VOLUME 18, N° 4 2024
Journal of Automation, Mobile Robotics and Intelligent Systems

Efficient Vehicle Detection and Classification Algorithm Using Faster R-CNN Models

Submitted: 16th March 2022; accepted: 3rd April 2024

Imad EL Mallahi, Jamal RIFFI, Hamid Tairi, Mohamed Adnane Mahraz

DOI: 10.14313/JAMRIS/4-2024/33

Abstract:
This study proposes an integrated framework for efficient traffic object detection and classification by leveraging advanced deep-learning techniques. The framework begins with the input of video surveillance, followed by an image-acquisition process to extract the relevant frames. Subsequently, a Faster R-CNN (ResNet-152) architecture was employed for precise object detection within the extracted frames. The detected objects are then classified using deep reinforcement learning, specifically trained to identify distinct traffic entities, such as buses, cars, trams, trolleybuses, and vans. The UA-DETRAC dataset served as the primary data source for training and evaluation, ensuring the model's adaptability to real-world traffic scenarios. Finally, the performance of the framework was assessed using key metrics, including precision, recall, and F1 score, providing insights into its effectiveness in accurately detecting and classifying traffic objects. This integrated approach offers a promising solution to enhance traffic surveillance systems and facilitate improved traffic management and safety measures in urban environments.

Keywords: *classification, deep learning, vehicle detection video surveillance, traffic estimation, efficient R-CNN, traffic detection, convolutional neural network*

1. Introduction

In recent years, both industry and academia have witnessed significant advancements in vehicle-detection technologies [1]. However, many state-of-the-art-image identification algorithms have struggled to meet stringent standards for vehicle detection. The primary challenges in automobile identification include substantial differences in object size, significant occlusions, and fluctuations in lighting conditions [2–5].

Sensor-based algorithms have been employed to address certain surveillance tasks in urban traffic systems, including vehicle counting, license plate identification, incident detection, driver facial emotion identification, and Internet of Things (IoT) source location and identification. Conversely, vision-based approaches offer the advantage of leveraging visual patterns to distinguish target objects in a manner similar to human perception.

While radar-sensor-based methods may be limited to relatively small areas, vision-based systems utilizing cameras can detect vehicles across vast regions and provide additional information about each detected vehicle simultaneously [6].

Consequently, researchers have extensively explored various computer vision and machine learning models to address a range of challenges in intelligent transportation systems. Classic vehicle detection algorithms, from early developments to contemporary approaches, have often relied on handcrafted features such as Haar-like features and histogram of oriented gradient (HOG) features. Notably, the cascaded detector has emerged as a pioneering real-time detection system, while methods such as Support Vector Machines (SVM) and deformable part-based models (DPM) have addressed issues such as heavy occlusion and significant variations in object sizes [7–13].

To address challenges, such as light variance, occlusion, and size variations, researchers have proposed innovative solutions. For instance, a strong CNN model was developed for traffic light recognition, and strategies for handling heavy occlusions caused by fixed surveillance cameras were devised [14–16]. In addition, a scale-aware Region Proposal Network (RPN) has been introduced to effectively identify vehicles of various sizes effectively [17, 18].

Among the various techniques explored, region-based convolutional neural networks (R-CNNs) have demonstrated promising performance in vehicle identification. The Faster R-CNN (Resnet-152), which integrates a Region Proposal Network (RPN), has shown competitive performance in object recognition tasks. Notably, recent works have reported impressive mean average accuracy on benchmark datasets such as KITTI and COCO, showcasing the effectiveness of Faster R-CNN (Resnet-152) in vehicle detection [19–22].

Despite its competitive performance, there remains room for improvement in Faster R-CNN (Resnet-152), particularly in addressing wide-scale variance in vehicle detection. Accurate classification of vehicles into distinct categories is crucial not only for transportation management but also for efficient damage detection in insurance solutions. Therefore, there is a pressing need for advancements in automatic damage assessment procedures to mitigate work accidents and ensure a comprehensive vehicle assessment [23–25].

Currently, most cities worldwide have several video surveillance systems [26–30]. They have grown rapidly, and now they have heterogeneous cameras with various resolutions [1]. Today, closed-circuit television works all times a day and week to produce a huge amount of data, mainly big and huge frames of videos. The data visualization presented in Figure 1 offers a comprehensive view of the input from video surveillance systems, enabling analysts to effectively monitor, analyze, and interpret activities captured within the monitored environment.

Faster R-CNN, an abbreviation for "Region-based Convolutional Neural Network," marks a significant advancement in computer vision's object detection capabilities. It introduces a streamlined approach, condensing object localization and classification into a single-step process. This breakthrough simplifies detection tasks, enabling a Faster R-CNN to swiftly analyze images and videos with exceptional accuracy. Additionally, the evolution of this technology, exemplified by YOLOv8, continues to push the boundaries between object detection and image segmentation. YOLO's real-time performance distinguishes it as the top choice for applications demanding both rapid and precise object identification.

The primary focus of this research is the detection and classification of traffic, which is a pivotal aspect of urban traffic management systems. The objective is to precisely categorize vehicles within defined regions in every frame to accurately evaluate traffic density. This critical information plays a key role in recognizing peak traffic times and congested areas and contributes significantly to urban planning efforts. Through this endeavor, our aim is to construct an extensive toolkit capable of offering nuanced analyses of traffic flow and trends, thereby bolstering traffic management strategies and urban planning initiatives. Figure 2 shows the structured classification and categorization of deep learning methods employed in vehicle detection and classification tasks. The methods are organized into several categories based on their underlying techniques and approaches.

Moreover, this video data can be used as a principal for automated vehicle control systems. There are many issues when working with Big Data traffic surveillance. To realize an intelligent system for traffic vehicle surveillance, we have an efficient hard disk system for storing, moving forward and back and analyzing videos. In this study, we focused on analyzing videos for traffic surveillance, which remains limited in terms of real-time data analysis.

Some representative papers in this theme used heterogeneous low-resolution data to estimate traffic density and count vehicles.

Many efforts have been made to provide vehicle analysis, counting, classification, detection, and satisfactory results in specific tasks. The remainder of this paper is organized as follows.

Section 1 introduces related studies on traffic classiffication and dominant feature selection methods. In Section 2, we discuss traffic classiffication and dominant feature selection methods. Section 3 presents the methodology used in this study. Section 4 presents experimental results and performance evaluation. Finally, in Section 5, we provide concluding remarks.

2. Classification and Recognition of Traffic Video

2.1. Related Works

Detecting and categorizing objects within video footage represents a critical challenge in the development of autonomous surveillance systems. Numerous algorithms have been proposed to address this challenge, ranging from background-subtraction-based techniques to classifier-based methods for object detection and classification in videos. Each approach to this problem presents its own set of advantages and drawbacks, necessitating careful consideration of the algorithm type best suited to the specific task at hand [31, 32].

Background subtraction methods involve the detection of new objects in an image that are absent from a reference background image. The fundamental principle involves subtracting a new image containing multiple objects to be detected from the reference image to yield a difference-encoded image. A threshold value was applied to enhance the background subtraction's tolerance to potential noise within the video. Subsequently, a blob detector was employed to identify and count objects. Each identified blob was then treated as a single object and subjected to classification algorithms for refinement. A more intricate background subtraction method based on a mixture of Gaussian (MoG) models, capable of detecting not only foreground object pixels but also the shadows they cast, has been proposed in previous studies. This MoG-based approach has also been utilized to detect human movements in videos [33–36].

The detection of specific objects within images poses challenges due to variations in object size, orientation, and instances of overlapping objects that cause occlusions. Addressing these challenges requires a detection algorithm that possesses certain properties including translation, rotation, and scale invariance. Typically, machine learning methods are employed to learn representations directly from available data to train models. Popular approaches involve utilizing low-level features, such as Scale-Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG), and Haar-like features, combined with machine learning techniques for object classification. This methodology is commonly referred to as the "Feature + Classifier" approach. For tracking or object detection in video surveillance, efficient convolutional neural networks (FCNN), are divided into two principal classes: single-phase captors and second phase.

Figure 1. Data visualization

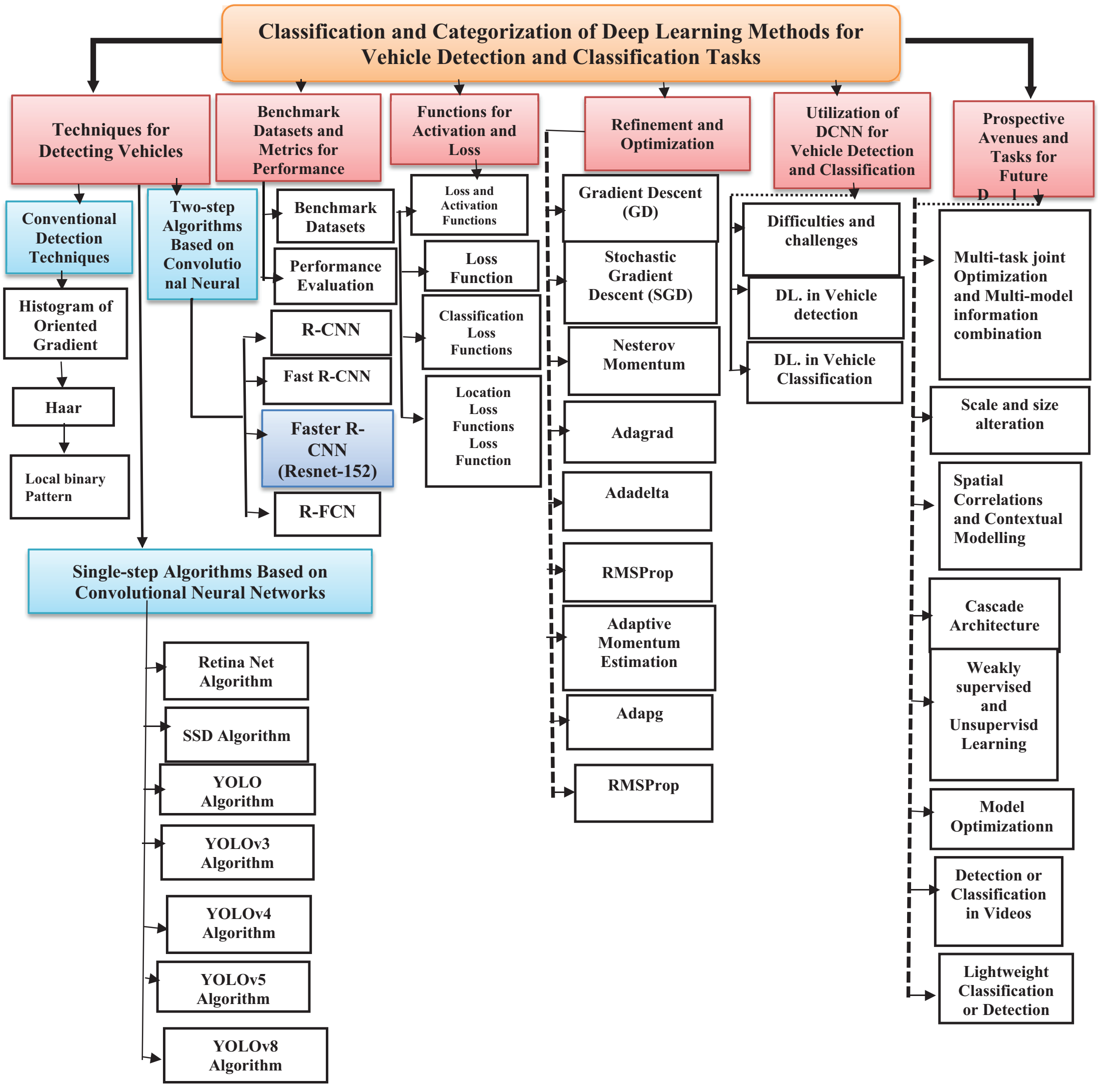

Figure 2. Classification and categorization of deep learning methods for vehicle detection and classification tasks

2.2. Objects Detection in Video Surveillance

The primary-phase captor is typically very fast and can be used to predict objects in video skipping boxes with classes within a simple network. The conventional experiments of the first captor were SSD and YOLO. In addition, the application of good speed-precision trade-of is version 2 for the YOLO method, particularly for vehicle detection via anchor clustering, multi-layer feature fusion strategy, and loss normalization. The R-CNN family of captors is used in many representative two-stage captor methods [37].

The comparison between the first- and second-stage captors began with the prediction of regions, and then classified and refined each of them during the two stages.

The first study on R-CNN employed a simple approach: a region was generated with many selective research algorithms and then the recognition the object was implemented. The global Vitesse of the efficient R-CNN is feasible for computing the time for each region. To resolve this issue, we propose a Fast R-CNN. Instead of executing a CNN for each region. Modifying selective research in the proposed Efficient R-CNN with this method, called region-suggested network, also computes the accuracy and Vitesse of the captor [38, 39].

3. Research Methodology

The process commences with the extraction of frames from the input video stream and is subsequently forwarded to the Faster R-CNN model for detection. Faster R-CNN encompasses two distinct models within its architecture. Initially, a region proposal (RP) augmented with Deformable Convolutional Networks (DCN) is executed to generate potential object regions. This stage efficiently identifies candidate regions within frames that are likely to contain the objects of interest. Subsequently, in the second model, a Fast R-CNN was employed to conduct object detection within the proposed regions, leveraging the refined region proposals obtained from the preceding stage. The architecture of Faster R-CNN, delineating the integration of region proposal and object detection stages, is shown in Figure 3, offering a visual representation of the model's intricate design and operational flow.

The workflow of the detection and classification models is shown in Figure 3. The diagram illustrates the initial conversion of input surveillance videos into a sequence of frames, followed by the application of the Faster R-CNN model for detection of each frame individually.

3.1. Dataset of Detection of the Video Surveillance

In this application, the number of surveillance cameras is more than forty surveillance cameras, and most surveillance cameras are fixed. In this dataset, most cameras had 25 frames per second, with a resolution of 960 × 540. Moreover, the streams of these videos are not very good because of hardware faults, blurring, and compression artifacts. Figure 1 shows the video from one of these cameras, and Figure 4 shows some experimentations cameras of UA-DETRAC dataset. Algorithm 1 leverages the efficiency of the Faster R-CNN framework, enhanced by the ResNet-152 backbone, to achieve state-of-the-art performance in object-detection tasks. Owing to its intricate architecture and feature extraction capabilities, it enables the accurate and efficient detection of objects in diverse real-world scenarios.

We focused on datasets such as UA-DETRAC and KITTI, noting key differences across various aspects: the total occurrences of the same grid, significant scale variations, viewing angles, and occlusion levels. The primary challenge with this dataset lies in the images themselves, as illustrated in Figure 4. Our study utilized a small camera monitoring one of the many traffic types, aiming to achieve optimal detection accuracy while providing a viable alternative to fixed cameras. To this end, we developed specific solutions and annotated over 60,000 polygons across 982 images from the selected camera, using the COCO Annotator utility [14]. Annotating video sequences can be highly time-consuming, especially in crowded scenes; therefore, we focused on traffic scenarios, weather conditions, and different times of the day. Additionally, we meticulously annotated individual vehicles with a high degree of confidence, particularly in dense traffic conditions. Table 1 provides an overview of the dataset distribution.

Algorithm 1. Object detection using Faster R-CNN (ResNet-152).

1. **Input Video Surveillance Data:**
   - Receive input from video surveillance systems that capture traffic scenes.
2. **Image Acquisition and Frame Extraction:**
   - Extract the relevant frames from the input video.
3. **Object Detection Using Faster R-CNN (ResNet-152):**
   - Employ Faster R-CNN (ResNet-152) architecture for precise object detection within extracted frames.
   - Detect and localize traffic objects, such as buses, cars, trams, trolleybuses, and vans.
4. **Object Classification using Deep Reinforcement Learning:**
   - Utilize deep reinforcement learning techniques for object classification.
   - Train the model to identify distinct traffic entities including buses, cars, trams, trolleybuses, and vans.
5. **Dataset Utilization:**
   - Utilize the UA-DETRAC dataset as the primary data source for training and evaluation.
   - Ensure adaptability of the model to various real-world traffic scenarios present in the dataset.
6. **Performance Evaluation:**
   - Assess the performance of the framework using key metrics such as precision, recall, and F1 score.
   - Evaluate the effectiveness of the framework in accurately detecting and classifying traffic objects.

   Output:
   - Generate insights into the effectiveness of the integrated approach for traffic surveillance.
   - Provide a promising solution for enhancing traffic management and safety measures in urban environments.

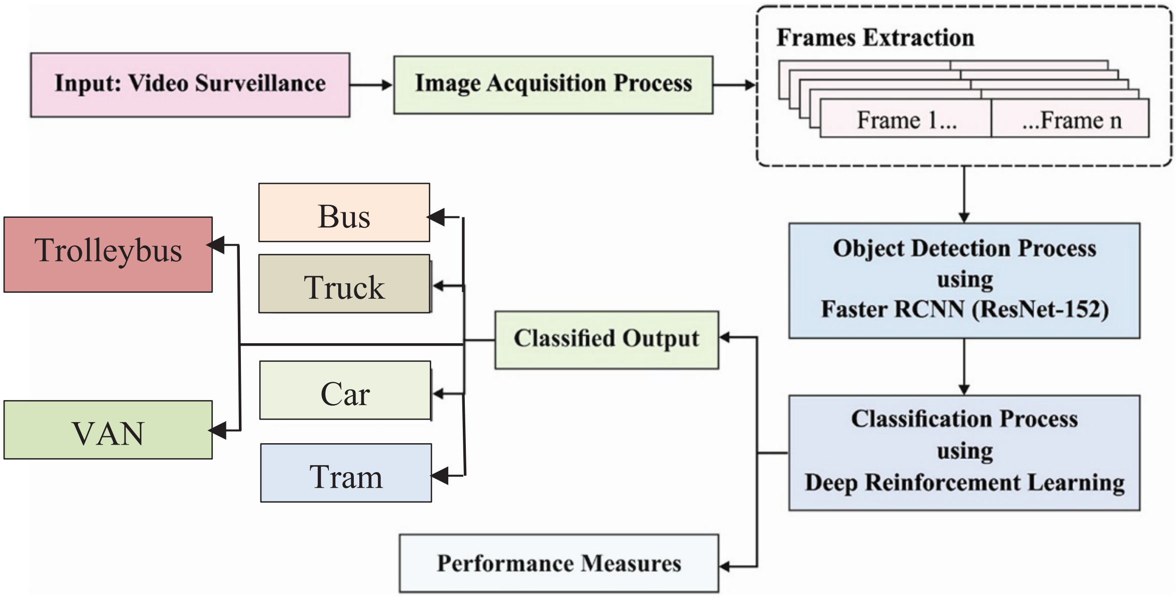

Figure 3. Intelligent video anomaly detection and classification using Faster R-CNN

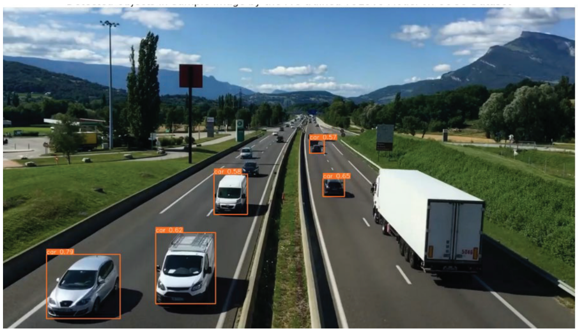

Figure 4. Some experimentations cameras of UA-DETRAC dataset

Table 1. Distribution of dataset

Type of vehicle	Number of instances	Mean instance per frame
Bus	1,234	1.26
Truck	2,415	2.46
Car	53,083	4.06
Trolleybus	611	0.62
TRAM	1,298	1,298
VAN	2,783	2.83

4. Experimentation

4.1. Numerical Result and Experimentations

To ensure robust evaluation, we utilized the publicly available UA-DETRAC dataset to construct our training dataset. This dataset has been extensively employed in previous studies focusing on traffic classification tasks, thus facilitating comparability with existing research outcomes. Despite the rich diversity of images contained within the UA-DETRAC dataset, the number of labeled instances per frame, which is crucial for training deep-learning models, is insufficient.

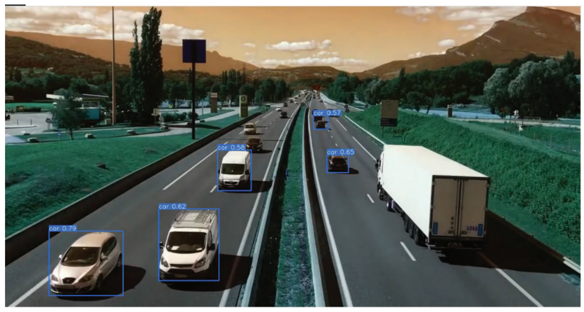

Figure 5. Object detection using Faster R-CNN

In response to this limitation, we augmented the training data by gathering supplementary UA-DETRAC data, thereby enriching our dataset with a more comprehensive representation of traffic flows and scenarios.

In this experiment, as shown in Figure 5, we performed object detection using Faster R-CNN on an HP EliteBook ×360 1130 G5 running Windows 11. The setup was based on a Python 3.8 environment, utilizing an Intel Core i7—8400 2.80 GHz CPU and an Nvidia GeForce RTX 2160 GPU with 8 GB of memory. We employed Darknet-53 as the convolutional neural network within the deep learning framework.

Figure 6 illustrates several experiments in object detection using Faster R-CNN. The confusion matrix, depicted in the figure, is an essential tool for evaluating classification performance, offering insights into the accuracy of the model's predictions. It quantifies correct and incorrect classifications across various classes and enables the calculation of key metrics, including true positives, true negatives, false positives, and false negatives. As illustrated in Figure 7, the confusion matrix delineates the classification outcomes for each class.

Figure 6. Some experimentations of object detection using Faster R-CNN

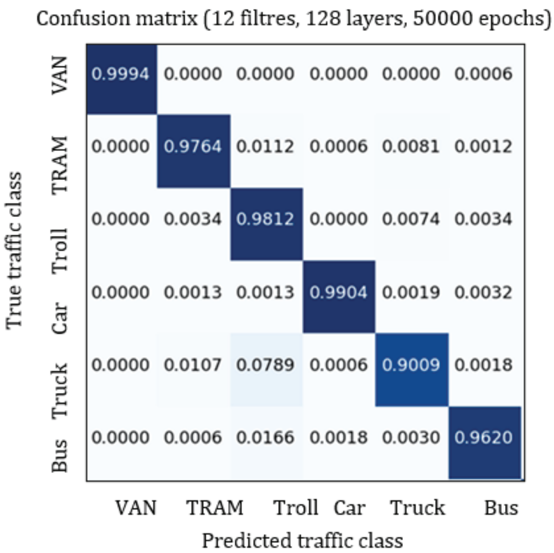

Figure 7. Traffic class confusion matrix: buses, cars, trams, trolleybuses, and vans

Notably, the model achieved an impressive overall classification accuracy of approximately 96.67%. Furthermore, by leveraging the information from the confusion matrix, metrics such as precision, recall, and F1-score for each service can be meticulously calculated, as shown in Figure 8. These metrics offer a comprehensive evaluation of the model's performance, capturing its effectiveness in correctly identifying instances belonging to specific classes.

5. Conclusion

This study presents an integrated framework that combines Faster R-CNN and deep reinforcement learning for traffic object detection and classification. By sequentially processing video surveillance data, extracting frames, and utilizing advanced deep learning techniques, the framework demonstrated promising results in accurately identifying and categorizing various traffic entities.

The utilization of Faster R-CNN (ResNet-152) ensures robust object detection capabilities, enabling the system to localize and identify traffic objects with high precision. Subsequently, deep reinforcement learning enhances the classification process,

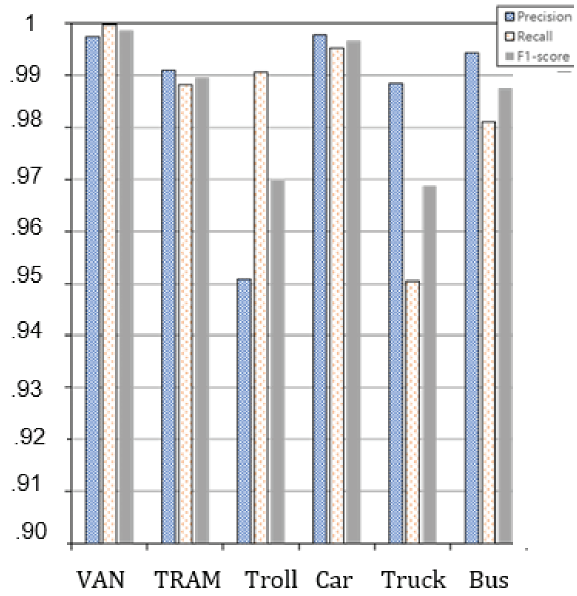

Figure 8. Precision, recall, and F1-score analysis for traffic classes: buses, cars, trams, trolleybuses, and vans

enabling the model to autonomously learn and classify objects, such as buses, cars, trams, trolleybuses, and vans.

The framework's performance was evaluated using key metrics such as precision, recall, and F1 score, providing comprehensive insights into its effectiveness. Through experimentation with the UA-DETRAC dataset, the framework shows its adaptability to diverse traffic scenarios, highlighting its potential for real-world applications.

AUTHORS

Imad EL Mallahi* – https://orcid.org/0000-0001-9735-6799 Sidi Mohamed Ben Abdellah University, Faculty of Sciences, Department of Computer Sciences, Fez, Morocco, e-mail: imade.elmallahi@usmba.ac.ma.

Jamal RIFFI – https://orcid.org/0000-0003-0818-7706 Sidi Mohamed Ben Abdellah University, Faculty of Sciences, Department of Computer Sciences, Fez, Morocco, e-mail: jamal.riffi@usmba.ac.ma.

Hamid Tairi – https://orcid.org/0000-0002-5445-0037 Sidi Mohamed Ben Abdellah University, Faculty of Sciences, Department of Computer Sciences, Fez, Morocco, e-mail: hamid.tahiri@usmba.ac.ma.

Mohamed Adnane Mahraz – https://orcid.org/0000-0002-0966-9654 Sidi Mohamed Ben Abdellah University, Faculty of Sciences, Department of Computer Sciences, Fez, Morocco, e-mail: hamid.tairi@usmba.ac.ma.

*Corresponding author

ACKNOWLEDGEMENTS

We gratefully acknowledge the support provided by the LISAC Laboratory, Department of Computer Sciences, Faculty of Science, Sidi Mohamed Ben Abdellah

University of Fez, Morocco, for their assistance in facilitating this study. Their support was instrumental in the successful completion of this work.

References

[1] R. Shafique et al., "Advancing Autonomous Vehicle Safety: Machine Learning to Predict Sensor-Related Accident Severity", *IEEE Access*, vol. 12, 2024, pp. 25933–25948.

[2] S. Ahn et al., "Explaining Deep Learning-Based Traffic Classification Using a Genetic Algorithm," *IEEE Access*, vol. 9, 2021, pp. 4738–4751; doi: 10.1109/ACCESS.2020.3048348.

[3] M.A. Berwo et al., "Deep Learning Techniques for Vehicle Detection and Classification from Images/Videos: A Survey" *Sensors*, vol. 23, no. 10, 2023: 4832; doi: 10.3390/s23104832.

[4] Y. Ghasemi et al., "Deep Learning-Based Object Detection In Augmented Reality: A Systematic Review," *Computers in Industry*, vol. 139, 2022, p. 103661, ISSN 0166-3615; doi: 10.1016/j.compind.2022.103661.

[5] X. Zhang et al., "Video Anomaly Detection And Localization Using Motion-Field Shape Description And Homogeneity Testing," *Pattern Recognition*, vol. 105, 2020, p. 107394.

[6] S Veluchamy, L R Karlmarx, K Michael Mahesh, Detection and Localization of Abnormalities in Surveillance Video Using Timerider-Based Neural Network, The Computer Journal, Volume 64, Issue 12, December 2021, Pages 1886–1906, https://doi.org/10.1093/comjnl/bxab002.

[7] Y. Fan et al., "Video Anomaly Detection and Localization via Gaussian Mixture Fully Convolutional Variational Autoencoder," *Computer Vision and Image Understanding*, vol. 195, 2020, p. 102920.

[8] A. Alam et al., "Intellibvr-Intelligent Large-Scale Video Retrieval for Objects and Events Utilizing Distributed Deep-Learning and Semantic Approaches," *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, IEEE, 2020, February, pp. 28–35.

[9] S. Liu and J. Tang, "Modified Deep Reinforcement Learning with Efficient Convolution Feature for Small Target Detection in Vhr Remote Sensing Imagery," *ISPRS International Journal of Geo-Information*, vol. 10, no. 3, 2021, p. 170.

[10] S.K. Lakshmanaprabu et al., "Online Clinical Decision Support System using Optimal Deep Neural Networks," *Applied Soft Computing*, vol. 81, 2019, p. 105487.

[11] Uthayakumar, J., Metawa, N., Shankar, K. et al. RETRACTED ARTICLE: Intelligent hybrid model for financial crisis prediction using machine learning techniques. Inf Syst E-Bus Manage 18, 617–645 (2020). https://doi.org/10.1007/s10257-018-0388-9.

[12] R. Hinami, T. Mei, and S. Satoh, "Joint Detection and Recounting of Abnormal Events by Learning Deep Generic Knowledge," *IEEE International Conference on Computer Vision (ICCV)*, 2017.

[13] R.T. Ionescu et al., "Object-Centric Auto-Encoders and Dummy Anomalies for Abnormal Event Detection in Video," *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[14] W. Luo, W. Liu, and S. Gao, "A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework," *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 341–349.

[15] M. Sabokrou, M. Fathy, and M. Hoseini, "Video Anomaly Detection and Localization Based on the Sparsity and Reconstruction Error of Auto-Encoder," *Electronics Letters*, vol. 52, no. 13, 2016, pp. 1122–1124.

[16] M. Ravanbakhsh et al., "Plug-and-Play CNN for Crowd Motion Analysis: An Application in Anomalous Event Detection," *WACV*, 2017.

[17] M. Sabokrou et al., "Deepanomaly: Fully Convolutional Neural Network for Fast Anomaly Detection in Crowded Scenes," *Computer Vision and Image Understanding*, vol. 172, 2018, pp. 88–97.

[18] M. Hasan et al., "Learning Temporal Regularity in Video Sequences," *CVPR*, 2016.

[19] D. Xu et al., "Learning Deep Representations of Appearance and Motion for Anomalous Event Detection," *BMVC*, 2015, pp. 1–12.

[20] M. Bellver et al., "Hierarchical Object Detection with Deep Reinforcement Learning," *Proceedings of the Conference on Neural Information Processing Systems*, Barcelona, Spain, December 2016, pp. 5–20.

[21] X. Kong et al., "Collaborative Deep Reinforcement Learning for Joint Object Search," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA 21–26 July 2017, pp. 7072–7081.

[22] B. Uzkent, C. Yeh, and S. Ermon, "Efficient Object Detection in Large Images using Deep Reinforcement Learning," *Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision*, Snowmass Village, CO, USA 1–5 March 2020, pp. 1824–1833.

[23] S. Liu, D. Huang, and Y. Wang, "Pay Attention to Them: Deep Reinforcement Learning-Based Cascade Object Detection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, 2020, pp. 2544–2556.

[24] S. Ren et al., "Faster RCNN (Resnet-152): Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, 2015, pp. 1137–1149.

[25] A.A. Micheal and K. Vani, "Automatic Object Tracking in Optimized UAV Video," *Journal of Supercomputing*, vol. 75, no. 8, 2019, pp. 4986–4999.

[26] X. Lei and Z. Sui, "Intelligent Fault Detection of High Voltage Line Based on the Faster RCNN," *Measurement*, vol. 138, 2019, pp. 379–385.

[27] Y. Ding et al., "Intelligent Fault Diagnosis for Rotating Machinery using Deep Q-Network Based Health State Classification: A Deep Reinforcement Learning Approach," *Advanced Engineering Informatics*, vol. 42, 2019, p. 100977.

[28] B.S. Murugan et al., "Region-Based Scalable Smart System for Anomaly Detection in Pedestrian Walkways," *Computers and Electrical Engineering*, vol. 75, 2019, pp. 146–160.

[29] L. Zadeh, "Probability Measures of Fuzzy Events", *Journal of Mathematical Analysis and Applications*, vol. 23, no. 2, 1968, pp. 421–427.

[30] I. El Mallahi et al., "A Distributed Big Data Analytics Models for Traffic Accidents Classification and Recognition Based SparkMlLib Cores," *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 16, no. 4, 2023, pp. 62–71. doi: 10.14313/JAMRIS/4-2022/34.

[31] R. Shafique et al., "Advancing Autonomous Vehicle Safety: Machine Learning to Predict Sensor-Related Accident Severity", *IEEE Access*, vol.12, 2024, pp. 25933–25948.

[32] N. Sohaee and S. Bohluli, "Nonlinear Analysis of the Effects of Socioeconomic, Demographic, and Technological Factors on the Number of Fatal Traffic Accidents", *Safety*, vol.10, no.1, 2024, p. 11.

[33] I.E. Mallahi et al., "Prediction of Traffic Accidents using Random Forest Model," *2022 International Conference on Intelligent Systems and Computer Vision (ISCV)*, Fez, Morocco, 2022, pp. 1–7; doi: 10.1109/ISCV54655.2022.9806099.

[34] A. Nasry, A. Ezzahout, and F. Omary, "People Tracking in Video Surveillance Systems Based on Artificial Intelligence," *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 17, no. 1, 2023, 59–68; doi: 10.14313/JAMRIS/1-2023/8.

[35] A. Ndayikengurukiye et al., "Resource Optimisation in Cloud Computing: Comparative Study of Algorithms Applied to Recommendations in a Big Data Analysis Architecture," Journal of Automation, Mobile Robotics and Intelligent Systems, vol. 15, no. 4, 2022, pp. 65–75; doi: 10.14313/JAMRIS/4-2021/28.

[36] A. Nasry, A. Ezzahout, and F. Omary, "People Tracking in Video Surveillance Systems Based on Artificial Intelligence," *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 17, no. 1, 2023, pp. 59–68; doi: 10.14313/JAMRIS/1-2023/8.

[37] N. Ferhane et al., "Data Driven Decision Making with Intelligent CCTV," In: M. Ezziyyani, J. Kacprzyk, V.E. Balas, eds., *International Conference on Advanced Intelligent Systems for Sustainable Development (AI2SD'2023). AI2SD 2023*. Lecture Notes in Networks and Systems, vol. 905. Cham: Springer, 2024; doi: 10.1007/978-3-031-52385-4_154.

[38] A. Ezzahoutz, H.M. Youssef and R.O.H. Thami, "Detection Evaluation and Testing Region Incoming People's in a Simple Camera View," *Second International Conference on the Innovative Computing Technology (INTECH 2012)*, Casablanca, Morocco, 2012, pp. 179–183; doi: 10.1109/INTECH.2012.6457804.

[39] A. Nasry, A. Ezzahout, and F. Omary, "People Tracking in Video Surveillance Systems Based on Artificial Intelligence," *Journal of Automation, Mobile Robotics and Intelligent Systems*, vol. 17, no. 1, 2023, pp. 59–68; doi: 10.14313/JAMRIS/1-2023/8.