A New Indonesian Traffic Obstacle Dataset and Performance Evaluation of YOLOv4 for ADAS

Agus Mulyanto[1*], Wisnu Jatmiko[2], Petrus Mursanto[2], Purwono Prasetyawan[3] & Rohmat Indra Borman[1]

[1]Faculty of Engineering and Computer Science, Universitas Teknokrat Indonesia, Bandar Lampung, Indonesia
[2]Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia
[3]Department of Electrical Engineering, Institut Teknologi Sumatera, Lampung Selatan, Indonesia
*E-mail: agus.mulyanto@teknokrat.ac.id

Abstract. Intelligent transport systems (ITS) are a promising area of studies. One implementation of ITS are advanced driver assistance systems (ADAS), involving the problem of obstacle detection in traffic. This study evaluated the YOLOv4 model as a state-of-the-art CNN-based one-stage detector to recognize traffic obstacles. A new dataset is proposed containing traffic obstacles on Indonesian roads for ADAS to detect traffic obstacles that are unique to Indonesia, such as pedicabs, street vendors, and bus shelters, and are not included in existing datasets. This study established a traffic obstacle dataset containing eleven object classes: cars, buses, trucks, bicycles, motorcycles, pedestrians, pedicabs, trees, bus shelters, traffic signs, and street vendors, with 26,016 labeled instances in 7,789 images. A performance analysis of traffic obstacle detection on Indonesian roads using the dataset created in this study was conducted using the YOLOv4 method.

Keywords: *ADAS; convolutional neural network (CNN); Indonesian Traffic Obstacle Dataset; intelligent transport systems (ITS); YOLOv4.*

1 Introduction

Nowadays, intelligence transport systems (ITS) such as advanced driver-assistance systems (ADAS) for self-driving cars are widely used [1,2]. One of the challenges in ADAS implementation is obstacle detection, which should be done with high accuracy to ensure that the system works well. Prior research focused on obstacle detection in ADAS using a monocular camera and odometry [3], while other researchers used deep learning in obstacle detection, achieving high accuracy [4,5].

In the last few years, multiple deep learning algorithms, i.e. convolutional neural networks (CNN), have been applied to ITS, especially in traffic obstacle detection systems. Object detection using deep learning has been done since 2013, when

Received December 10th, 2019, 1st Revision November 20th, 2020, 2nd Revision January 14th, 2021, 3rd Revision February 19th, 2021, Accepted for publication March 23rd, 2021.
 ISSN: 2337-5787, DOI: 10.5614/itbj.ict.res.appl.2021.14.3.6

Alex, *et al.* proposed using convolutional neural networks in object detection [6]. It was continued in 2015, when He, *et al.* [7] applied residual neural networks (Resnet) to improve using plain CNN. Meanwhile, the development of Fast RCNN [8], Faster RCNN [9], SSD [10], and YOLO [11] improved the performance of object detection methods, such as their accuracy and time of inference.

The availability of datasets is one of the determining factors in object detection performance. Large-scale image datasets such as ImageNet [12], PASCAL-VOC [13], MS COCO [14], GTSRB [15], KITTI [16], SYNTHIA [17], and Urban Object Detection [18] have been used by researchers for several purposes, including ITS, showing satisfactory performance. However, there remain some problems concerning the collection of data for deep learning. Existing datasets were built with various approaches and specific needs; hence, when applied to special needs it is necessary to use a customized dataset. For example, an urban object detection dataset [18] was developed with images from European roads, focusing on traffic conditions and seven classes of obstacles, i.e. cars, motorbikes, persons, traffic lights, buses, bicycles, and traffic signs. In the Indonesian road environment there are obstacles that are not the same as in other countries, such as street vendors, pedicabs, bus shelters, unique streets, and others. Therefore, a dataset representing Indonesian roads is needed, containing different traffic conditions, signs, and obstacles.

This study created a new traffic obstacle dataset consisting of objects and road obstacles in Indonesia to overcome deficiencies in existing datasets, especially concerning three new object classes: pedicabs, bus shelters, and street vendors. This dataset was used to evaluate the performance of a state-of-the-art deep learning method. More specifically the contributions of this study are as follows:

1. A new Indonesian traffic obstacle dataset was created for further research on ADAS with 26,016 labeled instances in 7,789 images. It is divided into eleven classes: cars, buses, trucks, bicycles, motorcycles, pedestrians, pedicabs, trees, bus shelters, traffic signs, and street vendors.
2. The proposed dataset was evaluated using YOLOv4 as a state-of-the-art of object detection technique based on a CNN one-stage detector.

The rest of the study is organized as follows: related work is described in Section 2; the creation of the dataset is explained in Section 3; a discussion of the experimental results for evaluation of the Indonesian Traffic Obstacle Dataset is presented in Section 4; Section 5 presents the conclusions and further study.

2 Related Works

Datasets are important in setting goals for models or methods in deep learning research to allow making performance comparisons. Datasets can be used to train and evaluate algorithms for more specific research, hence, they have to deal with several challenges. Many datasets have been built, such as ImageNet [12], PASCAL-VOC [13], COCO [14], GTSRB [15], KITTI [16], SYNTHIA [17], Urban Object Detection [18]. The datasets were developed and evaluated in the context of existing problems. ImageNet [12], PASCAL-VOC [13], and COCO [14] are large-scale datasets developed for various purposes. These datasets contain images with common objects such as animals, vehicles, plants, buildings, furniture, etc. in indoor or outdoor environments. Before the training process from the image dataset occurs, there are various pre-processing steps that should be conducted, such as morphological data filtering [19] and perceptual image adaptation [20].

Datasets such as GTSRB [15], KITTI [16], SYNTHIA [17], and Urban Object Detection [18] have been developed specifically for intelligent transport systems containing transportation-related objects such as vehicles, pedestrians, cyclists, traffic signs, traffic lights, and miscellaneous objects (e.g. trailers, Segways). The determination of objects contained in the dataset is based on the context of the problem to be solved and the research location. For example, the GTSRB dataset [15] contains 50,000 traffic sign images taken on different road types in Germany.

The Urban Object Detection Dataset [18] has seven traffic classes (cars, motorbikes, persons, traffic lights, buses, bicycles, and traffic signals), which were extracted from several different public datasets: PASCAL-VOC [13] provided 22%, Udacity [21] provided 65% and it was added with images captured in urban environments and on roads in Alicante, Spain. Another example is the Traffic Dataset from Linköping University (Sweden) [22]. The size and quality of the images in the different datasets is not the same. Therefore, it is necessary to balance data augmentation and size reduction, taking into account rotation or orientation problems, level of blur, image size (zoom in and zoom out), object transformation or position, and other factors.

Deep learning methods based on a convolutional neural network (CNN) with a two-stage detector approach, e.g. SPPNet [23], Pyramid Network [24], RCNN: Fast RCNN [8], Faster RCNN [25], or a one-stage detector approach, e.g. YOLO [11], YOLOv2 [26], YOLOv3 [27], YOLOv4 [28], SSD [10], RetinaNet [29], have begun to be widely applied in the world of computer vision. CNN-based deep learning methods are used explicitly for object detection and classification systems, including for intelligent transport systems, such as autonomous driving or self-driving cars [30], traffic monitoring [31,32], and advanced driver assistance systems [33,34].

One of the deep-learning functions in ITS is to detect objects in the traffic environment, such as obstacles, vehicles, pedestrians, traffic signs, and also to allow trajectory estimation of moving objects [35]. YOLOv4 is one of the state-of-the-art applications of CNN based a one-stage detector released in 2020. YOLOv4 improved FPS and average precision (AP) by 12% and 10% compared to its predecessor, YOLOv3 [28]. YOLOv4 showed 65.7% AP50 performance in training, using the MS COCO dataset, and it is capable of running at speed on a real-time system of ~65 FPS in a Tesla V100.

3 Constructing Indonesian Traffic Obstacle Dataset

The Indonesian Traffic Obstacle Dataset consists of eleven object classes: cars, buses, trucks, bicycles, motorcycles, pedestrians, pedicabs, trees, bus shelters, traffic signs, and street vendors with a total of 26,016 instances obtained from the labeling of 7,789 images.

Figure 1 shows the distribution of the number of image instances for each class in the Indonesian Traffic Obstacle Dataset, where each class contains 1,206 to 4,349 instances.

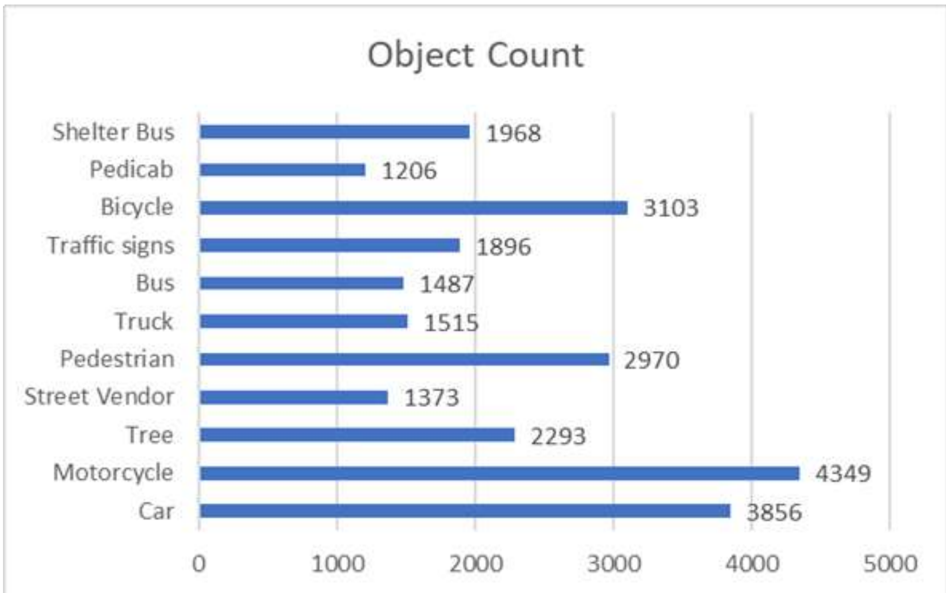

Figure 1 The Indonesian Traffic Obstacle Dataset contains eleven object classes: cars, buses, trucks, bicycles, motorcycles, pedestrians, pedicabs, trees, bus shelters, traffic signs, and street vendors.

The difference between the Indonesian Traffic Obstacle Dataset and other datasets lies in three additional object classes: pedicabs, bus shelters, and vendors, i.e. traffic obstacles in road conditions unique to Indonesia. Details regarding the object classes in existing datasets related to traffic obstacles only are shown in Table 1.

Table 1 Traffic obstacles around highways in existing datasets.

Datasets	Objects Related to Self-driving Car Obstacles										
	Pedestrians /persons	Bicycles/ cyclists	Motor-cycles / riders	Cars	Buses	Trucks	Trees	Traffic signs	Pedicabs	Bus shelters	Street vendors
ImageNet [12]	✓	✓	✓	✓	✓	✓	✓	✓	-	-	-
PASCAL-VOC [13]	✓	✓	✓	✓	✓	-	-	-	-	-	-
MS COCO [14]	✓	✓	✓	✓	✓	✓	-	✓	-	-	-
KITTI [16]	✓	✓	✓	✓	✓	✓	✓	-	-	-	-
SYNTHIA [17]	✓	✓	✓	✓	✓	-	✓	✓	-	-	-
Urban Object Detection [18]	✓	✓	✓	✓	✓	-	-	✓	-	-	-
Indonesian Traffic Obstacle Dataset (this work)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Creating the Indonesian Traffic Obstacle dataset for ITS was started by collecting images as the first step. The researchers collected images from various streets, highways, and public areas in Indonesia. The images were taken from the front and rear of a car's left and right viewpoints. Examples of images can be seen in Figure 2.

Figure 2 Object labeling of an image by an annotator.

The Indonesian Traffic Obstacle Dataset contains eleven classes: pedestrians, traffic signs, street vendors, vehicles (cars, trucks, buses, motorcycles), and other objects (bicycles, pedicabs, trees, bus shelters). As the second step, the images were pre-processed by doing alpha channel cleaning, making them the same size and cleaning them from blur and image damage using CAD tools. The images were then annotated using the RectLabel software, a powerful labeling software application for RCNN or YOLO. Every object inside the image was given a label by annotators [12]. As the third step, the researchers measured the quality of the annotations by applying a metric to evaluate the data's inter-annotator consistency [36]. The threshold metrics used were: accuracy, F1-score, and Cohen's kappa coefficient (or kappa in short). F1-score and accuracy disregarded chance agreements that are likely occur when people annotate instances.

We used kappa as a performance metric because of the expected chance agreement. Kappa is accepted as the de facto standard for the measurement of inter annotator agreement (IAA) [37] as the most well-known degree of rater agreement [38]. Cohen's kappa is defined as:

$$Kappa = \frac{P(A)-P(E)}{1-P(E)} \tag{1}$$

P(E) is the hypothetical probability of agreement by chance (with data labels randomly assigned) and P(A) is the observed relative agreement between two annotators. A kappa score of 0.81 to 1 indicates almost perfect agreement [19]. The researchers used a kappa score for each type of obstacle. As the final result, an overall kappa score of 0.853 was obtained, which is higher than the threshold. This means that the agreement between the annotators was valid and reflected almost perfect agreement. The result of the measurements can be seen in Table 2.

Table 2 Dataset evaluation using Kappa score.

Object type	**Both relevant**	**Both not relevant**	**Relevant – not relevant**	**Not relevant – relevant**	**Kappa**
Car	3777	53	11	15	0.799
Motorcycle	4259	59	23	8	0.788
Tree	2247	36	9	1	0.875
Street Vendor	1353	16	4	0	0.887
Pedestrian	2919	31	10	10	0.752
Truck	1493	19	3	0	0.925
Bus	1456	28	2	1	0.948
Traffic sign	1862	28	2	4	0.901
Bicycle	3033	46	15	9	0.789
Pedicab	1180	23	3	0	0.937
Shelter bus	1927	27	8	6	0.791
				Overall Kappa Score	**0.853**

Below is an example for calculating the kappa score (for the car class):

$$Kappa = \frac{P(A)-P(E)}{P(E)-1}$$

$$p_A = \frac{3777+53}{3856} = 0.99 \quad (2)$$

$$P(relevant) = \frac{(3777+11)}{3856} x \frac{(3777+15)}{3856} = 0.96606 \quad (3)$$

$$P(not) = \frac{(11+53)}{3856} x \frac{(15+53)}{3856} = 0,000293 \quad (4)$$

$$P(e) = P(relevant) + P(not) = 0.966606 + 0.000293 = 0.966353 \quad (5)$$

$$Kappa = \frac{P(a) - P(e)}{1 - P(e)} = \frac{0,993257 - 0,966353}{1 - 0,966353} = 0.7996034 \quad (6)$$

4 Experiments and Results

This study aimed to design a reliable advanced driver assistance system (ADAS) that can recognize objects around vehicles on roads in Indonesia to warn drivers. For this purpose, the researchers used the YOLOv4 model [28] as a state-of-the-art CNN-based one-stage detector to recognize eleven object classes: cars, buses, trucks, bicycles, motorcycles, pedestrians, pedicabs, trees, bus shelters, traffic signs, and street vendors. This study focused on the best performing model, YOLOv4, using the CSP-DarkNet53 framework.

In object or obstacle detection, high precision is not the only requirement. We need a model that can run on edge devices easily and processing input video in real-time with low-cost devices is also important. Thus, YOLOv4 was recently introduced for optimal speed (FPS) and accuracy (average precision) in object detection. It claims to have cutting-edge precision while keeping up high processing frame rates. Figure 3 shows the object detector architecture of YOLOv4.

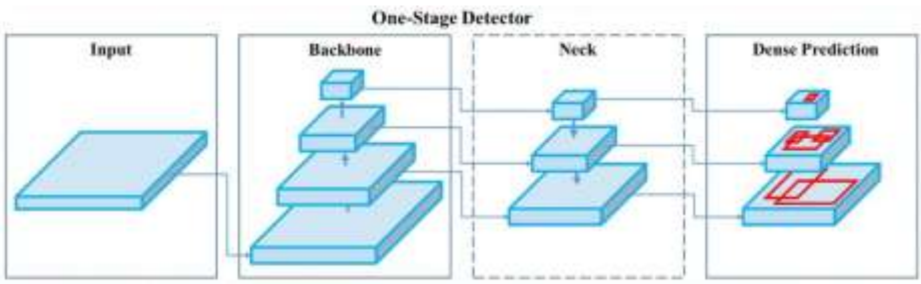

Figure 3 YOLOv4 object detector architecture, modified from [28].

This architecture contains CSP-DarkNet53 on the backbone, SPP and PAN on the neck, and YOLOv3 on the head:, which means that it performs dense prediction as in one-stage detectors. Cross-Part-Partial Connection (CSPNet) with DarkNet53, which is called the CSP-Darknet53 model, has higher precision in object detection compared to ResNet. It can partition the setting of any significant feature while maintaining the network operating speed.

The CNN assets built were tested with several parameters over 8000 iterations, 64 batches, and 16 subdivisions. In this study, 256 neurons were used in dense layers consisting of five convolutional layers, followed by max-pooling layers, and three fully-connected layers with 8-way softmax and 2000 epochs. In order to reduce overfitting on fully connected layers, the researchers used the dropout regularization method. To make the testing faster, non-saturating neurons were used with very efficient implementation of GPU convolution operations. This architecture was used based on the maximum values of precision, recall, F1-score, and mAP. The data training process was divided into three parts: 70% for training, 15% for testing, and 15% for validation. To find out the results of the YOLOv4 testing model with the Indonesian Traffic Obstacle Dataset (ITOD), the researchers used four measurement parameters, namely precision, recall, F1-score, IoU (intersection over union), and mean average precision (mAp). The results are presented in Table 3.

Table 3 Average evaluation if YOLOv4 on ITOD.

Average	YOLO v4
Precision	76%
Recall	82%
F1-score	79%
IoU (threshold = 0.5)	63.47%
mAP@0.50	81.41%

Table 3 above indicates that our datasets produced outstanding performance using deep learning for a CNN-based stage detector. The mAP50 was 81.41%, which is higher than the YOLOv4 baseline [27], whereas the MS COCO dataset achieved an AP50 of 65.7%. The distribution of AP per class is depicted in Table 4.

Table 4 shows the AP in detail for each class, indicating very robust classification for obstacle detection. The bus shelter, pedestrian, bicycle, and car classes had an AP of more than 85%, while the class with the lowest accuracy was the tree class, with an AP of 61.12%. After observing the classes, the researchers noticed that the mAP for the four classes of bus shelters, bicycles, pedestrians, and cars was better than that for the other classes because of better image quality, more varied poses or shooting angles, even object size, and better partiality factor. For the tree class, one problem encountered was that the dataset has a very large size, resulting in low accuracy. Tree images have a relatively large size and can even be exhaustive; hence the number of instances of this class has no significant positive effect on the accuracy.

Table 4 Accuracy achieved by YOLOv4 for 11 classes.

Class	AP %
Cars	85.1
Motorcycles	79.58
Trees	61.12
Street vendors	94.6
Pedestrians	87.89
Trucks	74.58
Buses	89.64
Traffic signs	76.28
Bicycles	94.78
Pedicabs	71.91
Bus shelters	93.72

Finally, to ensure that the YOLOv4 model built with the Indonesian Traffic Obstacle Dataset (ITOD) could be implemented in real-time for ADAS, this study conducted model testing with an on-road video captured in Bandar Lampung city, lasting for 39 minutes and 19 seconds. The instances of a low AP often tended to be related to hidden objects. Figure 4 shows the results of YOLOv4 model testing to detect obstacles on the road. Moreover, Figure 5 shows the false positives and false negatives from the YOLOv4 model for detection of the tree class in real-time video.

After observing the testing result, our YOLOV4 models based on CSP-DarkNet53 using the Indonesian Traffic Obstacle Datasets (ITOD) met the requirements of providing information on obstacles or objects around the vehicle for ADAS.

Figure 4 Accurate detection and recognition of all obstacles in a frame.

Figure 5 Detection results of tree class. Left: false negative; right: false positive.

5 Conclusion and Recommendation for Future Study

In this study, the researchers created the new Indonesian Traffic Obstacle Dataset (ITOD) for Intelligence Transport System (ITS), specifically for ADAS. The dataset consists of eleven classes, i.e. cars, buses, trucks, bicycles, motorcycles, pedestrians, pedicabs, trees, bus shelters, traffic signs, and street vendors. The dataset validity was measured using the kappa score with a result of 0.853, which is higher than the threshold. This study found that the dataset is valid and can be used in YOLO and PASCAL VOC format, which consist of more than one thousand objects per class.

The researchers tested a state-of-the-art CNN-based one-stage detector, namely YOLOv4, over DSP-DarknNet53 using ITOD, to determine this model's performance in detecting traffic obstacles on Indonesian roads. YOLOv4 achieved a sufficiently high mAP, estimated at 81.41; hence, this model can be utilized in real-time ADAS. Future study is recommended to enrich the dataset by adding obstacle images taken during rainy weather, the morning, evening and night time. The researchers plan to split the traffic signs dataset into separate datasets and will use the same process in this study and expand the dataset.

Acknowledgements

This research was supported by the Ministry of Research and Technology/National Research and Innovation Agency Republic of Indonesia under the funding scheme 'Collaborative Research between Universities' through LPPM Universitas Teknokrat Indonesia and LLDIKTI Region II with contract number 939/SP2H/LT/JAMAK/LL2/2020.

References

[1] Litman, T., *Autonomous Vehicle Implementation Predictions: Implications for Transport Planning*, Victoria Transport Policy Institute, 2018.

[2] Bojarski, M., Del-Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., Zhang, J. & Zhang, X., *End to End Learning for Self-driving Cars*, arXiv preprint, arXiv:1604.07316, Apr. 2016.

[3] Häne, C., Sattler, T. & Pollefeys, M., *Obstacle Detection for Self-driving Cars Using Only Monocular Cameras and Wheel Odometry*, International Conference on Intelligent Robots and Systems (IROS), pp. 5101-5108, 2015.

[4] Ramos, S., Gehrig, S., Pinggera, P., Franke, U. & Rother, C., *Detecting Unexpected Obstacles for Self-Driving Cars: Fusing Deep Learning and Geometric Modeling*, arXiv preprint, arXiv:1612.06573v1, Dec 2016.

[5] Tian, Y., Pei, K., Jana, S. & Ray, B., *DeepTest: Automated Testing of Deep-Neural-Network-driven Autonomous Cars*, IEEE/ACM 40th International Conference on Software Engineering (ICSE), pp. 303-314, 2018.

[6] Krizhevsky, A., Sutskever, I. & Geoffrey, H.E., *ImageNet Classification with Deep Convolutional Neural Networks*, Adv. Neural Inf. Process. Syst., **25**, pp. 1-9, 2012.

[7] He, K., Zhang, X., Ren, S. & Sun, J., *Deep Residual Learning for Image Recognition*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.

[8] Girshick, R., *Fast R-CNN*, Proc. IEEE Int. Conf. Comput. Vis., pp. 1440-1448, 2015.

[9] Ren, S., He, K., Girshick, R. & Sun, J., *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **39**(6), pp. 1137-1149, 2017.

[10] Liu, W., Anguelov, D., Erhan, D. Szegedy, C., Reed, S., Fu, C.Y. & Berg, A.C., *SSD: Single Shot MultiBox Detector*, arXiv preprint arXiv:1512.02325, Dec. 2015.

[11] Redmon, J., Divvala, S., Girshick, R. & Farhadi, A., *You Only Look Once: Unified, Real-time Object Detection*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.

[12] Deng, J., Dong, W., Socher, R., Li, L., Li, K. & Fei-Fei, L., *ImageNet: A Large-scale Hierarchical Image Database*, IEEE Conference on Computer Vision and Pattern Recognition, pp. 248-255, 2009.

[13] Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J. & Zisserman, A., *The Pascal Visual Object Classes (VOC) Challenge*, Int. J. Comput. Vis., **88**(2), pp. 303-338, 2009.

[14] Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. & Zitnick, C.L. *Microsoft COCO: Common Objects in Context*, Computer Vision – ECCV 2014, Fleet D., Pajdla T., Schiele B., Tuytelaars

T. (eds), Lecture Notes in Computer Science, Vol. 8693, Springer Cham, 2014.

[15] Stallkamp, J., Schlipsing, M., Salmen, J. & Igel, C., *The German Traffic Sign Recognition Benchmark: A Multi-class Classification Competition*, Proceedings of the International Joint Conference on Neural Networks, pp. 1453-1460, 2011.

[16] Geiger, A., Lenz, P., Stiller, C. & Urtasun, R., *Vision Meets Robotics: The KITTI Dataset*, Int. J. Rob. Res., **32**(11), pp. 1231-1237, 2013.

[17] Ros, G., Sellart, L., Materzynska, J., Vazquez, D. & Lopez, A.M., *The SYNTHIA Dataset: A Large Collection of Synthetic Images for Semantic Segmentation of Urban Scenes*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 3234-3243, 2016.

[18] Dominguez-Sanchez, A., Cazorla, M. & Orts-Escolano, S., *A New Dataset and Performance Evaluation of a Region-based CNN for Urban Object Detection*, Electron., **7**(11), 301, 2018.

[19] Landis, J.R. & Koch, G.G., *The Measurement of Observer Agreement for Categorical Data*, Biometrics, **33**, pp. 159-174, 1977.

[20] Khosravy, M., Gupta, N., Marina, N. & Member, S., *Perceptual Adaptation of Image Based on Chevreul – Mach Bands Visual Phenomenon*, IEEE Signal Process. Lett., **24**(5), pp. 594-598, 2017.

[21] Udacity, Inc., *An Open Source Self-Driving Car*, Udacity, https://github.com/udacity/self-driving-car (10 November 2020).

[22] Larsson, F. & Felsberg, M., *Using Fourier Descriptors and Spatial Models for Traffic Sign Recognition*, Springer: Berlin, pp. 238-249, 2011.

[23] He, K., Zhang, X, Ren, S. & Sun, J., *Spatial Pyramid Pooling in Deep Convolutional Networks*, IEEE Trans. Pattern Anal. Mach. Intell., **37**(9), pp. 1904-1916, 2015.

[24] Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B. & Belongie, S., *Feature Pyramid Networks for Object Detection*, IEEE Conf. Comput. Vis. Pattern Recognit., pp. 936-944, 2017.

[25] Ren, S., He, K., Girshick, R. & Sun, J., *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*, IEEE Trans. Pattern Anal. Mach. Intell., **39**(6), pp. 1137-1149, 2017.

[26] Redmon, J. & Farhadi, A., *YOLO9000: Better, Faster, Stronger*, 30th IEEE Conf. Comput. Vis. Pattern Recognition, pp. 6517-6525, 2017.

[27] Redmon, J. & Farhadi, A., *YOLOv3: An Incremental Improvement*, arXiv preprint, arXiv:1804.02767, 2018.

[28] Bochkovskiy, A., Wang, C.Y. & Liao, H.Y.M., *YOLOv4: Optimal Speed and Accuracy of Object Detection*, arXiv preprint, arXiv: 2004.10934, 2020.

[29] Lin, T.Y., Goyal, P., Girshick, R., He, K. & Dollar, P., *Focal Loss for Dense Object Detection*, IEEE Int. Conf. Comput. Vis., pp. 2999-3007, 2017.

[30] Strickland, M., Fainekos, G. & Ben-Amor, H., *Deep Predictive Models for Collision Risk Assessment in Autonomous Driving*, IEEE Int. Conf. Robot. Autom., pp. 4685-4692, 2018.

[31] Mandal, V., Mussah, A.R., Jin, P. & Adu-gyamfi, Y., *Sustainability Artificial Intelligence-enabled Traffic Monitoring System*, Sustainability, **12**(21), 9177, 2020.

[32] Mulyanto, A., Borman, R.I., Prasetyawan, P., Jatmiko, W., Mursanto, P. and Sinaga, A., *Indonesian Traffic Sign Recognition For Advanced Driver Assistent (ADAS) Using YOLOv4*, 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), pp. 520-524, 2020.

[33] Elyasi-Pour, R., *Simulation Based Evaluation of Advanced Driver Assistance Systems*, Thesis, Department of Science and Technology, Linköping University, Sweden, 2015.

[34] Aranjuelo, N., Unzueta, L., Arganda-Carreras, I. & Otaegui, O., *Multimodal Deep Learning for Advanced Driving Systems*, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10945 LNCS, pp. 95-105, 2018.

[35] Ball, J.E. & Tang, B., *Machine Learning and Embedded Computing in Advanced Driver Assistance Systems (ADAS)*, Electron., **8**(7), pp. 2-5, 2019.

[36] Di Eugenio, B. & Glass, M., *Squibs and Discussions: The Kappa Statistic: A Second Look*, Comput. Linguist., **30**(1), pp. 95-101, 2004.

[37] Passonneau, R.J. & Carpenter, B., *The Benefits of a Model of Annotation*, The 7th Linguistic Annotation Workshop & Interoperability with Discourse, pp. 187-195, 2013.

[38] Tang, W., Hu, J., Zhang, H., Wu, P. & He, H., *Kappa Coefficient: A Popular Measure of Rater Agreement*, Shanghai Arch. psychiatry, **27**(1), pp. 62-67, 2015.