

Review

Not peer-reviewed version

---

# A Review of Passenger Counting in Public Transport Concepts based on Image Processing and Machine Learning

---

[Aleksander Radovan](#)<sup>\*</sup>, Goran Đambić, Branko Mihaljević

Posted Date: 3 July 2024

doi: 10.20944/preprints202407.0263.v1

Keywords: public transport; passenger counting; image processing; machine learning



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

# A Review of Passenger Counting in Public Transport Concepts Based on Image Processing and Machine Learning

Aleksander Radovan <sup>1,\*</sup>, Goran Đambić <sup>1</sup> and Branko Mihaljević <sup>2</sup>

<sup>1</sup> Algebra University College; goran.dambic@algebra.hr

<sup>2</sup> RIT Croatia; bxmcada@rit.edu

\* Correspondence: aleksander.radovan@algebra.hr

**Abstract:** The accurate counting of passengers in public transport systems is crucial for optimizing operations, improving service quality, and planning infrastructure. It can also contribute to reducing the number of public transport lines where a high frequency of vehicles is not needed in certain periods during the year, but also by increasing the number of lines where the need is increased. This paper provides a comprehensive review of current methodologies and technologies used for passenger counting, with a focus on image processing and machine learning techniques and concepts. The paper explores various algorithms, including traditional computer vision methods and advanced deep learning approaches, highlighting their strengths and limitations. Additionally, the paper examines the integration of these technologies with other sensing modalities such as infrared sensors, Wi-Fi, and RFID. By analyzing recent advancements and case studies, this review aims to offer insights into the effectiveness, scalability, and practicality of different passenger counting solutions and offers a solution proposal. The paper also analyzed the current GDPR regulation that applies to the European Union and how it affects the use of systems like this. Future research directions and potential areas for technological innovation are also discussed to guide further developments in this field.

**Keywords:** public transport; passenger counting; image processing; machine learning

## 1. Introduction

Public transport is expected to grow significantly during the period from 2024 to 2032 [1]. Optimizing public transport management and determining the frequency of lines that can include buses, trams, trains, and subways has an important impact on reducing the carbon footprint, reducing traffic congestion, waiting for passengers and contributing to improving the user experience. The existence of cameras for the surveillance of vehicles and the transport space facilitates the provision of the necessary infrastructure for the inevitable implementation of passenger counting based on image processing and the use of machine learning algorithms. There are several research and published papers about passenger counting in public transport, which refer to the analysis of situations when there are few and many passengers in the vehicle, at night, during the day and during rainy days when passengers must carry umbrellas, which makes it difficult to count passengers in other demanding situations. situations that require adaptation of algorithms and approaches in analyzing and processing images of the interior of the vehicle.

One approach based on a single camera with an image resolution of 640x480 pixels mounted over access gate in the bus and 40-minutes video [2] proved that it could be used in cases when only one door on the vehicle exists. It was based on C++ programming language and OpenCV library [3] with two artificial neural networks, one for detecting a person and second that detects if the person enters or leaves the public transport vehicle. The implemented system doesn't require additional

mechanical parts, but the processing unit proves to be expensive. The challenges in this research related to detecting passengers were present in situations when the position of the passengers was static, when the color of the background and person clothing were similar, when the noise and background in the image didn't allow highlighting the object of interest and questionable behavior of the system in the event of a large crowd at the entrance.

Besides using only available video material obtained from public transportation vehicle cameras for artificial neural network training, the research [4] proved that a prepared dataset like PAMELA (Pedestrian Accessibility Movement Environment Laboratory of University College London) [5] metropolitan train and bus dataset that contained videos of the passengers getting in and out of the train could be used to train the machine learning model. The described approach was implemented in MATLAB and used edge analysis-based techniques for tracking and detection of objects, spatiotemporal techniques for selection the line from the region of interest, motion detection-based techniques to detect direction of moving person and model-based techniques for detecting the region of interest of the images using prior knowledge. The accuracy of the implemented system was affected by passengers who were moving toward or away from the top-mounted camera and changing dimensions of the passenger body, that resulted in multiple counting in some situations. There were situations where the system didn't detect passengers due to occlusion, complex interaction or shadow, that was sometimes detected as human. The opening of door gave false detection also as the foreground region is extracted on the motion-based method.

Simulation of public transport environment like performed in PAMELA dataset configured as a London Underground carriage to study effect of door width and vertical gap was employed in the research [6]. The PAMELA-UANDES dataset [7] contains a total of 14.834 training images and 13.237 testing images. A video dataset of 348 sequences captured by a standard CCTV-type camera was made publicly available, as the main contribution of this paper. Three deep learning object detectors (EspiNet, Faster-RCNN, and Yolo v3), and three benchmark trackers Markov Decision Processes (MDP), SORT, and D-SORT were evaluated. The models were trained from scratch. The main challenge in the used dataset is the angled camera view that defeats people detectors trained with popular datasets, but corresponds with typical sensor position in this kind of environment. This illustrates the fact that although data-driven learning approaches can produce adequate results, their generalization capabilities are still below that of human observers. The position of the only camera outside the public transport vehicle represents a limitation of this approach to be utilized in different scenarios, as well as the scenarios which only apply to the London Underground.

Instead of using only one surveillance camera in the public traffic vehicle, better accuracy can be achieved using rear and front door camera like in the research [8]. This research employed the training data that were extract with the time intervals set to 3 second to avoid capturing too similar images. The 3000 images of training data were augmented to 15.000 by Gaussian blur, Gaussian noise, mirror operation, and brightness adjustment. In the system testing, 500 images extracted from 100 other videos were used as testing data for the passenger detection stage. The passenger counting process was divided into three main parts: door state estimation (to engage counting of the passenger only if the doors are open), passenger detection (SSD – Single Shot Detector network was built for pre-training a passenger detection model), and passenger tracking and counting (a particle filter with a three-step cascaded data association algorithm is used to track each person). The counting process was performed in three different types of scenarios: daytime, nighttime and rainy days, because the image parameters and detection challenges were different. The passenger detection was focused on counting heads in the region of interest, as well as hats on the passenger heads. The tracking process was based on a sequence of images and using centroid of a bounding rectangle that tracked the passenger head. Using multiple cameras and image datasets that cover different scenarios like day, night, rain, etc. proved to have a significant effect on achieving better results in passenger detection in public transport.

A solution with RGB-D sensor, located over each bus door, for counting adults and children is described in the research paper [9]. It represents an approach an inexpensive and flexible solution to obtain, in real time, statistical measures on the amount of people present in the bus, with the use of

an Analytical Processing System that accesses the data stored in the database and extract statistical data and knowledge about the bus passengers. The solution uses a camera for each door facing downward with an angle of 15 degrees. The RGB-D sensor that was used in Asus Xtion Pro Live and is based on principle of structured light allows the construction of 3D depth maps of a scene in real-time. Structured near-infrared light is directed towards a region of space and a standard CMOS image sensor is used to receive the reflected light. The function of user tracking, in which an algorithm processes the depth image to determine the position of all the joints of any user within the camera range, is provided by an OpenNI compliant module called NiTE (Natural Interface Technology for End-User). To avoid false alarms during the route, the RGB-D system is activated, for allowing people counting, only when the bus arrives at the bus stop and opens its doors, and it is automatically deactivated when the driver closes the doors and restarts the autobus. The skeleton-tracking algorithm provided by NiTE the system can identify people in the scene at the entrance of the bus. Then it monitors the position and the orientation of the 3D coordinates of some joints of the skeleton tracked, so that it can understand if the person is getting in or off the bus. The system was tested only during the day with the total recording time for each camera about 10 minutes and didn't include situation in which the bus is full of passengers, so the depth sensor does not have a clear path to detect separate complete skeletons of every person.

One example of a passenger counting algorithm based on a hybrid machine learning approach by using Histogram of Oriented Gradients (HOG) [10] feature of passenger's heads on the extracted images from recorded video is described in the research [11]. Classification of head features is done by using Support Vector Machine (SVM) [12] as a classifier for the linear model. Heads are detected successfully after performing all steps. In the next step, Kanade-Lucas-Tomasi (KLT) [13] is used to realize head tracking, the multiple target tracking is achieved, and the head motion trajectory of passenger target is captured. In the last step, the proposed algorithm is moved to embedded system ADSP-BF609 for practical implementation. Embedded systems proved to be adjusted for general purpose passenger detection and cannot be easily upgraded to include situations in which passengers use umbrellas because it's raining or in similar situations.

Various automatic passenger counting were analyzed and presented in the research [14]. Video-based sensing based on two cameras mounted on the front and rear doorways of buses. RGB image frames were captured at 640x480 pixel resolution, and 30 frames per second utilizing software running on an adjacent Raspberry Pi 3B+. Passenger detection is achieved using the pre-trained convolutional neural network, and MobileNet SSD (Single Shot Detector) [15]. The second approach analyzed was 3D Infrared sensing that counts passengers from the infrared and depth data. The third approach was based on mobile Wi-Fi sensing that counts the number of devices in a defined space based on probe request messages the devices send when not connected to network. The decision whether a device is inside this space is made using the spatial and temporal overlap of the probe requests sent by this device. The spatial overlap is achieved by positioning several mobile Wi-Fi sensors around the bus. Between four and five sensors were used during testing period. The third approach was related to a sensor-grid mat that was placed at the rear door of the bus. Data recorded by the sensor was computed using an algorithm that calculates the center of pressure movement when passengers step on and off the mat. Two sensor mats with 24 sensing nodes each were taped down on the entrance floor. With the video-based Automatic Passenger Counting solution the achieved results were more accurate on the front door but had a clear decline in accuracy for the rear door. 3D Infrared Sensing achieved similar results to video-based overall. Mobile sensing accuracy during weekdays was considerably lower than during the weekends, likely a large difference in travel time (between one- and two-minutes during weekdays compared to seven minutes during weekends) results. The approach with sensor map resulted with generally strong results on the rear door, but because of an installation error of the sensor map on the front door. The greatest impediment to counting accuracy with the sensor map was people standing on the mat while the bus was in motion. This was most prominent when the bus was full, and standing room was limited. The measurements were taken during six days in December in 2018, so to obtain the most accurate results, it's required to perform the measurements during the whole year.



A system that employs a distributed people-counting approach in the metro based on IoT (Raspberry Pi) and ESP-32 Wi-Fi camera and is based on the YOLO v8 algorithm [16][17] in real-time is described in paper [18]. The dataset that was used was COCO (Common Objects in Context) dataset [19] that contains very useful resources for establishing of proof-of-concept in people-counting. The developed website enables commuters to access real-time information on compartment occupancy. The described approach proves that the provided data about the number of passengers in certain public transport vehicles with an addition of congestions on the streets during rush hour can improve user experience and decision-making when and where to use the public transport services.

A comparison between commercial and custom-made solutions for passenger counting proved that better results can be achieved [20]. The accuracy and precision of the different optical-based solutions are claimed to be between 98 and 99 percent in every case, but the indicators are deceptive, since they are usually obtained under ideal conditions (depot or in a laboratory test). The situation is very often completely different in real-world situations. The accuracy of an estimation is sensitive to camera position and angle, passenger-flow density, and lighting conditions. The cost of a commercial APC (Automatic passenger counting) is between EUR 1500 and EUR 3000 per door, and generally it is supplied as a part of a service with a monthly fee. There are systems that feature Raspberry Pi with camera with a total cost of approximately AUD 560. Considering the significant cost and the limitations of customizing of commercial APC systems, which directly impact the finances of public transport companies, and considering the accuracy of the systems, custom low-cost alternatives are representing a suitable option.

An innovative approach was deployed in research [21] where an evaluation of live data from a co-deployed light sensor system and WLAN probe sensor system, which allows for an in-depth analysis and comparison of the two bus occupancy estimation approaches. The datasets were collected on the same bus route going from one side of a medium-sized German city, through the city center, and to the other side. All datasets were collected at the same time of day but across three different days, and on the same physical bus. The WLAN probe dataset was passively collected from a sensor placed on the bus under the roof. The WLAN probes were emitted by WLAN enabled devices carried by passengers (like smart phones) in and outside the bus. For each WLAN probe collected, the time stamp and GPS location is recorded along with the recorded signal strength of the received probe. The light sensor recorded by light sensors placed in the entrance and exit of the bus. The events are recorded per bus stop on the bus route. Ground truth dataset consists of enter and exit events manually counted by a person per door riding the bus. It was proven that light sensor performs better in following ground truth data in detail, at a short-term, while suffering in the long-term. The WLAN estimator is not good at following ground truth in detail at a short term, but at a long-term it follows the general tendencies of ground truth and there is still a lot of room for improvement.

A Wi-Fi-based automatic bus passenger counting system, named iABACUS, was used to observe and analyze urban mobility by tracking passengers throughout their journey on public transportation vehicles [22]. It counts the number of active Wi-Fi interfaces of mobile devices carried by passengers. It is based on a de-randomization mechanism, which overcomes the issue of not being able to attribute two or more random MAC addresses to the same device and since the original MAC address is kept unknown, the identity of passengers cannot be inferred, and their privacy is preserved. The described system also tracks passengers throughout their journey on public transportation vehicles, by providing when they board or align from the bus. The disadvantages of this approach are that passengers can carry more than one device connected to Wi-Fi, which would disturb the accuracy of the results. Sometimes the devices outside the vehicles were counted as present in the vehicle and needed to be discarded. Randomization of the MAC addresses presented a serious problem, since the presence of only three devices can lead to a count six times higher in a 15-minute time interval.

The paper presents a review of passenger counting systems in public transports based on image processing and machine learning that aims to address prerequisites and challenges of such a system. It analyzes several different approaches, highlights pros and cons of every of the approaches and suggests best practices that were used in previously published papers. After explaining the concepts,

techniques and challenges that need to be considered when planning the implementation of a system for counting passengers in public transport, the proposed solution for such an implementation with all the key parts of such a system is presented and described.

The paper is structured as follows. The following section presents the concepts and techniques needed to implement an automatic passenger counting and challenges that need to be considered before implementing such a solution. Section 3 describes a solution proposal of a deployed passenger counting system. After that, in section 4, the impact of the GDPR on systems like this and the limitations it sets are described. Finally, in the Discussion and Conclusions section, the most important facts are described, and the most important conclusions are given, as well as the possibilities of expansion and improvement of the system for counting passengers in public transport in the future.

## 2. Concepts, Techniques and Challenges

The common concepts of counting passengers in public transport vehicles were based on defining areas of interest for counting or line of interest where the passenger need to pass when entering or leaving the vehicle. The images extracted from the video taken by the camera are then transformed to an appropriate format, analyzed by using artificial neural networks to detect the object of interest (for example, person, vehicle doors, etc.), and determine if the object is involved in the situation which means that the passenger counter needs to be updated.

Image processing and transformation includes foreground and background subtraction (for example, by using the Hough method [23]), features extraction [24], image segmentations also known as thresholding [25] used to create binary images to improve the accuracy of the object detection, performing morphological operations like dilation and erosion [26], color conversion (from RGB to HSV [27], for example), Histogram equalization, clustering (compacting pixels in regions), passenger detection and tracking counting the objects (passengers) by using the artificial neural networks based on trained model.

The number of cameras deployed in public transport vehicles plays an important role in avoiding false detection of passengers in the vehicle. Especially if the vehicle contains multiple entrance and exit doors. It is highly recommended to employ at least one camera that covers each of the doors. Research like [8] proved that the detection of open doors improves the precision, so the counting process starts when the doors are open, and it does change the number of passengers while driving.

During the year, the seasons, weather conditions, clothes worn by travelers' change, days and nights, different platform configuration, different types of lightning, different camera views, sunny and rainy days and the like change as well. Because of this, image processing algorithms as well as machine learning models must be adjusted and trained with various image datasets recorded in different scenarios to achieve robustness and return expected and reliable results. Many existing passengers counting in public transport performed the data collection only during the day, and not during the nights and rainy days [8]. For achieving the best results and to constantly improve the detection precision, additional image dataset needs to be collected, and with manual labelling of objects (passengers) on images, enhance and enrich the training dataset with new scenarios that weren't recognized as expected.

Depth sensors were also utilized like in the research [9], but in situations where the public transfer vehicle is full of passengers, the detected person skeletons can overlap which can result in wrong detection results. In combination with other sensors or cameras present in the public transport vehicle and using obtained data to analyze the situation in interest (like counting heads of passengers), the results could be improved in such cases.

One of the biggest challenges in establishing a system that will reliably count passengers in public transport is the acquisition of an adequate dataset of videos that include onboarding and offboarding passengers. The slowest way consists of collecting your own data set by storing videos from surveillance cameras throughout the year, processing and training models to detect and count passengers. In this way, it is possible to precisely adjust the detection system in real scenarios in

different situations in vehicles, and it represents an optimal way. Alternatively, it is possible to use an existing publicly available dataset like PAMELA mentioned in [4,7] as an initial version and improve it with additional videos if the results turn out not to be good enough. The size of the image dataset is also important, previous papers that provided the best detection precision included hundreds of videos and thousands and tens of thousands of images on which to train the model. The image datasets must be adequately divided into a set for training and a set for validation, usually divided in the proportion of 80% - 20%.

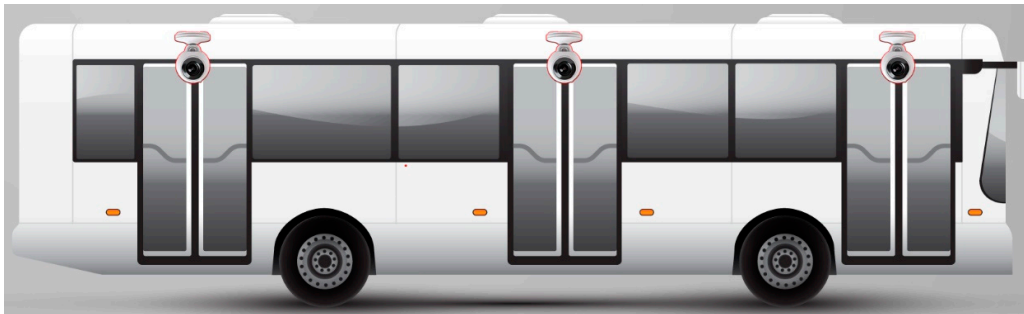
In cases where it is necessary to start the development of the system from the very beginning, it is recommended to use existing models that have already been sufficiently trained, such as YOLO in its latest version (at the time of writing this paper it was version v9 released on 02/21/2024) [28][29] and the COCO dataset [30], which contains a large the number of images of objects that can be tracked, such as people (limited person dataset only is available on [31]).

### 3. Solution Proposal

This chapter will describe a proposal for the implementation of a solution for counting passengers in public transport, from how the cameras should be placed, how to connect them to the common infrastructure, how to enable communication with the monitoring center, and how to collect video materials based on which prepare images for training a machine learning model for passenger recognition. The section also covers the best practices how to approach image processing and machine learning model training. The best results were achieved with image processing and machine learning techniques based on CCTV cameras located at the front and the rear door in the public transport vehicle.

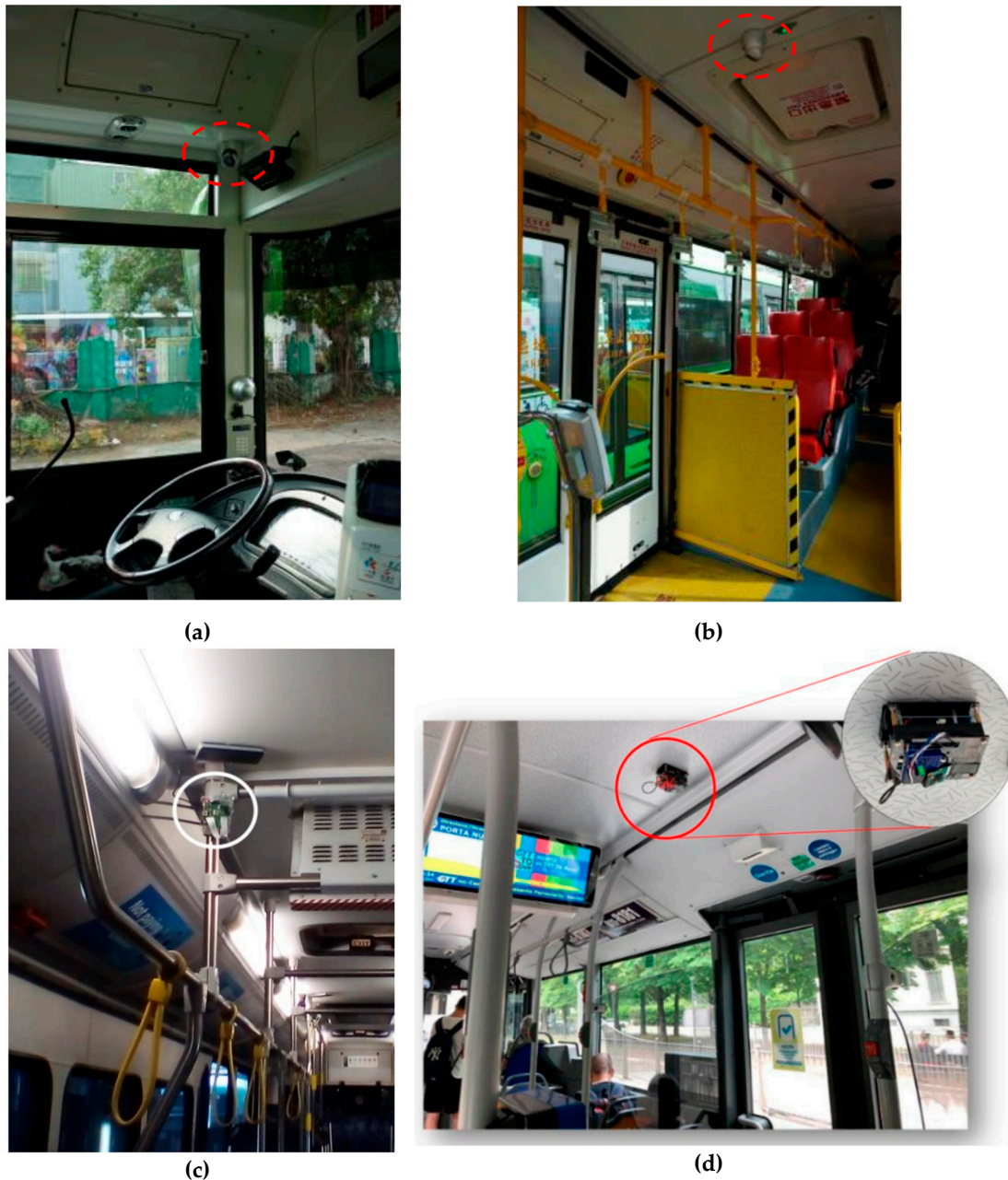
#### 3.1. Camera Locations

From the analysis of published research, it is concluded that to achieve the best results, it is necessary to install a CCTV IP (Closed-Circuit Television Internet protocol) camera at each door of the public translation vehicle as shown in Figure 1. In case of a bus or similar public transport vehicle, cameras must be located on the roof to be able to capture video about the passengers' heads.



**Figure 1.** Proposed locations of cameras in the public transport vehicles.

Example of camera locations are shown in Figure 2. The first two images, (a) and (b) show surveillance camera locations, and (c) and (d) show custom-made cameras in case of proof-of-concept solutions.



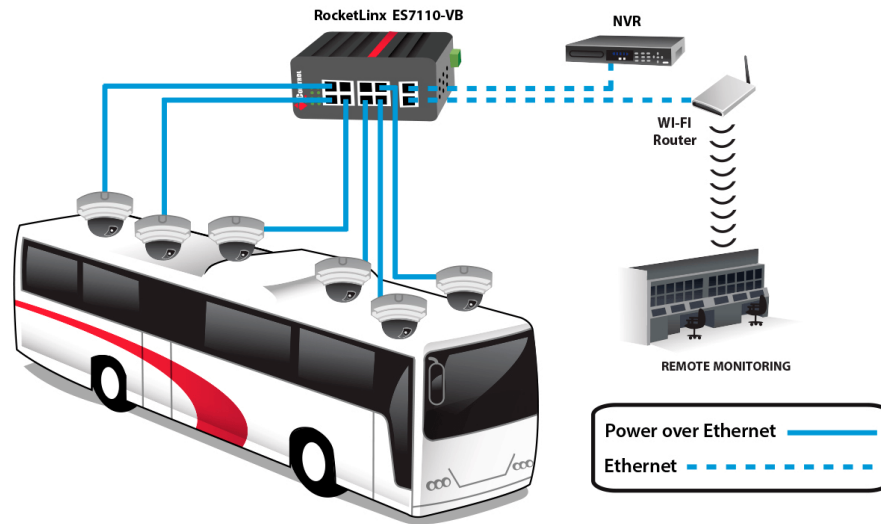
**Figure 2.** Positions of the cameras: (a) surveillance camera at the front door [8]; (b) surveillance camera at the rear door [8]; (c) custom-made camera at the front door [11]; (d) low-cost camera installation based on Raspberry Pi at the rear door [17].

### 3.2. Public Transport Vehicle External and Internal Network Infrastructure

As part of the process of installing cameras in vehicles, to be able to access recorded videos and enable real-time communication, it is necessary to use a network connection with the cameras using PoE technology [32]. Besides cameras an embedded computer functioning as a Network Video Recorder (NVR) is required to allow to access the recorded video materials. Bus surveillance systems must be wirelessly enabled with Wi-Fi router and cellular compatibilities. After the IP cameras send video to the onboard embedded computer, the computer streams live video through a secure cellular connection to a control center, where managers can access video images and monitor the bus



remotely. Figure 3 shows an example of components and internal and external network infrastructure in a bus [33].



**Figure 3.** Public transport vehicle external and internal network infrastructure [34].

### 3.3. Gathering the Image Dataset

Before starting the process of training a machine learning model based on the processing of images taken from videos recorded in public transport vehicles, it is necessary to prepare enough images, which should be several thousand or better several tens of thousands to achieve the expected results and adequate precision.

If it is feasible to collect images throughout the year, during and after hours, during the day and night, during all seasons and all-weather conditions, that would be ideal. If this is not possible and it is necessary to start from the beginning, the prepared YOLO v9 [28][29] model that can recognize people in pictures and the COCO [30][31] data set that contains many pictures of people can also be used.

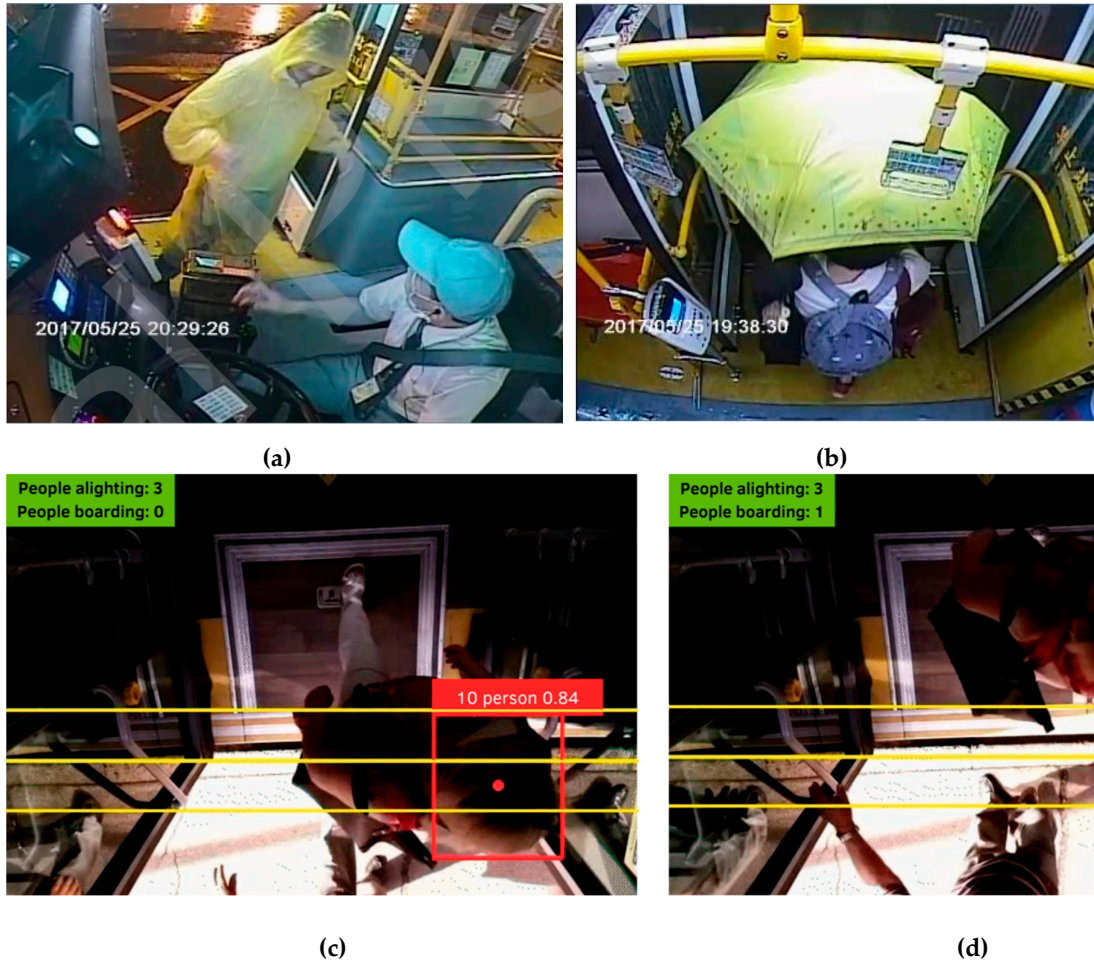
The YOLO (You Only Look Once) family of models has a significant impact in real-time object detection due to its balance of performance, especially speed and accuracy. The latest iteration, YOLO v9, released in February 2024, builds upon its predecessors with improvements. It leverages a refined architecture that enhances feature extraction and object localization capabilities. The latest version integrates advanced techniques such as adaptive anchor boxes, enhanced path aggregation networks, and transformer-based components, which contribute to its superior detection accuracy and robustness, especially in complex and dynamic environments like public transport vehicles [35][36].

The COCO (Common Objects in Context) dataset is a widely used benchmark in the field of object detection, segmentation, and captioning. It contains over 200,000 labeled images with more than 80 object categories, including people, which are critical for passenger counting applications. The COCO dataset's nature ensures that models trained on it can generalize well to various real-world scenarios. This is particularly important for passenger counting systems, as they must adapt to different lighting conditions, passenger densities, and environmental changes throughout the year [30].

Using YOLO v9 and the COCO dataset in the implementation of passenger counting systems offers real-time processing capabilities needed for reliable passenger detection and counting. The pre-trained weights on the COCO dataset provide an adequate starting point for fine-tuning to custom datasets collected from public transport environments. In situations where passengers overlap or when vehicles are crowded, the detection algorithms of YOLO v9 can more effectively distinguish individual passengers by counting individual passengers by counting heads since the camera is mounted on the roof of the public transport vehicle. By continuously updating the training datasets

with new scenarios and employing robust models, public transport systems can maintain improved accuracy and reliability in passenger counting, after manual labeling of passengers with a tool like Label Studio [37].

The door and passenger images represent regions where objects of interest are defined and lines of interest where object recognition is performed. To develop a system that does not depend on ideal conditions and covers different periods of the day and year and weather conditions, it is necessary to collect images in such specific scenarios and manually label it if needed. Figure 4 shows some of such situations, which refer, for example, to situations when it is raining, and passengers enter and exit public transport vehicles holding umbrellas and their heads are not fully visible in the images, or the lighting of the picture is not ideal.



**Figure 4.** Images in specific scenarios: (a) passenger wears a raincoat at night [8]; (b) passenger uses an umbrella while it's raining at night [8]; (c) passenger offboards from the bus and crosses the line of interest (yellow) in dark conditions [20]; (d) two passengers are boarding in dark conditions [20].

### 3.4. Image Processing and Machine Learning Model Training

The initial step in developing a robust passenger counting system in public transport includes image processing techniques to prepare the data for machine learning models. The raw video footage collected from cameras installed in public transport vehicles must undergo several preprocessing stages to enhance the quality and utility of the images. These stages include foreground and background subtraction, often achieved through methods like the Hough Transform [38], which isolates moving objects (passengers) from static backgrounds. Further, feature extraction techniques are employed to identify and highlight crucial elements within the images, such as edges, shapes, and textures.

Once the images are preprocessed, segmentation techniques, particularly thresholding, are applied to create binary images that simplify the detection process by distinguishing the object of interest from the rest of the image. Morphological operations like dilation and erosion [26] are then used to refine the shapes of detected objects, removing noise and filling gaps. Color conversion, such as converting from RGB to HSV color space [27], can also be beneficial, as it helps in distinguishing objects under varying lighting conditions. Histogram equalization is another important step that improves the contrast of the images, making the features more discernible for detection algorithms.

Following image preprocessing, the next crucial phase involves training machine learning models to accurately detect and count passengers. The collected dataset, ideally encompassing a wide range of scenarios—including different lighting conditions, weather conditions, and times of day—provides the necessary variability for training robust models. Deep learning techniques [39], particularly those based on convolutional neural networks (CNNs) [40], are widely used due to their performance in image recognition tasks. Models like YOLO [28][29] are particularly suitable for real-time applications due to their high speed and accuracy.

Training these models involves splitting the dataset into training and validation sets, typically in an 80-20 ratio. The training set is used to teach the model to recognize patterns and features associated with passengers, while the validation set is used to evaluate the model's performance and fine-tune its parameters. Data augmentation techniques, such as flipping, rotating, and scaling images, are often applied to increase the diversity of the training data and improve the model's generalization capabilities. Additionally, transfer learning [41] from pre-trained models on large datasets like COCO [30][31] can improve the training process and enhance the model's performance.

Finally, continuous retraining and validation are essential to maintain the accuracy and reliability of the passenger counting system. As new data is collected, particularly under previously unseen conditions, it is added to the training dataset, and the model is retrained to adapt to these new scenarios. This iterative process ensures that the system remains robust and capable of delivering accurate counts, which are crucial for optimizing public transport operations.

#### **4. GDPR Compliance and Passenger Counting Systems**

The General Data Protection Regulation (GDPR), enacted by the European Union in May 2018 [42–44], has significantly impacted how data, especially personal data, is collected, processed, and stored. Passenger counting systems in public transport vehicles, which often rely on video surveillance and image processing, must comply with GDPR to ensure the privacy and protection of individuals' data. This regulation mandates strict guidelines on data handling practices to safeguard individuals' privacy rights and prevent unauthorized access or misuse of personal data [45].

One of the primary GDPR concerns for passenger counting systems is the identification and processing of personally identifiable information (PII). Images and videos collected for counting passengers can inadvertently capture faces and other identifying features, thus classifying this data as PII under GDPR. To mitigate these risks, public transport operators must implement robust anonymization techniques, such as blurring or masking faces, before storing or processing the data. Additionally, clear signage informing passengers about the presence of surveillance cameras and the purposes of data collection is essential for maintaining transparency and obtaining informed consent, a core GDPR requirement.

The European Union has also introduced specific regulations to enhance the protection of personal data in the context of smart cities and public transport systems. These regulations emphasize the need for data minimization, ensuring that only the data necessary for the specific purpose of passenger counting is collected and retained. Public transport authorities are encouraged to conduct Data Protection Impact Assessments (DPIAs) to systematically analyze and mitigate potential privacy risks associated with their data processing activities. Furthermore, employing encryption and other security measures to protect data during transmission and storage is mandated to prevent data breaches.

Compliance with GDPR and these new regulations requires a multifaceted approach. Public transport operators must ensure that their data processing agreements with third-party service

providers, such as those supplying image processing and machine learning technologies, include stringent data protection clauses. Regular audits and updates to privacy policies and practices are also necessary to adapt to evolving legal requirements and technological advancements. Training employees on data protection principles and fostering a culture of privacy within the organization further strengthens compliance efforts.

While passenger counting systems provide valuable data for optimizing public transport operations, adherence to GDPR and EU regulations is imperative to protect passengers' privacy. By implementing anonymization techniques, conducting DPIAs, ensuring data minimization, and securing data through encryption, public transport operators can achieve the dual goals of operational efficiency and regulatory compliance [46].

## 5. Discussion

The implementation of passenger counting systems in public transport vehicles based on image processing and machine learning represents a suitable option for enhancing operational efficiency and service quality. This review has highlighted various methodologies, including defining areas or lines of interest for counting, advanced image preprocessing techniques, and the application of neural networks for object detection and tracking. However, several key issues and challenges must be addressed to realize the full potential of these technologies.

One of the primary challenges is ensuring accuracy and reliability in diverse and dynamic environments. Public transport vehicles operate under varying conditions such as different lighting, weather, and passenger densities, all of which can affect the performance of image processing algorithms. The need for robust preprocessing techniques, including foreground and background subtraction, feature extraction, and image segmentation, is critical to maintaining high detection accuracy. Additionally, the integration of morphological operations, color conversion, and histogram equalization further enhances the quality of the images for subsequent analysis.

Machine learning models, particularly those based on deep learning architectures like YOLO, have demonstrated significant advancements in object detection capabilities. However, training these models requires large, diverse, and well-labeled datasets. The availability and quality of these datasets are important for developing models that can generalize well across different scenarios encountered in public transport. The incorporation of transfer learning, using pre-trained models on extensive datasets like COCO, can accelerate the training process and improve initial performance. Continuous retraining with new data, including edge cases and rare events, is necessary to adapt to evolving operational conditions and to improve model robustness.

GDPR compliance and data privacy are critical considerations in deploying passenger counting systems. The collection and processing of video data raise concerns about the protection of personally identifiable information (PII). Anonymization techniques, data minimization, and conducting Data Protection Impact Assessments (DPIAs) are essential steps to ensure that passenger data is handled responsibly and legally. Public transport operators must implement stringent data protection measures, both in collaboration with third-party technology providers and within their own data management practices.

The deployment of multiple cameras to cover all entry and exit points in a vehicle is recommended to enhance detection accuracy and reduce false positives. The strategic placement of cameras can ensure adequate coverage, especially in vehicles with multiple doors. Research has shown that detecting the opening and closing of doors can significantly improve the precision of passenger counting, as it allows the system to focus on periods of high passenger movement.

While significant progress has been made in the development of image processing and machine learning-based passenger counting systems, several challenges remain. Addressing these challenges through preprocessing techniques, comprehensive training datasets, adherence to data privacy regulations, and strategic system design will be crucial for the successful deployment and operation of these systems in public transport. Continued research and innovation in this field will contribute to more efficient and responsive public transport services, ultimately benefiting both operators and passengers.



## 6. Conclusions

The adoption of image processing and machine learning techniques for passenger counting in public transport vehicles offers benefits for optimizing operational efficiency and improving service quality. This review has explored various methodologies, including defining areas or lines of interest for counting, applying advanced image preprocessing techniques, and employing neural networks for object detection and tracking. Despite the advancements, several challenges, including accuracy in diverse environments, data privacy concerns, and the need for comprehensive datasets, remain critical to address.

Ensuring high accuracy and reliability in passenger counting systems under varying operational conditions is paramount. Image preprocessing techniques, such as foreground and background subtraction, feature extraction, and image segmentation, are essential for enhancing detection accuracy. Moreover, leveraging deep learning models like YOLO, which can process images in real-time with high precision, has proven to be highly effective. The use of pre-trained models and continuous retraining with new datasets can further improve the performance and adaptability of these systems.

Adhering to GDPR and data privacy regulations is crucial for the ethical deployment of passenger counting systems. Implementing anonymization techniques, conducting Data Protection Impact Assessments (DPIAs), and ensuring data minimization are necessary to protect passengers' personal data. Public transport operators must collaborate with technology providers to incorporate stringent data protection measures and foster a culture of privacy within their organizations.

Future research should explore the integration of additional sensors, such as LIDAR, beside depth sensors, infrared sensors, with traditional camera-based systems. Additional sensors can provide complementary data that enhances the accuracy and reliability of passenger counting, especially in challenging conditions such as overcrowded vehicles.

With introduction of improved anonymization techniques protect passenger identities while retaining the necessary data for accurate counting is an important area for innovation. Techniques such as differential privacy and secure multi-party computation could be investigated to enhance data privacy.

Research into adaptive machine learning models that can dynamically adjust to varying conditions and continuously learn from new data will be valuable. This includes developing models that can handle different lighting conditions, weather scenarios, and passenger densities more effectively.

Implementing edge computing solutions to process data in real-time on-board public transport vehicles can reduce latency and enhance the responsiveness of passenger counting systems. This approach can also alleviate data privacy concerns by minimizing the transmission of sensitive data to central servers.

Designing scalable and modular passenger counting systems that can be easily adapted to different types of public transport vehicles and varying operational requirements is another area for technological innovation. This includes creating flexible architectures that can integrate with existing infrastructure and future technologies.

Developing and sharing comprehensive and diverse datasets that include various operational scenarios is important for training robust machine learning models. Public-private partnerships and collaborative research initiatives can play a significant role in creating and maintaining these datasets.

Incorporating user-centric design principles and feedback mechanisms to ensure that the systems meet the needs of both operators and passengers. This includes considering user privacy preferences, ease of deployment, and maintenance.

**Author Contributions:** Conceptualization, A.R.; Methodology, A.R.; validation, G.D., B.M.; formal analysis, G.D., B.M.; resources, A.R.; writing – original draft preparation, A.R., G.D., B.M.; writing – review and editing, A.R., G.D., B.M.; visualization, A.R.; supervision, G.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Public Transport Insights **2024**: Analyzing Trends, Size, Share, Demands and Growth Opportunities to 2023. Available online: <https://www.linkedin.com/pulse/public-transport-insights-2024-analyzing-trends-v4yne/> (accessed on 29 June 2024).
2. P. Chato, D. J. M. Chipantasi, N. Velasco, S. Rea, V. Hallo and P. Constante, "Image processing and artificial neural network for counting people inside public transport" **2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM)**, Cuenca, Ecuador, 2018, pp. 1-5, doi: 10.1109/ETCM.2018.8580287.
3. OpenCV library. Available online: <https://opencv.org/> (accessed on 30 June 2024).
4. S.H. Khan, M.H. Yousaf, F. Murtaza, S. Velastin, "Passenger detection and counting for public transport system", *NED University Journal of Research*, **2020**, Issue: XVII, Volume 2, pp. 35-46, doi: 10.35453/NEDJR-ASCN-2019-0016.
5. Velastin, S.A.; Fernández, R.; Espinosa, J.E.; Bay, A. Detecting, Tracking and Counting People Getting On/Off a Metropolitan Train Using a Standard Video Camera. *Sensors* **2020**, *20*, 6251, doi: 10.3390/s20216251.
6. Pedestrian Accessibility and Movement Environment Laboratory. Available online: <https://discovery.ucl.ac.uk/id/eprint/1414/> (accessed on 30 June 2024).
7. PAMELA UANDES dataset. Available online: [http://velastin.dynu.com/PAMELA-UANDES/whole\\_data.html](http://velastin.dynu.com/PAMELA-UANDES/whole_data.html) (accessed on 30 June 2024).
8. Ya-Wen Hsu, Ting-Yen Wang, Jau-Woei Perng, "Passenger flow counting in buses based on deep learning using surveillance video", *Optik*, Volume 202, **2020**, 163675, ISSN 0030-4026, 10.1016/j.jileo.2019.163675.
9. Liciotti, D., Cenci, A., Frontoni, E., Mancini, A., Zingaretti, P. (2017). An Intelligent RGB-D Video System for Bus Passenger Counting. In: Chen, W., Hosoda, K., Menegatti, E., Shimizu, M., Wang, H. (eds) *Intelligent Autonomous Systems 14. IAS 2016. Advances in Intelligent Systems and Computing*, vol 531. Springer, Cham, doi: 10.1007/978-3-319-48036-7\_34.
10. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177.
11. Haq, E.U., Huarong, X., Xuhui, C. et al. A fast hybrid computer vision technique for real-time embedded bus passenger flow calculation through camera. *Multimed Tools Appl* **79**, 1007–1036 (2020), doi: 10.1007/s11042-019-08167-y.
12. Derek A. Pisner, David M. Schnyer, Chapter 6 - Support vector machine, Editor(s): Andrea Mechelli, Sandra Vieira, *Machine Learning*, Academic Press, 2020, Pages 101-121, ISBN 9780128157398, doi: 10.1016/B978-0-12-815739-8.00006-7.
13. C. Zhang, J. Xu, A. Beaugendre and S. Goto, "A KLT-based approach for occlusion handling in human tracking," 2012 Picture Coding Symposium, Krakow, Poland, 2012, pp. 337-340, doi: 10.1109/PCS.2012.6213360.
14. Moser, I., McCarthy, C., Jayaraman, P.P., Ghaderi, H., Dia, H., Li, R., Simmons, M., Mehmood, U., Tan, A. M., Weizman, Y., Yavari, Al, Georgakopoulos, D., Fuss, F. K., A Methodology for Empirically Evaluating Passenger Counting Technologies in Public Transport, 41<sup>st</sup> Australasian Trasport Research Forum **2019**.
15. Debojit Biswas, Hongbo Su, Chengyi Wang, Aleksandar Stevanovic, Weimin Wang, An automatic traffic density estimation using Single Shot Detection (SSD) and MobileNet-SSD, *Physics and Chemistry of the Earth, Parts A/B/C*, Volume 110, 2019, Pages 176-184, ISSN 1474-7065, doi: 10.1016/j.pce.2018.12.001.
16. YOLOv8. Available online: <https://yolov8.com/> (accessed on 30 June 2024).
17. Ultralytics YOLOv8 Github repository. Available online: <https://github.com/ultralytics/ultralytics> (accessed on 30 June 2024).
18. Yoshida. T., Kihore, N. A., Thapaswi, A., Venkatesh, P., Ponderti, R.K., Smart metro: Real-time passenger counting and compartment occupancy optimization using IoT and Deep Learning, *International Research Journal of Modernization in Engineering Technology and Science* **2024**, Volume: 06/Issue:03/March-2024, e-ISSN: 2582-5208, doi: 10.56726/IRJMET550986.

19. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In European Conference on Computer Vision (pp. 740-755). Springer, Cham.
20. Pronello, C.; Garzón Ruiz, X.R. Evaluating the Performance of Video-Based Automated Passenger Counting Systems in Real-World Conditions: A Comparative Study. *Sensors* 2023, 23, 7719. <https://doi.org/10.3390/s23187719>.
21. Madsen, T.; Schwefel, H.-P.; Mikkelsen, L.; Burggraf, A. Comparison of WLAN Probe and Light Sensor-Based Estimators of Bus Occupancy Using Live Deployment Data. *Sensors* 2022, 22, 4111. <https://doi.org/10.3390/s22114111>.
22. Nitti, M.; Pinna, F.; Pintor, L.; Pilloni, V.; Barabino, B. iABACUS: A Wi-Fi-Based Automatic Bus Passenger Counting System. *Energies* 2020, 13, 1446. <https://doi.org/10.3390/en13061446>.
23. Tao Xu, Hong Liu, Yueliang Qian and Han Zhang, "A novel method for people and vehicle classification based on Hough line feature," International Conference on Information Science and Technology, Nanjing, 2011, pp. 240-245, doi: 10.1109/ICIST.2011.5765245.
24. Wamidh K. Mutlag et al., Feature Extraction Methods: A Review, Journal of Physics: Conference Series, Volume 1591, The Fifth, International Scientific Conference of Al-Khwarizmi Society (FISCAS) 2020, doi: 10.1088/1742-6596/1591/1/012028.
25. H.D. Cheng, X.H. Jiang, Y. Sun, Jingli Wang, Color image segmentation: advances and prospects, Pattern Recognition, Volume 34, Issue 12, 2001, Pages 2259-2281, ISSN 0031-3203, [https://doi.org/10.1016/S0031-3203\(00\)00149-7](https://doi.org/10.1016/S0031-3203(00)00149-7).
26. Hirata, N.S.T.; Papakostas, G.A. On Machine-Learning Morphological Image Operators. *Mathematics* 2021, 9, 1854. <https://doi.org/10.3390/math9161854>.
27. Vladimir Chernov, Jarmo Alander, Vladimir Bochko, Integer-based accurate conversion between RGB and HSV color spaces, Computers & Electrical Engineering, Volume 46, 2015, Pages 328-337, ISSN 0045-7906, <https://doi.org/10.1016/j.compeleceng.2015.08.005>.
28. Chien-Yao Wang, I-Hau Yeh, Hong-Yuan Mark Liao, YOLOv9: Leaning What You Want to Learn Using Programmable Gradient Information, Computer Vision and Pattern Recognition, 2024, <https://doi.org/10.48550/arXiv.2402.13616>.
29. YOLOv9 Github repository, Available online: <https://github.com/WongKinYiu/yolov9> (accessed on 30 June 2024).
30. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In European Conference on Computer Vision (pp. 740-755). Springer, Cham.
31. COCO Dataset Limited (PersonOnly), Available online: <https://universe.roboflow.com/shreks-swamp/coco-dataset-limited--person-only> (accessed on 30 June 2024).
32. F. G. Osorio, Ma Xinran, Yuan Liu, P. Lusina and E. Cretu, "Sensor network using Power-over-Ethernet," 2015 International Conference and Workshop on Computing and Communication (IEMCON), Vancouver, BC, 2015, pp. 1-7, doi: 10.1109/IEMCON.2015.7344442.
33. Bus Surveillance with Axiomtek's tBOX810-838-FL, Available online: <https://www.axiomtek.com/ArticlePageView.aspx?ItemId=1909&t=27> (accessed on 30 June 2024).
34. Transportation and bus surveillance: mobile security, Available online: <https://iebmedia.com/applications/transportation/transportation-and-bus-surveillance-mobile-security/>, (accessed on 30 June 2024).
35. YOLOv9: SOTA Object Detection Model Explained, Available online: <https://encord.com/blog/yolov9-sota-machine-learning-object-detection-model/>, (accessed on 30 June 2024).
36. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.
37. Label Studio Documentation, Available online: <https://labelstud.io/guide/>, (accessed on 30 June 2024).
38. Chakravartula Raghavachari, V. Aparna, S. Chithira, Vidhya Balasubramanian, A Comparative Study of Vision Based Human Detection Techniques in People Counting Applications, Procedia Computer Science, Volume 58, 2015, Pages 461-469, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2015.08.064>.
39. Hossen, M.A.; Naim, A.G.; Abas, P.E. Deep Learning for Skeleton-Based Human Activity Segmentation: An Autoencoder Approach. *Technologies* 2024, 12, 96. <https://doi.org/10.3390/technologies12070096>.

40. Z. Li, F. Liu, W. Yang, S. Peng and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 6999-7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.
41. Weiss, K., Khoshgoftaar, T.M. & Wang, D. A survey of transfer learning. J Big Data 3, 9 (2016). <https://doi.org/10.1186/s40537-016-0043-6>.
42. General Data Protection Regulation – GDPR, Available online: <https://gdpr-info.eu/>, (accessed on 30 June 2024).
43. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), Available online: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>, (accessed on 30 June 2024).
44. Voigt, Paul & Bussche, Axel. (2017). The EU General Data Protection Regulation (GDPR): A Practical Guide. 10.1007/978-3-319-57959-7.
45. Benyahya, M.; Kechagia, S.; Collen, A.; Nijdam, N.A. The Interface of Privacy and Data Security in Automated City Shuttles: The GDPR Analysis. Appl. Sci. 2022, 12, 4413. <https://doi.org/10.3390/app12094413>.
46. Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01), Available online: <https://ec.europa.eu/newsroom/article29/items/611236/en>, (accessed on 30 June 2024).

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.