omarov

Acting Associate Professor of the Department of Information
Systems
batyahan@gmail.com, orcid.org/0000-0002-8341-7113
Al-Farabi Kazakh National University, Kazakhstan

TRAFFIC SIGN RECOGNITION WITH CONVOLUTIONAL NEURAL NETWORK

Abstract: Road sign recognition is one of the most important steps drivers can take to avoid dangerous roads or accidents. The purpose of the research work is to develop a recognition system, increasing the classification accuracy of the model, using deep learning methods of Road sign recognition system for drivers in real time on the road. Stages of road sign image classification were carried out, and other authors' solutions were analyzed. In addition, in this work, the convolutional neural network (CNN) method with the help of deep learning was used to create a system for identifying and recognizing road signs. The proposed road sign recognition system can work in real time to recognize the image of road signs. In this paper, a model is trained using deep learning of 43 different road signs using existing datasets and collected local road signs. A traffic sign detection and recognition system is presented using an 8-layer convolutional neural network, which acquires different functions by training different types of traffic signs.

In previous studies, models were trained using simple machine learning algorithms, but the relevance of this study is that a CNN model was trained for a classification task based on convolutional neural networks using deep learning. As a result of the study, classification accuracy of 95% was obtained using deep learning methods. As a novelty of the work, it is possible to note the diversity of the convolutional network methods used to increase the efficiency of the used data set and model training algorithms, the variety of received road signs and algorithms for its recognition, as well as the achievement of a high accuracy rate. This allowed the system to overcome the limited accuracy and performance issues caused by environmental factors, and to be more versatile and accurate than most modern systems.

Keywords: CNN, neural network, traffic sign, deep learning, machine learning.

Introduction

A traffic accident is an event that occurs during the movement of vehicles on the road and with its participation, causing harm to human health or causing serious death, as well as causing damage to vehicles, structures, cargo or other material losses [1].

Data on traffic accident victims in our country exceeds the statistics of CIS countries such as Armenia, Kyrgyzstan, and Tajikistan. This is the result of events that have been carried out for more than 20 years. One of the main elements of the infrastructure is the road. Many factors (such as rain, snow, high temperatures, darkness, glare, and trucks) cause various damages that significantly affect traffic efficiency, driver safety, and vehicle costs.

According to the official statistics of the National Statistics Bureau of the Strategic Planning and Reforms Agency of the Republic of Kazakhstan, the number of traffic accidents in 2020 is 13,515.

The following tables show statistics on road accidents in Kazakhstan for 2020 (Figure 1).

The most traffic accidents are observed in Almaty city (3245), Almaty region (2150) and Turkestan region (1042).

The number of people who died as a result of the incident is 1,687, the number of injured is 12,417 [2]. The main causes of traffic accidents:

- increasing speed in prohibited areas;
- when crossing a pedestrian crossing;
- non-observance of the requirements determined by the signs on the carriageway [3].

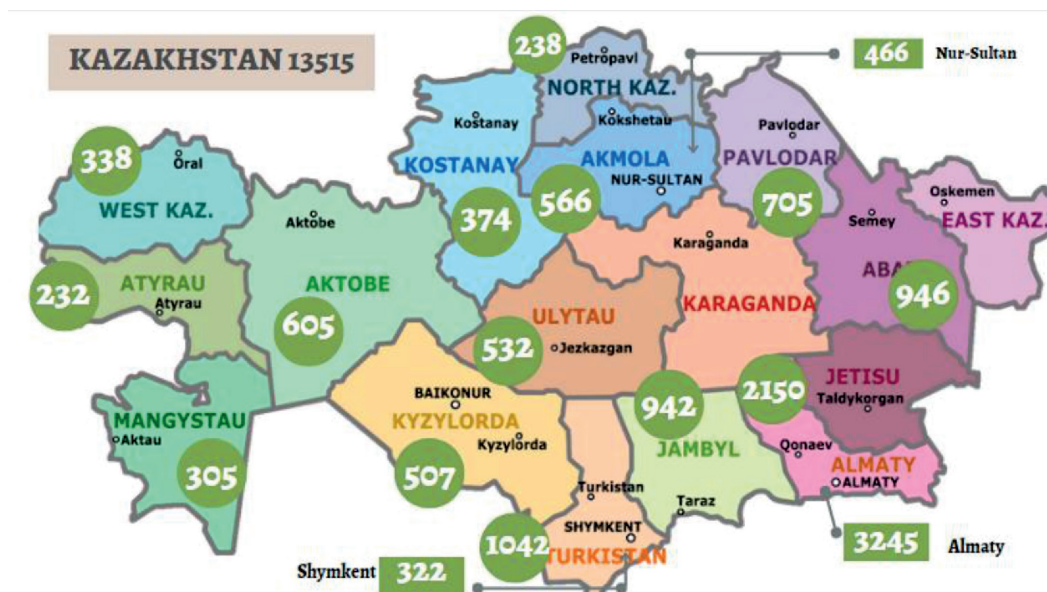


Figure 1. Number of road accidents in 2020 according to official statistics from the National Bureau of Statistics

The country needs a road sign recognition system that will help save lives. Road signs help prevent road accidents, ensuring the safety of both drivers and pedestrians. In addition, traffic lights guarantee compliance with certain laws by road users, which reduces the likelihood of traffic violations. Road signs should be a high priority for drivers or pedestrians. For various reasons, it is more likely to go unnoticed by road signs due to factors such as concentration, fatigue and sleep, for example. Other reasons that contribute to the fact that road signs go unnoticed include factors such as poor vision, exposure to the outside world and environmental conditions, weather changes, sun reflection, etc. The real-time road sign recognition system analyzes the images taken by the front camera of the car in real time to recognize the signs. They help increase safety by issuing warnings to the driver.

Literature review

Traffic sign detection systems are the most popular systems for detecting small objects (signs). Road sign detection typically involves using color or geometric features to generate candidate regions in a given image that may contain road signs.

Some approaches directly use RGB color space and color threshold for image segmentation and feature detection. Bouti A. et al. researchers divided the road sign recognition system into

two parts. The first part is real-time feature detection, and the second part is a feature detection method using datasets. For this, researchers used convolutional neural networks [4]. And the proposed network in this study uses neural network architecture. In particular, the CNN 2d-image recognition method is used.

Sun Y. et al. researchers [5] propose a new framework for deep learning consisting of two components, including a fully convolutional network (FCN) for controlled road sign recognition and a deep convolutional neural network (CNN) for object classification. The proposed approach [5] is experimentally compared with R-CNN [6].

E. Peng et al. researchers [7] use CNN [8] to perform road sign recognition and the results show that this approach is promising. However, accuracy and speed indicators showed very low accuracy. Therefore, based on these shortcomings, we proposed the accuracy method in this study in order to increase the accuracy of network recognition.

Wu Yiqiang et al. researchers [9] introduce a real-time road sign recognition algorithm that is tolerant to small objects and can detect all categories of road signs. (advanced feature selection [10, 11] can be achieved using recent algorithms such as spectral clustering). In particular, a two-level detection structure consisting of a regional recommendation module (RPM) responsible for object detection and a classification module (CM) for classifying detected objects is proposed. Information about color and shape was used in the identification of traffic lights. For example, to recognize traffic lights, RGB or HSV color space was used [12]. In addition to colors, geometric information is widely used. Geometric features of road signs, Hough transformation, angle detection and projections are used to determine the exact location of road signs [13].

Tasks set

Having reviewed the research results presented by the previous authors and taking into account the mentioned shortcomings, the following tasks were set to perform the work:

1. development of a data set to create a road sign recognition system;
2. classification of data sets into two classes;
3. differentiate the methods of classification and use an effective method;
4. model training based on convolutional network.

As the first step, the development of the data set is one of the most important steps, because the accuracy of the results of the algorithms, the accuracy of the results of the algorithms, directly depends on the performance of the model. A set of images converted to jpg format is considered as the initial data of the research. By simplifying the classification task into two classes, i.e. training and testing class, we proceed to the classification method. In order to choose the most useful and efficient method, the algorithm's accuracy indicators and execution time were taken into account.

The presented system uses a convulsive neural network. In this regard, the use of deep learning, the creation of a recognition system within the framework of its various methods, makes a great contribution to both science and society. Considering these comments, the following studies are proposed in this paper:

1) we provide a deep convulsive neural network for fast and efficient detection of road signs. Their implementation can simultaneously classify and identify small objects, such as road signs.

2) we reduce the model size and working time without loss of accuracy, using the network cutting and core stacking method to ensure the accuracy of the model.

3) further, in order to quickly and accurately identify road signs, we create a convolutional neural network consisting of 8 layers, and the accuracy of the model was 95%.

Materials and methods

The figure below shows the architecture of the traffic sign recognition system based on the CNN model (Figure 1).

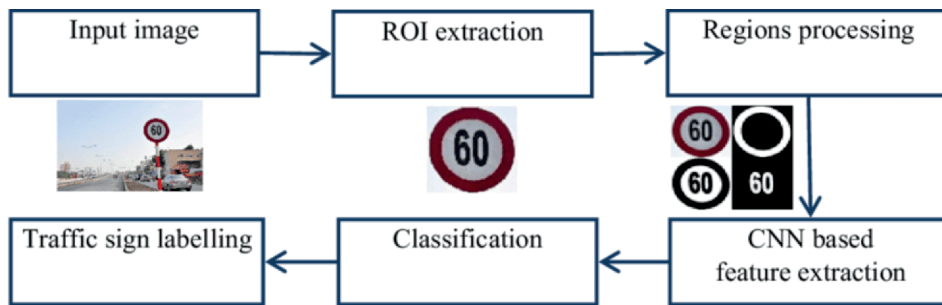


Figure 2. Architecture of the road sign detection and recognition system

The road sign detection and recognition system consists of a number of components, including:

1. a complex process of obtaining data that allows obtaining the necessary information;
2. classification of received data stream into frames;
3. identification of received signs;
4. neural network training, as well as evaluation of its efficiency and accuracy on test data;
5. a data set where road signs are stored;
6. identification of objects in frames;
7. get results.

Data set

The dataset used to train the traffic sign classifier is the Traffic Sign Recognition Test – GTSRB [14]. The GTSRB dataset consists of 43 road sign classes and about 50,000 images. The dataset used can be seen in Figure 3.



Figure 3. The GTSRB dataset

However, it is necessary to supplement the data in this standard. Therefore, by conducting additional experiments with the Tsinghua-Tencent 100K dataset [15], which contains many road scenes with large differences in many aspects, such as light and weather, as well as occlusion examples, methods are used.

Training Neural Network

The main goal of this research is to create a system for detecting and recognizing traffic signs using deep learning. First, a road sign recognition model was prepared using deep learning and the prepared model was tested. The deep learning process to train the model takes a lot of time, but the classification itself takes a relatively short time.

The purpose of the task of determining each part of the sign for determining road signs is to determine whether any one pixel belongs to one part of the sign. To solve this problem, the proposed solution is based on a convolutional neural network for classifying objects with geometric shapes. The sampling method is used between two fully connected layers to reduce overfitting by preventing complex coadaptations in the training data [16]. The output signal of each neuron is set to zero with a probability of 0.5. CNN learning is accelerated using graphics processing units (GPUs). A further speedup is achieved by using rectified linear units (ReLU) as an activation function, which is more efficient than $\tanh(x)$ hyperbolic tangent functions and is a sigmoidal function used in both the training phase and the evaluation phase in traditional neural models. A convolutional neural network, due to the use of a special operation – convolution itself – allows you to simultaneously reduce the amount of information stored in memory, due to which it copes better with higher-resolution images, and highlight the supporting features of the image, such as edges, contours or faces. It takes less than 20 periods to reach the minimum in the test set.

To process test images, CNN can provide the probability that each point in the center of the image is a label or not. This procedure yields a probability map. The probability of a point can be calculated by averaging the probability $\{P_1, \dots, P_N\}$ of each patch generated by random rotation around its center pixel, where i is the CNN classification probability calculated for an individual patch, and the probability number is equal to 5 for computational efficiency.

CNNs have more degrees of freedom and therefore exhibit greater variance and less bias. This fact causes CNN to overestimate the feature detection capability. Therefore, an appropriate threshold should be used. Determining precision and recall:

$$P = \frac{\text{true_positive}}{\text{true_positive} + \text{false_positive}} \quad (1)$$

$$R = \frac{\text{true_positive}}{\text{true_positive} + \text{false_negative}} \quad (2)$$

Then the measure of F is:

$$F_1 = \frac{2PR}{P+R} \quad (3)$$

The threshold value used to re-estimate the final probability is determined to give the highest score in the validation data. In this study, the threshold value of t is equal to 0.64, where the indicator is maximal.

Evaluation criteria

Different criteria were used to evaluate the proposed model. The purpose of the assessment is to identify as many cases as possible from the population for the screening method. Therefore, it is necessary to reduce false negative results by increasing the number of false

positive results. As a result, three main indicators should be identified: the frequency of true positive results (TPR), the frequency of false positive results (FPR) and accuracy (ACC). A different name for the first parameter is called sensitivity (SEN) and is written as equation (4) :

$$TPR = SEN = \frac{TP}{P} \quad (4)$$

Where the number of true positive values is TP, and the number of positive instances is P.

Evaluation of the second term, the frequency of false positive results expressed as an equation (8):

$$FPR = \frac{FP}{N} \quad (5)$$

The total number of negative cases in the population is equal to N, and the proportion of false positive results is equal to FP, and the number of true negative samples is equal to N. This is best interpreted as the ratio of true negative results to known real negative results. The specification is called SPEC, which is given as equation (6):

$$TNR = SPEC = \frac{TN}{N} = 1 - FPR \quad (6)$$

Thus, accuracy determines the balance between real positives and true negatives. This can be a very useful statistic when the number of positive and negative events is not the same. This is expressed as equation (7):

$$ACC = \frac{TP + TN}{P + N} \quad (7)$$

Research results

The dataset was classified using the classification function included in the python scikit-learn library. 80% of the data was used for training and 20% for testing. Accuracy functions were optimized using the Adam optimization algorithm. The proposed CNN model was tested using 15 epoch iterations.

After conducting successful training of the neural network, it is necessary to check its performance. Figure 4 below provides an example of the training phase of the CNN convulsive neural network model. This is the stage of implementation of the architecture of our model. The model architecture provides an 8-layer convulsive neural network for detecting and recognizing road signs. On the first layer of the model, a picture is obtained in 30x30.3 size, that is, 30x30 size and color RGB format. In the next layer, Conv2D is implemented using the ReLU activation function. On the third layer, the maxpool2d Association is applied. The MaxPool2D layer accepts various parameters, including extension, kernel size, step, extension, fill, and return indexes. The fourth Conv2D layer uses the ReLU activation function. The fifth layer again uses the MaxPool2D Association. In the sixth layer, a flatten layer is applied to smooth the inputs. The next layer is called the dense layer or dense layer. It is a layer deeply connected to the anterior layer, that is, the neurons of the stratum communicate with each neuron of its anterior layer. To bring the maximum accuracy obtained in the last layer closer to the probability of the result to 1, and the probability of error classes to 0, the Softmax function is used in the output dense layer. As a result, a recognized road sign is obtained.

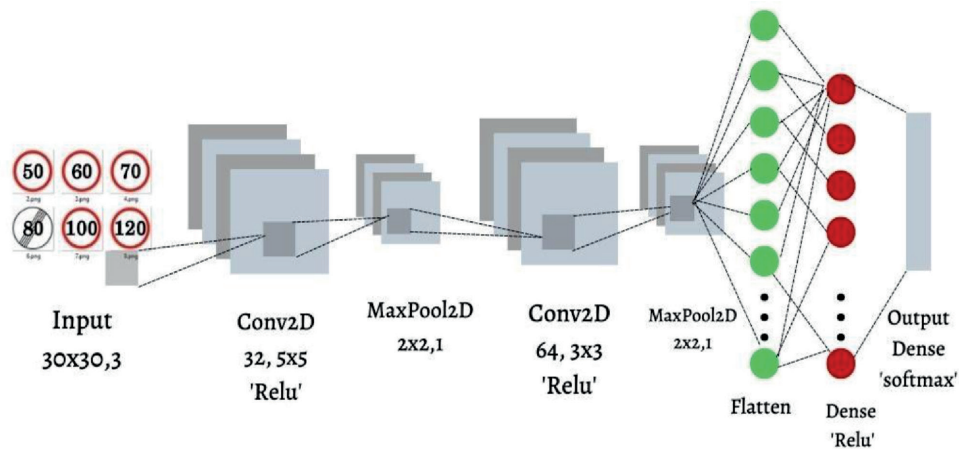


Figure 4. An example from the training phase of the CNN model

Matplotlib is a Python programming language library for data visualization with two-dimensional (2D) graphics (3D graphics are also supported). Using Matplotlib, we plotted the training and fitting accuracy of the model and their error. The figure below shows the accuracy of the model defined using Matplotlib (Figure 5). Accuracy is a discharge characteristic of the mechanical type of a moving point number. This is especially useful for determining the percentage if different classes are the same. It is calculated as the ratio of the number of correct forecasts to their total number and is expressed as a percentage. From the obtained data results, we calculate the error matrix with accuracy based on the Scikit-learn library:

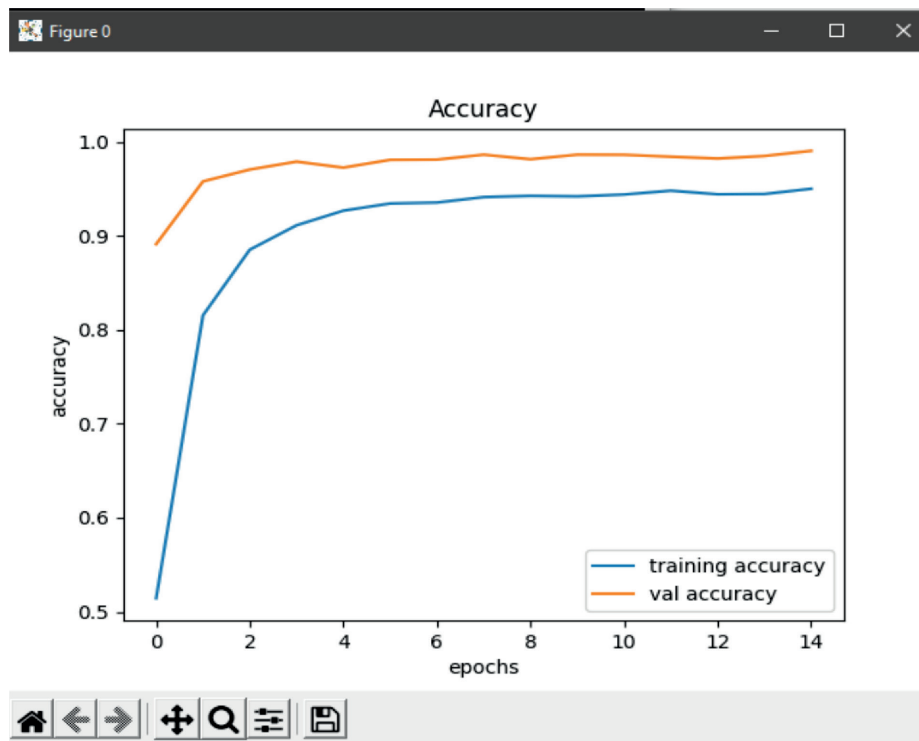


Figure 5. Model training and matching accuracy

In classification problems, the most natural choice is the threshold loss function. Such a loss function is discontinuous, minimizing the empirical risk turns out to be a difficult task of

combinatorial optimization. Therefore, all possible continuous approximations are used. One of the most important steps in model training is the loss function. The loss function is at the heart of the neural network. It is used to calculate the error between the actual and accepted responses (Figure 6).

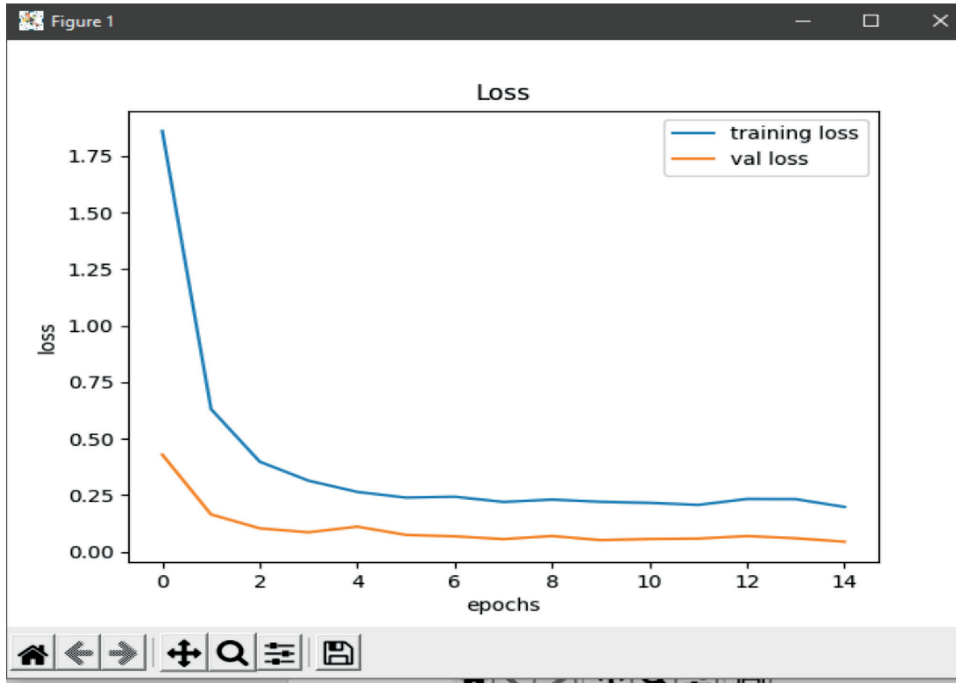


Figure 6. Model training and identification error

As you can see from the images above, the neural network has been successfully trained and is ready to process new images.

The image below shows an example of the feature recognition result of this system (Figure 7).



Figure 7. Traffic signs identified as a result of model training

Based on the results, 95% accuracy of road sign recognition in real time was achieved using in-depth learning in the best defined category. This measure of accuracy is 5.7% higher compared to other studies.

The results of the proposed work were compared with the results of other authors (Table 1). The comparison was carried out taking into account the results of the method used, the data set and the accuracy of the classification. The methods proposed in the latest literature in most cases are based on deep learning. For example, CR-CNN, SR-CNN, C-CNN, R-CNN are based on convolutional neural networks. The 8-layer convolutional neural network we presented showed higher results compared to the latest work. The Adam optimization method used and the proposed model architecture are of great importance in achieving this result.

Table 1. Comparison of the result obtained.

Team	Method	Data sets	Accuracy
Ours	8-layer convolutional neural network	GTSRB	95%
Igel, C [17]	Viola-Jones	GTSDb	87.23%
Z. Cai and N. Vasconcelos [18]	CR-CNN	COCO	87.4%
P. Sun et al. [19]	SR-CNN	COCO	83.9%
W. Hai et al. [20]	C-CNN	GTSDb	85.4%
Z. Zhao et al. [21]	R-CNN	MS COCO	86.9%
J. Cao et al. [22]	Faster R-CNN	TT100K	91%
X. Kuang et al. [23]	HSV-HOG	GTSDb	92.67%

Conclusion

The tasks set during the research were completed. A new way of identifying and recognizing traffic signs was proposed. A solution has been found to increase the size of the dataset. To improve network performance, focus loss is used to control the network of regional proposals. In addition, 3 convolutional and one fully connected layers were used for the detection of road signs, which is very useful in case of large road signs.

As the relevance of the work, it should be noted that the CNN convolutional network model was trained using deep learning, that is, based on neural networks. The novelty of the work is that the model obtains the highest value of the accuracy indicator using the new data set and the classification method.

In the works of other authors, the CNN model has an average accuracy of 90%, while our proposed network increased the accuracy rate of the retrained CNN model to 95% during the research. In addition, dense blocks are used in the classification network to increase accuracy in the recognition phase. The proposed approach performs very well in detecting and recognizing different categories of traffic signs.

Based on the results, 95% accuracy of real-time traffic sign recognition was achieved using deep learning in the best-defined category. The proposed research work can greatly help in reducing the number of traffic accidents.

References

1. Russian-Kazakh legal Explanatory Dictionary-reference book. (2008). Almaty: Zheti zhargy.
2. Number of road accidents. (2020). Retrieved from <https://stat.gov.kz/api/getFile/?docId=ESTAT419824>
3. State of accounting for the causes of road accidents in 2020. (2020). Retrieved from <https://stat.gov.kz/api/getFile/?docId=ESTAT101250>

4. Bouti, A., Mahraz, M.A., Riffi, J., & Tairi, H. (2020). A robust system for road sign detection and classification using LeNet architecture based on convolutional neural network. *Soft Computing*, 24(9), 6721–6733. <https://doi.org/10.1007/s00500-019-04307-6>
5. Sun, Y., Ge, P., & Liu, D. (2019, November). Traffic sign detection and recognition based on convolutional neural network. In *2019 Chinese Automation Congress (CAC)* (pp. 2851-2854). IEEE. <https://doi.org/10.1109/CAC48633.2019.8997240>
6. Han, C., Gao, G., & Zhang, Y., (2019). Real-time small traffic sign detection with revised faster-RCNN. *Multimedia Tools and Applications*, 78(10), 13263-13278. <https://doi.org/10.1007/s11042-018-6428-0>
7. Devyatkin, A.V., & Filatov, D.M. (2019, May). Neural network traffic signs detection system development. In *2019 XXII International Conference on Soft Computing and Measurements (SCM)* (pp. 125-128). IEEE. <https://doi.org/10.1109/SCM.2019.8903787>
8. Shao, F., Wang, X., Meng, F., Zhu, J., Wang, D., & Dai, J. (2019). Improved faster R-CNN traffic sign detection based on a second region of interest and highly possible regions proposal network. *Sensors*, 19(10). <https://doi.org/10.3390/s19102288>
9. Wu, Y., Li, Z., Chen, Y., Nai, K., & Yuan, J. (2020). Real-time traffic sign detection and classification towards real traffic scene. *Multimedia Tools and Applications*, 79(25-26), 18201-18219. <https://doi.org/10.1007/s11042-020-08722-y>
10. Zheng, W., Zhu, X., Wen, G., Zhu, Y., Yu, H., & Gan, J. (2018). Unsupervised feature selection by self-paced learning regularization. *Pattern recognition letters*, 132(4-11). <https://doi.org/10.1016/j.patrec.2018.06.029>
11. Zheng, W., Zhu, X., Zhu, Y., Hu, R., & Lei, C. (2018). Dynamic graph learning for spectral feature selection. *Multimedia tools and applications*, 77(22), 29739-29755. <https://doi.org/10.1007/s11042-017-5272-y>
12. Li, X., Ma, H., Wang, X., & Zhang, X. (2018). Traffic light recognition for complex scene with fusion detections. *IEEE Transactions on Intelligent Transportation Systems*, 19(1), 199-208. <https://doi.org/10.1109/tits.2017.2749971>
13. Hechri, A., & Mtibaa, A. (2020). Two-stage traffic sign detection and recognition based on SVM and convolutional neural networks. *IET Image Processing*, 14(5), 939-946. <https://doi.org/10.1049/iet-ipr.2019.0634>
14. Lin, Z., Yih, M., Ota, J.M., Owens, J., & Muyan-Ozcelik, P. (2019). Benchmarking Deep Learning Frameworks and Investigating FPGA Deployment for Traffic Sign Classification and Detection. *IEEE Transactions on Intelligent Vehicles*, 4(3), 385-395. <https://doi.org/10.1109/tiv.2019.2919458>
15. Tian, Y., Gelernter, J., Wang, X., Li, J., & Yu, Y. (2020). Traffic Sign Detection Using a Multi-Scale Recurrent Attention Network. *IEEE Transactions on Intelligent Transportation Systems*, 20(12), 4466-4475. <https://doi.org/10.1109/tits.2018.2886283>
16. Krizhevsky, A., Sutskever, I., & Hinton, G. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
17. Igel, C. (2013). Detection of traffic signs in real-world images: The German traffic sign detection benchmark. In *Proceedings of the International Joint Conference on Neural Networks, Dallas, TX, USA*, 4–9.
18. Cai, Z., & Vasconcelos, N. (2018). Cascade R-CNN: Delving into high quality object detection, in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 6154–6162. <https://doi.org/10.1109/CVPR.2018.00644>
19. Sun, P., Zhang, R.Y., Jiang, T., Kong, C., Xu, W., Zhan, M., Tomizuka, L., Li, Z., Yuan, C., Wang & Luo, P. (2020). Sparse R-CNN: End-to-end object detection with learnable proposals. arXiv:2011.12450
20. Hai, W., Kuan, W., Yingfeng, C., Ze, L. & C. Long. (2020). Traffic sign recognition based on improved cascade convolution neural network. *Automot. Eng.*, 42, 1256–1262.
21. Zhao, Z., Li, X., Liu, H. & Xu, C. (2020). Improved target detection algorithm based on libra R-CNN. *IEEE Access*, 8, 114044–114056. <https://doi.org/10.1109/ACCESS.2020.3002860>
22. Cao, J., Zhang, J. & Huang, W. (2021). Traffic sign detection and recognition using multi-scale fusion and prime sample attention. *IEEE Access*, 9, 3579–3591. <https://doi.org/10.1109/ACCESS.2020.3047414>
23. Kuang, X., Fu, W., & Yang, L. (2018). Real-time detection and recognition of road traffic signs using MSER and random forests. *Int J Online Eng*, 14(03), 34–51. <https://doi.org/10.3991/ijoe.v14i03.7925>