SAVUNMA BİLİMLERİ DERGİSİ
The Journal of Defense Sciences
https://dergipark.org.tr/tr/pub/khosbd

A Traffic Sign-Aware Artificial Intelligence Model For Advanced Driver Assistance Systems

Gelişmiş Sürücü Destek Sistemleri İçin Trafik İşaretlerine Duyarlı Yapay Zekâ Modeli

Seyfettin VADİ [1*] Simge KOÇAK [2]

[1,2] *Gazi University, Technology Faculty, Department of Electrical and Electronics Engineering, Ankara, Türkiye*

Makale Bilgisi

Araştırma makalesi
Başvuru: 14.09.2025
Düzeltme:02.11.2025
Kabul: 02.12.2025

Keywords

ADAS
Autonomous Vehicle
Traffic Sign Recognition System

Anahtar Kelimeler

ADAS
Otonom araç
Trafik işaretlerine duyarlı sistem

Highlights

A YOLOv8-based approach was utilized for traffic light detection and color recognition. The Ultralytics YOLOv8 library was installed in a GPU-supported environment. A carefully prepared traffic light dataset was created, reflecting various lighting and angle conditions. The model achieved a very high mean Average Precision (mAP) value by the 80th epoch, approximately 0.98. In addition to the mAP success, a balanced Precision-Recall performance was achieved.

Graphical Abstract

Özet

Günümüzde otonom sistemlerin giderek yaygınlaşması, sürüş ortamının algılanması ve sürücülere ya da otonom sistemlere gerçek zamanlı bilgi sunan Gelişmiş Sürücü Destek Sistemleri (ADAS) ile bütünleştiğinde, yol güvenliği açısından büyük bir dönüşüm sağlamaktadır. Özellikle bilgisayar teknolojilerindeki ilerlemeler ve derin öğrenme tabanlı yöntemlerin başarısı, hem insan faktöründen kaynaklanan hataları en aza indirme hem de trafiği akıcı, güvenli ve konforlu hale getirme potansiyeli taşımaktadır. Bu çalışmada, otoyollardaki trafik akışı sorunlarını, zaman kaybını ve maliyet artışlarını azaltmak amacıyla trafik işaretlerine, özellikle de trafik ışıklarına duyarlı bir yapay zekâ modeli geliştirilmiştir. Çalışmada, geleneksel otomatik olay tespit (AID) yöntemlerinde sıkça karşılaşılan düşük tespit oranları ve yüksek yanlış alarm sorunlarını aşmak üzere, sinir ağları ve derin öğrenme teknikleri kullanılmıştır. Geliştirilmiş YOLOv8 algoritması kullanılarak ileri seviye nesne tespiti algoritmaları yardımıyla, trafik ışıkları da dâhil olmak üzere çeşitli yol işaretleri ve şeritler, farklı çevresel koşullar ve görüş açıları altında yüksek doğrulukla saptanmıştır. Böylece yalnızca trafik akışına yönelik veriler değil, aynı zamanda işaretlerin anlamları, konumları ve yönleri de modele entegre edilerek, sürücülere ve otonom araçlara daha kapsamlı bilgi sunulmuştur.

Abstract

The increasing prevalence of autonomous systems today, when integrated with Advanced Driver Assistance Systems (ADAS) that perceive the driving environment and provide real-time information to drivers or autonomous systems, is bringing about a significant transformation in terms of road safety. Specifically, advancements in computer technology and the success of deep learning-based methods hold the potential to both minimize errors stemming from human factors and make traffic smoother, safer, and more comfortable. In this study, an artificial intelligence model sensitive to traffic signs, particularly traffic lights, has been developed to mitigate traffic flow issues, reduce time loss, and prevent cost increases on highways. The study employs neural networks and deep learning techniques to address the low detection rates and high false alarm issues commonly encountered in traditional automatic incident detection (AID) methods. Advanced object detection algorithms, the enhanced YOLOv8 algorithm, have been utilized to accurately identify various road signs and lanes, including traffic lights, under diverse environmental conditions and viewing angles. Thus, not only traffic flow data but also the meanings, locations, and directions of signs were integrated into the model, providing drivers and autonomous vehicles with a more comprehensive information environment.

*Corresponding author, e-mail: seyfettinvadi@gazi.edu.tr
DOI: 10.17134/khosbd.1783706
Elektronik/Online ISSN: 2148-1776
Basılı/Printed ISSN: 1303-6831

1. INTRODUCTION

Throughout history, the automobile has been a central element that facilitates human mobility, provides individuals with a sense of freedom, and shapes the physical and economic structure of cities. Although driving is considered a mundane routine today, even a moment of driver distraction can have serious consequences. Indeed, traffic accidents are one of the leading causes of death in many geographies. World Health Organization data reveal that an average of 1.35 million people worldwide lose their lives in traffic accidents each year. It is known that the rapidly growing urban population and the number of motor vehicles, especially in low- and middle-income countries where adequate traffic safety measures are not yet fully established, create increasing risks. These conditions elevate traffic accidents beyond a mere public health issue, making them a problem with a direct impact on economic development as well. Furthermore, financial losses from traffic accidents in these countries can reach up to 3% of the gross domestic product (GDP) [1].

The fact that countries have different traffic accident statistics on an international scale stems from a wide variety of factors, including road infrastructure, vehicle safety standards, driver behavior, alcohol and drug use, compliance with speed limits, education level, and the quality of emergency response services [2]. Advanced traffic safety policies, adequate inspections, high-tech vehicles, and informed drivers are often associated with lower fatality rates. In contrast, fatality rates remain higher in areas where inadequate measures are taken. In recent years, European Union (EU) countries have taken significant steps to enhance traffic safety, resulting in a noticeable reduction in fatalities caused by accidents [3]. However, achieving the 2025 and 2030 targets still seems challenging, highlighting the need for continuous strengthening of holistic strategies such as increasing infrastructure investments, expanding educational programs, supporting technological innovations, and raising public awareness.

While the "human factor" is at the heart of traffic accidents, the rapidly increasing urban population and growing transportation demand are putting additional pressure on existing transportation networks. United Nations projections indicate that by 2050, approximately 68% of the world's population will live in cities. This demographic shift makes sustainable and efficient transportation management even more critical, exacerbating issues such as heavy traffic, congestion, and the risk of accidents. New-generation technologies are of great importance in addressing these problems. Especially, deep learning-based approaches make it possible to automatically process visual data, quickly and accurately detect traffic signs, thereby strengthening real-time decision support systems. Thus, drivers and autonomous vehicles can operate based on real-time and reliable information, and cities are moving toward more innovative, safer, and more dynamic transportation systems [4]. Autonomous driving technologies, which have emerged due to advancements in technology, are attracting attention as a dynamic field encompassing a broad ecosystem of infrastructure, hardware, scenarios, and services. It is not possible to adhere to a specific timeline in this field. At the

*Corresponding author, e-mail: seyfettinvadi@gazi.edu.tr
DOI: 10.17134/khosbd.1783706
Elektronik/Online ISSN: 2148-1776
Basılı/Printed ISSN: 1303-6831

same time, some innovations have already entered daily life, and others will occur in the future as a result of overcoming technical obstacles or shaping political decisions. This automation process is part of a larger transformation resulting from the integration of technologies such as computers, mobile phones, and the Internet, which interact to create machines that understand, perceive, and process the physical environment. These machines are not only making their mark in the automotive industry but also in various other fields such as uncrewed aerial vehicles, maintenance robots, 3D printers, and security applications [5].

As shown in Figure 1, autonomous vehicles collect and analyze information from the environment, just like a human driver, and make informed decisions based on this information. Components such as sensors, radar, lidar, GPS, camera systems, and high-resolution maps enable the vehicle to perform essential functions successfully, including route planning, obstacle detection, and traffic sign interpretation. Today, vehicle control can be achieved without driver intervention in specific test scenarios and under suitable driving conditions. The operating principles of autonomous vehicles are defined by their levels of autonomy. According to the 6-level classification defined at the IEEE conference, Level 0 contains no autonomy, while Levels 1 and 2 provide driver support functions with limited assistance. Level 3 offers conditional autonomy; the vehicle can control its speed and direction independently, but driver intervention is expected in certain situations. Level 4 autonomy means the vehicle can perform many driving actions without human intervention, but the driver has the option to take control at any time. The final level, level 5, is full autonomy; at this level, driver input is not required, the vehicle performs every action a human can, and driver intervention can even be prevented. This level will go beyond established car design patterns, introducing new cabin concepts that focus on comfort and social interaction [6]. These developments also raise many ethical and legal questions. A real-life example came to light in 2018 when an autonomous vehicle in the testing phase caused the death of a pedestrian [7]. This incident has raised significant questions about how autonomous vehicles' decision-making mechanisms function in emergencies. While human drivers often make decisions based on intuition and emotion, autonomous vehicles react within milliseconds by calculating within the framework of predefined algorithms. This also leads to moral dilemmas, such as the traditional "Trolley Problem," being brought to the forefront [8]. All this uncertainty can make it difficult for consumers to trust technology. [9] emphasizes that the widespread adoption of a new technology depends on convincing consumers that the product performs its functions safely and reliably. In the autonomous vehicle ecosystem, not only end-users but also vehicle manufacturers, software developers, and infrastructure providers can be positioned as "consumers."

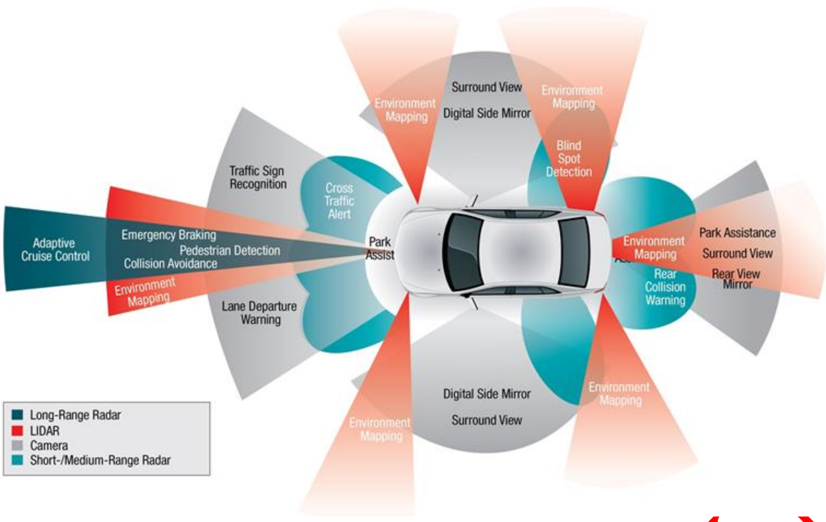

Figure 1. Sensor Technologies and Detection Areas Used in Autonomous Vehicles [6]

Establishing trust, defining ethical codes, and ensuring societal acceptance of these technologies are vital for autonomous vehicles to be fully integrated into the transportation ecosystem in the future. In the short term, we might see a limited number of test drives; however, in the long run, a fully autonomous transportation network could bring about a new way of traveling and interacting that surpasses today's imagination.

The history of work on autonomous vehicles and related technologies dates back to the mid-20th century, when the first concepts were proposed. One of the earliest known ventures in this field was the demonstration performed by General Motors at the 1939 World's Fair [10]. Subsequently, the development of a vehicle capable of reaching speeds of 20 km/h and following the white line on the road at the Mechanical Engineering Laboratory in the Tsukuba region of Japan in 1977 was an interesting step in autonomous driving. By the 1980s, not only the vehicle's own movements but also the perception of surrounding elements had become a topic of discussion. During this period, the first basic research on the processes of perceiving and recognizing road and traffic signs was conducted, particularly in a study carried out in Japan in 1984. This study highlights the importance of computer vision techniques in autonomous vehicles for the coming decades, emphasizing fundamental image processing methods. In 1989, Pomerleau conducted the first road tests of the ALVINN land vehicle, supported by DARPA in the United States. He successfully processed data from camera and laser sensors using neural network-based artificial intelligence techniques, enabling the vehicle to reach a speed of 30 km/h [10]. Subsequently, the "European Prometheus Project" conducted in Europe during the 1990s carried out the longest robotic car experiment to date, with the vehicle reaching a speed of 96 km/h on empty streets and demonstrating a successful drive. Around the same time, Dickmann's famous S-class autonomous car covered long distances without human intervention and at a speed of 158 km/h from Munich to Denmark. These developments

*Corresponding author, e-mail: seyfettinvadi@gazi.edu.tr
DOI: 10.17134/khosbd.1783706
Elektronik/Online ISSN: 2148-1776
Basılı/Printed ISSN: 1303-6831

have demonstrated the technical feasibility of the autonomous vehicle concept; however, the full-scale adoption of these systems has been awaiting advancements in cost, sensor quality, processing power, and data processing techniques [11].

By the 2000s, the variety of autonomous vehicles increased due to lower costs, increased wireless communication possibilities, and the emergence of more flexible processing platforms. In 2008, Çayıroğlu and Şimşir transferred wireless camera recordings to a computer using a robot they designed with a PIC16F877A [12], while Fong and colleagues experimented with Wi-Fi-based real-time control [13]. Kuo and his colleagues added wireless control functionality to an Arduino-based system [14]. In the same year, Espes and his colleagues developed a mobile robot that operates over the internet, utilizing Raspberry Pi and Arduino Mega. In 2017, Çetinkaya provided vehicle control via Bluetooth using fuzzy logic-based position control [15]. These examples demonstrate that not only the hardware but also the methods used in autonomous systems have undergone significant evolution.

On the other hand, following the first traffic sign detection studies in 1984, the literature has expanded rapidly, especially since the 2000s. In early approaches, techniques such as color segmentation, geometric moments, decision trees, and statistical methods were prominent [16]. Yuan and colleagues found that CNN-based deep learning methods were becoming widespread [17], and Stallkamp and colleagues demonstrated that CNN models could outperform human performance [18]. Xu and his colleagues, Arcos and his colleagues, Torres and his colleagues, Rahman and his colleagues, Mathias and his colleagues, Jung and his colleagues, Li and his colleagues, Abedin and his colleagues, and Song and his colleagues have made significant progress in traffic sign detection and classification using different deep learning architectures, region proposal networks (Faster R-CNN, R-FCN, SSD, YOLO), color segmentation, HSI-based thresholding, geometric moments, and methods like SURF [19-27].

In recent years, it has been observed that deep learning-based methods have replaced traditional algorithms. [28] The PRM and ABC algorithms were used to generate a turning route, demonstrating a classic approach. Meanwhile, in the same year, [29] utilized deep learning to ensure the autonomous vehicle reached its destination without human intervention. For parking problems, [30] achieved successful autonomous parking experiments using a machine learning approach, while [31] used reinforcement learning. [32], While implementing autonomous parking applications on embedded automotive platforms using end-to-end learning and Deep Neural Networks, [33] developed a successful automatic parking controller with a dual neural network model.

This chronological progression illustrates the historical development and evolution of autonomous vehicle technologies, as well as the methods for detecting and interpreting traffic signs. Studies that initially focused on basic image processing and sensor integration have made significant advancements in both accuracy and efficiency today, thanks to the introduction of deep learning and artificial intelligence-based models. Thus, autonomous driving is constantly improving and enriching not only mechanical autonomy but also visual perception, object recognition, and decision-making abilities.

In this study, a traffic light-sensitive artificial intelligence model has been developed to mitigate traffic flow issues, reduce time loss, and prevent cost increases on highways. In this context, the study employs neural networks and deep learning techniques to address the low detection rates and high false alarm rates frequently encountered in traditional automatic incident detection (AID) methods. Traffic lights were detected with high accuracy under different environmental conditions and viewing angles using advanced object detection algorithms, particularly with the help of the improved YOLOv8 algorithm. Thus, not only traffic flow data but also the meanings, locations, and directions of signs were integrated into the model, providing drivers and autonomous vehicles with a more comprehensive and up-to-date information environment. As a result, the developed model is expected to increase the operational efficiency of traffic management centers, improve driver safety and comfort, and enable autonomous vehicles to make more effective decisions in constantly changing driving environments.

2. ADAS AND ARTIFICIAL INTELLIGENCE-BASED TRAFFIC SIGN RECOGNITION

Advanced Driver Assistance Systems (ADAS) have become a significant part of the technological transformation in the automotive industry. These systems are designed to minimize the challenges drivers may face, reduce accident risks, enhance road safety, and improve the overall driving experience. Initially developed solely to guide the driver, these systems have evolved into a comprehensive technology package encompassing various functions, such as lane-keeping assist, adaptive cruise control, automatic emergency braking, blind spot warning, and traffic sign recognition. Artificial intelligence-based methods play a key role in the vehicle's perception of its surroundings and in conveying meaningful information to the driver by quickly and reliably recognizing traffic signs. Thus, the driver can instantly view speed limits, hazard warnings, or road regulations even in changing weather and lighting conditions, and take appropriate measures when necessary. From a historical perspective, the foundations of ADAS were laid with studies aimed at reducing the driver's workload. Pioneering projects like DRIVE programs and GIDS (Generic Intelligent Driver Support) have focused on developing an "intelligent" support system that can make sense of the dense data stream from in-vehicle sensors and applications and present it to the driver. This approach, on the one hand, has reduced information overload targeting the driver. On the other hand, it has enabled the development of automation-based functions that support timely and accurate decision-making. However, these developments have not eliminated the human factor; on the contrary, human-machine interaction has become even more critical with the advancement of technology. Human-related issues, such as distracted driving, fatigue, or incorrect decisions, remain relevant. As automation levels increase, the transfer of some control to the device directly affects drivers' trust, acceptance, and usage habits.

In conclusion, ADAS systems hold a vital position in the transportation ecosystem, both today and in the future. Functions such as automatic speed adjustment, autonomous driving, smart navigation, and collision avoidance are gradually transforming the driver's role while enhancing the safety and efficiency of the driving environment. In this transformation, AI-powered traffic sign recognition technologies are reducing errors caused by human factors, offering a safer, more comfortable, and predictable driving experience. Thus, ADAS contributes to the future of transportation by minimizing human intervention on the road to a fully automated highway system.

3. METHODOLOGY AND SIMULATION RESULTS

In this study, a YOLOv8-based model was developed for the automatic detection and color identification of traffic lights. For this model, a Jupyter computing environment with GPU (Graphics Processing Unit) support was preferred to train deep learning models effectively. Within the scope of the study, the necessary libraries (e.g., Ultralytics YOLOv8 and related Python libraries) were installed, thereby preparing the tools for use in the training and inference stages of the model. Specifically, the fast and user-friendly structure of YOLOv8 has provided a suitable foundation for accurately learning both the position and color information of objects, such as traffic lights. The presence of an NVIDIA GPU was checked, and the suitability of relevant system components (CUDA, PyTorch version, etc.) was verified to ensure the model could be trained at the highest possible speed during deep learning processes. Modeling was performed at epochs 40 and 80 to detect traffic lights and automatically identify their colors, and a comparison was made. Additionally, the project directory has been organized into a structured format, ensuring that the dataset, weight files, and model outputs are stored in a manageable manner. The YOLOv8 algorithm has been selected for its ability to provide fast, accurate, and comprehensive data in complex and constantly changing driving environments.

The primary reasons for evaluating the 40th and 80th epochs in the study are to monitor the model's progress in the learning process, verify the stability of the results, and demonstrate, with concrete data, how performance changes as the training duration increases. The study used the Confusion Matrix, F1-Confidence Curve, Precision-Confidence Curve, and Recall-Confidence Curve to compare the evaluations made at the 40th and 80th epochs.

3.1. Modeling by 40. Epoch

The dataset, comprising traffic light images and corresponding label information, was obtained from an internet-based platform (e.g., Roboflow). This dataset is organized to include color labels (e.g., red, yellow, green) in addition to the location tags for traffic lights. In the study, various images were combined to enhance data quality and represent different conditions (such as different angles and light intensities), and training, validation, and test sections were created. By integrating Ultralytics' YOLOv8 version into the project, it has become possible to customize the model. In this process, the library was checked for compatibility with the system. Then the basic functions provided by YOLOv8 (training, prediction, model evaluation), as shown in Figure 2, were tested for review and adaptation purposes. After all these steps are completed, the YOLOv8 model is ready for training in traffic light detection and color recognition, as per the created method. The training and evaluation stages of the model have been completed; the performance of traffic light detection and color classification has been analyzed based on the results obtained.

Figure 2. Results of the model at the 40th epoch

The Confusion Matrix is one of the fundamental tables used to evaluate the performance of a classification model. This table clearly shows which examples the model classified correctly, which ones it made mistakes on, and between which classes these errors are concentrated. Typically, the columns in the matrix represent the "true/actual classes," while the rows represent the classes "predicted by the model." Thus, it can be clearly seen which class is recognized with what accuracy or which classes are frequently confused with each other [33]. The diagonal cells in the confusion matrix (e.g., correctly predicting "green" for samples that are actually "green") reflect the model's successful classification rate; the height of these values indicates that the relevant class is well-learned. On the other hand, off-diagonal cells contain classes that the model confuses (for example, predicting an instance that is "actually green" as "red"), and the growth of values in these cells indicates difficulty in distinguishing between those classes. Additionally, some complexity matrices present ratios (normalized data) instead of raw numbers, allowing for the proportional observation of the model's error density. In this way, the weaknesses of a classification model that are not apparent from its overall accuracy rate can be identified; performance can be improved through steps such as necessary data improvement, additional label correction, or changes in model settings. Furthermore, being able to predict in advance where false positive or false negative errors are likely to occur contributes to the creation of safer and more effective decision-making processes in real-world usage scenarios (e.g., autonomous systems or medical diagnostic applications). The F1-Confidence curve in Figure 3 shows the relationship between the model's F1 score and the confidence score threshold. The model marks a prediction as "positive" if the confidence score it assigns to each prediction exceeds a certain threshold, thus increasing the Precision value by accepting only the most confident examples as positive as the threshold rises. In contrast, the Recall (the rate of capturing true positives) tends to decrease. On the other hand, lowering the threshold value leads to an increase in Recall, as more examples are considered positive, but a decrease in Precision. This graph shows the performance captured at different threshold values for the F1 score (the balance between Precision and Recall). For example, a peak obtained around the 0.3–0.4 threshold indicates that the model reaches its highest F1 score in this range. Additionally, since the different colors on the graph (e.g., orange: "red", blue: "green", etc.) reflect the F1 performance of the respective classes, it is possible to see on the same axis which class is more successful up to which threshold. Thus, it becomes possible to select the optimal confidence threshold value based on application requirements (e.g., "reducing false alarms" or "minimizing missed positive examples") and ensure that the model is adjusted to best suit the real-world usage scenario.

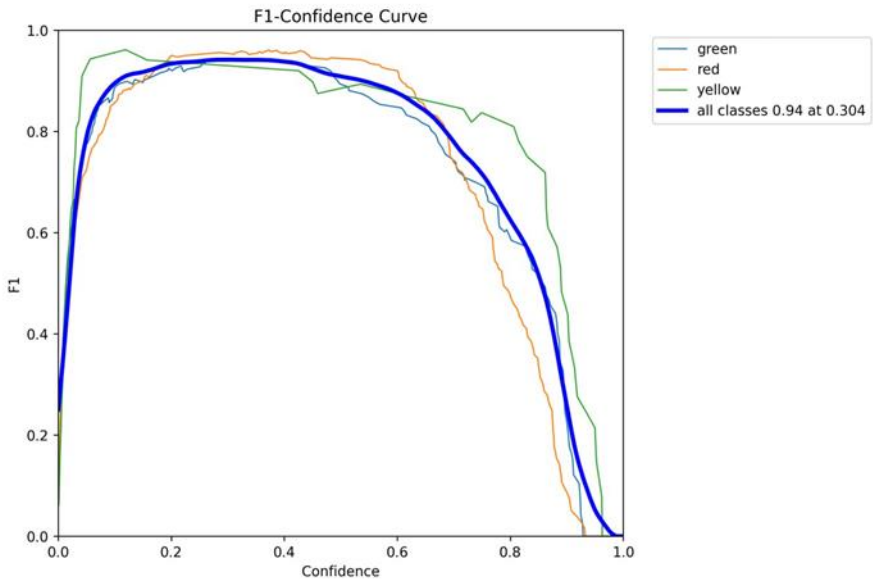

Figure 3. F-1 Confidence Curve

The Precision-Confidence Curve in Figure 4 is a valuable analysis tool that illustrates how the precision level of a classification model varies with different confidence score threshold values. The model assigns a specific confidence score to each prediction and considers predictions with a score above that threshold to be "positive." Therefore, as the threshold value increases, only predictions with the highest confidence are deemed positive, resulting in a higher overall precision. However, this approach can lead to a decrease in the Recall value (the rate at which true positives are captured) because the model, being more cautious, may miss some examples.

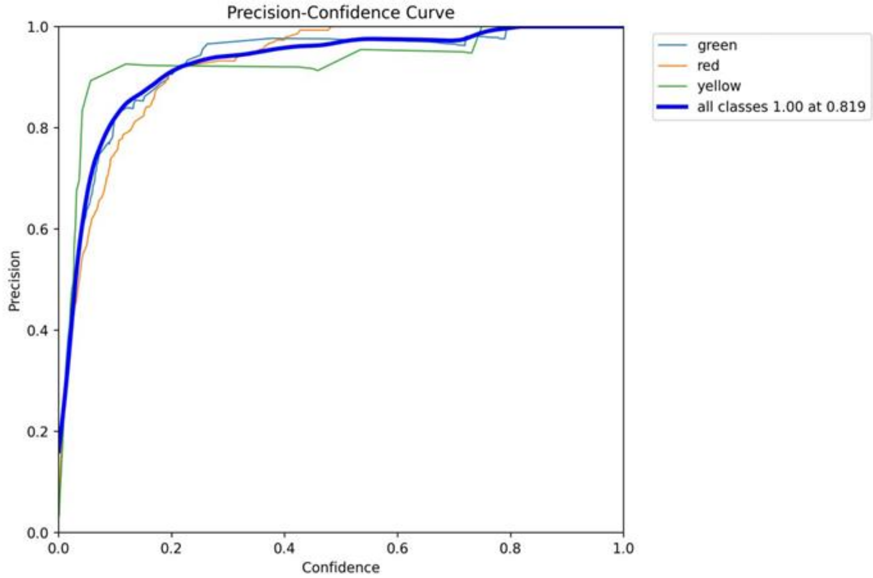

Figure 4. Precision-Confidence Graph

As seen in Figure 4, the horizontal axis represents the confidence score, and the vertical axis represents the precision value. All classes (thick blue line) and each class (blue: green, orange: red, green: yellow) are shown in different colors. At low threshold values, Precision is initially lower because the model labels relatively more instances as positive. As the threshold increases, the number of false positives (false alarms) decreases, and the curves approach 100% (1.00). It is particularly noteworthy that the thick blue line (around the ~0.82 threshold) reaches a precision of 1.00, meaning the model rarely produces false positives by accepting only "very confident" predictions as positive. However, while using a high threshold might be sensible if the application prioritizes "minimizing false alarms," a lower threshold should be chosen if the "missing no positive examples" approach is adopted. Therefore, the Precision-Confidence curve serves as a fundamental guide in determining the optimal threshold value for each application.

The Recall-Confidence curve in Figure 5 illustrates how the sensitivity (Recall) level of a classification model varies with specific confidence score threshold values. When the model assigns a confidence score to the prediction that exceeds the chosen threshold, it marks this example as "positive," thereby counting more examples as positive at lower threshold values and increasing Recall. However, as the threshold increases, only very certain cases are considered positive, leading to a rise in the number of missed samples (false negatives) and a decrease in Recall. Therefore, this curve serves as a critical guide in the process of determining the optimal threshold based on the application's priorities, such as "missing no positive examples" or "minimizing false positives as much as possible."

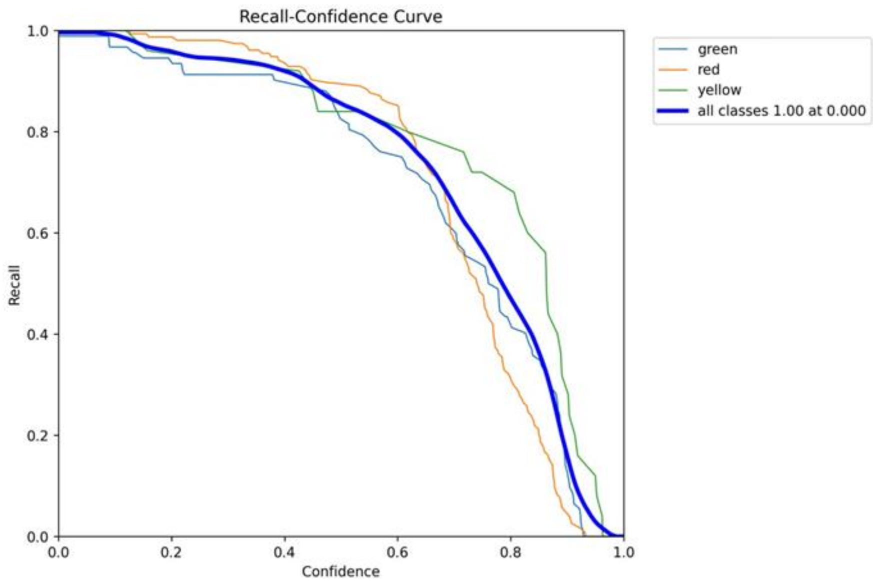

Figure 5. Recall-Confidence Graph

The results obtained when the proposed model is applied to real-life scenarios are presented in Figure 6. Accordingly, it is observed that the model's sensitivity to perceiving the color in question remains stable despite environmental factors.

Figure 6. Application of the Model to Real Life

3.2. Modeling by 80. Epoch

In the curve shown in Figure 7, the model's F1-Confidence performance at epoch 80 is evaluated in comparison to the previous results (epoch 40). Here, the highest F1 score for "all classes" (thick blue curve) increases from 0.94 to 0.96, and the peak (optimal threshold) shifts from 0.30 to 0.37, indicating that the model's overall accuracy level has improved and it achieves its best balance at a slightly higher confidence threshold. When examined on a class basis (blue: "green", orange: "red", green: "yellow"), it is observed that although the curves initially rise rapidly and then continue to decline, their peak points are higher or show a later decline compared to the 40th epoch. For example, the fact that the red (orange) line remained at a relatively high level around 0.8 suggests that the model may have learned color differentiation more effectively during the additional training process.

*Corresponding author, e-mail: seyfettinvadi@gazi.edu.tr
DOI: 10.17134/khosbd.1783706
Elektronik/Online ISSN: 2148-1776
Basılı/Printed ISSN: 1303-6831

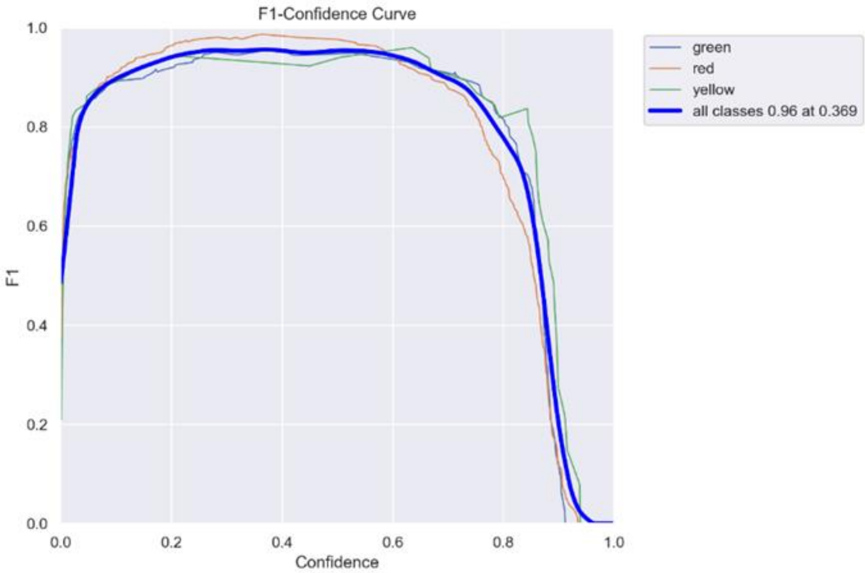

Figure 7. F-1 Confidence Graph (80. Epoch)

In conclusion, the F1-Confidence curve for the 80th epoch demonstrates a balanced Precision-Recall performance, exhibiting both a higher maximum F1 score and a slightly later decline compared to the previous epoch (40th Epoch), particularly at threshold values within the 0.3–0.4 range.

The Precision-Confidence curve for the 80th epoch in Figure 8 shows that the model achieved more stable and higher Precision levels, especially around high threshold values (e.g., 0.7–0.9), compared to the previous (40th epoch) results. While the "All classes" (thick blue line) curve is seen to fluctuate within the 95-100% range over a longer interval compared to the 40th epoch, a similar improvement is also noticeable for individual classes (blue: green, orange: red, green: yellow). For example, although the "red" (orange) class had a high Precision value in the previous period, it appears to have remained more stable in the 0.8-0.9 confidence region in the results of the 80th epoch. Similarly, the fact that the "green" and "yellow" classes are also positioned within the range of 0.5–0.8, at a higher Precision level compared to the 40th epoch, indicates that the model's ability to reduce the number of false positives improved during the additional training period.

Nevertheless, it should be noted that the optimal confidence threshold needs to be determined based on the level of false positives the application can tolerate, as a very high threshold could lead to a decrease in Recall (the rate of finding true positives).

*Corresponding author, e-mail: seyfettinvadi@gazi.edu.tr
DOI: 10.17134/khosbd.1783706
Elektronik/Online ISSN: 2148-1776
Basılı/Printed ISSN: 1303-6831

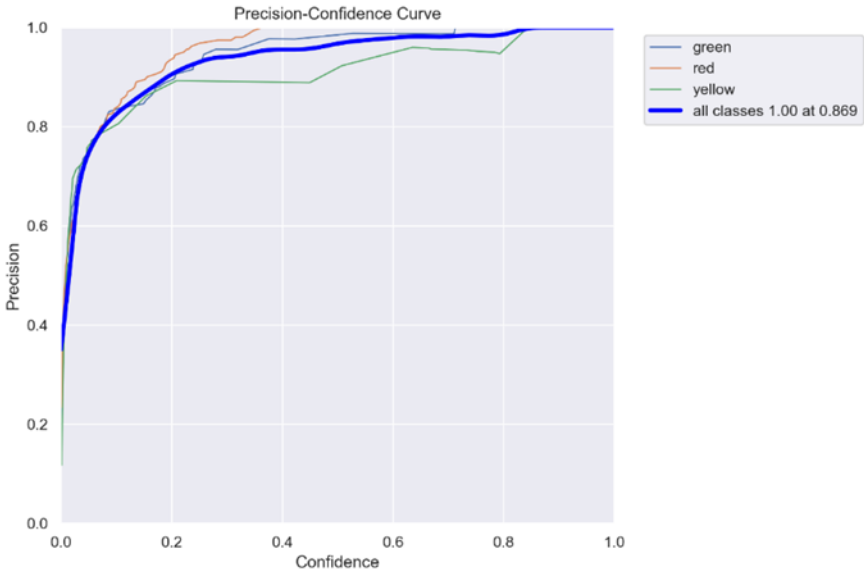

Figure 8. Precision-Confidence Graph (80th Epoch)

Figure 9 presents a comparison of the Precision-Recall performance obtained at epoch 80 with the data from epoch 40. Bu grafiğe bakıldığında, mAP@0.5 değerinin 0,978’den yaklaşık 0,982’ye yükselmesi, modelin genel doğruluk düzeyinde belirgin bir iyileşme olduğunu göstermektedir. Particularly noteworthy is the extremely high Precision-Recall level of the red class, which reached 0.994, while the green and yellow classes showed stable development, fluctuating between 0.96 and 0.98. However, the sudden drop in the curves at points where the Recall value approaches the range of 0.9–1.0 indicates that the number of false positives increases as the model labels more examples as "positive" to achieve high sensitivity.

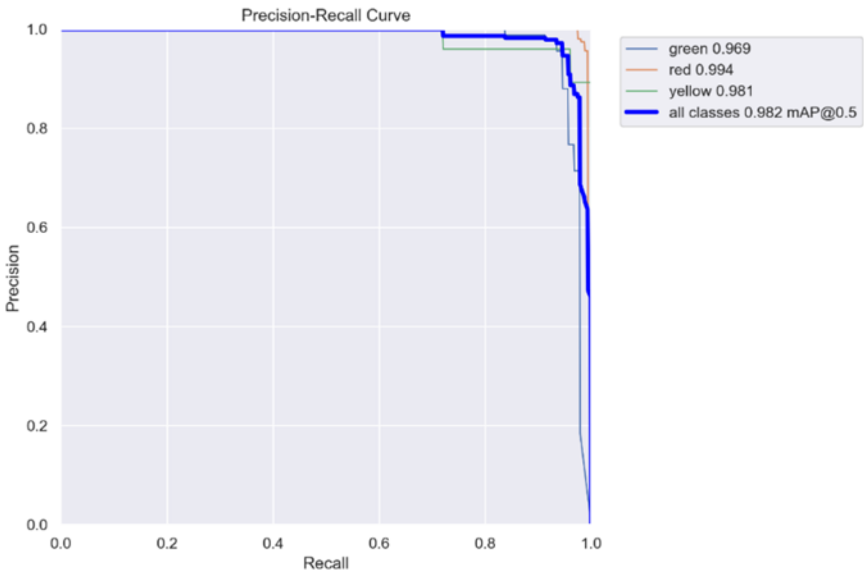

Figure 9. Precision-Recall Graph (80th Epoch)

*Corresponding author, e-mail: seyfettinvadi@gazi.edu.tr
DOI: 10.17134/khosbd.1783706
Elektronik/Online ISSN: 2148-1776
Basılı/Printed ISSN: 1303-6831

The data presented in Figure 10 for the 80th epoch shows that, compared to the previous results (40th epoch), the model was able to capture almost all positive examples within the 0.0–0.2 confidence interval (Recall ≈ 1.0). It is observed that as the threshold value increases (from 0.6 to 0.8), Recall gradually decreases, while the red (orange) and yellow (green) curves remain at high levels up to a threshold value of ~0.8. Similarly, green (blue) can also maintain a Recall value above 0.9, up to 0.6. This table shows that after the 80th epoch, the model is successful across a broader range in terms of "not missing positive examples," but experiences a sharp drop at thresholds above 0.8, similar to what was observed at the 40th epoch. Therefore, it can be said that the additional training process improves the model's recall performance at low and medium confidence thresholds, while at high threshold values, the "fewer positives" marking approach increases the risk of false negatives.

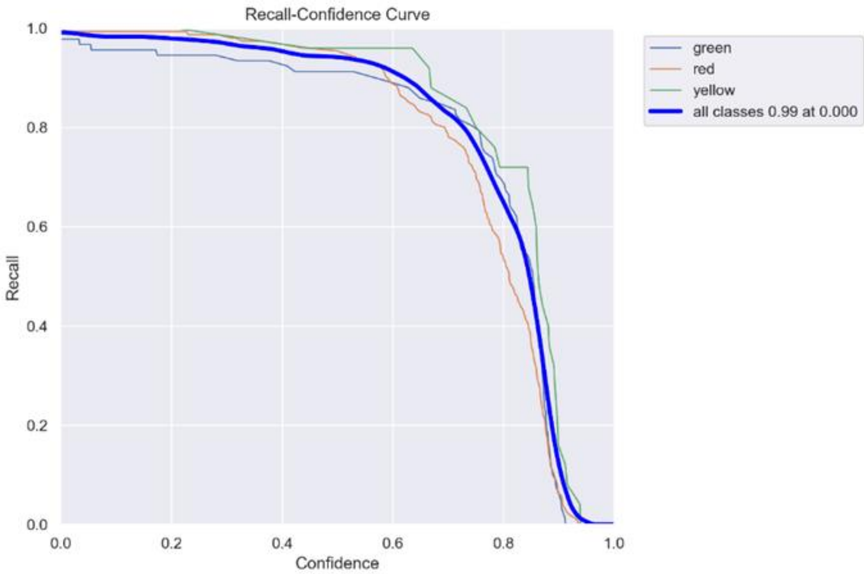

Figure 10. Recall-Confidence Graph (80th Epoch)

Figure 11 illustrates the changes in training and validation losses, as well as the basic performance metrics (Precision, Recall, mAP) of the model at the 80th epoch. Specifically, the train/box_loss, train/cls_loss, and train/dfl_loss graphs show a consistent decreasing trend from the first epochs, indicating that the model is learning increasingly well. The fact that these losses approached a plateau level by the end of the 80th epoch can be interpreted as the model's weights having largely stabilized, and that it was performing very little additional learning. It is also noteworthy that the validation losses (val/box_loss, val/cls_loss, val/dfl_loss) at the bottom of the image are almost zero. This situation could arise because the validation dataset is relatively easy for the model, or due to a difference in the visualization scale, resulting in losses being expressed with small values. Both possibilities suggest that the model needs to be evaluated with a different and more challenging test dataset to better understand its actual performance. The metric graphs in Figure 11 show that the model's Precision and Recall values

*Corresponding author, e-mail: seyfettinvadi@gazi.edu.tr
DOI: 10.17134/khosbd.1783706
Elektronik/Online ISSN: 2148-1776
Basılı/Printed ISSN: 1303-6831

significantly increased over time. The fact that Precision ranged between 0.9 and 1.0 indicates that the model significantly reduced its false positive rate. In contrast, the fact that Recall rose to around 0.8 suggests that the model correctly identified the vast majority of truly positive samples. Ayrıca mAP@0.5 değerinin 0,95–0,98 düzeylerine, mAP@0.5–0.95 değerinin ise 0,6–0,7 bandına ulaşması, modelin farklı IoU (Intersection over Union) eşiklerinde de tutarlı biçimde yüksek başarı gösterdiğini ortaya koymaktadır. However, the very low validation losses and the fact that the metrics are at relatively high levels indicate that there is not much difference in difficulty between the training data and the validation data for the model.

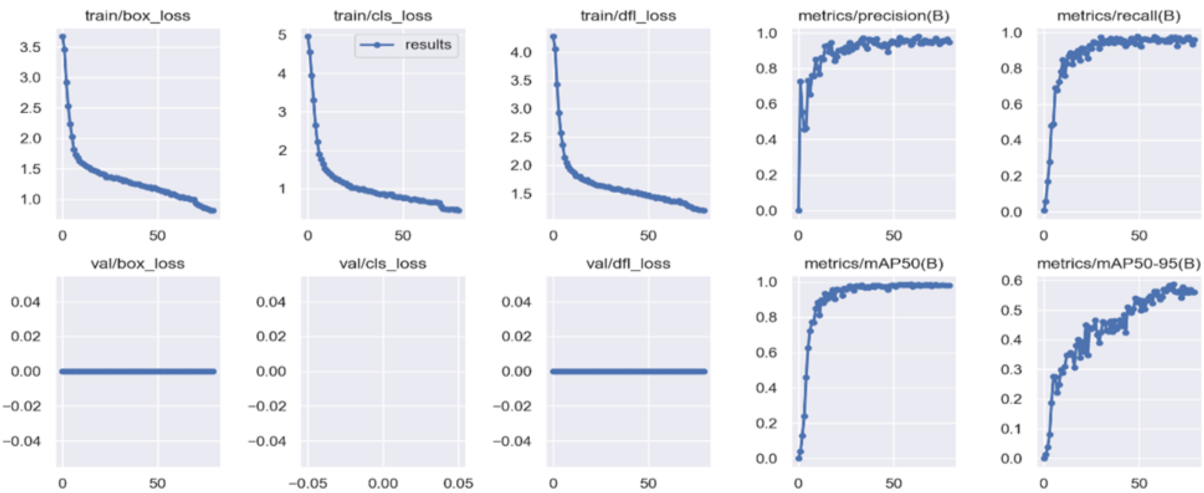

Figure 11. Changes in basic performance metrics with training and validation losses in 80. Epoch

4. ASSESSMENT

The application of the model in a real-world setting is shown in Figure 13. As can be seen from the results obtained from the field application, the model's generalization performance has been verified, and it is observed that the persistence and reliability of the results obtained have been increased.

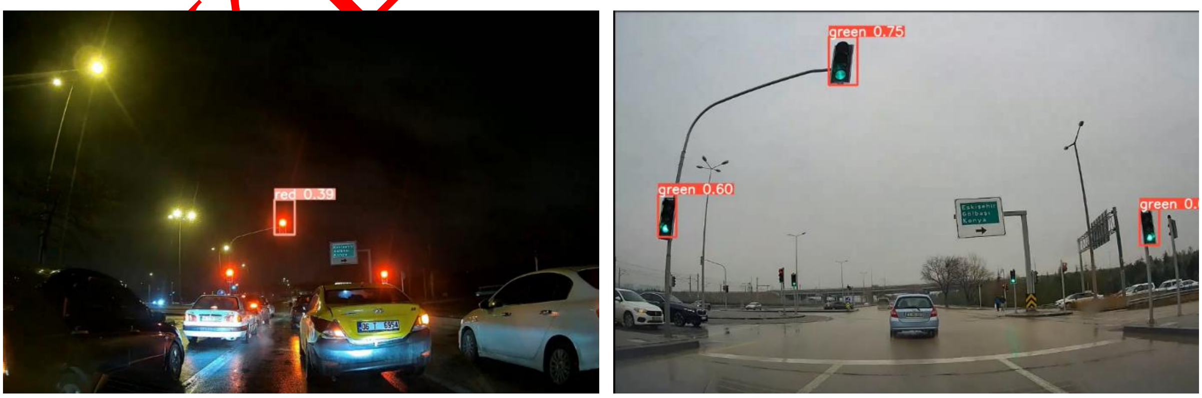

Figure 12. Testing the Model Under Different Conditions

Figure 13 shows that the complexity matrix data indicate significant differences when comparing the model's outputs at the 40th and 80th epochs. For example, the red class was already classified with 99% accuracy at the 40th epoch, and not only did it maintain this value at the 80th epoch, but also reduced the background mixing ratio (red → background) from approximately 64% to 36%. A similar improvement was observed in the Yellow class, with the correct recognition rate increasing from 0.92 to 0.96. In contrast, the confusion between green and background (green → background) appears to have risen from 0.29 at epoch 40 to 0.55 at epoch 80. This situation indicates that, while the overall accuracy of the model has increased, it is particularly prone to confusion with background objects in green light, suggesting the need for additional data diversification or improvements in model settings (e.g., threshold values). The results presented in Figure 13 confirm that the developed YOLOv8-based model is not only a high-accuracy laboratory output but also a practical tool that can be used in real-world applications such as ADAS or fully autonomous vehicle systems.

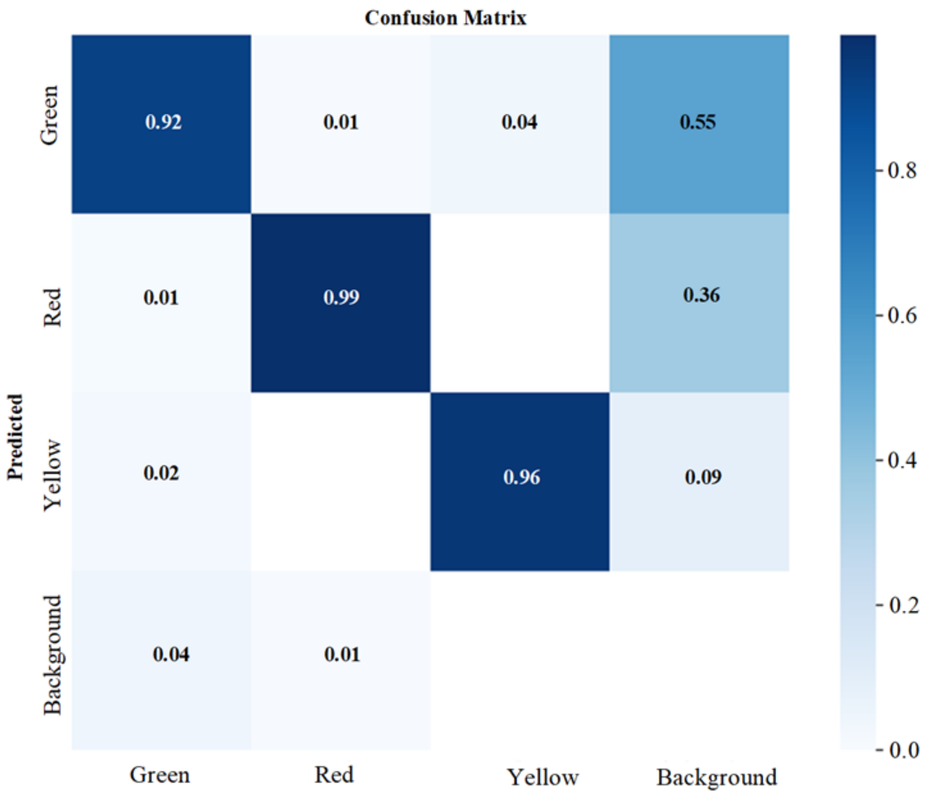

Figure 13. Confusion Matrix Table

When compared to similar studies in the literature, it is observed that detection rates are low and error rates are high [34]. In this study, the YOLOv8 algorithm, enhanced with preferred neural networks and deep learning techniques, was developed to detect traffic signs with high accuracy under different environmental conditions and viewing angles. To summarize the contribution of this study in comparison to other studies in the literature, it surpasses traditional traffic detection methods, providing a highly accurate, reliable, and practical solution for the autonomous systems ecosystem. In the study, organizing the directory structure ensured that the project was stored in a manageable manner, preventing the mixing of data, model parameters, and results, thereby facilitating transparent and reproducible analysis processes. Reproducibility is of great importance in clarifying which data individuals who will repeat the study will use in model training and performance measurement. The model's overfitting or underfitting conditions have been evaluated. The results obtained highlight the model's high generalization performance.

5. CONCLUSION

In this study, a YOLOv8-based approach was developed and comprehensively evaluated for the task of traffic light detection and color recognition. At the beginning of the research, the Ultralytics YOLOv8 library was installed in a GPU-supported environment, followed by the creation of a carefully prepared traffic light dataset to reflect different lighting and angle conditions.

Throughout the training process, performance metrics such as loss, accuracy, F1 score, and mAP of the model were monitored at regular intervals; by the 80th epoch, a high mAP value (≈0.98) and balanced Precision-Recall performance were achieved. It was observed that the model yielded near-perfect results, especially in detecting red lights. However, some false detections occurred due to background objects that were similar in color to green lights.

The findings indicate that YOLOv8 possesses strong capabilities in both spatial and color discrimination for traffic light detection and color classification tasks. However, more extensive field tests are needed to fully assess how well the model generalizes when faced with real-world diversity (such as night shots, different camera angles, and climate conditions). Future studies could enhance the model's generalization ability by increasing the dataset's diversity (e.g., traffic light designs from different countries or challenging weather conditions). Additionally, integrating the model with advanced techniques such as multi-output architectures or domain adaptation can provide valuable contributions, particularly in meeting the safety and accuracy requirements of autonomous vehicles. As part of future research, developing the necessary regulations and standards for the legal and social acceptance of autonomous vehicles is crucial. In this context, ethical problems and security issues must be addressed, and the system's transparency and accountability must be ensured. Additionally, focus can be placed on optimizing deep learning techniques and exploring new architectures to develop more successful algorithms. It is also critically important to create mechanisms that enable the system to detect such situations and hand control over to the driver in cases where the algorithm fails. These types of approaches not only increase the reliability of autonomous vehicles but also build user trust.

6. REFERENCES

[1] WHO (World Health Organization) (2018). Global Status Report on Road Safety 2018. World Health Organization, Geneva.

[2] Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A.A., Jarawan, E., & Mathers, C. (2004). World Report on Road Traffic Injury Prevention. World Health Organization, Geneva.

[3] European Commission (2022). EU Road Safety Policy. European Commission, Mobility and Transport.

[4] Smiley, A., Brookhuis, K., & De Waard, D. (1987). Human Behavior and Traffic Safety. In: Evans, L. & Schwing, R.C. (Eds.), Human Behavior and Traffic Safety (pp. 51–70). Springer, Boston, MA.

[5] Y. Zein, M. Darwiche, O. Mokhiamar, GPS tracking system for autonomous vehicles. Alexandria Engineering Journal (2018) 57, 3127–3137

[6] Zanchin, B. C., Adamshuk, R., Santos, M. M., & Collazos, K. S. (2017, October). On the instrumentation and classification of autonomous

cars. In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 2631-2636). IEEE

[7] Wakabayashi, D. (2019). Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam. Retrieved from https://www.nytimes.com/2018/03/19/technology/uber-driverlessfatality.html

[8] Chitradevi, B., ve P. Srimathi. "An overview on image processing techniques." International Journal of Innovative Research in Computer and Communication Engineering 2.11."(2014).

[9] Cowan, R. S. (1987). The consumption junction: A proposal for research strategies in the sociology of technology. The social construction of technological systems: New directions in the sociology and history of technology, 26180.

[10] Ozguner U., Acatman T. ve Redmil K., (2011), "Autonomous Ground Vehicles", USA, Artech House, 3-5.

[11] Pomerleau D.A., (1989), "ALVINN: An Autonomous Land Vehicle in a Neural Network", In Advances in Neural Information Processing Systems, Morgan Kaufmann, San Mateo.

[12] Çayıroğlu İ., Şimşir M., (2008), "PIC ve Step Motorla Sürülen Bir Mobil Robotun Uzaktan Kamera Sistemi ile Kontrolü", Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 24(1-2), 1-16.

[13] Fong P. E., Yusoff M. A., (2011), "Real-Time Control of Wi-Fi Surveillance Robot", Proceeding of the International Conference on Advanced Science Engineering and Information Technology, Malasia.

[14] Kuo G., Cheng C., Wu C., (2014), "Design and Implementation of a Remote Monitoring Cleaning Robot", International Automatic Control Conference, Taiwan.

[15] Espes D., Autret Y., Vareille J., Le Parc P., (2014), "Designing a Low-Cost Web-Controlled Mobile Robot for Home Monitoring", The Eighth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, Italy.

[16] Çetinkaya A., (2017), "Otonom Bir robotun bulanık kontrolör yaklaşımı ile konum kontrolü", Karatay Üniversitesi, YL Tezi.

[17]Ellahyani, A., Ansari, M. E., ve Jaafari, I. E. (2016). Traffic sign detection and recognition based on random forests. Applied Soft Computing, 46, 805-815.

[18]Stallkamp, J., Schlipsing, M., Salmen, J., ve Igel, C. (2012). Man vs. Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition. Neural networks, 323–332.

[19] Xu, X., Jin, J., Zhang, S., Zhang, L., Pu, S., ve Chen, Z. (2019). Smart data driven traffic sign detection method based on adaptive color threshold and shape symmetry. Future Generation Computer Systems, 94, 381-391.

[20] Arcos-García, A., Álvarez-García, J. A., ve Soria-Morillo, L. M. (2018). Evaluation of deep neural networks for traffic sign detection systems. Neurocomputing, 316, 332-344.

[21] Torres, L. T., Paixao, T. M., Berriel, R. F., De Souza, A. F., Badue, C., Sebe, N., ve Oliveira-Santos, T. (2019). Effortless Deep Training for Traffic Sign Detection Using Templates and Arbitrary Natural Images. International Joint Conference on Neural Networks. Budapest.

[22] Rahman, Q. M., Sünderhauf, N., ve Dayoub, F. (2019). Did You Miss the Sign? A False Negative Alarm System for Traffic Sign Detectors. arXiv:1903.06391.

[23] Mathias, M., Timofte, R., Benenson, R., ve Van Gool, L. (2013). Traffic sign recognition — How far are we from the solution? International Joint Conference on Neural Networks (IJCNN), (s. 1-8). Dallas.

[24] Jung, S., Lee, U., Jung, J., ve Shim, D. H. (2016). Real-time Traffic Sign Recognition System with Deep Convolutional Neural Network. 13th International Conference on Ubiquitous Robots and Ambient Intelligence. Urai.

[25] Li, Y., Møgelmose, A., ve Trivedi, M. M. (2016). Pushing the "Speed Limit": HighAccuracy U.S. Traffic Sign Recognition with Convolutional Neural Networks. IEEE Transactions on Intelligent Vehicles, 167-176.

[26] Abedin, Z., Dhar, P., Hossenand, M. K., ve Deb, K. (2017). Traffic Sign Detection and Recognition Using Fuzzy Segmentation Approach and Artificial Neural Network Classifier Respectively. International Conference on Electrical, Computer and Communication Engineering, (s. 518-523). Cox's Bazar.

[27] Song, S., Que, Z., Hou, J., Du, S., ve Song, Y. (2019). An efficient convolutional neural network for small traffic sign detection. Journal of Systems Architecture, 97, 269-277.

[28] Alpkıray N., (2019), "Otonom Keşif Amaçlı Robot Sistemleri İçin Geri dönüş Rotası Hesaplama Algoritması Geliştirilmesi", Cumhuriyet Üniversitesi, YL Tezi.

[29] Lin, P., Abney, K., & Jenkins, R. (Eds.). (2017). Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence. Oxford University Press.

[30] Notomista G., Ve Botsch M., (2016), "A Machıne Learnıng Approach For The Segmentation Of Driving Maneuvers And Its Application In Autonomous Parking", JAISCR, 2017, Vol. 7, No. 4, pp. 243 – 255.

[31] Zhang P., Xiong L., Yu Z., Fang P., Yan S., Yao S., Zhou Y., (2019), "Reinforcement Learning-Based End-to-End Parking for Automatic Parking System", sensors 2019, 19, 3996.

[32] Lin, P., Abney, K., & Jenkins, R. (Eds.). (2017). Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence. Oxford University Press.

[33] Moon J., Bae I., Kim S., (2019), "Automatic Parking Controller with a Twin Artificial Neural Network Architecture", Hindawi 2019, Article ID 4801985.

[34] Ayachi, R., Afif, M., Said Y., and Abdelali, A. B, "Traffic Sign Recognition Based On Scaled Convolutional Neural Network For Advanced Driver Assistance System," 2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS), Genova, Italy, 2020, pp. 149-154.