

Review

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

Aleksander Radovan ^{1,*}, Leo Mršić ², Goran Đambić ¹ and Branko Mihaljević ³

¹ Department of Software Engineering, Algebra University College, 10000 Zagreb, Croatia; goran.dambic@algebra.hr

² Department of Program Engineering, Algebra University College, 10000 Zagreb, Croatia; leo.mrsic@algebra.hr

³ Department of Information Sciences & Technologies, Rochester Institute of Technology (RIT Croatia), 10000 Zagreb, 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 number 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, without the actual implementation of the automatic passenger counting system (APC), but with a proposal based on image processing and machine learning techniques and concepts, since it represents one of the most used approaches. The research explores various technologies and algorithms, like card swiping, infrared, weight and ultrasonic sensors, RFID, Wi-Fi, Bluetooth, LiDAR, thermos cameras, including CCTV cameras and traditional computer vision methods, and advanced deep learning approaches, highlighting their strengths and limitations. 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 research also analyzed the current General Data Protection Regulation (GDPR) 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



Citation: Radovan, A.; Mršić, L.; Đambić, G.; Mihaljević, B. A Review of Passenger Counting in Public Transport Concepts with Solution Proposal Based on Image Processing and Machine Learning. *Eng* **2024**, *5*, 3284–3315. <https://doi.org/10.3390/eng5040172>

Academic Editors: Juvenal Rodriguez-Resendiz, Akos Odry, José Manuel Álvarez-Alvarado and Marco Antonio Aceves-Fernandez

Received: 27 October 2024
Revised: 1 December 2024
Accepted: 3 December 2024
Published: 10 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Public transport systems worldwide are facing increasing demands to improve efficiency, reduce operational costs, and enhance the overall quality of service. Counting passengers is a key component in achieving these goals, as it enables transport operators to optimize vehicle deployment, adjust schedules, and manage resources more effectively. Traditional manual methods of counting passengers have proven to be labor-intensive, inaccurate, and insufficient for large-scale operations. Over the past decade, advancements in image processing and machine learning have opened new avenues for automating passenger counting processes with higher accuracy and scalability. These technologies offer the ability to extract valuable real-time insights, aiding in the dynamic adjustment of transport services according to demand patterns. Additionally, integrating passenger counting systems with other sensor technologies, such as infrared sensors, Wi-Fi, LiDAR, and RFID, provides a multi-modal approach that enhances system reliability and precision. However, despite their promise, many of these technologies face challenges in real-world applications, particularly regarding cost, scalability, and privacy concerns under regulations like GDPR. This research paper addresses the need for a comprehensive review of

current methodologies, highlights their strengths and limitations, and proposes pathways for innovation in future passenger counting systems. In this way, this review contributes valuable knowledge for both academia and industry, fostering the development of more effective, privacy-conscious, and adaptable solutions in public transport management.

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. In this paper, the greatest emphasis will be on the analysis of solutions for counting passengers in public transport based on machine learning, a technology that is becoming more and more available and enables the use of existing infrastructure in public transport vehicles, such as surveillance cameras.

The paper is structured as follows: The Section 2 presents the introduction to passenger counting technologies, from card swiping and ticketing systems to ultrasonic sensors, with a focus on their strengths and limitations. Section 3 gives a review of the literature dealing with the counting of passengers in public transport based on the technologies described in the second chapter. Section 4 covers the methodology used in the analysis of state-of-the-art implementation of the passenger counting technologies, with comparison among the technologies regarding precision, computational complexity, real-time suitability, and hardware requirements. Section 5 describes concepts and techniques needed to implement the APC system and challenges that need to be considered before implementing such a solution. Section 6 describes a solution proposal for a deployed passenger counting system. After that, in Section 7, the impact of the GDPR on systems like this and the limitations it sets are described. Finally, in Sections 8 and 9, 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. Passenger Counting Technologies

The accurate counting of passengers is essential for optimizing public transportation services, ensuring operational efficiency, and improving service quality. Over the years, a wide range of technologies has emerged to facilitate passenger counting, ranging from basic systems such as card swiping data to sophisticated sensing and machine learning-based approaches. This section provides a comprehensive overview of the technologies currently used for passenger counting, highlighting their functionalities, strengths, and limitations.

2.1. Card Swiping and Ticketing Systems

One of the earliest and most used technologies for passenger counting involves data from card swiping and ticketing systems. These systems, often implemented in urban transit networks, require passengers to swipe a card [2], insert a paper ticket [3] upon boarding, or tap an RFID-enabled ticket [4]. By comparing smart card data with manually collected data from the Seoul Metro Company (Seoul, Republic of Korea), it was proved that smart card data can reliably reflect the number of users at metro stations; this was used for validating counting systems [2]. In the study, [3], the conclusions suggest that where there is a trend towards digital ticketing systems, certain demographics still favor the use of paper tickets that needs to be considered as one of the segments related to passenger counting in the public transport vehicles. An option to combine RFID and mobile technology to transform traditional ticketing processes into a digital experience for passengers is highlighted and researched in [4]. Integrating RFID technology with Android applications proved to be a robust solution for ticketing and announcement systems in public transport.

The strengths of these systems are based on providing an accurate record of ticketed entries and are relatively easy to implement in closed-fare environments. On the other hand, card swiping systems only count paying passengers and do not account for fare evasion, system malfunctions, or passengers who may board without swiping (e.g., in open fare systems or through multiple doors). Card swiping also has the limitation of accurately capturing passenger travel patterns, particularly for non-pendular trips. Pendular travel patterns are repetitive, such as daily commuting between home and work or school, but non-pendular travel patterns are irregular, and occasional trips with varying routes and destinations and are less predictable. If the APC only requires swiping at boarding, not at alighting, it creates challenges in providing accurate passenger counting.

2.2. RFID

RFID (Radio Frequency Identification) systems utilize radio frequency tags embedded in passenger cards or devices that are detected when passengers pass through an RFID reader at entry points [5]. Unlike traditional card-swiping systems, RFID does not require active swiping, making the process more seamless.

RFID offers fast, contactless interaction and can be integrated with existing fare collection systems to provide automated passenger counts. On the other hand, RFID still requires passengers to carry specific RFID-enabled cards, and the system cannot track fare evasion or passengers without tags. In situations where multiple passengers enter a vehicle simultaneously, overlapping responses can overwhelm the RFID reader, where the system might fail to determine the exact number of passengers or differentiate between closely spaced tags.

2.3. Infrared Sensors

Infrared (IR) sensors are commonly installed near doors or on turnstiles to detect when passengers pass through an entry or exit point [6,7]. These sensors work by detecting the interruption of infrared beams as passengers move through the sensor field.

IR sensors are cost-effective and can provide a reasonable estimate of passenger flow. Regarding limitations, infrared sensors are susceptible to inaccuracies when passengers pass through in groups or when objects (e.g., luggage, strollers, or umbrellas) block the sensors. They also cannot differentiate between boarding and alighting passengers without directional sensors.

Besides infrared sensors, infrared cameras provide reliability in counting passengers in public transport under various environmental conditions, because they detect the heat emitted by objects and people in their field of view. Since people have a body temperature higher than the surrounding environment, they stand out clearly in infrared imaging. It can be used for tracking the movement trajectory of passengers so it can differentiate between boarding and alighting passengers and, therefore, detect the change in passenger number within the vehicle. Since infrared cameras do not rely on capturing visual or personal data like regular cameras, they are more privacy friendly. They detect only the heat signatures of individuals, making them compliant with privacy regulations, such as GDPR.

Infrared cameras may not perform as well in environments where the temperature variation is minimal, for example, in extremely hot weather when the contrast between passengers and the environment may be minimal. In situations where passengers are packed closely together, such as during rush hours, multiple people's heat signatures may overlap and produce inaccurate passenger detection.

2.4. Wi-Fi and Bluetooth Tracking

Wi-Fi and Bluetooth-based systems have been developed to track passengers by detecting signals from their mobile devices. When passengers with Wi-Fi [8] or Bluetooth-enabled devices [9] are near sensors, their presence is registered. If it is based on a Wi-Fi probe request, it sends out probe requests containing its MAC (Media Access Control) address and signal strength to identify the presence of a passenger and approximate their

location relative to the receiver of the signal. Bluetooth-enabled devices periodically emit signals when they are set to a discoverable mode or when scanning for other devices, and receivers capture these signals to detect passengers. To protect user privacy, modern devices use randomized MAC addresses in their probe requests. Instead of broadcasting the actual hardware MAC address, the device sends out a pseudo-randomly generated MAC address during probe requests. These randomized addresses change frequently, for example, every few minutes, and are unique to each network or detection session, making them difficult to track across multiple locations.

The strength of this system is that it provides continuous, non-intrusive data and can estimate crowd sizes in real-time, making it suitable for tracking passengers without direct interaction. However, since not all passengers carry devices with enabled Wi-Fi or Bluetooth, the results are often inaccurate. Moreover, concerns around privacy and data security have emerged, particularly in regions with stringent data protection regulations like GDPR.

2.5. LiDAR

LiDAR (Light Detection and Ranging) technology uses laser pulses to detect and map objects in each area. In passenger counting applications, LiDAR is used to create 3D models of passenger movements, allowing for highly accurate detection of individuals entering or exiting a vehicle [10].

LiDAR offers high precision in complex and crowded environments and is less susceptible to occlusion compared to infrared systems. It can also differentiate between individuals and objects. The high cost of LiDAR systems makes them less suitable for large-scale deployments, and they may require significant processing power to analyze the data in real-time.

2.6. CCTV with Image Processing

Closed-circuit television (CCTV) cameras equipped with image processing algorithms have become a prominent solution for passenger counting [11]. These systems use video feeds to detect and count passengers, employing computer vision techniques to identify individuals within the camera's field of view.

CCTV-based systems provide a non-intrusive and scalable solution for passenger counting, allowing for real-time monitoring and high accuracy when combined with machine learning algorithms. Challenges include ensuring accuracy in crowded environments, handling occlusions, and complying with privacy regulations, especially in regions governed by strict data protection laws like GDPR. Frequent problems with the use of CCTV cameras are related to the angles where the cameras are directed, so to cover the entire space of the vehicle, multiple cameras must be used, or wide-angle cameras must be used.

2.7. Machine Learning and Deep Learning Approaches

Recent advances in machine learning, particularly deep learning, have revolutionized passenger counting by enabling systems to automatically learn and adapt to different environments. Neural networks, especially convolutional neural networks (CNNs), are used to analyze video data from CCTV cameras and other imaging systems, providing highly accurate passenger counts even in complex, dynamic settings [12].

Deep learning algorithms can handle crowded environments and complex passenger behaviors, making them more robust than traditional methods. They also improve over time as they process more data. These systems require significant computational resources and large datasets for training. Moreover, privacy concerns and data protection regulations are critical issues when deploying AI-driven surveillance systems in public spaces.

2.8. Thermal Cameras

Thermal cameras detect heat signatures emitted by passengers, allowing for counting without the need for visible light. These systems are particularly useful in low-light or

night-time conditions and can differentiate passengers from inanimate objects based on heat [13].

Thermal cameras are highly effective in varying lighting conditions and can maintain privacy by capturing only heat data without identifying individual passengers. However, thermal cameras may struggle in extremely crowded environments or when passengers are close together, making accurate differentiation more difficult.

2.9. Ultrasonic Sensors

Ultrasonic sensors use sound waves to detect the presence and movement of passengers. These sensors are often installed in doorways and can be used in conjunction with other systems, such as infrared sensors, to improve counting accuracy [14].

Ultrasonic sensors are affordable and reliable in simple environments where there is limited congestion. Their accuracy diminishes in crowded areas or when passengers move unpredictably. They also may fail to detect passengers carrying large objects that obscure the sensor field.

2.10. Weight Sensors and Sensor-Grid Mat

Weight sensors built within public transport vehicles can monitor passenger numbers by measuring changes in a vehicle's weight as passenger board or alight. They function by detecting the total weight variance in a vehicle. Combining two types of sensors, infrared and weight, proved to be effective and can reduce the error rate to 17.5% [15]. The actual weight of the vehicle can be calculated from the driving dynamics, traction, and energetic data of the vehicle and the track geometry. Calculation of the total weight and estimation of the number of passengers at specific capacity utilization with sufficient accuracy was performed on a tramline in Budapest [16]. An alternative way of using the existing weighing system installed in modern trains, primarily used for braking control, was in research that offers a comparative analysis of passenger counts obtained through the weighing technique [17]. This technique was compared to infrared sensor accuracy specific to urban trains in Copenhagen and proved to be more accurate. A method that leverages human body kinematics to analyze passenger movement patterns was proposed in the research that combines Support Vector Machines [18] and proves that this combination of concepts reaches the precision of 93.98% to recognize if the passenger is getting off or getting on the train. Weight sensors do not differentiate between various factors such as the weight of luggage or goods, which can negatively affect the accuracy of occupancy count.

Sensor-grid mats utilize an array of sensors arranged in a grid pattern to detect the presence of passengers as they step on the mat. They offer collecting data on passenger volumes and entry and exit patterns without requiring extensive infrastructure changes [19].

Sensor-grid mats are cost-effective and simple to install, offer real-time data collection and are relatively compact and can fit into limited spaces at the entrances of public transport vehicles. On the other hand, in cases when a public transport vehicle is full, passengers standing on the mat can lead to inaccurate readings. Vibrations during travels and cases when there are multiple people stepping on and off the mat simultaneously can result in erroneous counts. Since sensor-grid mats rely on pressure sensing, one of the challenges is the detection of passenger movement direction, which is related to counting boarding and alighting cases and has a negative effect on the precision of detection.

Each technology employed for passenger counting presents unique advantages and challenges, with some being more suitable for specific environments or use cases. Traditional methods like card swiping and infrared sensors provide basic functionality but face limitations in accuracy and comprehensiveness. In contrast, modern approaches, including image processing and machine learning techniques, offer more sophisticated and scalable solutions. However, these advanced systems raise concerns related to cost, complexity, and privacy. As public transport systems continue to evolve, the integration of multiple sensing modalities, such as LiDAR, thermal cameras, and RFID, coupled with advances in artificial

intelligence, holds promise for creating highly accurate, reliable, and compliant passenger counting systems.

3. Literature Review

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.

3.1. Manual Counting of Passengers

The basic form of passenger counting is performed by observers at defined locations, or it is carried out based on the methods that depend on the number of tickets sold [20]. Identifying the number of passengers can be performed by extracting card swiping data as well. Neither of the two methods can give accurate data due to the traffic dynamics as well as the large number of passengers since they depend on factors such as fatigue of the observer, their expertise and experience, engagement, weather conditions, part of the day, or the fact that some passengers do not use cards for public transport. To maximally reduce the human factor, APC systems are being increasingly used.

3.2. IR Sensors

A wide range of technologies is already available for APC systems, like IR sensors in combination with the already installed sensors on suspensions in buses [16]. The data about the pressure from the vehicle suspension are gathered to calculate the additional vehicle mass, thus, assuming the number of passengers. The disadvantage of this method is the necessity of assuming the average weight of passengers, which can result in wrong results if children and people traveling with heavier things are considered, but in combination with an IR sensor this error can be reduced. If only IR sensors and machine learning techniques are used [7], the achieved passenger counting accuracy is around 80%, and there is still a lot of room for improvement. Another case of measurement of pressure inside the vehicle suspension system in trams to determine or estimate the number of passengers is described in paper [17]. The estimate of the passenger number according to the mass in trams with many passengers and uncoordinated entry/exit from the tram was determined. The mass of a single passenger is taken as an average value per gender. The concept is that the total mass of the vehicle is calculated via dynamic ride, haul, and vehicle energy data. This is a concept based on an existing system of measuring pressure in the railway vehicles' braking system [18] or measuring pressure between the floor and the foot of the passengers [20].

3.3. CCTV Cameras

One approach based on a single camera with an image resolution of 640×480 pixels mounted over an access gate in the bus and 40 min video [21] proved that it could be used in cases when only one door on the vehicle exists. It was based on the C++ programming language and the OpenCV library [22] with two artificial neural networks, one for detecting a person and the second that detects if the person enters or leaves the public transport vehicle. The implemented system does not require additional mechanical parts, but the processing unit proves to be expensive. The challenges in this research were related to detecting passengers in situations when the position of the passengers was static, when the color of the background and person's clothing were similar, when the noise and background in the image did not allow for highlighting of the object of interest, and in the event of a large crowd at the entrance.

An alternative solution based on one camera installed in a bus was presented in paper [23]. The counting system was based on feature-point tracking and an online clustering framework to count passengers in complex situations with heavy occlusion and crowding.

It employs a Kanade–Lucas–Tomasi (KLT) tracker [24] as an approach to feature extraction aimed at addressing the traditional image registration techniques by using spatial intensity information for efficient matching and for feature detection, using a unique clustering algorithm based on the appearance and disappearance of feature points. Potential issues with this approach include cases when people are carrying children together, which require an improved camera setup and calibration.

Another paper, Ref. [25] presents a real-time passenger counting system for buses using dense stereovision. It extracts information through a stereo-matching procedure and tracks passengers' heads to count them accurately. This approach proved to be effective in crowded situations, but it requires having many cameras installed to cover the whole space for passengers and can produce inaccurate results in cases where passengers wear hats, etc.

A combination of top-view vision and depth images is captured by a video-plus-depth camera mounted on the ceiling as described in paper [26]. Although the system was not tested in public transport vehicles, in areas where passengers are moving, the concepts used showed that they can be applied to vehicles as well. With a morphological operator for the elimination of the optical noise, human objects are drawn out, and the trajectory of the detected object is determined by the application of the algorithm that uses the mentioned combination of top-view vision and depth images. A similar method was described in paper [27], which uses an RGB-D video and zenithal camera in the bus door, capturing the passenger's flow. By combining the RGB image and depth image, it was possible to detect the heads of passengers and track their movements. Very crowded environments and wearing hats proved to be a challenge with this system and requires optimization of the algorithm.

A single camera approach mounted at the front of the bus and directed towards the area where the passengers are sitting is described in paper [28]. The system counts passengers by detecting skin color and performing morphological conversion operations from RGB to an HSV color model, using segmentation and thresholding techniques, removing noise and unwanted objects and smoothing of the image. The disadvantages of this solution are the difficulty of counting passengers at the back of the bus, but also the impossibility of detecting passengers in the case of a full bus and the overlapping of passengers, which makes it impossible to correctly detect passengers.

3.4. Wi-Fi Tracking

A Wi-Fi-based bus passenger counting system was described in paper [29]. The system utilizes Wi-Fi signals to estimate the number of passengers on a bus by tracking their devices. It overcomes the challenges posed by MAC address randomization and does not require passengers to take any action or connect to an access point, thus, preserving their privacy. The system's accuracy is highly dependent on the number of devices with active Wi-Fi interfaces.

Another Wi-Fi-based solution for counting passenger load in a bus was presented in paper [30], where detection of the periodic network probing activity from Wi-Fi devices built into smart phones was performed. Although the experiment ran for approximately 70 min, 1390 unique MAC addresses were recorded, and a total of 12,639 packages were monitored during the experiment. The paper proved the existence of Wi-Fi activity that relates with the number of waiting passengers and the number of riding passengers. As with similar solutions, the problem with this approach is that some passengers may not be using mobile devices or may not have Wi-Fi connectivity enabled. A similar approach was described in the paper [8], but with an estimation of the number of passengers based on the EM (Expectation–Maximization) algorithm. Data privacy was handled by using randomization of the MAC address of the device. An alternative approach to the detection of mobile devices was described in paper [31]. The detection was performed by recording the version of the operating system and producer with other information about the device, like the MAC address. The device is detected upon entry into the observed area, and it is performed in real-time. An example with four Wi-Fi sensors deployed on the bus was

used in paper [32]. The results proved that the overestimation problem due to interference from external devices can be mitigated considerably by combining probe request data from several Wi-Fi sensors. An alternative estimation based on Wi-Fi is described in paper [33], since estimations can never be totally accurate due to situations where some people may not carry devices with a Wi-Fi interface or the interfaces are not enabled, some people may have more than one device, devices located in nearby spaces can be detected, the length of stay in the room may be insufficient for the detection of a number of devices, and some transmissions may not be detected. The method indicated a certain correlation with the ground truth number of people in the environment.

An innovative approach was deployed and researched [34], where an evaluation of live data from a co-deployed light sensor system and WLAN probe sensor system 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 are 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. The ground truth dataset consists of enter and exit events manually counted by a person per door riding the bus. It was proven that a light sensor performs better in following ground truth data in detail in the short term while suffering in the long term. The WLAN estimator is not good at following ground truth in detail in the short term, but in the long term it follows the general tendencies of ground truth, and there is still a lot of room for improvement.

A Wi-Fi-based APC system, named iABACUS, was used to observe and analyze urban mobility by tracking passengers throughout their journey on public transportation vehicles [26]. 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 system described also tracks passengers throughout their journey on public transportation vehicles, providing information regarding 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 that is six times higher in a 15 min time interval.

3.5. CCTV with Machine Learning

Besides using only available video material obtained from public transportation vehicle cameras for artificial neural network training, research in [10] proved that a prepared dataset like the PAMELA (Pedestrian Accessibility Movement Environment Laboratory of University College London) [35] metropolitan train and bus dataset that contained videos of the passengers getting in and out of the train could be used to train a machine learning model. The described approach was implemented in MATLAB (version 2022a) and used edge analysis-based techniques for tracking and detecting objects, spatiotemporal techniques for selecting the line from the region of interest, motion detection-based techniques to detect the of a 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, which resulted in multiple counting in

some situations. There were situations where the system did not detect passengers due to occlusion, complex interaction, or shadow, which were sometimes detected as humans. The opening of the door gave false detection, also as the foreground region is extracted on the motion-based method.

A simulation of the public transport environment, like that performed in the PAMELA dataset configured as a London Underground carriage to study the effect of door width and vertical gap, was employed and researched in [36]. The PAMELA-UANDES dataset [37] 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 that only apply to the London Underground.

Instead of using only one surveillance camera in public traffic vehicles, better accuracy can be achieved using rear and front door cameras as researched in [38]. This research employed training data that were extracted with the time intervals set to 3 s 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 testing system, 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 (an 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 the 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 an RGB-D sensor (Asus Xtion Pro Live, Taipei, Taiwan), located over each bus door, for counting adults and children, is described and researched in [39]. It represents an approach and an inexpensive and flexible solution to obtain, in real time, statistical measures on the number of people present in the bus with the use of an analytical processing system that accesses the data stored in the database and extracts 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, which was used in Asus Xtion Pro Live and is based on the 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 (Asus Xtion Pro Live, Taipei, Taiwan) 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, allowing people to count only when the bus arrives at the bus stop and the doors open, and it is automatically deactivated when the driver closes the doors and restarts the autobus. The skeleton-tracking algorithm provided by NiTE means 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 on or off the bus. The system was tested only during the day with the total recording time for each camera of about 10 min and did not include situations in which the bus is full of passengers, and so the depth sensor does not have a clear path to detect separate complete skeletons of every person.

An example of a passenger counting algorithm that uses a hybrid machine learning approach and utilizes the Histogram of Oriented Gradients (HOG) [40] is described and analyzed in [41]. The images were extracted from the recorded video from the public transport vehicles. Classification of head features is performed by using a Support Vector Machine (SVM) [19] as a classifier for the liner model. Heads are detected successfully after performing all steps. In the next step, Kanade–Lucas–Tomasi (KLT) [24] is used for reality head tracking; the multiple target tracking is achieved, and the head motion trajectory of the passenger target is captured. In the last step, the proposed algorithm is moved to the embedded system, ADSP-BF609, for practical implementation. Embedded systems were proven to be adjusted for general purpose passenger detection and could not be easily upgraded to include situations in which passengers use umbrellas in instances of rain or in similar situations.

Various APC systems were analyzed and presented and researched in [42]. Video-based sensing is based on two cameras mounted on the front and rear doorways of buses. RGB image frames were captured at a 640×480 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) [43]. 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, which estimates the number of devices in a defined space by analyzing the probe request messages sent by devices when they are not connected to a 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 the 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 APC solution, the achieved results were more accurate on the front door but had a clear decline in accuracy for the rear door. Three-dimensional 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 the sensor map resulted in generally strong results on the rear door, but this was 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 the standing room was limited. The measurements were taken during six days in December in 2018, so to obtain the most accurate results, it is 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 an ESP-32 Wi-Fi camera and is based on the YOLO v8 algorithm [44,45] in real-time is described and researched in [46]. The dataset that was used was the COCO (Common Objects in Context) dataset [47] that contains very useful resources for establishing 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 data provided about the number of passengers in certain public transport vehicles with an

addition of congestion 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 [48]. 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 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 a Raspberry Pi with a camera with a total cost of approximately AUD 560. Considering the significant cost and the limitations of customizing commercial APC systems, which directly impact the finances of public transport companies, and considering the accuracy of the systems, custom low-cost alternatives are a suitable option.

One of the options for counting passengers is the focus on counting heads of the passengers. Head count methods were tested on density maps with 10,800 images [49]; the results show a mean absolute error of 1 head per frame, equivalent to 11% relative error. Researchers [50] managed to achieve 99% accuracy and 0.041 s recognition speed based on three public head datasets, even though only the CPU (without the GPU) was used. With this simplification of the detection algorithm, focusing only on the part of the object being detected, the performance and computational complexity of the algorithm can be reduced without significantly reducing accuracy.

One additional solution [51] is based on tracking trajectories of the passengers and a large-scale bus passengers' dataset, which was annotated for detection and tracking, and counting proved to demonstrate performance, meeting the requirements in terms of counting precision and speed. The resulting precision was improved by the addition of DeepSORT tracker and YOLOv5 detector and reached 96.6% and proved that combining different modules and technologies positively affects the quality of passenger detections.

Table 1 contains an overview of 43 existing studies related to passenger counting systems in public transport, from buses, metros, trains, and general concepts tested in laboratory conditions. The table includes card swiping technologies, infrared technologies (IR), weight change detection, RFID sensors, systems using Wi-Fi technology, Bluetooth technology, CCTV cameras, LiDAR technology, machine learning (ML), thermal cameras (TC), or ultrasonic sensors (US). The 'X' symbol represents the usage of a specific concept in the APC system analyzed. According to Table 1, it can be concluded that, considering the development of machine learning technology and the presence of cameras in public transport vehicles, an approach with image processing was often used, mostly in buses, which will be the focus of the rest of the research. In addition to the analysis of existing state-of-the-art research, the continuation of the research also describes the proposal for the implementation of a system for counting passengers in public transport based on concepts that achieved the best counting accuracy in other studies. Although this proposal has not yet been implemented, the research contains guidelines and recommendations for implementation, including the best practices, with an emphasis on avoiding problems encountered in other system implementations.

The following chapter provides information on the concepts, techniques, and challenges involved in planning, implementing, and testing a public transport passenger counting system based on the use of CCTV cameras and machine learning. It analyzes several different approaches, highlights pros and cons of each 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. This is followed by chapters related to privacy issues and restrictions related to the GDPR. The last two chapters are related to the discussion and conclusion.

Table 1. Overview of papers (reference number) related to APC systems.

Ref. nr.	PT Vehicle Type	Used Technology										
		Card Swiping	IR	Weight	RFID	Wi-Fi	Bluetooth	Cameras	LiDAR	ML	TC	US
[2]	Bus and Metro	X										
[3]	Metro	X										
[4]	Bus				X							
[5]	-				X							
[6]	Bus, Subway trains		X									
[7]	-		X									
[8]	Bus					X						
[9]	Bus						X					
[10]	Train								X	X		
[11]	Train, Bus							X		X		
[12]	Bus							X		X		
[13]	Car							X		X	X	
[14]	Bus											X
[15]	Bus		X	X				X				
[16]	Tram			X								
[17]	Train		X	X						X		
[18]	Bus			X						X		
[19]	Bus			X				X		X		
[21]	Bus							X		X		
[23]	Bus							X				
[24]	Bus							X				
[25]	-							X				
[26]	Bus							X				
[27]	Bus							X				
[28]	Bus							X				
[29]	Bus					X						
[30]	Bus					X				X		
[8]	-					X						
[31]	Bus					X						
[32]	-					X						
[33]	-					X						
[34]	Bus					X		X		X		
[26]	Bus					X						
[35]								X		X		
[36]	Bus							X		X		

Table 1. Cont.

Ref. nr.	PT Vehicle Type	Used Technology										
		Card Swiping	IR	Weight	RFID	Wi-Fi	Bluetooth	Cameras	LiDAR	ML	TC	US
[37]	Bus							X				
[38]	-							X		X		
[39]	Bus							X		X		
[47]	Metro							X		X		
[48]	Bus							X		X		
[49]	Bus							X		X		
[50]	Bus							X		X		
[51]	Bus							X		X		

4. Methodology

4.1. Introduction

The methodology of this research involved analyzing existing implementations of passenger counting systems in public transport by examining the technologies used in the studies and the types of vehicles where the counting was implemented, which is shown in Table 1. The analysis also included detection using multiple technologies at the same time, such as image processing and machine learning or combining several different sensors at once, which resulted in improved precision in the detection and counting of passengers. Each of the existing 43 state-of-the-art studies is listed as a separate line in Table 1, with the reference number, type of public transport vehicle, and the technology used. In these studies, card swiping technologies were used (on two occasions [2,3]), infrared technologies (IR—on four occasions [6,7,15,17]), weight detection sensors (on five occasions [15–19]), RFID chips (on two occasions [4,5]), a Wi-Fi network on which connected passenger mobile devices (in 9 papers [8,26,29–34]), short-range Bluetooth technology (in one paper [9]), CCTV cameras (in half of the papers, twenty in total [11–13,15,19–28,34–39,47–51]), LiDAR sensors (in one paper [10]), thermal cameras (in one paper [13]) and ultrasonic sensors (in one paper [14]). By far most works described the use of CCTV cameras, 23, of which 15 combined with machine learning, and in 29 to 43 cases it was implemented in buses, based on which the focus of this research was determined. Using the conclusions and best practices from those works, recommendations for the implementation of systems in buses based on the use of cameras and machine learning are given in this paper.

4.2. Precision Comparison

Table 2 shows a comparison of the precision of each of the mentioned technologies. The best results are achieved by public transport passenger counting systems based on CCTV cameras that use machine learning. Some technologies such as RFID, Wi-Fi, or Bluetooth require passengers to wear certain types of devices or sensors, which can significantly affect accuracy. Adding the use of machine learning has a positive effect on accuracy and can solve problems in challenging situations such as overlapping passengers or congestion in public transport vehicles. Environmental factors such as weather, lighting, sound, and passenger density also affect measurement accuracy, which must be considered when designing a system for APC in public transport vehicles. For this reason, in the newly implemented systems, several different technologies and sensors are used simultaneously, which are partially complementary, to achieve an even higher level of precision in passenger counting.

Table 2. Comparison of different technologies used in APC systems.

Technology	Typical Precision Range	Description
Card swiping data [2]	70–90%	It achieves the best results when passengers enter the vehicle. In most cases it is not used when passengers leave the vehicle and in cases where passengers use cards.
Infrared Sensors [15,16]	80–95%	The best precision is achieved by placing the sensors in positions near the vehicle doors where it is possible to best detect the movement of passengers. It achieves the best results in situations where several passengers overlap, but inadequate results in large crowds.
Weight Sensors [19]	80–90%	It achieves the best results when calculating the total weight of passengers, but it achieves inadequate results when individual passengers enter and exit. It is based on the average weight of passengers, which sometimes varies and is not always reliable.
RFID Technology [9]	82–98%	It achieves high precision if all passengers use RFID devices. Accuracy is worse in the case when device signals overlap, or passengers use RFID devices.
Wi-Fi Technology [8,24]	75–94%	It provides a rough estimate of the number of passengers, as it is assumed that passengers will have devices that connect to the Wi-Fi network and only have one device. Inadequate results are also achieved due to signal interference.
Bluetooth technology [7]	73–77%	Like Wi-Fi technology, accuracy depends on the number of Bluetooth devices and the number of devices per user.
CCTV Cameras (without ML) [22,27]	75–97%	Accuracy depends on manual counting or motion detection, which can result in low accuracy in case of overlapping passengers or objects, and problems such as insufficient image quality.
CCTV Cameras (with ML) [11,26,32]	92–99%	In the case of using CCTV cameras and machine learning, a significant improvement in accuracy is achieved. The biggest challenges are based on the lighting of the image, the camera angle or large crowds of vehicles.
LiDAR sensors [8]	95–96%	A very high level of accuracy is achieved in counting and tracking passengers, even in large crowds or in different lighting conditions inside the vehicle.
Thermo Cameras [12]	70–98%	It is based on the detection of the heat of the passenger's body and achieves the best results when visibility is reduced, but most challenges exist when there are large crowds in vehicles, during which the bodies of the passengers overlap.
Ultrasonic Sensors [13]	88–89%	It achieves good results in detecting the entry and exit of passengers and has the most challenges in crowded and noisy environments.

4.3. Computation and Real-Time Performance Comparison

Table 3 shows a comparison and analysis of the requirements for computing resources, real-time performance, and computational complexity of each of the technologies used in APC systems.

The least computational complexity in data processing algorithms during passenger detection is required by systems based on infrared and ultrasonic sensors and systems that support card swiping. These systems are suitable for use in real-time systems.

Moderate requirements for computer resources are present with RFID, Wi-Fi, and Bluetooth technologies, because they are more sensitive to interference and overlapping with other devices, which can result in the necessary fine-tuning of algorithms and additional time required for data processing. The additional time required for data processing results in a potential delay in passenger count results, which can generate delays in real-time systems as well.

The most demanding technologies in terms of computer complexity and hardware are LiDAR and thermal cameras that use machine learning, because they require specialized hardware such as GPUs to be able to process a large amount of data in a short time, which enables the use of such systems in real time.

In the case of using CCTV cameras and machine learning, it also requires considerable computing resources, such as GPUs or specialized hardware accelerators such as TPU (Tensor Specific Unit). Given that the accuracy of passenger counting depends on image quality and external influences such as lighting or crowd levels, it is necessary to further adjust the developed algorithms to achieve better accuracy.

Table 3. Computational and real-time performance of technologies used in APC systems.

Technology	Computational Complexity	Real-Time Performance	Hardware Requirements
Card swiping data [2]	Low	High	Low (simple card reader and server)
Infrared sensors [5]	Low-Medium	High	Moderate (infrared sensors and microcontroller)
Weight sensors [15,17–19,40]	Low	Medium	Moderate (weight sensor array)
RFID technology [3,9]	Low-Medium	Medium-High	Moderate (RFID readers and tags)
Wi-Fi technology [8,24,30,31]	Medium	Medium	Moderate (Wi-Fi receivers and data processing units)
Bluetooth Technology [7]	Medium	Medium	Moderate (Bluetooth receivers and processors)
CCTV cameras (without ML) [20,22,35]	Low	Medium	High (basic CCTV camera and server)
CCTV cameras (with ML) [11,26,40,46]	High	Medium	High (GPU-enabled server or edge AI hardware)
LiDAR sensors [8]	High	High	High (LiDAR sensor and GPU or high-end processor)
Thermo cameras [12]	High	Medium-High	High (thermal camera and GPU or real-time analysis)
Ultrasonic sensors [13]	Low	High	Low-Moderate (ultrasonic sensor array)

According to the data in Table 3, technologies such as LiDAR and CCTV with the use of machine learning are confirmed as the most robust and effective for use in a system for counting passengers in public transport, but at the same time they require the most computing resources and the most powerful hardware, which also affects the cost of implementation.

5. Concepts, Techniques, and Challenges in Case of Using Cameras and Machine Learning

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, which isolates moving objects (passengers) from static backgrounds. Furthermore, feature extraction techniques are employed to identify and highlight crucial elements within the images, such as edges, shapes, and textures.

Machine learning (ML) in image processing for passenger counting is a powerful approach, especially when high accuracy is essential in dynamic or crowded environments. ML techniques and models can interpret complex patterns, identify individual passengers in various settings, and adapt to changing conditions [38]. There are several machine learning concepts suitable for use in APC systems described in this section.

Convolutional Neural Networks (CNNs) are designed for image and video data processing that use convolutional layers that apply filters to detect features like edges, shapes, and textures [52]. They support detecting and classifying individuals even with background noise or occlusions. Using a trained CNN, each frame captured by a camera is processed independently to detect and count passengers. Detection of individual passengers in complex scenes is accomplished by using models like Faster R-CNN use RPNs (Region Proposal Network) to locate people in crowded environments [53]. Pre-trained

CNN models (e.g., ResNet, VGG) represent fine-tuned datasets of specific public transport settings to enhance accuracy with limited data [54].

Recurrent Neural Networks (RNNs) [55] and Long Short-Term Memory (LSTM) [56] networks are capable of handling sequential data and learning temporal dependencies. They are valuable when analyzing sequences of frames in video data to track the flow of passengers. RNNs and LSTMs are useful in video-based passenger counting where frame sequences are analyzed to monitor passenger movements, such as entry and exit patterns, rather than isolated frame counting. LSTMs are used to analyze the temporal relationship between frames, capturing patterns of passenger movement. Integrating LSTM with CNN layers allows the passenger counting system to predict and adjust to crowd patterns over time, aiding in real-time crowd management.

YOLO (You Only Look Once) [44,45] is a real-time object detection model that processes images as a single pass, making it much optimized than region-based methods like R-CNN [53]. YOLO's grid-based approach predicts bounding boxes and class probabilities simultaneously, optimizing speed and accuracy. That makes it highly suitable for real-time passenger counting in public transportation systems and when a high frame rate is necessary. It can be used to detect passengers in live video feeds providing near-instantaneous counts.

Semantic segmentation [57] divides an image into different segments, assigning each pixel to a class (like passengers or background). Models like U-Net [58] and DeepLab [59] can be used for pixel-wise classification, which is useful in highly dense and occluded environments when individuals need to be segmented precisely and distinguish people from backgrounds at a pixel level. Therefore, semantic segmentation can estimate the number of passengers even if they are tightly packed. It can be used for heatmaps that highlight high-density zones within a vehicle.

Optical flow analysis algorithms [60] analyze the motion between consecutive frames in a video and compute the velocity of objects, which can be used to track passenger movements and count entries or exits. They are useful in applications focused on movement patterns, such as monitoring crowd flows in and out. By tracking motion near entry and exit points, optical flow can help identify irregular movement patterns, which can be used in high-density transportation systems.

Background subtraction algorithms [61] separate the foreground (like passengers) from the background by learning the static background and highlighting dynamic elements. When integrated with ML, they can adapt to gradual changes in lighting and background shifts. They are suitable for fixed-point passenger counting where the camera is stationary, and background subtraction is effective in relatively stable environments like bus stops or entry gates. The system continually learns the background, which allows to detect passengers as they move through the frame. Combined with CNNs [52] or other classifiers, background subtraction can improve accuracy by reducing the false positives of non-passenger objects.

Transfer learning with pre-trained models uses large pre-trained models like ResNet [54], VGG [54], or MobileNet [43], which have learned basic image features. Fine-tuning these models on smaller datasets specific to public transport can yield high accuracy with reduced training data. It represents a good choice when dataset size is limited, such as in smaller public transport systems. Pre-trained models are adjusted to focus on unique environmental factors, like lightning or passenger density in a specific transport context. It can be used to adapt models to different public transport scenarios, like buses, trains, or stations with minimal retraining.

Support Vector Machines (SVMs) [62] and Decision Trees (like Random Forests) [63] are effective in simpler passenger counting scenarios where image features, like colors or texture are extracted first and then classified. Used methods are more appropriate for systems with lower computational resources or simpler counting tasks, such as distinguishing between "empty" and "occupied" section of a public transport vehicle. Extracted image features are classified using SVMs or Random Forests for low-complexity passenger

detection. It is used in simpler applications, such as determining if a bus is above or below a certain capacity threshold that can be combined with weight sensors.

The common concepts of counting passengers in public transport vehicles were based on defining areas of interest for counting or lines of interest where the passenger needed to pass when entering or leaving the vehicle. The images extracted from the video taken by the camera are then transformed into an appropriate format and analyzed by using artificial neural networks to detect the object of interest (for example, a person, vehicle doors, etc.) and to 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 [64,65]), features extraction [66], image segmentations, also known as thresholding [67], are used to create binary images to improve the accuracy of the object detection, performing morphological operations like dilation and erosion [68], color conversion (from RGB to HSV [69], 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 models.

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 [38] proved that the detection of open doors improves 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, and clothes worn by travelers change; moreover, day and night, different platform configuration, different types of lightning, different camera views, sunny and rainy days 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 datasets need to be collected, and with manual labeling of objects (passengers) on images, enhance and enrich the training dataset with new scenarios that were not recognized as expected.

Depth sensors were also utilized, like researched in [27], 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 boarding and alighting passengers. The slowest way consists of collecting your own dataset by storing videos from surveillance cameras throughout the year and 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 [10,34] 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%. The reason for such a division of the dataset lies in the fact that when training the model, it is necessary to include as many edge cases as possible so that the passenger detection results are as good as possible, and on the other hand, part of the dataset must be of adequate size to avoid overfitting.

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 v11 released on 30 September 2024) [70] and the COCO dataset [47], which contains a large number of images of objects that can be tracked, such as people (a limited person dataset is only available in [71]).

6. 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 to prepare images for training a machine learning model for passenger recognition. The section also covers the best practices for 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 rear door of the public transport vehicle.

Figure 1 shows the process that includes all the necessary steps to implement a system for APC in public transport. At the beginning, it is necessary to determine the positions of the cameras that will adequately cover the doors of public transport vehicles to detect passengers in the boarding and disembarking phase. The next phase is related to setting up the network infrastructure that will connect the cameras and the rest of the computer equipment. At this stage, it is necessary to pay special attention to security aspects and the preservation of data privacy. The cycle that can be repeated whenever it is necessary to improve the system includes the stages of collecting a dataset with passenger images and scenarios to improve the accuracy of the APC system as much as possible. This is followed by the stage of image or video processing and model training to improve the functionality of counting passengers as many times as possible until the desired precision is achieved.

6.1. Camera Locations

Capturing an optimal Field of View (FOV) is important for reliable passenger counting. If cameras are placed too high, they may miss smaller passengers (like children), if placed too low, they risk occlusion from other passengers, objects, or seats. High-traffic times and dense crowding can lead to occlusions, making it difficult for the camera to detect each passenger. Public transportation environments, such as buses, often have variable lighting due to natural light changes or low-light conditions at night, which can affect image clarity. In-vehicle cameras are subject to vibrations, acceleration, and deceleration, which can blur images or misalign cameras over time.

Using multiple cameras at different angles or strategic positions, like at entry and exit points, helps minimize occlusions and provides a view of the area, increasing the accuracy of passenger detection. Wide-angle lenses or fish-eye cameras can capture a broader area, reducing blind spots and helping to identify passengers across a wider FOV, even in crowded situations. Cameras with automatic calibrations can adjust for small shifts in position or orientation caused by vehicle movement. Infrared or low-light cameras can operate in a variety of lighting conditions, allowing for consistent passenger counting day and night.

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 transportation vehicle, as shown in Figure 2. In the case of a bus or similar public transport vehicle, cameras must be located on the ceiling to be able to capture video of the passengers' heads.

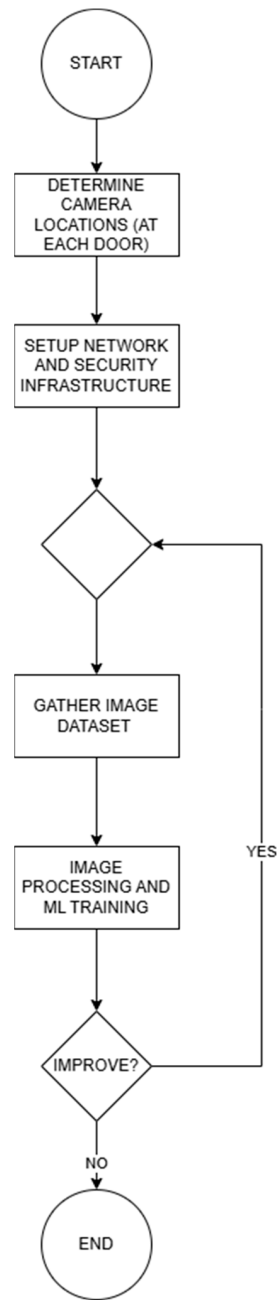


Figure 1. Process diagram for setting up and upgrading the APC system.

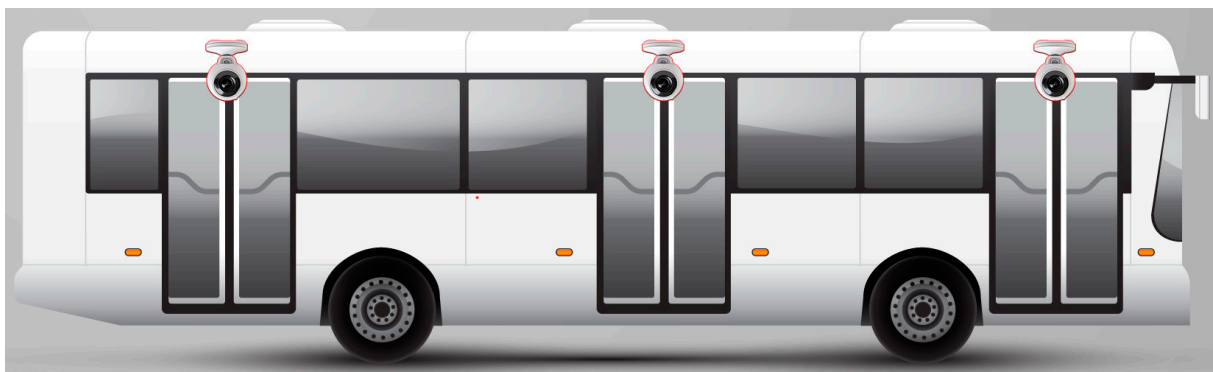
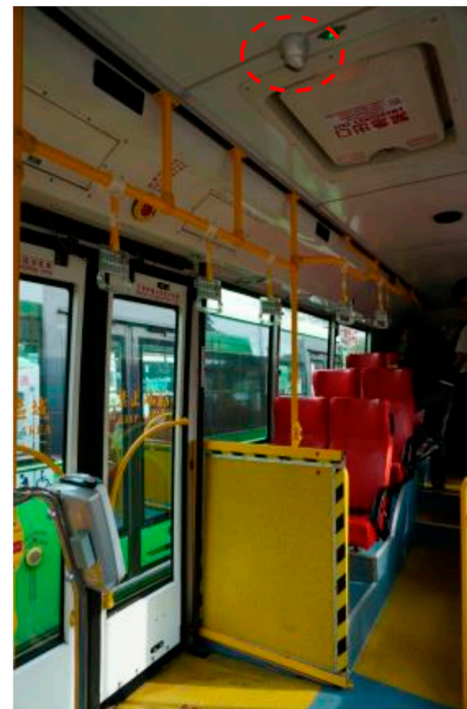


Figure 2. Proposed locations of cameras in public transport vehicles.

Examples of camera locations are shown in Figure 3. The first two images (a,b) show surveillance camera locations, and (c,d) show custom-made cameras in case of proof-of-concept solutions. The circles in the pictures indicate the positions of the cameras.



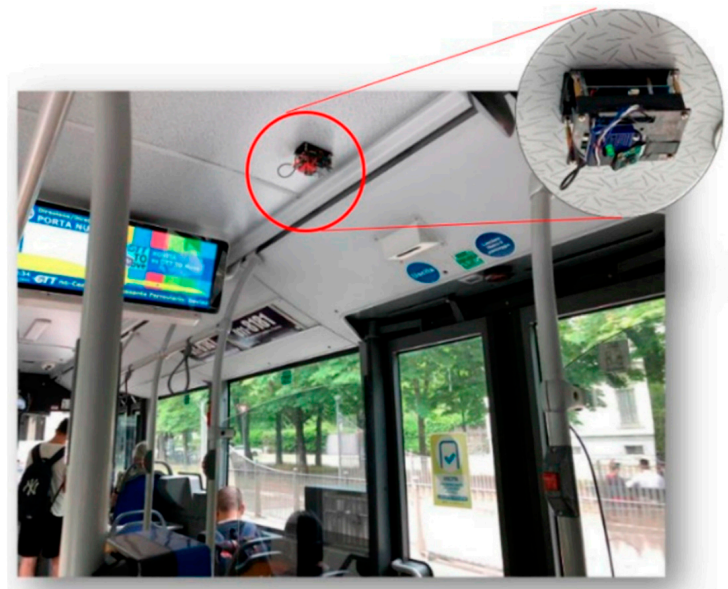
(a)



(b)



(c)



(d)

Figure 3. Positions of the cameras: (a) surveillance camera at the front door [26]; (b) surveillance camera at the rear door [26]; (c) custom-made camera at the front door [19]; (d) low-cost camera installation based on Raspberry Pi at the rear door [46].

6.2. Public Transport Vehicle External and Internal Network Infrastructure

High-resolution image data required bandwidth to be transmitted in real time, which can be challenging in areas with inadequate network connectivity or in moving vehicles with inconsistent connections, that can affect the timeliness of decision making. With passenger data being transmitted over networks, data security and privacy is critical, especially in jurisdictions with strict privacy regulations (like GDPR). Transmitting large volumes of data consumes significant power, which can be an issue in battery powered or low-power devices.

By processing data locally on edge devices, like onboard servers or AI-enabled cameras, only relevant information like passenger counts need to be transmitted, reducing bandwidth usage and enhancing data security. Edge processing also minimizes latency and is essential in real time for APC systems. Video compression algorithms, like H.256, reduce file sizes while maintaining quality, making it easier to transmit data without overloading network resources. APC systems can be set up to transmit only when connectivity is available or during low network traffic times, storing data locally otherwise. In those cases, Wi-Fi offloading in stations or depots can be used. Uploading data can be performed only when vehicles are idle and in a stable network environment. Secure encryption protocols, like AES-256, can protect data during transmission, ensuring privacy compliance and security. Authentication mechanisms like VPNs or private APNs (Access Point Names) are also used in more secure public transport setups.

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 [72,73]. Besides cameras, an embedded computer functioning as a Network Video Recorder (NVR) is required to allow access to recorded video materials. Bus surveillance systems must be wirelessly enabled with Wi-Fi routers and cellular compatibilities. After the IP cameras send video to the onboard embedded computer, the computer streams a live video through a secure cellular connection to a control center, where managers can access video images and monitor the bus remotely. Figure 4 shows an example of components and internal and external network infrastructure in a bus [73].

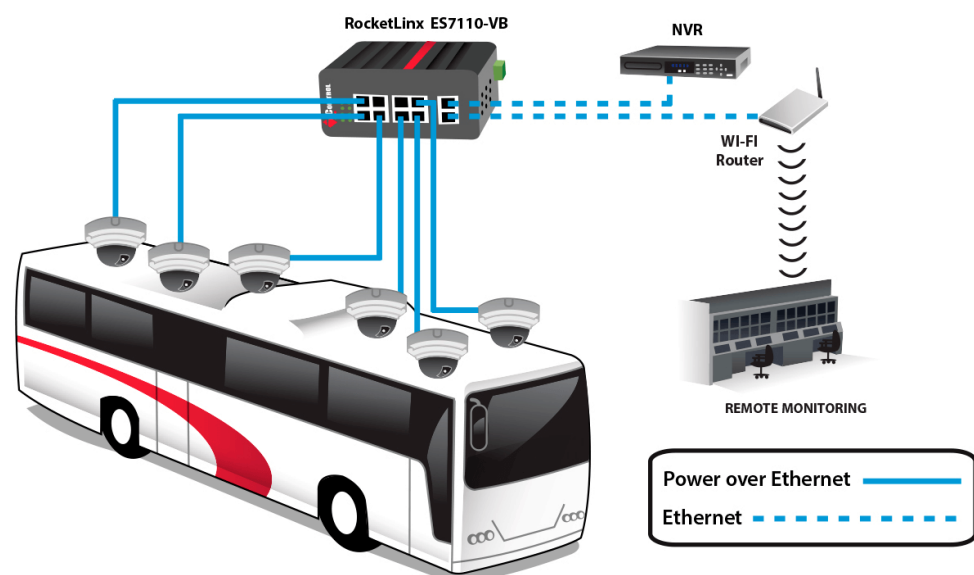


Figure 4. Public transport vehicle external and internal network infrastructure [73].

Besides using Wi-Fi, 5G technology as one of the mainstream options is also an alternative for establishing the infrastructure [74,75], due to its high speed, low latency, and enhanced connectivity. The APC implementation based on 5G technology enables optimized data transmission from sensors, like CCTV cameras, to a central system, where the information is processed and analyzed.

6.3. 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, which would be ideal. If this is not possible and it is necessary to start from the beginning, the prepared YOLO v11 [54] model that can recognize people in pictures and the COCO [47,71] dataset 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 v11, released on 30 September 2024, builds upon its predecessors with improvements. It leverages a refined architecture that enhances feature extraction and object localization capabilities. The latest version introduces significant improvements in architecture and training methods, making it a versatile choice for a wide range of computer vision tasks, especially in complex and dynamic environments like public transport vehicles [69,76].

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 [71].

Using YOLO v11 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 custom datasets collected from public transport environments. In situations where passengers overlap or when vehicles are crowded, the detection algorithms of YOLO v11 can more effectively distinguish individual passengers by counting their 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 [77].

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 5 shows some of such situations, which refer, for example, to situations when it is raining and when 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.

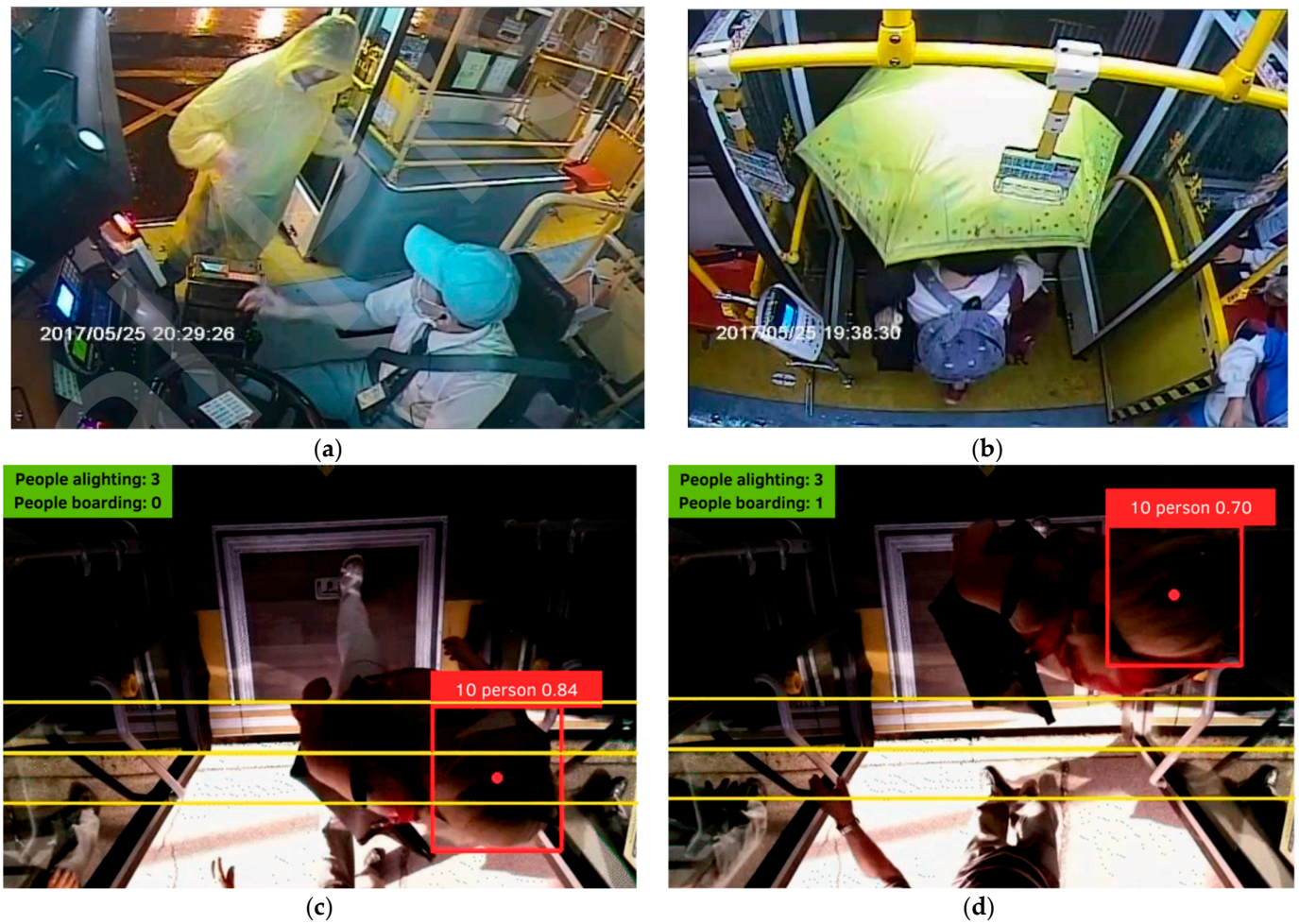


Figure 5. Images in specific scenarios: (a) passenger wears a raincoat at night [26]; (b) passenger uses an umbrella while it is raining at night [26]; (c) passenger alights from the bus and crosses the line of interest (yellow) in dark conditions [46]; (d) two passengers are boarding in dark conditions [46].

6.4. Image Processing and Machine Learning Model Training

Storing large volumes of high-resolution image or video data can strain onboard memory, especially when regulations require retraining data for specified periods for verification or analysis. Image-based systems inherently collect identifiable information, making passenger privacy a significant issue. Regulations in many areas require anonymizing or blurring faces to protect individuals' identities. APC systems often rely on machine learning models trained on labeled data. Collecting labeled image data is time consuming, particularly in environments with variable lighting, crowding, and different passenger demographics. Cameras need to perform well in varied environments, such as different vehicle types or station layouts. Constant changes in lighting, weather, and seasonal clothing can affect image clarity and model accuracy.

To protect passenger privacy, APC systems need to focus on detecting human shapes and counting individuals rather than capturing identifiable details. Techniques like face blurring or using thermal or silhouette imaging also enhance privacy. APC systems based on adaptive learning can continually improve their accuracy by periodically updating the model with new, labeled data, which can be useful in adapting to different transport environments and varying crowd densities. Combining image data with non-visual data, such as RFID or weight sensors, enables redundancy, allowing the system to fill in gaps when images are unclear or occluded. These hybrid systems provide robust, multi-faceted passenger counts. Onboard systems store and process data temporarily, purging it after

processing to reduce storage demand. Data can be helped only for a limited time to comply with privacy regulations and then can be automatically deleted once processed.

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 [68] are then used to refine the shapes of detected objects, removing noise and filling gaps. Color conversion, such as converting from RGB to an HSV color space [66], 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 [78], particularly those based on convolutional neural networks (CNNs) [79], are widely used due to their performance in image recognition tasks. Models like YOLO [70] 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 to maximize covering as many boundary cases as possible and to avoid overfitting like described before. 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 [80] from pre-trained models on large datasets like COCO [47,71] 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 are collected, particularly under previously unseen conditions, they are 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.

7. GDPR Compliance and Passenger Counting Systems

The General Data Protection Regulation (GDPR), enacted by the European Union in May 2018 [81–83], has significantly impacted how data, especially personal data, are 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 [84].

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 these 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 are collected and retained. Public trans-

port 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 [85].

The impact of the GDPR on the conditions for using CCTV cameras in public places such as public transport vehicles is significant in areas where it is applied, such as the area of the European Union, is significant and needs to be considered before implementing APC. The EU AI Act [86] defines regulations that ban information systems that scrape facial images from the internet or CCTV footage, infer emotions in the workplace or educational institutions, and categorize people based on their biometric data. Due to this fact, APC systems must not save any private biometric data about passengers in public transport, blur faces during video processing, use "privacy-friendly" datasets [8] for training, and anonymize collected data, which would avoid any violation of the rules described in the regulation.

8. 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.

First, it is necessary to emphasize the difference between the terms used to determine the quality of the APC system. The accuracy and precision represent two crucial factors that represent the quality of the APC system [87]. The analysis of the accuracy evaluated the capability of the APC system to count passengers well. The accuracy measures the systematic over or undercounting of passenger values relative to the "ground truth". Precision measures the distribution of errors between the measured and true value of passenger activity and is always evaluated according to a level of confidence. Although accuracy and precision refer to errors in the measurement, the fundamental difference is in the nature of the error. It was proved that the APC based on the IR sensor accurately measures alighting passengers, while it presents a slight tendency to systematically undercount boarding passengers.

One of the primary challenges is ensuring accuracy and reliability in diverse and dynamic environments [88,89]. 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 [90]. The need for robust preprocessing techniques, including foreground and background subtraction, feature extraction, and image segmentation, is critical to maintaining high detection accuracy [91]. Additionally, the integration of morphological operations, color conversion, and histogram equalization further enhances the quality of the images for subsequent analysis [92].

Machine learning models, particularly those based on deep learning architectures like YOLO, have demonstrated significant advancements in object detection capabilities [93]. However, training these models requires large, diverse, and well-labeled datasets [94]. The availability and quality of these datasets are important for developing models that can generalize well across different scenarios encountered in public transport [95]. The incorporation of transfer learning, using pre-trained models on extensive datasets like COCO [47], can accelerate the training process and improve initial performance [96]. 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 [84,86]. 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 are 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.

Based on the best practices that result with the highest precision, it is recommended to use 2 cameras installed in corners of the bus and more fisheye cameras in the central part of the bus [97], depending on the total size of the bus. The deployment of multiple cameras to cover all entry and exit points in a public transport vehicle is recommended to enhance detection accuracy and reduce false positives [26]. The strategic placement of cameras can ensure adequate coverage, especially in vehicles with multiple doors. Research [26] 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. The vertical and horizontal angle between the camera and the object depends on the resolution of the camera. The YOLO v11 model [69] provides a compromise between precision and time of evaluation. The dataset that consists of 470,000 unique human instances with an average of 23 persons per image and different levels of occlusions presents an optimal starting point for training the YOLO v11 model—The CrowdHuman benchmark [98]. Key Computer Vision tasks supported by YOLO v11 that fit the requirements of the APC system are object detection tasks that excels in identifying and localizing objects within images or video frames and pose estimation that is used to track movements or poses. Oriented object detection (OBB) allows more precise localization of rotated objects and objects tracking that traces the path of objects in a sequence of images or video frames that can be applied to passengers in public transport vehicles [99].

Benchmark analysis of APC solutions based on ten criteria: technology, accuracy, environment, coverage, interface, interference, robustness (for devices), price, pricing model, and system integration for companies [100] showed that most APC systems advertise accuracies of 95% or higher according to company datasheets. However, the interviews with local transport companies reveal a significant disparity between the accuracy of APC systems as advertised in product datasheets and their actual performance under real conditions, with a drop of 20–30% in real accuracy. Since the process of verifying the accuracy of APC systems is mainly conducted in the laboratory, with simulated passenger counts that may require passengers to behave specifically to achieve the best theoretical precision. Therefore, there is still room to improve the precision and accuracy of the APC system, and this research offers a proposal for implementing a system that can be continuously improved.

Crowded boarding situations represent one of the most challenging problems in APC systems [101]. The proposal uses the YOLO v8 [44,45] algorithm and modified ByteTrack algorithm for improved multi-tracking of passengers. It is built upon the strengths of SORT (Simple Online and Realtime Tracking) and provides the advantage of addressing the issue of tracking fragmentation in crowded scenes by effectively utilizing low-confidence detections. The training was performed on a CrowdHuman dataset [98]. The results proved to achieve high precision of passenger detection even in crowded situations. In future

works, the newest version of YOLO v11 [67] with compatible tracking algorithms will be utilized to target this challenge and provide an improved concept based on the most recent version of libraries, datasets, and best practices described in this manuscript.

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 [102], comprehensive training datasets [103], adherence to data privacy regulations [104], 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.

9. 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 use cases that combine several different sensors like CCTV cameras, Wi-Fi probes, LiDAR technology, etc. Additional sensors can provide complementary data that enhance the accuracy and reliability of passenger counting, especially in challenging conditions such as overcrowded vehicles.

With the introduction of improved anonymization techniques, protecting 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, like 5G, 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 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 are 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 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. and L.M.; methodology, A.R.; validation, L.M., G.Đ. and B.M.; formal analysis, L.M., G.Đ. and B.M.; resources, A.R.; writing—original draft preparation, A.R., G.Đ. and B.M.; writing—review and editing, A.R., L.M., G.Đ. and B.M.; visualization, A.R.; supervision, L.M. and G.Đ. 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; further inquiries can be directed to the corresponding author.

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

References

- 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).
- Park, J.Y.; Kim, D.-J.; Lim, Y. Use of Smart Card Data to Define Public Transit Use in Seoul, South Korea. *Transp. Res. Rec.* **2008**, *2063*, 3–9. [CrossRef]
- Mohammed, T.; Fujiyama, T. Investigating Paper Ticket Usage on London Underground’s Network. In Proceedings of the 14th Conference on Advanced Systems in Public Transport (CASPT), Brisbane, Australia, 23–25 July 2018.
- Chowdhury, P.; Bala, P.; Addy, D.; Giri, S.; Chaudhuri, A.R. RFID and Android based smart ticketing and destination announcement system. In Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 21–24 September 2016; pp. 2587–2591. [CrossRef]
- Lale, D.; Săvescu, C.; Gogoneată, S.; Nechifor, C.; Vasile, M.; Niculescu, F. Passengers Monitoring System with Infrared Sensors and Microcontroller. *TURBO* **2021**, *VIII*, 4–11.
- Cardoso, D.T.; Manfroi, D.; de Freitas, E.P. Improvement in the Detection of Passengers in Public Transport Systems by Using UHF RFID. *Int. J. Wirel. Inf. Netw.* **2020**, *27*, 116–132. [CrossRef]
- Mathews, E.; Poigne, A. An Echo State Network based pedestrian counting system using wireless sensor networks. In Proceedings of the 2008 International Workshop on Intelligent Solutions in Embedded Systems, Regensburg, Germany, 10–11 July 2008; pp. 1–14. [CrossRef]
- Myrvoll, T.A.; Håkegård, J.E.; Matsui, T.; Septier, F. Counting public transport passenger using WiFi signatures of mobile devices. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 1–6. [CrossRef]
- Kostakos, V. Using Bluetooth to capture passenger trips on public transport buses. *CoRR* **2008**. Available online: https://www.researchgate.net/publication/220484433_Using_Bluetooth_to_capture_passenger_trips_on_public_transport_buses (accessed on 29 June 2024).
- Seidel, R.; Jahn, N.; Seo, S.; Goertler, T.; Obermayer, K. NAPC: A Neural Algorithm for Automated Passenger Counting in Public Transport on a Privacy-Friendly Dataset. *IEEE Open J. Intell. Transp. Syst.* **2022**, *3*, 33–44. [CrossRef]
- Khan, S.H.; Yousaf, M.H.; Murtaza, F.; Velastin, S. Passenger Detection and Counting for Public Transport System. *NED Univ. J. Res.* **2020**, *XVII*, 35–46. [CrossRef]
- Hsu, Y.-W.; Chen, Y.-W.; Perng, J.-W. Estimation of the Number of Passengers in a Bus Using Deep Learning. *Sensors* **2020**, *20*, 2178. [CrossRef]
- Kuchár, P.; Pirník, R.; Tichý, T.; Rástočný, K.; Skuba, M.; Tettamanti, T. Noninvasive Passenger Detection Comparison Using Thermal Imager and IP Cameras. *Sustainability* **2021**, *13*, 12928. [CrossRef]
- Sutjarittham, T.; Gharakheili, H.H.; Kanhere, S.S.; Sivaraman, V. Estimating Passenger Queue for Bus Resource Optimization Using LoRaWAN-Enabled Ultrasonic Sensors. *IEEE Syst. J.* **2022**, *16*, 6265–6276. [CrossRef]
- Kotz, A.J.; Kittelson, D.B.; Northrop, W.F. Novel Vehicle Mass-Based Automated Passenger Counter for Transit Applications. *Transp. Res. Rec.* **2016**, *2563*, 37–43. [CrossRef]
- Kovacs, R.; Nadai, L.; Horvath, G. Concept validation of an automatic passenger counting system for trams. In Proceedings of the 2009 5th International Symposium on Applied Computational Intelligence and Informatics, Timisoara, Romania, 28–29 May 2009; pp. 211–216. [CrossRef]
- Nielsen, B.F.; Frølich, L.; Nielsen, O.A.; Filges, D. Estimating passenger numbers in trains using existing weighing capabilities. *Transp. A Transp. Sci.* **2013**, *10*, 502–517. [CrossRef]

18. Zhu, F.; Gu, J.; Yang, R.; Zhao, Z. Research on Counting Method of Bus Passenger Flow Based on Kinematics of Human Body and SVM. In Proceedings of the 2008 Second International Symposium on Intelligent Information Technology Application, Shanghai, China, 20–22 December 2008; pp. 14–18. [\[CrossRef\]](#)
19. Moser, I.; McCarthy, C.; Jayaraman, P.P.; Ghaderi, H.; Dia, H.; Li, R.; Simmons, M.; Mehmood, U.; Tan, A.M.; Weizman, Y.; et al. A Methodology for Empirically Evaluating Passenger Counting Technologies in Public Transport. In Proceedings of the 41st Australasian Transport Research Forum, Canberra, Australia, 30 September–2 October 2019.
20. Grgurević, I.; Juršić, K.; Rajič, V. Review of Automatic Passenger Counting Systems in Public Urban Transport. In *5th EAI International Conference on Management of Manufacturing Systems*; Knapčiková, L., Peraković, D., Behúnová, A., Periša, M., Eds.; EAI/Springer Innovations in Communication and Computing; Springer: Cham, Switzerland, 2022. [\[CrossRef\]](#)
21. Chato, P.; Chipantasi, D.J.M.; Velasco, N.; Rea, S.; Hallo, V.; Constante, P. Image processing and artificial neural network for counting people inside public transport. In Proceedings of the 2018 IEEE Third Ecuador Technical Chapters Meeting (ETCM), Cuenca, Ecuador, 15–19 October 2018; pp. 1–5. [\[CrossRef\]](#)
22. OpenCV Library. Available online: <https://opencv.org/> (accessed on 30 June 2024).
23. Yang, T.; Zhang, Y.; Shao, D.; Li, Y. Clustering method for counting passengers getting in a bus with single camera. *Opt. Eng.* **2010**, *49*, 037203. [\[CrossRef\]](#)
24. Zhang, C.; Xu, J.; Beaugendre, A.; Goto, S. A KLT-based approach for occlusion handling in human tracking. In Proceedings of the 2012 Picture Coding Symposium, Krakow, Poland, 7–9 May 2012; pp. 337–340. [\[CrossRef\]](#)
25. Khoudour, L.; Yahiaoui, T.; Meurie, C. Real-time passenger counting in buses using dense stereovision. *J. Electron. Imaging* **2010**, *19*, 031202. [\[CrossRef\]](#)
26. Nitti, M.; Pinna, F.; Pintor, L.; Piloni, V.; Barabino, B. iABACUS: A Wi-Fi-Based Automatic Bus Passenger Counting System. *Energies* **2020**, *13*, 1446. [\[CrossRef\]](#)
27. Dan, B.-K.; Kim, Y.-S.; Suryanto; Jung, J.-Y.; Ko, S.-J. Robust people counting system based on sensor fusion. *IEEE Trans. Consum. Electron.* **2012**, *58*, 1013–1021. [\[CrossRef\]](#)
28. Li, F.; Yang, F.-W.; Liang, H.-W.; Yang, W.-M. Automatic Passenger Counting System for Bus Based on RGB-D Video. In Proceedings of the 2nd Annual International Conference on Electronics, Electrical Engineering and Information Science, Xi'an China, 2–4 December 2016.
29. Nasir, A.; Gharib, N.; Jaafar, H. Automatic Passenger Counting System Using Image Processing Based on Skin Colour Detection Approach. In Proceedings of the 2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA), Kuching, Malaysia, 15–17 August 2018; pp. 1–8. [\[CrossRef\]](#)
30. Oransirikul, T.; Nishide, R.; Piumarta, I.; Takada, H. Measuring Bus Passenger Load by Monitoring Wi-Fi Transmissions from Mobile Devices. *Procedia Technol.* **2014**, *18*, 120–125. [\[CrossRef\]](#)
31. Kalikova, J.; Krcaľ, J. People counting by means of Wi-Fi. In Proceedings of the 2017 Smart City Symposium Prague (SCSP), Prague, Czech Republic, 25–26 May 2017; pp. 1–3. [\[CrossRef\]](#)
32. Mehmood, U.; Moser, I.; Jayaraman, P.P.; Banerjee, A. Occupancy Estimation using WiFi: A Case Study for Counting Passengers on Busses. In Proceedings of the 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, 15–18 April 2019; pp. 165–170. [\[CrossRef\]](#)
33. Oliveira, L.; Schneider, D.; De Souza, J.; Shen, W. Mobile Device Detection Through WiFi Probe Request Analysis. *IEEE Access* **2019**, *7*, 98579–98588. [\[CrossRef\]](#)
34. 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. [\[CrossRef\]](#)
35. 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. [\[CrossRef\]](#)
36. Pedestrian Accessibility and Movement Environment Laboratory. Available online: <https://discovery.ucl.ac.uk/id/eprint/1414/> (accessed on 30 June 2024).
37. PAMELA UANDES Dataset. Available online: http://velastin.dynu.com/PAMELA-UANDES/whole_data.html (accessed on 30 June 2024).
38. Hsu, Y.-W.; Wang, T.-Y.; Perng, J.-W. Passenger flow counting in buses based on deep learning using surveillance video. *Optik* **2020**, *202*, 163675. [\[CrossRef\]](#)
39. Liciotti, D.; Cenci, A.; Frontoni, E.; Mancini, A.; Zingaretti, P. An Intelligent RGB-D Video System for Bus Passenger Counting. In *Intelligent Autonomous Systems 14 IAS 2016, Proceedings of the IAS 2016, Shanghai, China, 3–7 July 2016*; Chen, W., Hosoda, K., Menegatti, E., Shimizu, M., Wang, H., Eds.; Advances in Intelligent Systems and Computing; Springer: Cham, Switzerland, 2017; Volume 531. [\[CrossRef\]](#)
40. Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; Volume 1, pp. 886–893. [\[CrossRef\]](#)
41. Haq, E.U.; Xu, H.; Chen, X.; Zhao, W.; Fan, J.; Abid, F. A fast hybrid computer vision technique for real-time embedded bus passenger flow calculation through camera. *Multimed. Tools Appl.* **2020**, *79*, 1007–1036. [\[CrossRef\]](#)
42. Pisner, D.A.; Schnyer, D.M. Chapter 6—Support vector machine. In *Machine Learning*; Mechelli, A., Vieira, S., Eds.; Academic Press: Cambridge, MA, USA, 2020; pp. 101–121, ISBN 9780128157398. [\[CrossRef\]](#)

43. Biswas, D.; Su, H.; Wang, C.; Stevanovic, A.; Wang, W. An automatic traffic density estimation using Single Shot Detection (SSD) and MobileNet-SSD. *Phys. Chem. Earth Parts A/B/C* **2019**, *110*, 176–184. [CrossRef]
44. YOLOv8. Available online: <https://yolov8.com/> (accessed on 30 June 2024).
45. Ultralytics YOLOv8 Github Repository. Available online: <https://github.com/ultralytics/ultralytics/blob/main/docs/en/models/yolov8.md> (accessed on 30 June 2024).
46. 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. *Int. Res. J. Mod. Eng. Technol. Sci.* **2024**, *6*, 3684–3691. [CrossRef]
47. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; Springer: Cham, Switzerland, 2014; pp. 740–755.
48. 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. [CrossRef]
49. Rendon, W.D.M.; Anillo, C.B.; Jaramillo-Ramirez, D.; Carrillo, H. Passenger Counting in Mass Public Transport Systems using Computer Vision and Deep Learning. *IEEE Lat. Am. Trans.* **2023**, *21*, 537–545. [CrossRef]
50. Kim, H.; Sohn, M.K.; Lee, S.H. Development of a Real-Time Automatic Passenger Counting System using Head Detection Based on Deep Learning. *J. Inf. Process. Syst.* **2022**, *18*, 428–442. [CrossRef]
51. Labit-Bonis, C.; Thomas, J.; Lerasle, F. Visual and automatic bus passenger counting based on a deep tracking-by-detection system. *HAL Open Sci.* **2021**. Available online: <https://hal.science/hal-03363502> (accessed on 30 June 2024).
52. Bishop, C.M.; Bishop, H. Convolutional Networks. In *Deep Learning*; Springer: Cham, Switzerland, 2024. [CrossRef]
53. Suresh, K.; Bhuvan, S.; Palangappa, M.B. Social Distance Identification Using Optimized Faster Region-Based Convolutional Neural Network. In Proceedings of the 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 8–10 April 2021; pp. 753–760. [CrossRef]
54. Haque, M.F.; Lim, H.-Y.; Kang, D.-S. Object Detection Based on VGG with ResNet Network. In Proceedings of the 2019 International Conference on Electronics, Information, and Communication (ICEIC), Auckland, New Zealand, 22–25 January 2019; pp. 1–3. [CrossRef]
55. Bisong, E. Recurrent Neural Networks (RNNs). In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*; Apress: Berkeley, CA, USA, 2019. [CrossRef]
56. Graves, A. Long Short-Term Memory. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Studies in Computational Intelligence; Springer: Berlin/Heidelberg, Germany, 2012; Volume 385. [CrossRef]
57. Guo, Y.; Liu, Y.; Georgiou, T.; Lew, M.S. A review of semantic segmentation using deep neural networks. *Int. J. Multimed. Inf. Retr.* **2018**, *7*, 87–93. [CrossRef]
58. Yao, W.; Zeng, Z.; Lian, C.; Tang, H. Pixel-wise regression using U-Net and its application on pansharpening. *Neurocomputing* **2018**, *312*, 364–371. [CrossRef]
59. Niu, Z.; Liu, W.; Zhao, J.; Jiang, G. DeepLab-Based Spatial Feature Extraction for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 251–255. [CrossRef]
60. Kale, K.; Pawar, S.; Dhulekar, P. Moving object tracking using optical flow and motion vector estimation. In Proceedings of the 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Noida, India, 2–4 September 2015; pp. 1–6. [CrossRef]
61. Garcia-Garcia, B.; Bouwmans, T.; Silva, A.J.R. Background subtraction in real applications: Challenges, current models and future directions. *Comput. Sci. Rev.* **2020**, *35*, 100204. [CrossRef]
62. Salcedo-Sanz, S.; Rojo-Álvarez, J.L.; Martínez-Ramón, M.; Camps-Valls, G. Support vector machines in engineering: An overview. *WIREs Data Min. Knowl. Discov.* **2014**, *4*, 234–267. [CrossRef]
63. Becker, T.; Rousseau, A.-J.; Geubbelmans, M.; Burzykowski, T.; Valkenborg, D. Decision trees and random forests. *Am. J. Orthod. Dentofac. Orthop.* **2023**, *164*, 894–897. [CrossRef]
64. Xu, T.; Liu, H.; Qian, Y.; Zhang, H. A novel method for people and vehicle classification based on Hough line feature. In Proceedings of the International Conference on Information Science and Technology, Nanjing, China, 26–28 March 2011; pp. 240–245. [CrossRef]
65. Raghavachari, C.; Aparna, V.; Chithira, S.; Balasubramanian, V. A Comparative Study of Vision Based Human Detection Techniques in People Counting Applications. *Procedia Comput. Sci.* **2015**, *58*, 461–469. [CrossRef]
66. Mutlag, W.K.; Ali, S.K.; Aydam, Z.M.; Taher, B.H. Feature Extraction Methods: A Review. *J. Phys. Conf. Ser.* **2020**, *1591*, 012028. [CrossRef]
67. Cheng, H.D.; Jiang, X.H.; Sun, Y.; Wang, J. Color image segmentation: Advances and prospects. *Pattern Recognit.* **2001**, *34*, 2259–2281. [CrossRef]
68. Hirata, N.S.T.; Papakostas, G.A. On Machine-Learning Morphological Image Operators. *Mathematics* **2021**, *9*, 1854. [CrossRef]
69. Chernov, V.; Alander, J.; Bochkov, V. Integer-based accurate conversion between RGB and HSV color spaces. *Comput. Electr. Eng.* **2015**, *46*, 328–337. [CrossRef]
70. Jocher, G.; Qiu, J. Ultralytics YOLO11. 2024. Available online: <https://github.com/ultralytics/ultralytics> (accessed on 13 October 2024).

71. COCO Dataset Limited (PersonOnly). Available online: <https://universe.roboflow.com/shreks-swamp/coco-dataset-limited-person-only> (accessed on 30 June 2024).
72. Osorio, F.G.; Xinran, M.; Liu, Y.; Lusina, P.; Cretu, E. Sensor network using Power-over-Ethernet. In Proceedings of the 2015 International Conference and Workshop on Computing and Communication (IEMCON), Vancouver, BC, Canada, 15–17 October 2015; pp. 1–7. [CrossRef]
73. Transportation and Bus Surveillance: Mobile Security, 10 July 2018. Available online: <https://iebmmedia.com/applications/transportation/transportation-and-bus-surveillance-mobile-security/> (accessed on 13 October 2024).
74. 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).
75. Pons, M.; Valenzuela, E.; Rodríguez, B.; Nolzaco-Flores, J.A.; Del-Valle-Soto, C. Utilization of 5G Technologies in IoT Applications: Current Limitations by Interference and Network Optimization Difficulties—A Review. *Sensors* **2023**, *23*, 3876. [CrossRef]
76. Bochkovskiy, A.; Wang, C.Y.; Mark Liao, H.Y. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
77. Label Studio Documentation. Available online: <https://labelstud.io/guide/> (accessed on 30 June 2024).
78. Hossen, M.A.; Naim, A.G.; Abas, P.E. Deep Learning for Skeleton-Based Human Activity Segmentation: An Autoencoder Approach. *Technologies* **2024**, *12*, 96. [CrossRef]
79. Li, Z.; Liu, F.; Yang, W.; Peng, S.; Zhou, J. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 6999–7019. [CrossRef]
80. Weiss, K.; Khoshgoftaar, T.M.; Wang, D.D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 9. [CrossRef]
81. General Data Protection Regulation—GDPR. Available online: <https://gdpr-info.eu/> (accessed on 30 June 2024).
82. 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).
83. Voigt, P.; Bussche, A. *The EU General Data Protection Regulation (GDPR): A Practical Guide*; Springer: Cham, Switzerland, 2017. [CrossRef]
84. 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. [CrossRef]
85. 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).
86. EU AI Act: First Regulation on Artificial Intelligence. Available online: <https://www.europarl.europa.eu/topics/en/article/2023/06/01STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> (accessed on 13 September 2024).
87. Olivo, A.; Maternini, G.; Barabino, B. Empirical Study on the Accuracy and Precision of Automatic Passenger Counting in European Bus Services. *Open Transp. J.* **2019**, *13*, 250–260. [CrossRef]
88. Kuo, Y.-H.; Leung, J.M.; Yan, Y. Public transport for smart cities: Recent innovations and future challenges. *Eur. J. Oper. Res.* **2023**, *306*, 1001–1026. [CrossRef]
89. Lukic Vujadinovic, V.; Damjanovic, A.; Cacic, A.; Petkovic, D.R.; Prelevic, M.; Pantovic, V.; Stojanovic, M.; Vidojevic, D.; Vranjes, D.; Bodolo, I. AI-Driven Approach for Enhancing Sustainability in Urban Public Transportation. *Sustainability* **2024**, *16*, 7763. [CrossRef]
90. McCarthy, C.; Ghaderi, H.; Martí, F.; Jayaraman, P.; Dia, H. Video-based automatic people counting for public transport: On-bus versus off-bus deployment. *Comput. Ind.* **2025**, *164*, 104195. [CrossRef]
91. Huynh-The, T.; Banos, O.; Lee, S.; Kang, B.H.; Kim, E.-S.; Le-Tien, T. NIC: A Robust Background Extraction Algorithm for Foreground Detection in Dynamic Scenes. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *27*, 1478–1490. [CrossRef]
92. Kuchár, P.; Pirník, R.; Janota, A.; Malobický, B.; Kubík, J.; Šišmišová, D. Passenger Occupancy Estimation in Vehicles: A Review of Current Methods and Research Challenges. *Sustainability* **2023**, *15*, 1332. [CrossRef]
93. Diwan, T.; Anirudh, G.; Temburne, J.V. Object detection using YOLO: Challenges, architectural successors, datasets and applications. *Multimed. Tools Appl.* **2023**, *82*, 9243–9275. [CrossRef] [PubMed]
94. L'Heureux, A.; Grolinger, K.; Elyamany, H.F.; Capretz, M.A.M. Machine Learning With Big Data: Challenges and Approaches. *IEEE Access* **2017**, *5*, 7776–7797. [CrossRef]
95. Gong, Y.; Liu, G.; Xue, Y.; Li, R.; Meng, L. A survey on dataset quality in machine learning. *Inf. Softw. Technol.* **2023**, *162*, 107268. [CrossRef]
96. Tercan, H.; Guajardo, A.; Meisen, T. Industrial Transfer Learning: Boosting Machine Learning in Production. In Proceedings of the 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Helsinki, Finland, 22–25 July 2019; pp. 274–279. [CrossRef]
97. Marczyk, M.; Kempinski, A.; Socha, M.; Cogiél, M.; Foszner, P.; Staniszewski, M. Passenger Location Estimation in Public Transport: Evaluating Methods and Camera Placement Impact. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 17878–17887. [CrossRef]
98. Shao, S.; Zhao, Z.; Li, B.; Xiao, T.; Yu, G.; Zhang, X.; Sun, J. CrowdHuman: A benchmark for detecting human in a crowd. *arXiv* **2018**, arXiv:1805.00123.

99. Khanam, R.; Hussain, M. YOLOv11: An Overview of the Key Architectural Enhancements. *arXiv* **2024**, arXiv:2410.17725. [[CrossRef](#)]
100. Pronello, C.; Baratti, L.; Anbarasan, D. Benchmarking the Functional, Technical, and Business Characteristics of Automated Passenger Counting Products. *Smart Cities* **2024**, *7*, 302–324. [[CrossRef](#)]
101. Rawat, N.; Jeengar, K.; Agarwal, A.; Kaur Chahal, R.J. Boarding Alighting Counting in Different Transit Vehicles under Crowded Conditions. In Proceedings of the 2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 12–14 July 2024; pp. 1–6. [[CrossRef](#)]
102. Wang, P.; Chen, X.; Chen, J.; Hua, M.; Pu, Z. A two-stage method for bus passenger load prediction using automatic passenger counting data. *IET Intell. Transp. Syst.* **2021**, *15*, 248–260. [[CrossRef](#)]
103. Jahn, N.; Siebert, M. Engineering the Neural Automatic Passenger Counter. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105148. [[CrossRef](#)]
104. Shafaeipour, N.; Stanciu, V.-D.; van Steen, M.; Wang, M. Understanding the protection of privacy when counting subway travelers through anonymization. *Comput. Environ. Urban Syst.* **2024**, *110*, 102091. [[CrossRef](#)]

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.