

A Traffic Signal Recognition Algorithm Based on Self-paced Learning and Deep Learning

Tingmei Wang*, Haiwei Shen, Yuanjie Xue, Zhengkun Hu

College of Applied Science and Technology, Beijing Union University, Beijing 102200, China

Corresponding Author Email: yykjttingmei@buu.edu.cn



<https://doi.org/10.18280/isi.250211>

ABSTRACT

Received: 2 December 2019

Accepted: 25 January 2020

Keywords:

traffic signal recognition, self-paced learning (SPL), machine learning (ML), deep learning (DL), unmanned driving

Traffic signal recognition is a critical function of the intelligent vehicle system (IVS). Many algorithms can achieve a high accuracy in traffic signal recognition. But these algorithms have poor generalization ability, and their recognition rates vary greatly with datasets. These defects hinder their application in unmanned driving. To solve the problem, this paper introduces self-paced learning (SPL) to the image recognition of traffic signs. Based on complexity, the SPL automatically classifies samples into multiple sets. If machine learning (ML) algorithm is trained by the sample sets in ascending order of complexity, a universal computing model will be obtained, and the ML algorithm will have a better generalization ability. Here, the support vector machine (SVM) is adopted as the classifier for traffic sign detection, and the convolutional neural network (CNN) is employed as the classifier for traffic sign recognition. Then, the two classifiers were trained by the SPL on two public datasets: German Traffic Sign Detection Benchmark (GTSDB) and German Traffic Sign Recognition Benchmark (GTSRB). The model obtained through the training was tested on Belgium Traffic Sign Detection Benchmark (BTSDB) and KITTI datasets. The results show that the obtained computing model achieved similar accuracy on the training sets and test sets. Hence, the SPL can indeed enhance the generalization ability of ML algorithms, and promote the application of CNN, SVM, and other ML algorithms in unmanned driving.

1. INTRODUCTION

Traffic signal recognition is a critical function of the intelligent vehicle system (IVS). Since the 1990s, many machine learning (ML) algorithms have been applied to realize traffic signal recognition of the IVS, namely, principal component analysis (PCA) [1], random forest (RF) [2], support vector machine (SVM) [3], convolutional neural network (CNN) [4-6], and sparse representation [7]. These ML algorithms can recognize traffic signals accurately.

However, experiments show that many ML algorithms have a much lower testing accuracy than training accuracy, when they are trained on one dataset (e.g. Belgium Traffic Sign Detection Benchmark (BTSDB)) and tested on another dataset (e.g. German Traffic Sign Detection Benchmark (GTSDB)) [4].

Obviously, the computing models obtained through the training of the ML algorithms are overfitted. The overfitting indicates the weak generalization ability of these algorithms. As a result, the ML algorithms might make more misjudgments in actual scenarios than in training scenarios.

During unmanned driving, the accuracy of image recognition is affected by various complicated factors, namely, the complex background, vehicle jitter, light and shadow, and bad weather (e.g. rain, fog, and snow, see Figures 1-6). Sometimes, the traffic signs are damaged, polluted, blocked, or faded. It is impossible to cover all these situations in the training data. If the traffic signs are not identified stably by the image recognition algorithm, the control system of intelligent vehicles might make misjudgments, resulting in fatal traffic accidents. Therefore, it is a common goal among scholars

engaging in intelligent vehicles to improve the stability and generalization ability of traffic sign recognition algorithms.

Self-paced learning is an emerging method that can effectively enhance the generalization ability of ML algorithms. As a training method for ML algorithms, the self-paced learning (SPL) is developed from curriculum learning (CL). In 2009, Bengio et al. [8] introduced the concept of the CL into the ML, creating the CL training method. The CL is an ancient thought in pedagogy: Since human cognition is a process from simple to complex, the learning objectives of human learning should be divided into phased curriculums in ascending order of complexity. Under the CL training method, the samples are classified to multiple sets based on complexity, each of which is called a curriculum. Then, the ML algorithm learns the samples in each set from simple to complex. Experiments show that the computing models obtained by this training method is unlikely to overfit, and enjoys strong generalization ability.

In the original CL method, the complexity of each curriculum is evaluated manually. However, subjective errors might occur in curriculum classification during the application process. Moreover, a huge workload is incurred in the face of large sample sets. To solve these problems, Kumar et al. [9] proposed the SPL in 2010, aiming to automatically divide the samples into curriculums by complexity in the training process. Specifically, a regular term related to the samples is added to the objective function, such as to give a complexity score of each sample after the training. The samples will be grouped iteratively based on their complexity scores. After that, the model will be trained by the CL method, and new scores will be rated. This process continues until the model no longer

changes in the training. Eliminating manual intervention, the SPL has a great advantage over the CL. Recently, the SPL has been implemented to recognize multiple images and videos (Khan et al. [10]; Basu and Christensen [11]; Tang et al. [12]; Jiang and others [13-22]). Experiments have proved that the SPL training could improve the test accuracy of ML algorithms, and prevent the occurrence of overfitting. This means the SPL enhances the generalization ability of ML algorithms.

The image recognition of traffic signs involves two steps: traffic sign detection, and traffic sign recognition. For real-time performance, the SVM is often adopted for traffic sign detection, while the CNN is widely employed for traffic sign recognition [23-25]. Both SVM and CNN could be trained by the SPL to obtain better generalization ability [26]. Therefore, this paper introduces the SPL training to the image recognition of traffic signs. The SVM and CNN were selected as the classifiers for traffic sign detection and traffic sign recognition, respectively, and subjected to SPL training on two public datasets, namely, GTSDDB and German Traffic Sign Recognition Benchmark (GTSRB). The computing model obtained through the training was tested on BTSDB and KITTI datasets [27]. The test results show that the obtained computing model achieved similar accuracy on the training sets and test sets. Hence, the SPL can indeed solve the overfitting problem in image recognition of traffic signs, enhance the generalization ability of ML algorithms, and create highly applicable traffic sign recognition models.



Figure 1. Normal traffic sign



Figure 2. Faded traffic sign



Figure 3. Traffic sign with a complex background



Figure 4. Damaged traffic sign



Figure 5. Traffic sign in rain



Figure 6. Traffic sign in snow

2. ALGORITHM DESIGN

2.1 Traffic sign detection

For real-time performance, it is impossible to recognize every block in the original image with the traffic sign recognition algorithm. Hence, traffic sign detection algorithm should be adopted to extract the blocks that might be traffic signs from the original image in a quick and accurate manner. For the same reason, it is also impossible to detect every block in the original image with the traffic sign detection algorithm. There must be a step in the algorithm to reduce the number of candidate blocks, i.e. to extract the regions of interest (ROIs).

Following the same set of color and shape standards, traffic sign images from the same region carry obvious features. The main colors are red, yellow, blue, and white, and the typical shapes are triangle, rectangle, circle, and octagon. Therefore, the ROIs in the original image could be separated based on the hue in the hue-saturation-value (HSV) color space, and then denoised based on some shape features.

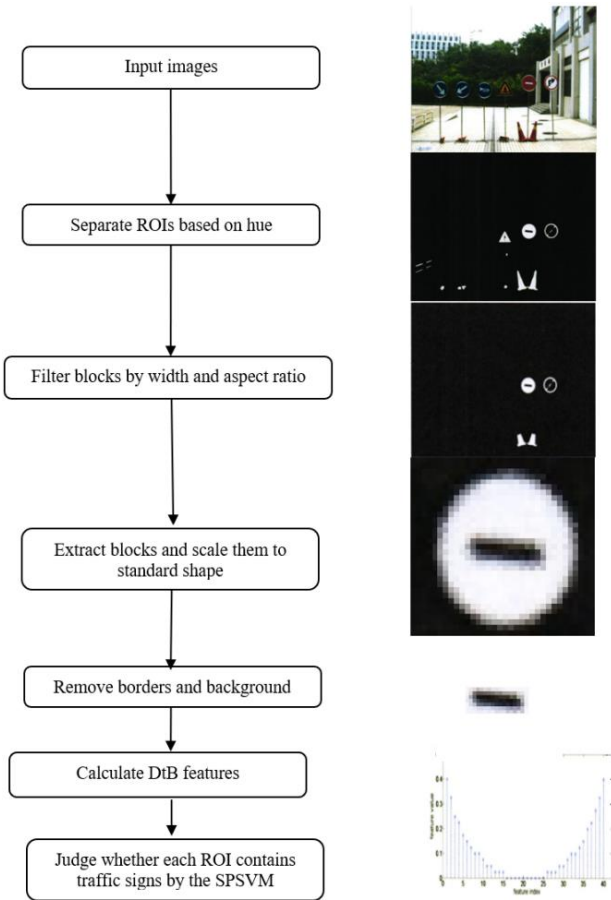


Figure 7. Workflow of traffic flow detection

Most traffic signs have borders and background color. The semantics of traffic signs mostly reside in the graphics within the borders. The borders and background color complicates the original image, reducing the detection accuracy. To solve the problem, the traffic sign detection algorithm should also remove the borders, convert the original image into a binary image, and eliminate noises, leaving only the important shape features in the ROIs.

For unified processing, the traffic sign detection algorithm should scale the ROIs to a standard size, and describe the shape features of the ROIs with distance to border (DtB). Hence, both hue and shape features could be employed to depict the ROIs. Figure 7 above illustrates the ROI extraction procedure.

It is a binary classification problem to judge whether an ROI is a traffic sign. The SVM is exceptionally good at obtaining the discriminant model of binary classification problems quickly and accurately. As a result, the SVM has frequently been adopted as a classifier in traffic sign detection algorithms.

Extracted from the camera data of intelligent vehicles, the traffic sign images may have abnormalities like occlusion, and color fading. These abnormalities will make the hue and shape features of the images different from expected. A good traffic sign detection algorithm should be able to accurately detect all traffic sign images with abnormalities.

Taking all normal images as simple samples and all abnormal images as complex samples, the classifier of the traffic sign detection algorithm could be trained by the CL strategy. Under this strategy, the characteristic parameters of simple samples are extracted from simple curriculums, while the extra characteristic parameters of complex samples are extracted from complex curriculums. With the extracted parameters, the obtained computing model could distinguish

as many samples as possible, and thus approximate the real model.

Of course, it is a heavy work to classify original images into curriculums, owing to the sheer number of traffic sign samples. Thus, this paper decides to train the SVM by the SPL, and integrates the training method with the learning method into a novel algorithm called SPSVM. The performance of the SPSVM was tested in three steps: Firstly, the ROI extraction method was implemented to extract the candidate blocks from the training set and the test set; Next, every block was given a label about whether it is a traffic sign; Finally, the discriminant model was obtained by applying the SPSVM on the training set, and used to discriminate the training set and the test set, producing the accuracy of the SPSVM on the two datasets.

2.2 Traffic sign recognition

The images confirmed by the traffic sign detection algorithm as traffic sign images are the input data of the traffic sign recognition algorithm, which determines the semantics of these images, i.e. allocate each image into its class of traffic signs. Considering the various types of traffic signs, the recognition process is a multi-classification problem.

The DL algorithms boast relatively high classification accuracy for multi-classification problems. Among them, the CNN has been widely reported as the most accurate traffic sign recognition algorithm. Therefore, this paper takes the CNN as the classifier of the traffic sign recognition algorithm.

Our CNN consists of three convolutional layers and a fully connected layer. The first convolutional layer has 100 7×7 filters, each of which convolve a 7×7 neighborhood in the input image. The second convolutional layer has 150 4×4 filters; The third convolutional layer has 250 4×4 filters; The fully connected layer contains 300 neurons, and outputs 43 features. The input data of the CNN are black and white traffic sign images scaled to the pixel size of 30×30 . The output data are the probability of each input image belonging to a type of traffic signs.

The input data of the traffic sign recognition algorithm contain both normal samples and abnormal samples. The abnormality of abnormal samples refers to the abnormal shape features induced by complex situations, such as occlusion, jitter, and noise. The CL strategy could also be applied to train the CNN. Under this strategy, the CNN could extract the characteristic parameters of simple samples from simple curriculums, and the extra characteristic parameters of complex samples from complex curriculums.

Therefore, this paper decides to train the CNN by the SPL, and integrates the training method with the learning method into a novel algorithm called SPCNN. The performance of the SPCNN was tested in three steps: Firstly, the images detected from training set and test set were all given labels about which type of traffic signs they belong to; Next, the traffic sign images detected from the training set were used to train the CNN; Finally, the trained CNN was adopted to recognize all the images detected from the training set and the test set, producing the accuracy of the SPCNN on the two datasets.

2.3 The SPL

The details of the SPL are given below. Let $D = \{(x_i, y_i)\}_{i=1}^n$ be the training set, where x_i is the i -th sample, and y_i is the label of the i -th sample. Suppose $L(y_i, g(x_i, w))$ is the loss of the prediction model brought by the i -th sample, where

g is the prediction function, and w is the parameter of the prediction model. Then, the objective function of the SPL optimization can be defined as:

$$\begin{aligned} & \min_{w, v \in \{0,1\}^n} E(w, v; \lambda) \\ & = \sum_{i=1}^n (v_i L(y_i, f(x_i, w)) + h(v_i, \lambda)) \end{aligned} \quad (1)$$

where, h is the self-paced regular function, whose attributes and properties are introduced by Meng et al. [28]. The following self-paced regular function is selected for this research:

$$h(v, \lambda) = -\lambda v \quad (2)$$

For the CNN, the loss function can be expressed as:

$$L(y_i, f(x_i, w)) = -\log p(y = y_i | x) \quad (3)$$

where, $p(y=y_i|x)$ is the value of CNN output corresponding to the class y_i .

For the SVM, the hinge loss function can be expressed as:

$$L(y_i, f(x_i, w)) = \max\{0, 1 - (w^T x_i + b) y_i\} \quad (4)$$

where, b is the bias.

Obviously, after each training, the loss of each sample can be obtained easily from the outputs of the CNN and the SVM.

Eq. (1) can be solved by the expectation-maximization (EM) algorithm (Algorithm 1).

In the EM algorithm, dataset D is the set of all samples; the array v is used to select the samples for each training; w is the parameter value of the ML model; L is the value of the loss function; λ is the threshold for sample groups; s is used to determine the increment of sample complexity in each training.

During the SPL, the training is divided into multiple rounds. In each round, the training samples include all the simple samples and part of the complex samples. The sample composition reflects the knowledge acquisition in education process: reviewing simple knowledge, and probing deep into complex knowledge.

Algorithm 1: The EM algorithm

Input: Input dataset D , w_0 , λ , and s

Output: w of $\min_v E(w, v; \lambda)$.

1 repeat

2 Calculate L of every sample by w ;

3 Sort the samples in ascending order of their loss values L ;

4 for $i=1$ to n

5 if $L(y_i, f(x_i, w)) < \lambda$ then $v_i=1$; // select this sample

6 else

7 $v_i=0$; // do not select this sample

8 end

9 end

10 Update w by solving $\min \sum_{i=1}^n v_i L_i(y_i, f(x_i, w)) + \frac{1}{2} \|w\|_2$

11 $\lambda = \lambda + s$

12 until w is not changed or $v_i=1, i=1, 2, \dots, n$.

3. EXPERIMENTS

3.1 Datasets

Four datasets were selected for our experiments, including GTSDDB, KITTI, BTSDB, and GTSRB. The former two were for training, and the latter two were for testing. The GTSRB contains 51,840 images containing traffic signs; the KITTI includes 400 1,242×375 images captured by two cameras (left and right); the GTSDDB offers 900 1,360×800 images; the BTSDB provides 3,133 images shot by eight cameras.

3.2 Traffic sign detection

The SVM is well known for its speed and effectiveness in binary classification problems. Therefore, the SPSVM was applied to detect the ROIs extracted by the classifier, and judge whether the ROIs contain traffic signs. The SPSVM was compared with ChnFtrs [29], 3D TS [30], Bayesian classifier, and the original SVM, which does not use the SPL.

The detection results were evaluated by four metrics, Accuracy, False Rate, Area under Curve (AUC), and Average Precision (AP). Accuracy is the ratio of the number of correctly recognized traffic signs to the number of all traffic signs; False Rate is the ratio of the number of incorrectly recognized traffic signs to the number of all traffic signs; AUC is the area under the receiver operating characteristic (ROC) curve; AP is the area under the Precision-Recall (PR) curve. The AUC has been adopted in many experiments to evaluate the performance of the classifier. However, the AP outperforms the AUC when the number of samples is unevenly distributed across different categories. This metric was selected for our experiments for the following reasons: In our samples, the number of images with traffic sign has a huge difference from that of images without traffic sign.

In our experiments, the images were not classified by the labels (i.e. mandatory, dangerous, and prohibited) of the datasets. Instead, every ROI extracted in the previous step was allocated one of the two categories: “Yes” (the ROI contains traffic sign) and “No” (the ROI does not contain traffic sign), making traffic sign detection a binary classification problem rather than a multi-classification problem.

The results (Table 1) show that the five algorithms had similar performance in traffic sign detection on the training set. However, the SPSVM outperformed the other algorithms on the test set. Besides, the Accuracy values of the SPSVM on the two datasets differed by less than 1%. This means the SPL indeed enhances the generalization ability of the SVM.

3.3 Traffic sign recognition

The SPCNN code was programmed on MATLAB, and compared with IDSIA DNN [12], IDSIA MCDNN [12], and Multi-Scale CNN [4]. The detection results were evaluated by three metrics: Accuracy, mean Area under Curve (mAUC), and mean Average Precision (mAP). According to the results of traffic sign recognition (Table 2), the proposed SPCNN achieved better performance than the contrastive methods on the test set. Therefore, the SPL training overcomes the overfitting problem, which is common to the DL algorithms, and improves the generalization ability of the CNN.

Table 1. The results of traffic sign detection

	Training				Test			
	Accuracy	False Rate	AUC	AP	Accuracy	False Rate	AUC	AP
ChnFtrs	98.53	3.21	0.9721	0.9272	94.32	9.21	0.9215	0.8927
3D TS	97.72	4.54	0.9613	0.9193	95.54	10.13	0.9223	0.8902
Bayesian	95.36	5.18	0.9549	0.9247	94.72	10.35	0.9107	0.8824
SVM	97.39	4.23	0.9687	0.9231	94.21	10.87	0.9114	0.8892
SPSVM	98.15	2.73	0.9837	0.9612	97.17	5.53	0.9822	0.9531

Table 2. The results of traffic sign recognition

	Training			Test		
	Accuracy	AUC	AP	Accuracy	AUC	AP
IDSIA DNN	98.52	0.9415	0.9012	88.27	0.8443	0.8093
IDSIA MCDNN	99.46	0.9512	0.9194	87.16	0.8310	0.8024
Multi-Scale CNN	99.17	0.9426	0.9163	88.92	0.8469	0.8102
SPCNN	99.35	0.9748	0.9284	95.78	0.9353	0.9017

4. CONCLUSIONS

In traffic sign detection and recognition, the recognition rates of many algorithms vary greatly from dataset to dataset. To solve the problem, this paper introduces the SPL to train the ML algorithms, and verifies the generalization ability of the trained model through numerical experiments. The experimental results show that our model could achieve similar Accuracies on training set and test set. This means the SPL could reduce the overfitting problem, and enhance the generalization ability of the ML algorithms. Thus, the SPL is suitable for training traffic sign recognition models with a high requirement on precision.

In this research, the SPL is innovatively applied to the detection and recognition of traffic signs. However, the datasets are not sufficiently large, the traffic sign images are not highly diversified, and the complex situations are rather limited. As a result, the obtained model cannot be directly applied to actual unmanned driving control systems. In the future research, more largescale datasets from different countries will be learned, and the scale of the CNN will be expanded, aiming to create a highly accurate recognition model for industrial use.

The SPL has not been extensively studied, when compared with immensely popular learning methods like reinforcement learning [31], active learning [32], and ensemble machine learning [33]. Many researches fail to consider the SPL, because this learning strategy, only capable of enhancing the generalization ability of the target algorithm, performs poorly on the training set. However, many experiments have proved that the SPL helps to develop highly universal models for the ML algorithms, making the learning results are applicable to real-world scenarios. As a result, it is of great value to promote the SPL.

ACKNOWLEDGMENT

The work was supported by The Research and Practice on The Through-Type Training Mode of High-End Technologies and Personnel with Technical Skills of Beijing, Beijing Municipal Education Commission, Beijing, China (Grant No.: 2018-54), The Scientific Research Program Project of Beijing Education Commission on Traffic Light Recognition Based on Deep Learning for Intelligent Vehicles (Grant No.:

KM201911417003), The Famous Teachers of Beijing and The Academic Research Project of Vocational Education and Industry Research Center, College of Applied and Science, Beijing Union University.

REFERENCES

- [1] Gao, H., Liu, C., Yu, Y., Li, B. (2014). Traffic signs recognition based on PCA-SIFT. In Proceeding of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, pp. 5070-5076. <https://doi.org/10.1109/WCICA.2014.7053576>
- [2] Zaklouta, F., Stanculescu, B., Hamdoun, O. (2011). Traffic sign classification using KD trees and random forests. In The 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, pp. 2151-2155. <https://doi.org/10.1109/IJCNN.2011.6033494>
- [3] Feng, G., Ma, L., Tan, X. (2014). Ground traffic signs recognition based on Zernike moments and SVM. In 2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference, Chongqing, China, pp. 478-481. <https://doi.org/10.1109/ITAIC.2014.7065096>
- [4] Jang, C., Kim, H., Park, E., Kim, H. (2016). Data debiased traffic sign recognition using MSERs and CNN. In 2016 International Conference on Electronics, Information, and Communications (ICEIC), Da Nang, Vietnam, pp. 1-4. <https://doi.org/10.1109/ELINFOCOM.2016.7562938>
- [5] Liu, S.Z. (2019). A traffic sign image recognition and classification approach based on convolutional neural network. 2019 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), Qiqihar, China, pp. 408-411. <https://doi.org/10.1109/ICMTMA.2019.00096>
- [6] How, D.N.T., Sahari, K.S.M., Hou, Y.C., Basubeit, O.G.S. (2019). Recognizing Malaysia traffic signs with pre-trained deep convolutional neural networks. In 2019 4th International Conference on Control, Robotics and Cybernetics (CRC), Tokyo, Japan, pp. 109-113. <https://doi.org/10.1109/CRC.2019.00030>
- [7] Li, C., Hu, Y., Xiao, L., Tian, L. (2012). Salient traffic sign recognition based on sparse representation of visual perception. In 2012 International Conference on

- Computer Vision in Remote Sensing, Xiamen, China, pp. 273-278. <https://doi.org/10.1109/CVRS.2012.6421274>
- [8] Bengio, Y., Louradour, J., Collobert, R., Weston, J. (2009). Curriculum learning. In Proceedings of the 26th Annual International Conference on Machine Learning, pp. 41-48. <https://doi.org/10.1145/1553374.1553380>
- [9] Kumar, M.P., Packer, B., Koller, D. (2010). Self-paced learning for latent variable models. In Advances in Neural Information Processing Systems, pp. 1189-1197.
- [10] Khan, F., Mutlu, B., Zhu, J. (2011). How do humans teach: On curriculum learning and teaching dimension. In Advances in Neural Information Processing Systems, pp. 1449-1457.
- [11] Basu, S., Christensen, J. (2013). Teaching classification boundaries to humans. In Twenty-Seventh AAAI Conference on Artificial Intelligence.
- [12] Tang, Y., Yang, Y.B., Gao, Y. (2012). Self-paced dictionary learning for image classification. In Proceedings of the 20th ACM International Conference on Multimedia, pp. 833-836. <https://doi.org/10.1145/2393347.2396324>
- [13] Jiang, L., Meng, D., Mitamura, T., Hauptmann, A.G. (2014). Easy samples first: Self-paced reranking for zero-example multimedia search. In Proceedings of the 22nd ACM International Conference on Multimedia, pp. 547-556. <https://doi.org/10.1145/2647868.2654918>
- [14] Jiang, L., Meng, D., Yu, S.I., Lan, Z., Shan, S., Hauptmann, A. (2014). Self-paced learning with diversity. In Advances in Neural Information Processing Systems, pp. 2078-2086.
- [15] Jiang, L., Meng, D., Zhao, Q., Shan, S., Hauptmann, A.G. (2015). Self-paced curriculum learning. In Twenty-Ninth AAAI Conference on Artificial Intelligence.
- [16] Zhao, Q., Meng, D., Jiang, L., Xie, Q., Xu, Z., Hauptmann, A.G. (2015). Self-paced learning for matrix factorization. In Twenty-ninth AAAI Conference on Artificial Intelligence.
- [17] Zhang, D., Meng, D., Li, C., Jiang, L., Zhao, Q., Han, J. (2015). A self-paced multiple-instance learning framework for co-saliency detection. In Proceedings of the IEEE International Conference on Computer Vision, pp. 594-602.
- [18] Gong, M., Li, H., Meng, D., Miao, Q., Liu, J. (2018). Decomposition-based evolutionary multiobjective optimization to self-paced learning. IEEE Transactions on Evolutionary Computation, 23(2): 288-302. <https://doi.org/10.1109/TEVC.2018.2850769>
- [19] Li, H., Gong, M. (2017). Self-paced Convolutional Neural Networks. In IJCAI, pp. 2110-2116.
- [20] Meng, D., Zhao, Q., Jiang, L. (2017). A theoretical understanding of self-paced learning. Information Sciences, 414: 319-328. <https://doi.org/10.1016/j.ins.2017.05.043>
- [21] Zhang, D., Meng, D., Han, J. (2016). Co-saliency detection via a self-paced multiple-instance learning framework. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(5): 865-878. <https://doi.org/10.1109/TPAMI.2016.2567393>
- [22] Lin, L., Wang, K., Meng, D., Zuo, W., Zhang, L. (2017). Active self-paced learning for cost-effective and progressive face identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(1): 7-19. <https://doi.org/10.1109/TPAMI.2017.2652459>
- [23] Yang, Y., Luo, H., Xu, H., Wu, F. (2015). Towards real-time traffic sign detection and classification. IEEE Transactions on Intelligent Transportation Systems, 17(7): 2022-2031. <https://doi.org/10.1109/TITS.2015.2482461>
- [24] CireşAn, D., Meier, U., Masci, J., Schmidhuber, J. (2012). Multi-column deep neural network for traffic sign classification. Neural Networks, 32: 333-338. <https://doi.org/10.1016/j.neunet.2012.02.023>
- [25] Sermanet, P., LeCun, Y. (2011). Traffic sign recognition with multi-scale convolutional networks. In The 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, pp. 2809-2813. <https://doi.org/10.1109/IJCNN.2011.6033589>
- [26] Graves, A., Bellemare, M. G., Menick, J., Munos, R., Kavukcuoglu, K. (2017). Automated curriculum learning for neural networks. In Proceedings of the 34th International Conference on Machine Learning, 70: 1311-1320. <https://doi.org/10.5555/3305381.3305517>
- [27] Geiger, A., Lenz, P., Stiller, C., Urtasun, R. (2013). Vision meets robotics: The Kitti dataset. The International Journal of Robotics Research, 32(11): 1231-1237. <https://doi.org/10.1177%2F0278364913491297>
- [28] Meng, D., Zhao, Q., Jiang, L. (2015). What objective does self-paced learning indeed optimize? arXiv preprint arXiv:1511.06049.
- [29] Mathias, M., Timofte, R., Benenson, R., Van Gool, L. (2013). Traffic sign recognition—how far are we from the solution? In The 2013 international joint conference on Neural networks (IJCNN), Dallas, TX, USA, pp. 1-8. <https://doi.org/10.1109/IJCNN.2013.6707049>
- [30] Timofte, R., Zimmermann, K., Van Gool, L. (2014). Multi-view traffic sign detection, recognition, and 3D localisation. Machine Vision and Applications, 25(3): 633-647. <https://doi.org/10.1007/s00138-011-0391-3>
- [31] Kaelbling, L.P., Littman, M.L., Moore, A.W. (1996). Reinforcement learning: A survey. Journal of Artificial Intelligence Research, 4: 237-285. <https://doi.org/10.1613/jair.301>
- [32] Settles, B. (2009). Active learning literature survey. University of Wisconsin-Madison Department of Computer Sciences.
- [33] Zhang, C., Ma, Y. (2012). Ensemble machine learning: Methods and applications. Springer Science & Business Media.