Düzce University
Journal of Science & Technology

Research Article

Deep Learning Based Traffic Sign Recognition Using YOLO Algorithm

Gökalp ÇINARER [a,*]

[a] *Department of Computer Engineering, Faculty of Engineering, Bozok University, Yozgat, TÜRKİYE*
**Correspending author's e-mail adress: gokalp.cinarer@bozok.edu.tr*
DOI: 10.29130/dubited.1214901

ABSTRACT

Traffic sign detection has attracted a lot of attention in recent years among object recognition applications. Accurate and fast detection of traffic signs will also eliminate an important technical problem in autonomous vehicles. With the developing artificial intelligency technology, deep learning applications can distinguish objects with high perception and accurate detection. New applications are being tested in this area for the detection of traffic signs using artificial intelligence technology. In this context, this article has an important place in correctly detecting traffic signs with deep learning algorithms. In this study, three model of (You Only Look Once) YOLOv5, an up-to-date algorithm for detecting traffic signs, were used. A system that uses deep learning models to detect traffic signs is proposed. In the proposed study, real-time plate detection was also performed. When the precision, recall and mAP50 values of the models were compared, the highest results were obtained as 99.3, 95% and 98.1%, respectively. Experimental results have supported that YOLOv5 architectures are an accurate method for object detection with both image and video. It has been seen that YOLOv5 algorithms are quite successful in detecting traffic signs and average precession.

Keywords: *Deep learning, traffic sign recognition, Yolo, Artificial intelligence*

Yolo Algoritması Kullanarak Derin Öğrenme Tabanlı Trafik İşareti Tanıma

ÖZ

Trafik işaretleri tespiti nesne tanıma uygulamaları arasında son yıllarda oldukça fazla ilgi görmektedir. Trafik işaretlerinin doğru ve hızlı şekilde algılanması otonom araçlarda önemli bir teknik sorunu da ortadan kaldıracaktır. Gelişen yapay zeka teknolojiyle beraber derin öğrenme uygulamaları yüksek algılama ve doğru tespitle objeleri ayırt edebilmektedir. Yapay zeka teknolojisi kullanarak trafik levhalarının tespiti için bu alanda yeni uygulamalar test edilmektedir. Bu kapsamda bu makale trafik işaretlerini derin öğrenme algoritmalarıyla doğru tespit etmek için önemli bir yere sahiptir. Bu çalışmada trafik işaretlerinin tespiti için güncel bir algoritma olan YOLOv5'in en yeni üç modeli kullanılmıştır. Derin öğrenme algoritmalarını temel alan bir trafik işaret algılama ve tanıma sistemi önerilmiştir. Önerilen çalışmada aynı zamanda gerçek zamanlı levha tespiti de gerçekleştirilmiştir. Modellerin precision, recall ve mAP50 değerleri karşılaştırıldığında en yüksek sonuçlar sırasıyla %99.3, %95 ve %98.1 olarak elde edilmiştir. Deneysel sonuçlar YOLOv5 mimarilerinin hem görüntü hem de video ile nesne tespiti için doğru bir yöntem olduğunu desteklemektedir. YOLOv5 algoritmalarının trafik işaretlerini ve ortalama hassasiyeti (mAP) algılamada oldukça başarılı olduğu görülmüştür.

Anahtar Kelimeler: *Derin öğrenme, Trafik işareti tanıma, Yolo, Yapay zeka*

Geliş: 05/12/2022, Düzeltme: 25/12/2022, Kabul: 10/05/2023

INTRODUCTION

Recently, the increasing use of artificial intelligence has included intelligent systems in many aspects of our lives. Service robots, smart homes and autonomous vehicles are autonomous systems that are increasing in number every day. The increase in the number of unmanned vehicles that recognize traffic signs and autonomous vehicles that follow lanes in recent years can be shown as the best example of this. In addition, the determination of traffic signs has an important place in the decision-making process of the vehicle[1]. For this reason, it is a very important problem for vehicles to detect traffic signs correctly and to enable vehicles to activate automatic software systems accordingly.

In the report given by the International Association for Safe Road Travel, while 1.1 million people lost their lives as a result of traffic accidents in 1990, this number increased to 1.4 million in 2013 [2]. For similar reasons, many large companies have increased their investments in autonomous vehicles and are working on the detection of objects in traffic. Although there is a standard size and shape of road signs in the world, many problems such as some environmental factors, inadequacy of lights and weather conditions sometimes force drivers to perceive the signs. In adverse conditions, traffic accidents can be prevented by making artificial intelligence-based traffic sign detection and ensuring that both drivers and autonomous vehicles carry out this process more healthily. Deep learning algorithms are used to solve the problem that arise to detection traffic signs. Researchers use object detection algorithms such as YOLO [3], SSD [4] and CNN [5] to detect traffic signs. There may be categories such as speed limit, turn prohibitions and warnings in the recognition of traffic signs [6]. Creating Artificial Intelligence models that consider all these categories will greatly ease the detection of signs and the implementation of the system.

In this study, three different YOLOv5 models will be used to detect traffic signs. With the comparisons between them, it will be possible to get the best result in YOLOv5.

The main contributions of the study are:
• Training the dataset with three different YOLOv5 models
• Analyzing the results obtained
• Determining at which values three different models are better
• Detection of real-time traffic signs with 3 different models
Each step of the study is designed as follows. Part 2 is divided into literature review, part 3 materials and methods, part 4 results, part 5 discussion and part 6 conclusion.

II. RELATED WORK

Sign detection in traffic is a new research topic that has been studied recently with the increase of autonomous vehicle number. In the literature, there are studies conducted in this field with different deep learning architectures and traditional methods. While in traditional methods, it is possible to classify with machine learning algorithms by obtaining the features of the image, classification studies obtained from instant data with deep learning are more limited. In addition, it is seen that traffic signs belonging to different countries are classified. Shao et al. [13] proposed the Faster R-CNN model for traffic sign detecetion. They increased the recognition speed of the network by using Gabor wavelets and reached an accuracy of 99.01% with their proposed method Zhanwen Liu et al. [14] created a training set by taking 9176 data from the Chinese traffic sign dataset TT100K. They divided the 221 different signals in the TT100K into 25 classes and found the accuracy rate as 89.7%. Lingcai Zeng et al. [15] have presented a model that proposes the addition of a contentious closure network (AON) to the CNN model in training and detection capacity of the traffic network. Zhang et al. [16] have changed the convolutional layers for YOLOV2 model and proposed an advanced single-stage traffic sign detector. They used the Chinese traffic sign dataset to make all these processes work better on Chinese roads. Belghaouti et al. studied a LeNet model for automatic traffic sign detection. In the study, they used the German traffic data set.

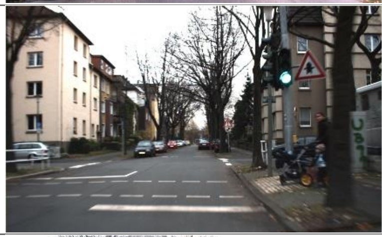
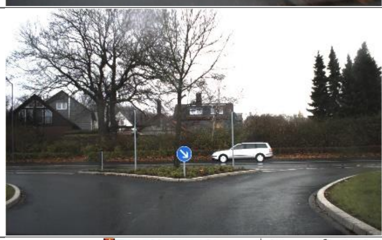
They obtained an accuracy of 99% with the method they proposed [17]. Tabernik et al. used a basic model and suggested some settings in the model for better traffic sign detection. This method reached 97.5% accuracy with Mask R-CNN [18].

III. MATERIAL AND METHOD

A. DATASET

A open source kaggle dataset consisting of 741 traffic signs was used for training the model [19]. Initial data is German Traffic Sign Detection Benchmark (GTDRB). These photos are categorized in 4 classes. These classes are set to prohibitory, danger, mandatory and other. The prohibition category consists of images that are prohibited from speeding, overtaking, traffic, and truck entry. Danger category includes priority at next intersection, left and right bend, bend, rough road, slippery road, road narrowing, construction, traffic signal, pedestrian crossing, school crossing, bicycle crossing, snow, animals traffic signsIn the mandatory category, there are images of right, left or straight going and different directions taken in the roundabout area. Finally, in the other category, the restriction ends, there are basic traffic signs such as priority pedestrian giving way, giving way, stopping and traffic-free area, no-traffic zone. 592 of 741 plates were set as trains and 149 as test. Some of the plates inside the classrooms are shown in Table 1.

Table 1. *Dataset image categories*

Id	Class Name	Image
0	prohibitory	
1	danger	
2	mandatory	
3	other	

B. DEEP LEARNİNG BASED MODEL

CNN, which is the most popular of deep learning applications, is the basic architecture used in different areas such as image processing, object detection, classification. This architecture also includes basic algorithms such as YOLO, one of the single-stage object detections in deep learning object detection. Its performance, especially in fast and single-stage detection, enables it to be used in traffic sign detection as well. You Only Look Once (YOLO) is an object recognition algorithms, has been widely used since its first day due to the simplicity of the network structure and high detection speed. The difference of YOLO from the detection algorithms developed in previous times is that it handles object detection as a single regression [20]. By optimizing YOLO, the loss function was reduced, as a result of which the detection speed of the object increased, but it started to give lower accuracy than other object recognition models [21].

The Yolo algorithm has developed continuously since its release and different versions have been published. In this context, the YOLOV5 model was introduced. The number of different studies with YOLOv5 is limited. With YOLO models, applications such as pedestrian detection [22], apple detection [23], license plate recognition [24], diseased tomato detection [25] and mask detection [26] have been made. In this study, the s, m and l models of YOLOv5 were compared.

In the basic structure, the backbone, neck and output are the parts that make up the YOLOv5 architecture [27]. The input images pass through the focus images in the backbone. Then, feature maps are created by extracting the features of the input images. YOLOv5 uses the CSPNet structure to extract features [28]. The advantage of the CSP structure is that iterative gradient information can be reduced in optimization. The backbone takes the base layer's replicated feature map and uses a dense block to pass it on to the next layer. The CBL structure in the backbone represents the combination of Conv2D, BatchNormal and LeakyRelu [29].

In the proposed model, the number of epochs of 200 is determined to show all the training data to the network. This value is limited to 200 in order to prevent excessive adaptation and memorization. The batch size value is set to 16. While the neural network was being weighted, the system learning step was started by using 16 images in each refresh. The learning rate was determined as 0.01.

All images are fixed at 608x608x3 size. The img values in which the training was carried out were also used in the test data. Hyperparameters were applied as standard for all models, thus making the comparative analysis more objective. In Table 2, parameter numbers for s, m and l models in YOLOV5 are given.

Table 2. *Parameter numbers of models*

Model Name	Parameter (million)
YOLOv5s	7.2
YOLOv5m	21.2
YOLOv5l	46.5

YOLOv5 uses (spatial pyramid pooling) SPP to provide better detection of objects with different scales, (Path aggregation network) PANET for feature aggregation, and neck section includes two network (Feature Pyramid Network) FPN and the other one is (Path Aggregation Network) PAN [30]. With this combination, the features in a small area are better extracted, increasing the detection sensitivity. Figure 1 shows the YOLOv5 architecture.

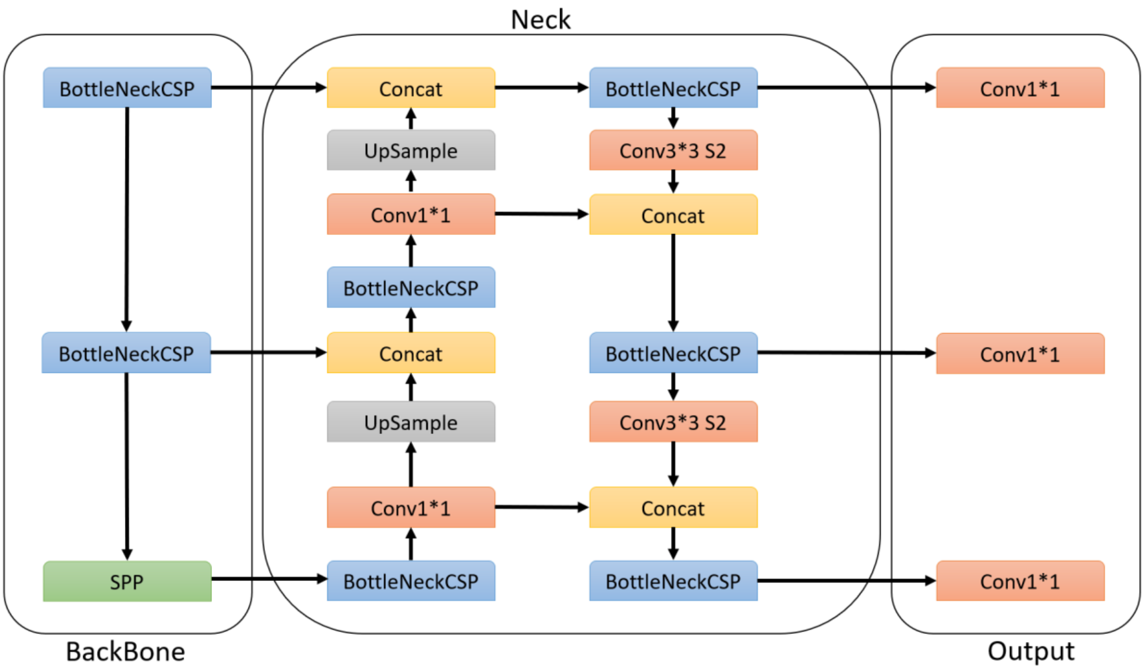


Figure 1. *YOLOv5 architecture*

C. PERFORMANCE METRICS

There ara many criteria to analyse the correct performance of an (Artificial intelligence) AI architecture model. Examples of these criteria are precision, recall, (Intersection over Union) IoU. In the formulas, TP true positive, FP false positive c define the number of categories in the system, n reference threshold number and k threshold value. The P(k) precision value is the R(k) recall value. Map value gives the average of multiple AP values. The map value takes a value in the range of 0-1, and the higher the value it takes, the better the result. Ground Truth is the actual location in the image where the object is tagged. The actual location is bounded by boxes. Intersection over Union (IoU) is one of the performance measure which finds the difference wit ground truth of the object and bound box predicted by the model. In object detection algorithms, many bounding boxes are drawn for the object and each bounding box has a confidence value. Bounding boxes that do not exceed the threshold value are removed. When the IoU value exceeds the threshold value, it is detected as object [31].

Precision: It is the evaluation of how close or scattered the measurements are.

$$Precision = \frac{TP}{\mathrm{TP} + \mathrm{FP}} \tag{1}$$

Recall: Indicates the true positive rate. It is used for probability of correctly detecting objects.

$$Recall = \frac{TP}{\mathrm{TP} + \mathrm{FN}} \tag{2}$$

mAP: Average of all AP scores

$$mAP = \frac{1}{c}\sum_{k=i}^{N} P(k) \triangleq R(k) \tag{3}$$

IV. RESULTS

The software tests of the system were created using the Python 3.7 programming language. Python libraries were used to evaluate and visualize the results obtained in the study. The performance of each model was run on google colab. Colab is a cloud-based, easy-to-use programming environment. All parameters are set the same in all models. For this reason, the models were compared under the same values.
The test phase of the study was also carried out with real-time mobile camera images of the traffic signs on the E-88 highway in the city center of Yozgat. Algorithm application results obtained from sample image sections of different signs on the route are shown in Figure 3.

In evaluating the performance of the network, studies were carried out by applying different parameter values. Each parameter value is fixed in all models for accurate determination of the performances of all models. Precision, recall and mAP (mean average precision) results for YOLOv5s, YOLOv5m and YOLOv5l models were compared. For the purpose of measuring the recognition quality of the models, the mAP metric, which references the 0 to 1 value range and determines the average precision of the results obtained in this range, was used. In training, precision, recall and map@0.5 values of each model were determined. The results of each model according to the classes are shown in detail in Tables 3, 4 and 5. When Table 3 is examined, it is seen that the Yolo V5s model reaches the highest map@0.5 value with 99.4% in speed limit images. When the precision and recall values were examined, the best results were obtained in the vield class with 100% precision and 97% recall values. When all three models are examined comparatively, it is seen that the "vield" class has the highest precison value. In addition, Map50 values in four classes in speed limit, vield, mandatory and other categories are over 92%, which is an indication of how excellent the performance of YOLO models is. In addition, the speed limit achieved a higher value than the others, with a recall value of 99.1%.

Table 3. *Yolov5s performance results*

Model	Class	Precison	Recall	mAP50
YOLOv5s	Prohibitory	0.969	0.984	0.994
	Danger	1	0.97	0.977
	Mandatory	0.975	0.881	0.938
	Other	0.977	0.885	0.974

Table 4. *Yolov5m performance results*

Model	Class	Precison	Recall	mAP50
YOLOv5m	Prohibitory	0.972	0.991	0.995
	Danger	0.995	0.964	0.985
	Mandatory	0.982	0.935	0.964
	Other	0.999	0.909	0.982

Table 5. *Yolov5l performance results*

Model	Class	Precison	Recall	mAP50
YOLOv5l	Prohibitory	0.986	0.991	0.995
	Danger	1	0.966	0.995
	Mandatory	0.986	0.833	0.924
	Other	1	0.954	0.96

Finally, the average results of the data set in all classes with the YOLOv5 models are given in Table 6.

Table 6. *Average performance results*

Model Adı	Precision	Recall	mAP50
YOLOv5s	0.98	0.93	0.971
YOLOv5m	0.987	0.95	**0.981**
YOLOv5l	0.993	0.936	0.968

The results show that the yolov5m model has the highest training quality, with the mAP metric of 98.1% at the end of 200 epochs, when all four different categories are considered. In addition, the YoloV5l model reached 99.3% precision in the average results. According to the Recall values, the YoloV5m model gave the highest result with 95%. Precision-sensitivity confidence curve plots of all models are shown in Figures 2, 3 and 4.

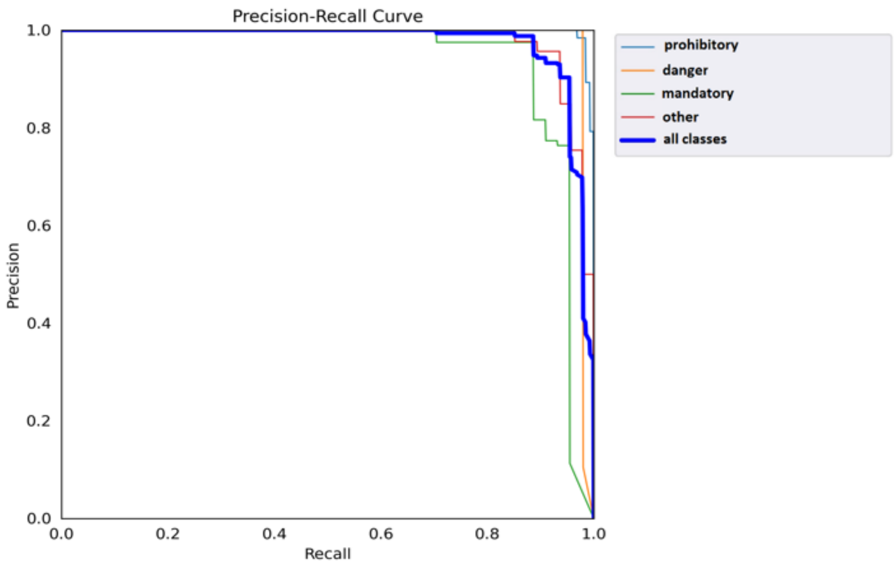


Figure 2. *YOLOv5s Confidence Curve Plots*

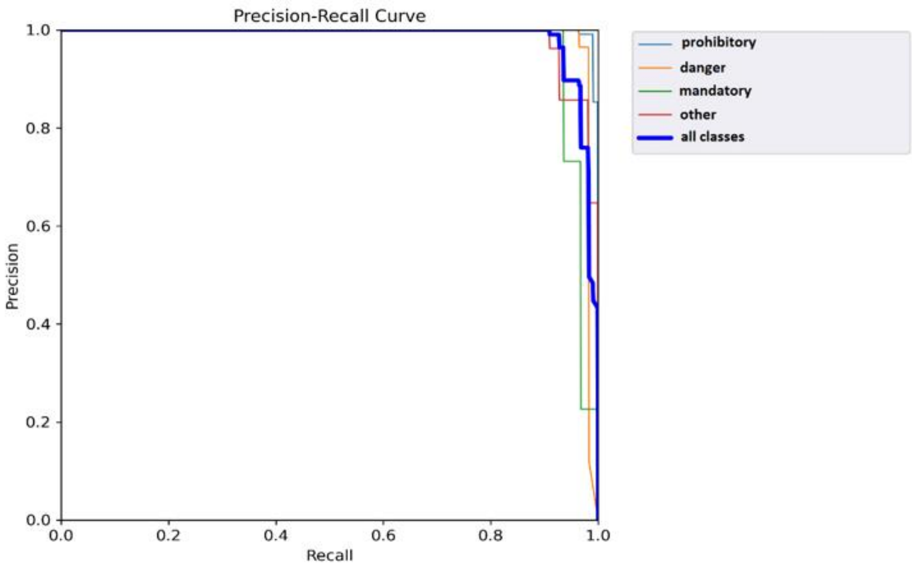


Figure 3. *YOLOv5m Confidence Curve Plots*

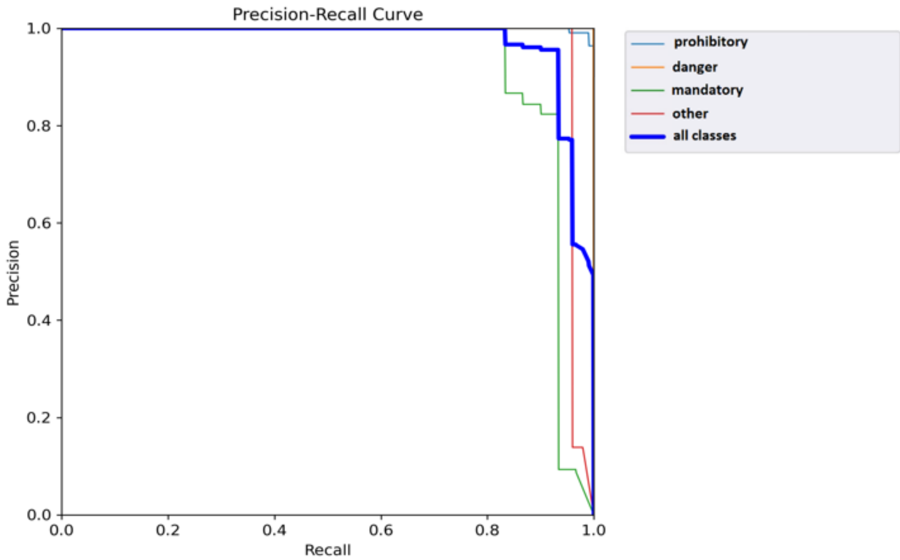


Figure 4. *YOLOv5l Confidence Curve Plots*

In Figure 5, the result sections of the traffic signs, which were determined from the real video image created for testing purposes, obtained in 3 models are given. The performances of the models on the real-time video created for traffic signs were examined comparatively. When Figure 5 is examined, the traffic sign in sample a was obtained by the yolov5 algorithm with the highest rate of 96%. When sample b was examined, it was seen that the yolov5m and yolov5l models reached 97% accuracy. Finally, in sample c, it was seen that the yolov5l model reached the highest accuracy with 97% in detecting the result of the traffic sign. Considering the average results for all three traffic signs, it was seen that the yolov5l model was the most successful algorithm with 96.3%. It has been observed that all three models, including different traffic sign images, detect them with very high success. Yolov5L model has achieved a very high accuracy in detecting these traffic signs.

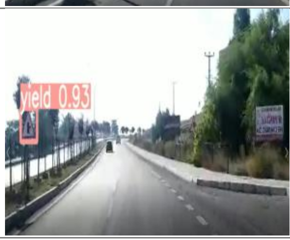
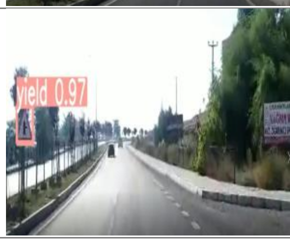
	Models		
Samples	**YOLOV5s**	**YOLOV5m**	**YOLOV5l**
(a)			
(b)			
(c)			

Figure 5. *Traffic Sign Recognition Results YOLOv5 Models*

V. DISCUSSION

The comparison of the Yolo models applied in the study with the literature and model precision results is shown in Table 7. Although the data sets and the number of images used in the studies differ, the precison values obtained by the algorithms were compared with the Yolo models. Considering the number of original images used and the results obtained, it can be said that the study showed a remarkable success in detecting traffic signs with different YOLO models.

The reason for choosing the YOLOv5 algorithm in mark detection is that it can detect objects hig accuracy and quick. In the other side YOLOv5 preferable for traffic sign detection from real-time videos. In this study, the s, m and l models of YOLOv5 were compared and the best model among them was determined. Studies in the literature are only for a certain model. Obtaining higher values than the results obtained in the literature also supports the success of the applied model.

As a result of the study, YOLOv5s reached 0.98 precision, YOLOv5m 98.7 and YOLOv5l 99.3 precision. When the results obtained are compared with the literature, it is seen that the study provides high success.

***Table 7**. Literature review*

#	Ref.	Author(s)	Dataset	Model	Results
1	[7]	Yanzhao Zhu, Wei Qi Yan	Custom Dataset	YOLOV5	0.97
2	[8]	Haifeng Wan et al.	GTSDB	TSYOLO	0.83
3	[9]	Eric Hsueh-Chan Lu et al.	Custom Dataset	YOLOV5	0.88
4	[10]	Li Yi et al.	TT100K	YOLOV4	0.92
5	[11]	Linfeng Jiang et al.	TT100K	YOLOV5	0.87
6	[12]	Ammar Aggar et al.	IQTSDB	YOLOV5	0.90
7		Our model	Traffic Signs	YOLOV5S	0.98
8		Our model	Traffic Signs	YOLOV5M	0.98
9		Our model	Traffic Signs	YOLOV5L	0.99

VI. CONCLUSION

Automatic detection of traffic signs has become very important in the world. In this study, traffic sign detection application was carried out with YOLO object detection models. Labeled traffic signs in the dataset were tested using Turkish traffic signs with three different YOLO models. In the study, the highest mAP50 value with 98.1% was reached with the YOLOv5m model. The mAP50 values of all models used in the study are over 96%, which is very promising in automatic detection of traffic signs.

It is seen that the models in the study can help experts in the use of unmanned vehicles. It is considered that the high accuracy obtained in real-time image detection will be applied at longer distances and different traffic signs in future studies.

In line with the successful results obtained in the study, it has been seen that the use of Artificial Intelligence technology and deep learning algorithms in traffic sign classification will become widespread in the coming years with the developing technology.

VI. REFERENCES

[1] R. Timofte, K. Zimmermann, and L. Van Gool, "Multi-view traffic sign detection, recognition, and 3D localisation," Machine vision and applications, vol. *25* no. 3, pp. 633-647, 2014.

[2] P. S. Zaki, M. M. William, B. K. Soliman, K. G. Alexsan, K. Khalil, and M. El-Moursy, "Traffic signs detection and recognition system using deep learning," arXiv Prepr. arXiv2003.03256, 2020.

[3] C. Dewi, R.C. Chen, Y.T. Liu, X. Jiang, and K. D. Hartomo, "Yolov4 for advanced traffic sign recognition with synthetic training data generated by various GAN," IEEE Access, vol. 9, pp. 97228-97242, 2021.

[4] S. You, Q. Bi, Y. Ji, S. Liu, Y. Feng, and F. Wu, "Traffic sign detection method based on improved SSD," *Information*, vol. 11, no. 10, pp. 475, 2020.

[5] A. Shustanov, and P. Yakimov, "CNN design for real-time traffic sign recognition," *Procedia Engineering*, vol. 201, pp. 718-725, 2017.

[6] Z. Liu, Y. Musha, and H. Wu, "Detection of traffic sign based on improved Yolov4," In 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP), 2022, IEEE, pp. 444-448.

[7] Y. Zhu, and . Q. W. Yan, "Traffic sign recognition based on deep learning," Multimedia Tools and Applications, vol. 81, no. 13, pp. 17779-17791, 2022.

[8] H. Wan, L. Gao, M. Su, Q. You, H. Qu, and Q. Sun, "A novel neural network model for traffic sign detection and recognition under extreme conditions," *Journal of Sensors*, 2021.

[9] E. H. C. Lu, M. Gozdzikiewicz, K. H. Chang, and J. M. Ciou, "A hierarchical approach for traffic sign recognition based on shape detection and image classification," *Sensors*, vol. 22, no 13, pp. 4768, 2022.

[10] L. Yi, L. Jinguo, Z. Yongjie, and M. Ping, "Detection of self-explosive insulators in aerial images based on improved Yolov4," In Journal of Physics: Conference Series vol. 2320, no. 1, pp. 012025, IOP Publishing, 2022.

[11] L. Jiang, H. Liu, H. Zhu, and G. Zhang, "Improved Yolov5 with balanced feature pyramid and attention module for traffic sign detection," In MATEC Web of Conferences vol. 355, EDP Sciences, 2022.

[12] A. Aggar, and M. Zaiter, "Iraqi traffic signs detection based on Yolov5," In 2021 International Conference on Advanced Computer Applications, 2021, IEEE, pp. 5-9.

[13] F. Shao, X. Wang, F. Meng, J. Zhu, D. Wang, and J. Dai, "Improved faster R-CNN traffic sign detection based on a second region of interest and highly possible regions proposal network," *Sensors*, vol. 19, no. 10, pp. 2288, 2019.

[14] Z. Liu, M. Qi, C. Shen, Y. Fang, X. Zhao, "Cascade saccade machine learning network with hierarchical classes for traffic sign detection," *Sustainable Cities and Society*, 67, 102700, 2021.

[15] L. Zeng, B. Sun, D. Zhu, "Underwater target detection based on Faster R-CNN and adversarial occlusion network," Engineering Applications of Artificial Intelligence, vol. 100, 104190, 2021.

[16] J. Zhang, M. Huang, X. Jin , and X. Li , "A real-time Chinese traffic sign detection algorithm based on modified Yolov2," *Algorithms*, vol. 10, no. 4, pp. 127, 2017.

[17] O. Belghaouti, W. Handouzi, and M. Tabaa, "Improved traffic sign recognition using deep ConvNet architecture," *Procedia Computer Science*, vol. 177, pp. 468–473, 2020.

[18] D. Tabernik, and D. Skočaj, “Deep learning for large-scale traffic-sign detection and recognition,” IEEE transactions on intelligent transportation systems, vol. 21, no. 4, pp. 1427-1440, 2019.

[19] V. Sichkar. (2020). *Traffic Signs Dataset in Yolo Format* [Online]. Available: https://www.kaggle.com/datasets/valentynsichkar/traffic-signs-dataset-in-yolo-format

[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: unified, real-time object detection,” In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779-788.

[21] M. Sozzi, A. Kayad, D. Tomasi, L. Lovat, F. Marinello, and L. Sartori, “Assessment of grapevine yield and quality using a canopy spectral index in white grape variety,” In Precision agriculture 19, pp. 111-129, Wageningen Academic Publishers, 2019.

[22] W. Lan, J. Dang, Y. Wang, and S. Wang, “Pedestrian detection based on Yolo network model,” In 2018 IEEE international conference on mechatronics and automation (ICMA), 2018, IEEE, pp. 1547-1551.

[23] Y. Tian, G. Yang, Z. Wang, H. Wang, E. Li, and Z. Liang, “Apple detection during different growth stages in orchards using the improved Yolo-v3 model,” *Computers and Electronics in Agriculture*, vol. 157, pp. 417-426, 2019.

[24] R. N. Babu, V. Sowmya, and K. P. Soman, “Indian car number plate recognition using deep learning,” 2nd international conference on intelligent computing, instrumentation and control technologies, 2019, IEEE, pp. 1269-1272.

[25] J. Liu, and X. Wang, “Tomato diseases and pests detection based on improved Yolov3 convolutional neural network,” *Frontiers in Plant Science*, vol. 11, pp. 898, 2020.

[26] J. Yu, and W. Zhang, “Face mask wearing detection algorithm based on improved Yolo-v4,” *Sensors*, vol. 21, no. 9, pp. 3263, 2021.

[27] S. Tan, G. Lu, Z. Jiang, and L. Huang, “Improved Yolov5 network model and application in safety helmet detection,” 2021 IEEE International Conference on Intelligence and Safety for Robotics, 2021, IEEE, pp. 330-333.

[28] J. Wan, B. Chen, and Y. Yu, “Polyp detection from colorectum images by using attentive Yolov5,” *Diagnostics*, vol. 11, no. 12, pp. 2264, 2021.

[29] F. Zhou, H. Zhao, and Z. Nie, “Safety helmet detection based on Yolov5,” In 2021 IEEE International Conference on Power Electronics, Computer Applications, 2021, IEEE, pp. 6-11.

[30] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “Path aggregation network for instance segmentation,” Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 8759-8768.

[31] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, S. Savarese, “Generalized intersection over union: a metric and a loss for bounding box regression,” Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 658-666.