Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi

International Journal of Sustainable Engineering and Technology

ISSN: 2618-6055 / Cilt: 8, Sayı: 1, (2024), 26 – 43

Araştırma Makalesi – Research Article

DETECTION OF LPG VEHICLES IN RISKY AREAS WITH COMPUTER VISION TECHNIQUES

Öznur SUÇEKEN[1], *Gül Fatma TÜRKER[1]

[1]Süleyman Demirel University, Faculty of Engineering, Department of Computer Engineering, Isparta

(Geliş/Received: 01.06.2024, Kabul/Accepted: 12.06.2024, Yayınlanma/Published: 26.06.2024)

ABSTRACT

Vehicle use is becoming more widespread day by day due to the world population growth. Within the scope of intelligent transportation systems, the information technologies sector and the transportation sector work in an integrated manner to solve the problems caused by the increasing number of vehicles. Data obtained with sensors and cameras are analyzed with artificial intelligence-based information technologies and used in autonomous vehicles, security, traffic management, navigation and passenger information systems. Computer vision enables machines to extract meaningful patterns and relationships from images by combining image processing and deep learning technologies. Computer vision techniques are applied in many fields such as tourism, health, industry, defense, transportation, service, e-commerce, etc. The applications developed provide solutions to various challenges in the transportation sector. For vehicles using Liquified Petroleum Gas (LPG) fuel, the gases in LPG tanks are flammable and pose a potential explosion hazard, especially in certain areas in cities. Entry of LPG vehicles is prohibited in institutions and organizations such as hospitals, shopping malls, hotels that have indoor parking services. The control method of the ban is carried out by assigning a personnel and checking the vehicle trunks. In this study, LPG fueled vehicles were automatically detected using computer vision techniques. Vehicle image data captured by mobile cameras in different provinces in Turkey were trained and compared with four different deep learning models. As a result of training and performance tests on the models, the YOLOv8 model was more effective than the other models with an accuracy of 0.994 mAP and a speed of 11.6 ms. It has been shown to be a stable model in terms of real-time monitoring in real life. It is envisaged that the developed system can contribute to the applications of computer vision techniques as well as benefit the national economy, public life safety and environmental protection.

Keywords: Computer vision, Deep learning, Image processing, LPG, Vehicle.

LPG' Lİ ARAÇLARIN RİSKLİ ALANLARDA BİLGİSAYARLI GÖRÜ TEKNİKLERİ İLE TESPİTİ

ÖZ

Dünya nüfus artışına bağlı olarak araç kullanımı gün geçtikçe yaygınlaşmaktadır. Akıllı ulaşım sistemleri kapsamında artan araç sayısının neden olduğu sorunları çözmek için bilişim sektörü ile ulaşım sektörü entegre bir şekilde çalışmaktadır. Sensörler ve kameralarla elde edilen veriler, yapay zeka tabanlı bilişim teknolojileriyle analiz edilerek otonom araçlar, güvenlik, trafik yönetimi, navigasyon ve yolcu bilgilendirme sistemlerinde kullanılmaktadır. Bilgisayarlı görü, görüntü işleme ile derin öğrenme teknolojilerinin birlikte kullanılması sonucu makinelerin, görüntülerden anlamlı örüntüler ve ilişkiler çıkarmasını sağlamaktadır. Bilgisayarlı görü teknikleri turizm, sağlık, sanayi, savunma, ulaşım, hizmet, e-ticaret vb. birçok alanda uygulanmaktadır. Geliştirilen

***Sorumlu Yazar/Corresponding Author:** gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

uygulamalar ulaşım sektöründe çeşitli zorluklara çözüm üretmektedir. Liquified Petroleum Gas (LPG) yakıtı kullanan araçlar için, LPG tanklarındaki gazların yanıcı olması ve patlama ihtimali yaratması nedeniyle, özellikle şehirlerdeki belirli alanlarda tehlike oluşturmaktadır. Kapalı otopark hizmeti bulunduran hastaneler, alışveriş merkezleri, oteller gibi kurum ve kuruluşlarda LPG' li araçların girişi yasaklanmıştır. Yasağın denetim yöntemi ise bir personelin görevlendirilmesi ve araç bagajlarının kontrol edilmesiyle gerçekleştirilmektedir. Bu çalışmada LPG yakıtıyla çalışan araçların bilgisayarlı görü teknikleri kullanılarak otomatik bir şekilde tespiti yapılmıştır. Türkiye'de farklı illerde mobil kameralar aracılığıyla çekilen araç görüntü verileri dört farklı derin öğrenme modeli ile eğitilerek karşılaştırılmıştır. Modeller üzerinde eğitim ve performans testleri sonucu YOLOv8 modelinde, 0.994 mAP doğruluk ve 11.6 ms hız değerleri ile diğer modellerden daha etkili sonuç elde edilmiştir. Güncel hayatta gerçek zamanlı izleme açısından kararlı bir model olduğu gösterilmiştir. Geliştirilen sistemin, bilgisayarlı görü tekniği uygulamalarına katkıda bulunmasının yanı sıra ulusal ekonomiye, toplum can güvenliğine ve çevrenin korunmasına fayda sağlayabileceği öngörülmektedir.

Anahtar kelimeler: Araç, Bilgisayarlı görü, Derin öğrenme, Görüntü işleme, LPG.

1. Introduction

The concept of Intelligent Transportation Systems (ITS) has entered our lives as a result of combining vehicles and transportation infrastructure with information technologies. ITS are advanced technological applications that include wireless and wired communication-based computing, control algorithms, electronics and other technologies. It aims to move the transportation sector forward with studies in many areas such as providing multi-directional data exchange between human-vehicle-infrastructure-center, regulating traffic flow, navigation services, optimum use of transportation facilities and ensuring travel safety [1].

Artificial Intelligence (AI) is the ability of computers to learn and make decisions by mimicking the working principles of the human brain. Computer vision is the process by which computers extract meaning from different types of visual data and is part of AI systems. This technique includes the concepts of image processing and deep learning. Image processing is when a computer evaluates visual data and makes sense of it. Deep learning is the computer's ability to learn and make decisions on its own by imitating the human brain. By using these two technologies together, computer vision technique was created. Computer vision applications especially in transportation systems have an increasing importance. Many applications such as monitoring traffic density, detecting traffic jams, license plate detection, traffic violation detection, occupancy rate data in parking lots, abnormal situation detection on highways, hazardous material detection, etc. help to solve transportation problems of cities by increasing the effectiveness of intelligent transportation systems [2].

Automatic license plate recognition systems, which provide traffic management and instant traffic monitoring and contribute to the collection of important statistics on road conditions, were among the first computer vision applications in the field of ITS. These systems offer various solutions such as automatic license plate recognition, tracking of highway or parking areas, traffic congestion prediction. In their study, Shashirangana et al. found that in the context of automatic license plate recognition system, single-stage object detection algorithms show high performance on various datasets, while multi-stage object detection algorithms have lower accuracy and computational efficiency [2]. In recent studies, it is accepted that Convolutional Neural Network (CNN) based artificial intelligence architectures are preferred in automatic license plate recognition systems and recognition performance is improved by using methods such as YOLOv3, Faster R-CNN [3].

Traffic sign recognition, used in autonomous vehicles and advanced driver assistance systems, is a type of computer vision application that aims to identify traffic signs in an image from a limited number of options. Sermanet and LeCun were among the first in the literature to use CNN-based deep learning for traffic sign recognition [3,4]. Thanks to computer vision models, automatic feature extraction can be performed and the accuracy rate of traffic sign recognition can be increased up to 99% [5]. Another area of study in the field of vehicle detection and classification is the extraction of class description information such as model and color of vehicles. These applications, which are especially vital for security systems, can identify vehicles with the desired characteristics among the big data obtained from hundreds of traffic camera images [6-8]. Jahongir Azimjonov processed online traffic images obtained from intersections and highways with image processing techniques to compute instantaneous data about

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

 ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

the traffic in selected areas of the city. These data include classification of vehicles, numerical values (direction and speed) of each vehicle. At the end of this study based on CNN-based YOLOv3 technique, 90% success rate was observed [9]. Abdullah Sökülmez obtained an accuracy rate of over 90% on the detection and location of vehicles with YOLOv3 on a 3-minute highway scene using security cameras [10]. Systems that learn from normal data are required to detect unusual events. In anomaly detection, the normal data distribution is learned and deviations from this distribution are considered as anomalies [11,12]. Nayak et al. used computer vision techniques in video anomaly detection to detect anomalous activities such as fights, traffic violations, and unusual objects [13]. Yarens J. Cruz successfully detected the misalignment of the parts to be welded with a 97.7% success rate by using CNN architecture and image processing techniques in LPG welding inspection processes. They showed that this technology can be used in the industrial industry due to its low cost and effective result [14]. When the studies in the literature are examined, it is seen that effective solutions have been developed in transportation systems using computer vision techniques. Studies on LPG have generally been carried out in the fields of fault or leak detection in the industrial sector. No study on LPG tank detection has been encountered.

The access of LPG tanks in vehicles to places that offer parking services or where there is a risk of explosion, such as hospitals, shopping malls, airports, ferries, tunnels, etc. is controlled by security policies. In case of any leakage in LPG tanks, the gases in LPG accumulate on the ground due to the fact that the gases in LPG are heavier than air and pose a serious risk of explosion. For this reason, security policies have a very important place. Inspection systems are usually implemented by an officer opening the vehicle trunks and checking the presence of the LPG tank. There is a great risk of danger if the person in charge has a margin of error such as overlooking or neglecting the controls. In addition, the process of stopping the vehicle, opening the trunk and checking the vehicle during the control process creates a waste of time for the drivers and the appropriate employment of the person in charge creates a cost for these organizations.

The aim of this study is to develop a new inspection system in areas where the entry of LPG vehicles poses a risk. In this system developed for the automatic detection of LPG vehicles, images obtained from cameras were processed using computer vision techniques and an automatic access permission system was created by detecting LPG entries in vehicles. Within the scope of the study, a unique dataset was created and the data was trained with deep learning models after preprocessing and augmentation operations. In the rest of the study, the values of the most successful model are presented as a result of the comparison.

2. Material and Method

2.1. Liquefied petroleum gas

Liquified Petroleum Gas (LPG) is a fuel consisting of a mixture of propane and butane gases. It is obtained during the processing of natural gas or oil refining. LPG-powered vehicles are widely used in Turkey, South Korea, Poland, Italy and Russia [15]. According to TÜİK (Turkey Statistical Institute) data, the number of LPG vehicles in Turkey was 5.11 million as of March 2024. Figure 1 shows the distribution of automobiles according to fuel types [16].

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

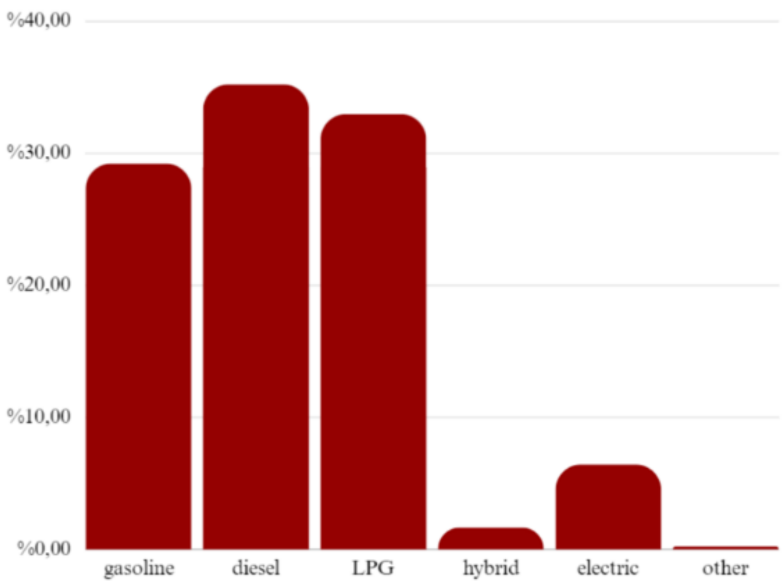

Figure 1. Number of land vehicles by fuel types (Yakıt türlerine göre kara araçları)

Although LPG fuel is more environmentally friendly and economical than gasoline and diesel fuels, it is subject to certain restrictions due to safety risks. LPG becomes an explosive gas when mixed with air at a certain rate. Especially in closed parking lots, the tanks in LPG vehicles are banned because they may pose a risk of explosion due to the possibility of leakage. In this study, computer vision techniques are used to effectively and quickly detect LPG vehicles in areas where they pose a risk.

2.2. Creation of dataset

The dataset is constructed from images taken from the rear bumper and side door alignments of vehicles in the cities of Aydın, Isparta and Muğla in Turkey, with the permission of the vehicle owners, so that the LPG inlets are visible. In parking garages, there are generally fewer restrictions on the entrances of vehicles with fabricated LPG (designed to run on LPG from the original production). These vehicles are equipped with safer tanks and sealing systems, with the LPG inlet usually under the gas cap. For this reason, no images of fabricated LPG vehicles were taken during the creation of the dataset.

2.3. Preprocessing and augmentation of dataset

The dataset, which contains a total of 939 LPG vehicle images, was labeled and the LPG input of each image was placed in a bounding box and assigned to the class named “lpg-var”. Preprocessing was performed on the images to make the images in the dataset more distinct.

Data augmentation improves the learning ability of the model by increasing the diversity of the training data. Models trained with data from different angles, positions or lighting conditions can better adapt to various real-world scenarios [16]. When the training data is limited, the model may memorize the training set (overfitting). The data augmentation methods used in this study and the ratios used are given in Table 1.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

 ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

Table 1. Data augmentation methods and rates used (Veri çoğaltma yöntemleri ve kullanım oranları)

Method	Rate
Grayscale	%20
Blur	1.7px
Noise	%2
Bounding Box Crop	Min %0 Max %20
Bounding Box Rotation	-20 / +20
Bounding Box Shear	±10 Horizontal / ± 10 Vertical

After the procedures were applied, the number of images in the dataset was increased to 2271 in total. The dataset was divided into 80% train, 10% valid, 10% test.

2.4. Computer vision

Computer vision is a combination of image processing and pattern recognition. The main purpose of computer vision is to process data, create models and extract information from images. It works by using analytics and optical sensors to automatically extract valuable information from an object. The development of this field is driven by the adaptation of human vision for information retrieval. Thus, it has become a branch of artificial intelligence and simulated human visualization [17,18]. Technological work with computer vision is expanding into other engineering fields such as geographic remote sensing, robotics, computer and human communication, healthcare and satellite communications.

Object detection is a fundamental research area in computer vision and artificial intelligence. Object detection aims to find the target of interest from the image, accurately determine the category and give the bounding box of each target. It is widely used in target tracking, event detection, behavior analysis, medical image analysis and many other fields [19].

2.4.1. Faster R-CNN

It is an R-CNN model proposed by Ren. The model is divided into two modules. One is a fully convolutional neural network used to generate all region recommendations, and the other is the Fast R-CNN detection algorithm. A set of convolutional layers is shared between these two modules. The input image is transmitted through the CNN network to the last shared convolutional layer. Although it is quite good in terms of detection accuracy, it still does not perform as well in real-time detection [20].

2.4.2. ResNet50

ResNet50 (Deep Residual Network) was developed by Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun in 2015. The ResNet network consists of 49 convolutional layers and one fully connected layer. The network structure can be divided into seven parts. The first part contains no residual blocks and mainly calculates the convolution, regularization, activation function and maximum pool of the input object. The second, third, fourth and fifth parts of the structure all contain residual blocks, which essentially solve the problem of training loss with the increase of network layers. Adding the full link layer results in a total of 50 layers, hence the name ResNet50. There are two basic blocks in ResNet50's network architecture. One is the identity block. Its input and output dimensions are the same. Therefore, they can be connected in series. The other basic block is the convolution block. Input and output sizes are different. Therefore it cannot be connected in series [21].

2.4.3. YOLO

In 2015, Joseph Redmon and colleagues introduced the YOLO (You Only Look Once) algorithm, an object detection system that for the first time performs all the steps necessary to detect an object using a single neural network. It reframes object detection as a single regression problem, directly from image pixels to bounding box coordinates and class probabilities. This unified model simultaneously estimates class probabilities for multiple bounding boxes and the objects covered by the boxes. At the time of its

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

 ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

release, the YOLO algorithm produced significant features for detecting and determining object coordinates, outperforming leading algorithms in both speed and accuracy [22,23]. In the following years, the YOLO algorithm has been upgraded to several versions for computer vision research. The first three versions were developed by Joseph Redmon, the author of the YOLO algorithm. However, after the release of YOLOv3, Redmon announced the end of his research. In early 2020, the official YOLO Github account published YOLOv4 by Alexey Bochkovskiy, based on Redmon's Darknet framework. A month after the release of YOLOv4, Glenn Jocher, a researcher developing YOLO algorithms on the Pytorch framework, and his Ultralytics research department released YOLOv5. Although not developed by the team of algorithm authors, YOLOv5 showed outstanding performance compared to previous versions [24].

Since the release of the first model in 2015, the YOLO family has continued to evolve with versions YOLO, YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, YOLOv7, YOLOv8, YOLO-NAS, YOLO-World, YOLOX and YOLOv9 [25]. The YOLO model series is used in many areas such as real-time target detection applications such as automated driving and security systems, object detection applications for inventory management, face recognition applications for secure access protocols, automatic text extraction requiring natural language processing with Natural Language Processing (NLP), robot controls [26].

The working logic of YOLO is based on the CNN algorithm. The first 20 convolutional layers of the model are pre-trained with ImageNet by adding a temporary mean pooling and fully connected layer. Since research has shown that adding convolution and connected layers to a pre-trained network improves the performance of the model, this pre-trained model is transformed to perform detection. The last fully connected layer of YOLO estimates both class probabilities and bounding box coordinates. The YOLO algorithm surrounds the objects it detects on the images with a bounding box. YOLO divides the input image into NxN grids. Each grid is responsible for determining whether there is an object in its area and if it thinks there is an object, whether the center point is in its area. Then, the grid that decides that the object has a center point estimates that bounding box with its parameters and confidence scores [27].

In this study, YOLOv5, YOLOv8, YOLOv9, Faster R-CNN and ResNet50 models described above are trained and their performances are compared. Computer vision algorithms are divided into two types, single shot and multi shot detectors. Single shot detectors work by the model processing the image only once, so they stand out in terms of speed. Multi shot detectors perform object detection in two or more stages. In the first stage, object regions are suggested and in the second stage, these regions are localized by classifying them in detail, which is why they stand out in terms of accuracy. The YOLO model family is preferred because of its high accuracy and speed performance in real-time applications. Thanks to its lightweight model structure, it is easy to implement in embedded systems. In YOLO versions, the most recent (v8, v9) and proven (v5) versions are preferred. The Faster R-CNN model is especially preferred when accuracy is critical. In terms of speed, it can underperform the YOLO family. The ResNet50 algorithm, a transfer learning model for the backbone, is usually used with Faster R-CNN model training. The system performs LPG vehicle detection by analyzing images obtained from mobile cameras. By using computer vision techniques, a new approach has been developed for automatic and stable vehicle safety inspections for risky areas. The workflow chart in Figure 2 shows the steps of the study.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

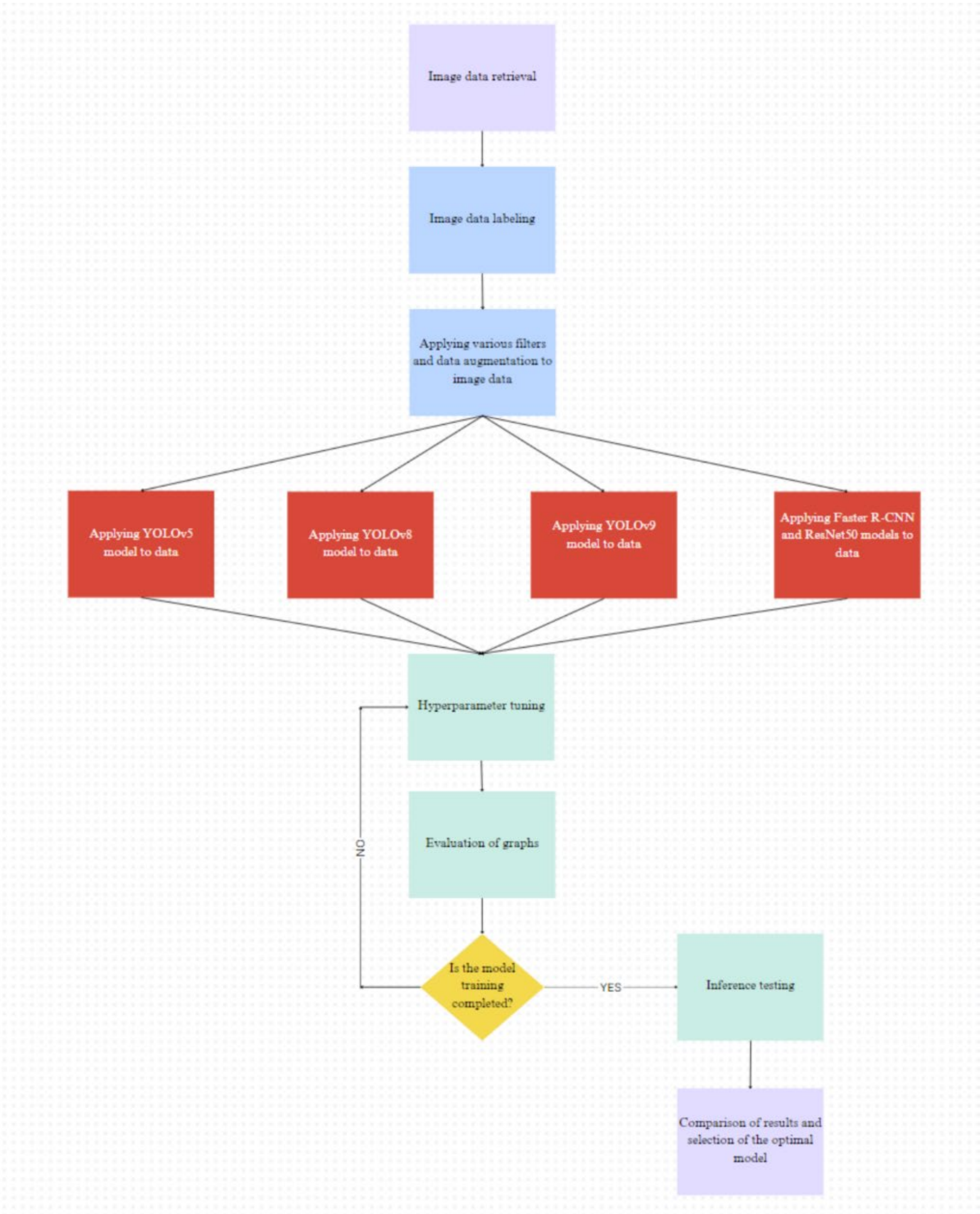

Figure 2. Work flow chart (İş akış şeması)

2.5. Performance Metrics

Mean squared error (MSE) is the average of the squared differences between the actual and predicted values of the regression model. The following formula shows the function that calculates the MSE value [28].

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x})^2 \tag{1}$$

The Region Proposal Networks (RPN) loss function is used to train the RPN network to generate proposals that are more likely to contain objects and have more accurate bounding box coordinates. The

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

RPN loss function consists of two terms. The first is the classification loss. This term measures how well the RPNs are able to classify each proposal as an object or background. Classification loss is usually calculated using the binary cross entropy loss function. The other is the bounding box regression loss. This term measures how well the RPN can predict the bounding box coordinates of each object proposal. The regression loss is typically calculated using a smooth L1 loss function. Below is the mathematical formula for the RPN loss function.

$$L = L_{cls} + L_{box} \tag{2}$$

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}}\sum_i \left(L_{cls}(p_i, p_i^*) + \frac{\lambda}{N_{box}}\sum_i p_i^* \,.\, L_1^{smooth}(t_i - t_i^*)\right) \tag{3}$$

L_{cls} is the logarithmic loss function between two classes. In the region proposal layer, regions are only divided into two classes, either background or object. The variable p_i^* is the confidence target. It is equal to 1 when there is an object. Otherwise it is 0 [29].

Object detection and segmentation metrics are used to analyze and evaluate the performance of object detection and segmentation models. Average precision ratio (mAP) and intersection and union ratio (IoU) are commonly used segmentation metrics. mAP is the average of the precision values for each computed object class. The value K represents the number of object classes. The function that calculates the mAP value is given in the formula below.

$$mAP = \frac{1}{k}\sum_i^k AP_i \tag{4}$$

IoU is the overlap between the actual region and the region predicted by the model. The function that calculates the IoU value is given in the following formula [30]

$$IoU = \frac{GP}{GP+YP+YN} \tag{5}$$

3. Research Findings

In this study, one-stage object detection algorithms YOLOv5, YOLOv8, YOLOv9 versions and two-stage object detection algorithms Faster R-CNN and ResNet50 algorithms are compared. The image data is divided into 1998 train, 180 valid and 93 test sets. After training the models with the training data in the dataset using computer vision techniques, mAP, loss and speed values were obtained. The performance results on the validation data in the dataset are given in separate charts for each model.

3.1. YOLOv5 model outputs

After preprocessing and data replication operations were performed on the dataset, training was performed with the YOLOv5 model architecture. The hyperparameters previously determined by the model were used in the training. An Early Stopping Callback with a patience of 100 was used during training. Training was performed with 500 epochs. According to the results obtained from the performance tests, the graphs in Figure 3 were obtained.

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

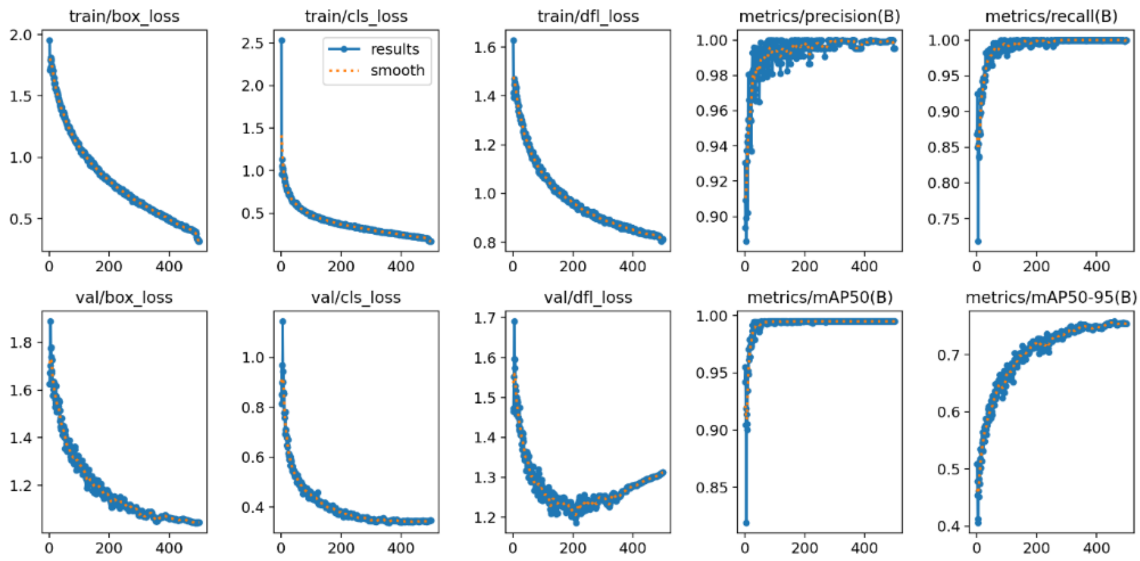

Figure 3. YOLOv5 graphics (YOLOv5 grafikleri)

In the performance test performed after the training, the performance values in Table 2 were obtained.

Table 2. YOLOv5 performance values (YOLOv5 performans çıktıları)

Model	mAP(50)	Loss (cls_loss)	Speed
YOLOv5	0.995	0.169	28.4 ms

The inference test was performed with the model obtained after training and the results are given in Figure 4.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

Figure 4. YOLOv5 test output (YOLOv5 test çıktıları)

3.2. YOLOv8 model outputs

Training was performed with the YOLOv8 model architecture. The hyperparameters predefined by the model were used in the training. An Early Stopping Callback with a patience of 100 was used during training. Training was performed with 500 epochs. The model completed training in 437 epochs. According to the results obtained from the performance tests, the graphs in Figure 5 were obtained.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

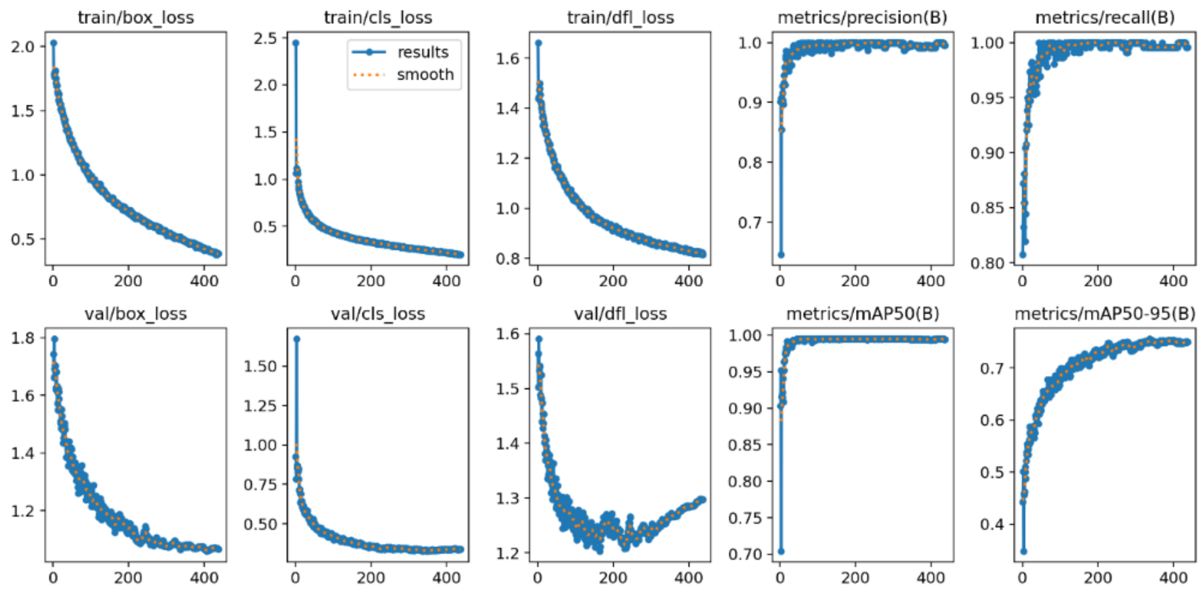

Figure 5. YOLOv8 graphics (YOLOv8 grafikleri)

In the performance test performed after the training, the performance values in Table 3 were obtained.

Table 3. YOLOv8 performance values (YOLOv8 performans değerleri)

Model	mAP(50)	Loss (cls_loss)	Speed
YOLOv8	0.994	0.200	11.6 ms

The inference test was performed with the model obtained after training and the results are given in Figure 6.

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

Figure 6. YOLOv8 test output (YOLOv8 test çıktıları)

3.3. YOLOv9 model outputs

Training was performed with the YOLOv9 model architecture. The hyperparameters predefined by the model were used in the training. An Early Stopping Callback with a patience of 100 was used during training. Training was performed with 500 epochs. The model completed training in 299 epochs. According to the results obtained from the performance tests, the graphs in Figure 7 were obtained.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

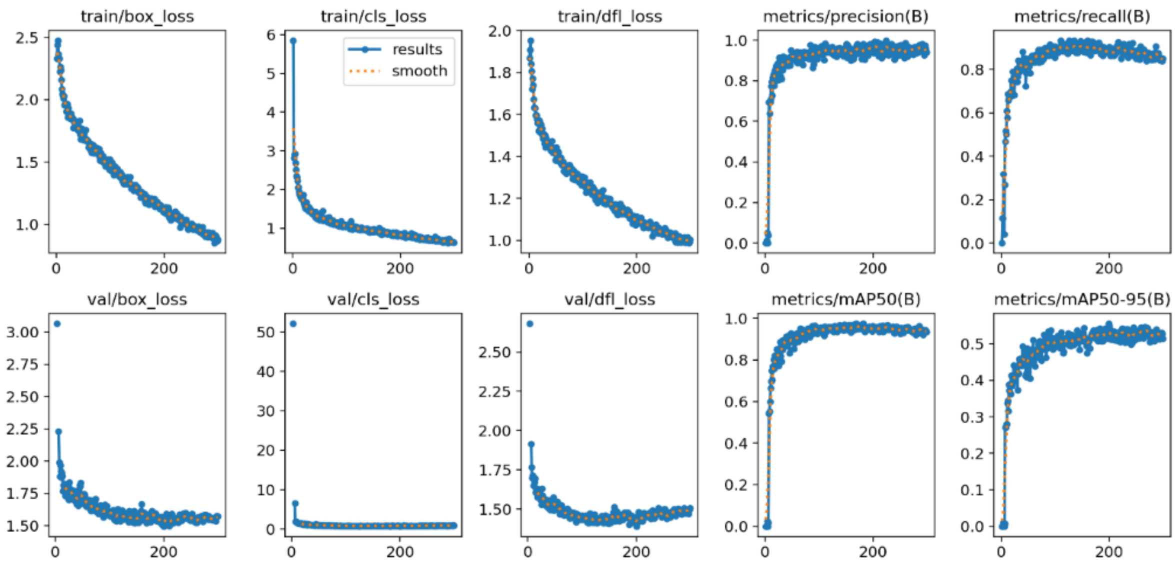

Figure 7. YOLOv9 graphics (YOLOv9 grafikleri)

In the performance test performed after the training, the performance values in Table 4 were obtained.

Table 4. YOLOv9 performance values (YOLOv9 performans değerleri)

Model	mAP(50)	Loss (cls_loss)	Speed
YOLOv9	0.935	0.628	32.7 ms

The inference test was performed with the model obtained after training and the results are given in Figure 8.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

Figure 8. YOLOv9 test output (YOLOv9 test çıktıları)

3.4. Faster R-CNN and ResNet50 model outputs

Training was performed by using FasterR-CNN and ResNet50 model architectures together. The hyperparameters previously determined by the models were used in the training. An Early Stopping Callback with a patience of 100 was used during training. Training was performed with 500 epochs. During the model training, it was observed that the training was completed at epoch 97. According to the results obtained from the performance tests, the graphs in Figure 9 were obtained.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

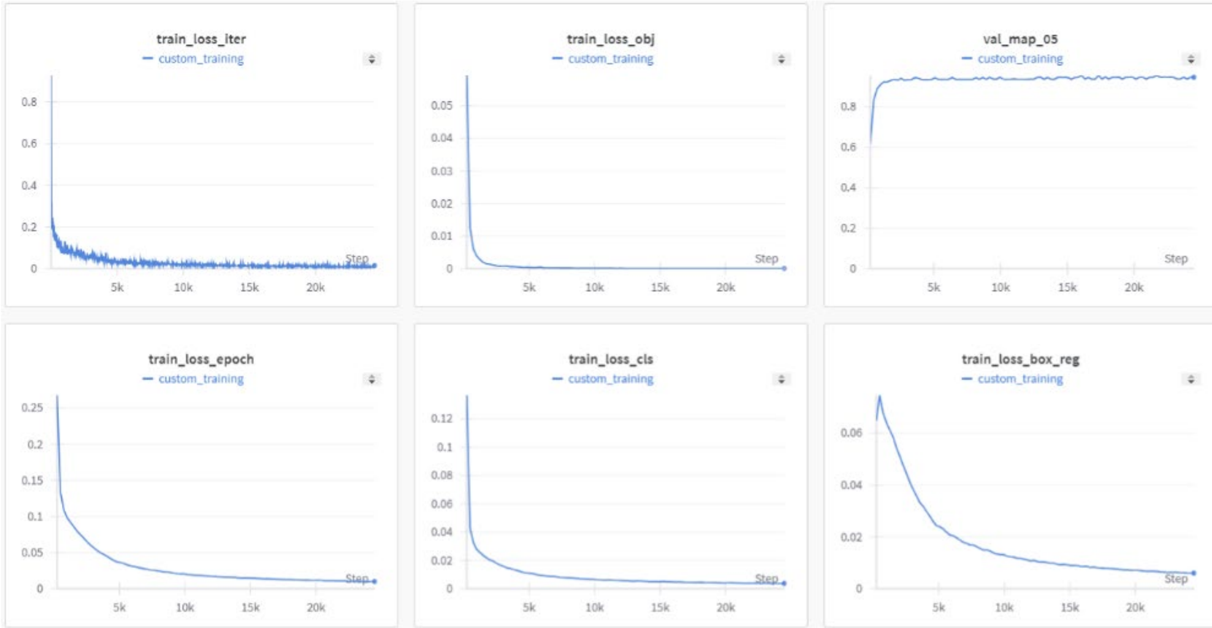

Figure 9. Faster R-CNN and ResNet50 graphics (Faster R-CNN ve ResNet50 grafikleri)

In the performance test performed after the training, the performance values in Table 5 were obtained.

Table 5. Faster R-CNN and ResNet50 performance values (Faster R-CNN ve ResNet50 performans değerleri)

Model	mAP(50)	Loss (cls_loss)	Speed
Faster R-CNN and ResNet50	0.947	0.011	151.4 ms

The inference test was performed with the model obtained after training and the results are given in Figure 10.

Figure 10. Faster R-CNN and ResNet50 test output (Faster R-CNN ve ResNet50 test çıktıları)

4. Results and Discussion

Studies in the literature show that YOLO models are widely used in intelligent transportation systems due to their speed and accuracy performance. In real-time applications such as traffic density analysis,

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

license plate detection and hazardous material detection, it has been shown that the performance increases with the preference of CNN-based architectures [9]. It has been observed that studies in the field of LPG are generally carried out on leak detection and industrial issues. In one study, Yarens J. Cruz et al. successfully detected the misalignment of the parts to be welded with a 97.7% success rate by using CNN architecture and image processing techniques in LPG welding inspection processes. Thanks to its low cost and effective result, it has shown that this technology can be used in the industrial industry [24]. Although similar applications have been carried out in terms of subject matter in the reviewed studies, there is no study on the detection of LPG vehicles using computer vision techniques. This study will contribute to the literature with this unique value. In addition to contributing to the application of computer vision techniques, the developed system is expected to benefit the national economy, public life safety and environmental protection.

This study was carried out to control the entry of LPG vehicles in areas where they pose a risk with an automatic and sustainable control system using computer vision techniques. In case of any leakage in LPG tanks, since the gases in LPG are heavier than air, they accumulate on the ground and create a serious risk of explosion. For this reason, security policies have a very important place. Image data captured by mobile cameras in Aydın, Isparta and Muğla cities were trained with four different models and compared.

The performance results of the trained models are presented in Table 6. In the performance tests performed on the validation dataset after training the models with different architectures, the highest accuracy value of 0.995 was obtained with the YOLOv5 model. The lowest loss was obtained with the Faster R-CNN-ResNet50 model with a value of 0.0118. In the performance tests, the fastest running model was YOLOv8 with 11.6 ms, while the slowest running model was Faster R-CNN and ResNet50.

Table 6. Performans values of models (Modellerin performans değerleri)

Method	mAP	Loss	Speed
YOLOv5	0.995	0.169	28.4 ms
YOLOv8	0.994	0.200	11.6 ms
YOLOv9	0.935	0.628	32.7 ms
Faster R-CNN and ResNet50	0.947	0.011	151.4 ms

After training and performance tests on YOLOv5, YOLOv8, YOLOv9 and Faster R-CNN, ResNet50 models, a comparison was made and YOLOv8 showed 0.001 lower performance than YOLOv5 with an accuracy of 0.994, but the model speed of 11.6 ms was higher than the other models. In this case, it has been revealed that YOLOv8 is a more preferable model in terms of applicability in real life and real-time monitoring. Thanks to the designed system, it is predicted that a more sustainable and stable system can be created as well as saving time by automatically performing environmental safety inspections in institutions and organizations. With this system, vehicles that have been retrofitted with LPG can be easily detected and their entry into enclosed parking areas can be prevented.

5. Conclusion

In this study, LPG fueled vehicles are automatically detected using computer vision techniques. Vehicle image data captured by mobile cameras in different provinces of Turkey were trained and compared with four different deep learning models. As a result of training and performance tests on the models, the YOLOv8 model was more effective than the other models with an accuracy of 0.994 mAP and a speed of 11.6 ms. It has been shown to be a stable model in terms of real-time monitoring in real life.

For future studies, the performance of the training can be improved with larger data sets and it can be made a sustainable system with security cameras or cameras to be placed at parking lot entrances. In addition, with the LPG detection system developed, it is possible to calculate analyses such as carbon footprint and fuel use by creating LPG vehicle statistics.

*Sorumlu Yazar/Corresponding Author: gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

Acknowledgements

I would like to thank Abdullah Fatih HÖCÜ, Ali Tezcan SARIZEYBEK for their support during my research and Yılmaz SUÇEKEN for his hard work in creating the dataset.

References

[1] Ş.G. Taç, Karayolu ulaşımında meydana gelen trafik kazalarının önlenmesinde akıllı ulaşım sistemlerinin etkisi. Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi, 1(2) (2018) 12-21.

[2] F. Coşkun, H.D. Gülleroğlu, Yapay zekânın tarih içindeki gelişimi ve eğitimde kullanılması. Ankara University Journal of Faculty of Educational Sciences (JFES), 54(3) (2021) 947-966.

[3] P. Sermanet, Y. LeCun, Traffic sign recognition with multi-scale convolutional networks. In The 2011 international joint conference on neural networks, IEEE, (2011) (pp. 2809-2813).

[4] S. Ren, K. He, R. Girschich, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, (2015) 28.

[5] S. Lange, F. Ulbrich, D. Goehring, Online Vehicle Detection Using Deep Neural Networks and Lidar Based Preselected Image Patches. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV) (2016).

[6] U.K.K. Pillai, D. Valles, An initial deep CNN design approach for identification of vehicle color and type for amber and silver alerts. In 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC) IEEE. (2021) pp. 0903-0908.

[7] A. Gholamhosseinian, J. Seitz, Vehicle classification in intelligent transport systems: An overview, methods and software perspective. IEEE Open Journal of Intelligent Transportation Systems, (2021) (2) 173-194.

[8] J .N. Niroomand, C. Bach, M. Elser, Robust vehicle classification based on deep features learning. IEEE (2021) Access, 9, 95675-95685.

[9] J. Azimjonov, A. Özmen, A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways. Advanced Engineering Informatics, (2021) (50) 101393.

[10] A. Sökülmez, Karayolu Bakış Açılı Görüntü İle Derin Öğrenme Tabanlı Araç Algılama- Sayma, Trafik Yoğunluğu Hesabı YApılması ve Google Haritalara Anlık Veri Sunma Sistemi. Yönetim Bilişim Sistemleri Ansiklopedi, 1.

[11] M.Z. Zaheer, J.H. Lee, S.I. Lee, B.S. Seo, A brief survey on contemporary methods for anomaly detection in videos. In 2019 International Conference on Information and Communication Technology Convergence (ICTC) IEEE (2019) (pp. 472-473).

[12] D.R. Patrikar, M.R. Parate, Anomaly detection using edge computing in video surveillance system. International Journal of Multimedia Information Retrieval, (2022) 11(2), 85-110.

[13] R. Nayak, U.C. Pati, S.K. Das, A comprehensive review on deep learning-based methods for video anomaly detection. Image and Vision Computing, (2021) (106) 104078.

[14] Y.J. Cruz, M. Rivas, R. Quiza, A. Villalonga, E.R. Haber, G. Beruvides, Ensemble of convolutional neural networks based on an evolutionary algorithm applied to an industrial welding process. Computers in Industry, (2021) (133) 103530.

[15] N. Turgut, Sıvılaştırılmış petrol gazı (LPG) sistemleri. Atatürk Üniversitesi Ziraat Fakültesi Dergisi, (2010) 14(1-2).

[16] https://www.verikaynagi.com/grafik/yakit-turlerine-gore-kara-tasit-sayisi/ (accessed 23.05.2024)

[17] J.S. Suri, Computer vision, pattern recognition and image processing in left ventricle segmentation: The last 50 years. Pattern Analysis & Applications, (2000) 3(3) 209-242.

[18] V. Wiley, T. Lucas, Computer vision and image processing: a paper review. International Journal of Artificial Intelligence Research, (2018) 2(1) 29-36.

***Sorumlu Yazar/Corresponding Author:** gulturker@sdu.edu.tr

ISSN: 2618-6055 / 8, (1), 26 – 43, 2024
DOI: 10.62301/usmtd.1493932

[19] J. Deng, X. Xuan, W. Wang, Z. Li, H. Yao, Z. Wang, A review of research on object detection based on deep learning. In Journal of Physics: Conference Series (2020) (Vol. 1684, No. 1, p. 012028). IOP Publishing.

[20]https://documentation.sas.com/doc/en/pgmsascdc/v_050/casdlpg/p1np8zbnoyd0brn1dhehthuuxj4q.htm#p01r9nxiv9wu0in150lnlp23bfwb (accessed 23.05.2024)

[21] P. Chu, Z. Li, K. Lammers, R. Lu, X. Liu, Deep learning-based apple detection using a suppression mask R-CNN. Pattern Recognition Letters, (2021) (147) 206-211.

[22] J. Redmon, S. Divvala, R. Girshick, 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).

[23] D. Thuan, Evolution of Yolo algorithm and Yolov5: The State-of-the-Art object detention algorithm (2021).

[24] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, Y. Kwon, J. Fang, M. Thanh Minh, ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpu and openvino export and inference. Zenodo (2022).

[25] https://blog.roboflow.com/guide-to-yolo-models/ (accessed 23.05.2024)

[26] X. Cong, S. Li, F. Chen, C. Liu, Y. Meng, A review of YOLO object detection algorithms based on deep learning. Frontiers in Computing and Intelligent Systems, (2023) 4(2) 17-20.

[27] https://www.v7labs.com/blog/yolo-object-detection (accessed 23.05.2024)

[28] D. Chicco, M. J. Warrens, G. Jurman, The Coefficient of Determination R-Squared is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. PeerJ Computer Science, (2021) (7) e623.

[29] P. Steno, A. Alsadoon, P.W.C. Prasad, T. Al-Dala'in, O.H. Alsadoon, A novel enhanced region proposal network and modified loss function: threat object detection in secure screening using deep learning. The Journal of Supercomputing, (2021) (77) 3840-3869.

[30] R. Padilla, S. L. Netto, E.A. Da Silva, A Survey on Performance Metrics for Object-Detection Algorithms. In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), (2020) 237-242.

***Sorumlu Yazar/Corresponding Author:** gulturker@sdu.edu.tr