DATA FUSION OF VIDEO AND LIDAR TRAFFIC SURVEILLANCE DATA: PRACTICAL ASSESEMENT OF IMPLEMENTED SOLUTION IN JELGAVA CITY

Nikolajs Bumanis[1], Gatis Vitols[1], Inga Meirane[2]
[1]Latvia University of Life Sciences and Technologies, Latvia; [2]WeAreDots Ltd, Latvia
nikolajs.bumanis@llu.lv, gatis.vitols@llu.lv, inga.meirane@wearedots.com

Abstract. The typical scenario of using only video data for traffic surveillance proves to be cheaper and easier to implement, but on the other hand requires complex setup to achieve state of the art performance. The integrated setup with different spectrum devices, such as LiDAR, alleviates the necessity to multiply the amount of video cameras and complexity of their location; especially considering busy intersections. The analysis of data acquired as a result of data fusing raw data or information provided by these devices leads to a multi-functional system. Our previous work addresses development of a surveillance solution that combines multiple video cameras with LiDAR sensors setup over the 4-lane street in Jelgava city, Latvia. The output of these devices was fused and analysed using Machine Learning methods in order to detect vehicles, including licence number plates' registration and recognition, and movement parameters. The dimensions of detected vehicles were also measured using LiDAR data in order to identify vehicles that violate local road traffic regulations or parking policy. The implemented solution was set up with consideration of various video surveillance affecting factors like variable vehicles' movement, weather conditions and object occlusion. However, real-world experience shows that it is almost impossible to predict every case and built-in measures to prevent certain errors. This paper aims to assess installed surveillance solutions: 1) identify factors and cases that affect overall performance, 2) provide measures to eliminate or alleviate the effect of identified factors, and 3) extend comprehensibility of the solution by incorporating these measures into processing flow. In result, the hardware and software setup, including the previously proposed surveillance model, was analysed versus achieved performance. The solutions to tackle each identified factor are provided and comprehensive guidelines and requirements for similar surveillance systems are developed.

Keywords: lidar, video data, traffic surveillance, intelligent traffic system.

Introduction

Today, the Internet of Things (IoT), as a set of sensors, hardware, and application software, is an essential component of a successful information system [1] that provides the ability to obtain information about an ongoing process using sensors both directly and from external and/or related systems. IoT can become one of today's most important technological advances if its potential is fully realised [2]. One of the applications of IoT is urban traffic surveillance using various video monitoring hardware in conjunction with object detection and tracking algorithms [3; 4].

The tasks of object detection and tracking are common in the field of computer vision [5] and while there are always novel algorithms and models proposed [6; 7], practical application, policies and objectives are main constraints for real world implementations. For instance, video surveillance is currently a go-to strategy for citywide security solution [8] and traffic surveillance [8]. In recent years, the interest towards real-world applications regarding pedestrian and vehicle detection and tracking is growing. The topic of traffic surveillance also falls under this trend as there are various new methods, approaches and applications proposed frequently [9-12], where the current trend is slowly switching from obtaining the optimal vehicle detection method to identification of traffic anomalies, accidents and road regulation offence detection and prediction [13-15].

There are various traffic surveillance solutions available on the market. For example, Vidar product from Adaptive Recognition [15] promises improved performance for vehicle's number plate recognition task in any conditions and uses a dual-lens camera with infra-red and visible light capabilities. However, the Vidar focuses on the number plate recognition, and cannot be used, for example, to detect lane change. On the other hand, Traffic Control & Surveillance System (TCSS) offered by EV-Dynamic [16] or Traffic Solution and City Surveillance Solution offered by SensorTec-EU [17] are sophisticated systems that aim to cover all components of the surveillance process, but these solutions are commercially locked hardware and software sets which prohibits any modifications or open source solution additions.

The performance of a surveillance system is dictated by multiple factors, i.e., internal factors - hardware and software, and external factors, such as illumination, weather and obstructions. Hardware tends to be selected during early stages of software development according to analysis of external factors [19; 20], but although it is commonly known [21; 22] that external factors affect video quality required for object detection and tracking, most solutions are developed without encapsulating all possible scenarios [22]. The degree of external factors' effect depends on the nature of the factor itself - rainfall and snowfall may have a higher degree of performance impact that is short term, whereas sunset may have an object detection impairing effect that lasts around an hour [21]. In addition, the placement of monitor devices dictates the scope of potentially achievable goals. For instance, He et al. [23] proposed five different video camera placement strategies, but focused on maximising the coverage over metropolis' intersections.

One of the solutions is to implement sophisticated object detection algorithms [24] where a large set of imperfect data is included, but this method can only be applied at the post-installation phase at which enough data is already gathered. Alternatively, it is possible to add a sensor type hardware like LiDAR, which is used to detect objects with a laser that constructs a 3D environment. This provides data necessary enough to determine objects dimensions that in turn can be used for object classification using existing classification methods, whereas correct classification is especially important to identify freight vehicles [25]. While LiDAR may be used for identification using high accuracy distance detection, it still lacks performance when it comes to classification; therefore, LiDAR is often used as a main or complementary data source in conjunction with video cameras [26].

The data from LiDAR and video cameras can be fused, alleviating the downsides of using LiDAR or video cameras separately [27]. There are various fusion algorithms, such as Dempster-Shafer evidential theory, Bayesian inference, Monte Carlo techniques and Kalman Filter (including extended Kalman Filter) that are commonly used for traffic related scenarios [28]. These algorithms were developed with consideration of tackling data fusion challenges like data imperfection, conflicting data, outliers and spurious records, etc. [29], and their implementation mainly focuses on traffic forecasting and travel time estimation. While traffic monitoring may not require full implementation of such algorithms, the principles of data science behind them can be used, whereas real-world applications still require a tailored development approach [24; 30]. The implementation of data fusion between LiDAR and video cameras into traffic monitoring solutions are often [31; 32] done without proper initial investigation of external factors and tend to focus on either one particular error inducing the factor in isolation or a particular task. The typical scenarios of LiDAR and video camera data fusion include pedestrian detection and autonomous vehicle driving. For instance, Wu et al. [33] proposed to use LiDAR 3D point cloud data to increase the quality of the detected object's shape, and Zhao et al. [34] fused LiDAR and video data to produce environmental situations for autonomous vehicles. While these proposals aim to improve the solution, they do not address the general framework and all possible real-world usage cases.

Our previous work [35] addresses development of a surveillance solution that combines multiple video cameras with LiDAR sensors setup over a 4-lane street in Jelgava city, Latvia. The implemented solution was set up with consideration of various surveillance affecting factors like dynamics of vehicle movement (speed, change of lanes, etc.), weather conditions and object occlusions in the range of sensor operation. However, real-world experience shows that it is almost impossible to predict every case and built-in measures to prevent those. This paper aims to assess an installed surveillance solution: 1) identify factors and cases that affect overall performance, 2) provide measures to eliminate or alleviate the effect of identified factors, and 3) extend comprehensibility of the solution by incorporating these measures into processing flow. In result, the hardware and software setup, including the previously proposed surveillance model, was analysed versus achieved performance. The solutions to tackle each identified factor are provided and comprehensive guidelines and requirements for similar surveillance systems are developed.

Materials and methods

Experimental setup. The traffic surveillance solution's validation was performed near the intersection of Liela and Pasta streets in Jelgava city, Latvia, (56° 39.1115' N, 23° 43.3802' E) (Fig. 1a) for the total duration of 6 months (183 days), from 01.04.2021 till 30.09.2021 (this excludes initial hardware installation and testing, both of which happened in March, 2021). Two traffic directions were monitored with video cameras and LiDAR sensor setup to run 24 h a day, 7 days a week (this includes periods, where data was unobtainable due to connection or power issues).

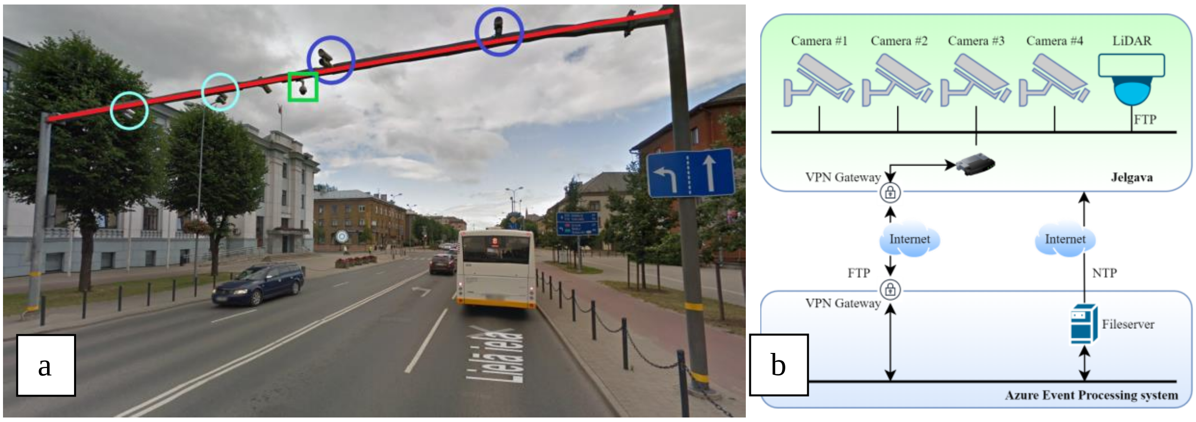

Fig. 1. **Experimental setup (a), where red line – structural support bar, blue circles – video cameras for traffic going into city and turning left, light blue circles – video cameras for traffic going out of city, green square – LiDAR sensor; visual representation of equipment topology (b)**

Weather conditions for the period of the experiment were obtained using webservice, which is provided by the Latvian Centre for Environment, Geology and Meteorology Ltd., available at https://www.meteo.lv/. The average temperature was 9.3 °C (Fig. 2, a). Maximum observed amount of precipitations in form of rain was registered on 12.09.2021 with the value of 9.4 mm * h^{-1} for a period from 17:00 to 18:00.

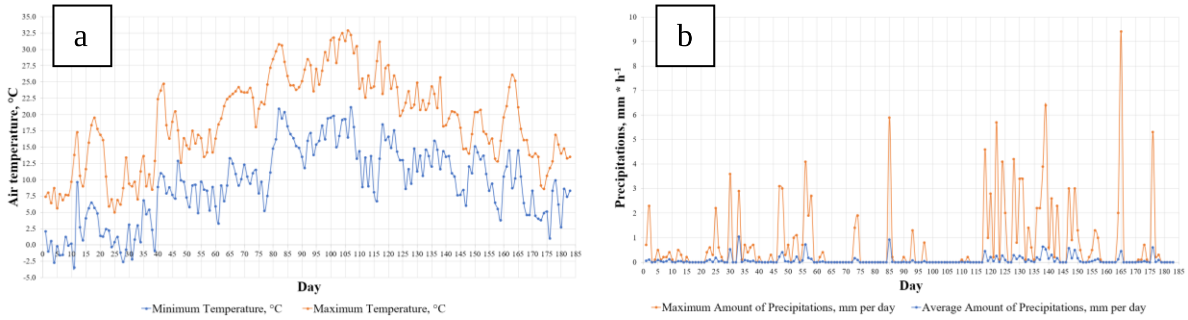

Fig. 2. **Weather conditions during experimental period**: a – minimum and maximum temperature, °C per day; b – maximum and average amount of precipitations, mm per day

Data. According to the experiment's goals and due to constant improvements in configuration and calibration no long-term statistical data has been gathered. The short-term and testing data included the following information: vehicles speed (km·h^{-1}), vehicle dimensions – width, height and length (m), vehicle category based on classification (car, car with trailer, van, pickup van, van with trailer, truck), driving lane (0-4), driving direction (forward, turning left, changing lane). In addition, number plate's recognition was performed using an algorithm, which was developed by WeAreDots Ltd [35]. This algorithm is constructed around two methods – Single Shot Multibox Detector for vehicle's detection, number plate's detection and cropping; and Siamese Neural Network for the number plate's identification (i.e., detection of separate symbols, outputting symbols as text string and comparing this string with those, previously detected). The vehicle's number plate was used to detect the violations of traffic rules. In consideration of GDPR and drivers' privacy, data gathering was supported by legal agreement with local authorities. All gathered data is first sent to Microsoft Azure Blob storage accessed through REST API. Microsoft Azure is a comprehensive data analytics solution that provides all necessary means for statistical analysis. Algorithms, developed by WeAreDots Ltd. were used in conjunction with built-in methods and algorithms to process the data.

The average throughput of data was around 20 GB per day provided by LiDAR sensory system (TIC501 controller with LMS511 2D sensor protected by weather protection hood 2063050 all by Sick AG) positioned over the middle line and four video cameras (four M16B bodies, each with Mx-O-SMA-

S-6L237 night lens and Mx-O-SMA-S-6D23 day lens all by Motobix and Infrared illuminator M-series 860nm by Emitlight) (Fig. 3). Video cameras were responsible for detecting driving or parked vehicles, but LiDAR for measuring the dimension of the detected vehicles. Due to changes of the video cameras' field of view and operating range no strict dimensions regarding the observation zones were taken, however, it was noted that under normal driving conditions, it takes approximately 200 ms to 1200 ms from the vehicle detection with the camera to the LiDAR detection zone (Fig. 3, Δt).

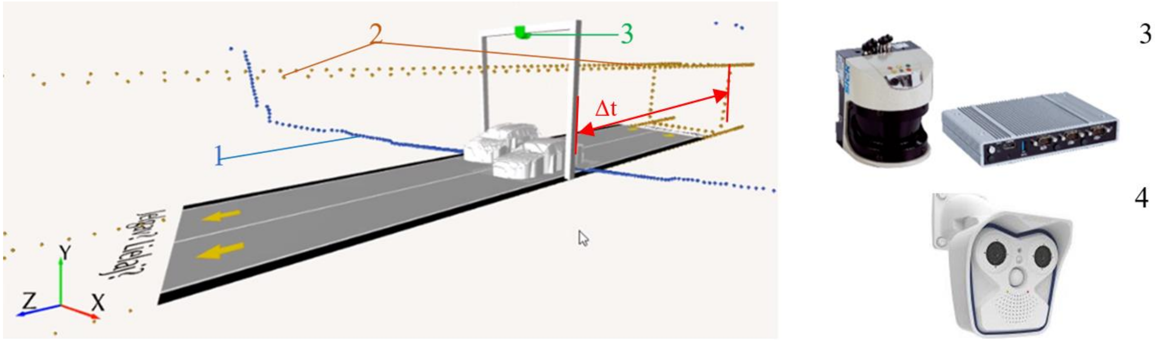

Fig. 3. **Area of interest of LiDAR sensor (1) and video cameras (2):** 3 – LiDAR sensor; 4 – video cameras; Δ*t* – delay between video camera detection and LiDAR detection

Data acquisition procedure. Data acquisition from video cameras was initiated based on multi-scale contrast changes in the configured area of interest. Images of a single vehicle were taken by at least one of the video cameras that had an overlapping Field of View. Images were uploaded into cloud storage (Microsoft Azure BLOB storage) to perform pre-processing and selection by removing duplicates and images unsuitable for vehicle number plate recognition. In order to increase the probability of the image being suitable two light spectres were used - infrared and visible light (multiple images in both spectres were taken). LiDAR scanned each passing vehicle and measured its dimensions and attempted to apply a label based on a built-in classification table provided by the LiDAR's manufacturer. The results of these scans and the applied class were encapsulated into an XML file that was saved to the FTP server. Both data inputs were protected by VPN gateway, where the data exchange protocol Site-to-site VPN IKEv2 (Internet Key Exchange v.2) was used for improved security. Each vehicle could get only a single XML file; that way for a single XML file a dynamic number of images were created. For more details regarding data, including examples of XML, please refer to [35].

Data fusion. In order to link video and LiDAR a novel method to data synchronisation was used. This method was introduced in [35] and is based on complementary feature-level data fusion, where vehicles are recognized and linked (between LiDAR and video source data) using the coordinate system, object's shape and timestamp. While implementing this method, the vehicle's recognition was specified as an event that includes data about a recognized vehicle (i.e., type and vehicle licence number) in the form of a single image (acquired after selection versus configured quality threshold value), this vehicle's dimensions, including the movement direction and speed, in the form a 3D point cloud, and metadata about the recognition of the event itself. In the framework of a single event the analysis regarding violation of traffic and parking rules in Jelgava city was performed.

Objective. The goal of practical experimentation was to prove the viability of the previously proposed data synchronisation method for three tasks: a) identifying parking violation (illegal parking or prolonged parking), b) controlling the movement of a particular vehicle based on its number plate, and c) notification to law enforcement establishments about freight vehicles entering the city. Therefore, the main objective was to determine the confidence level of the data synchronisation. The confidence level was calculated by the algorithm itself (i.e., no additional formula was used) as a performance metric for each particular case and in conjunction with the object detection results was used for validation. Validation was performed for various external conditions and with multiple hardware and software configurations and calibration by manipulating video camera parameters: contrast, sensitivity and field of view, and data related parameters - amount of data transmitted and network's throughput. Firstly, this led to adjusting and improving the data synchronisation method based on the acquired knowledge about the affecting factors, and, secondly, this uncovered previously unbeknownst factors and limitations. Nonetheless, any major change in configuration led to inability to compare data sets from different iterations; therefore, no overall statistical analysis was performed, and the results in the next section depict the various factor influence for a period from 20 minutes up to 1 to 7 days.

Results and discussion

During the validation period the following information was gathered: a) LiDAR point cloud data: vehicle's detection event timestamp, movement direction and lane, vehicle's category, speed and dimensions, b) video recognition data: vehicle's number plate, and c) performance results: LiDAR and video data synchronisation rate, vehicle's number plates recognition rate. An example with GUI screenshot can be seen in Fig. 4.


LiDAR notikumu saraksts

Kolonnas Filtri Kārtošanas secība Eksportēt Pārlādēt Noklusētais skats

Ieraksta nr.	Reg.numurs	TL tips	Josla	Ātrums	Platums	Gārums	Augstums	Apvienošan...	Atpazīš...	Izveidots	Nobīde no centra	Transporlīdze...
1		Car	0	34.42	1.653	3.878	0	0		14.09.2021 18:47:24	1.274	15.03.2021 08:1...
4		Car	0	34.68	1.691	3.823	1.761	0		14.09.2021 18:47:24	0.384	15.03.2021 08:0...

Ierakstu skaits 20 1 - 2 no 2 1

Fig. 4. **Screenshot of the event registration system**

Due to the approach of choosing only a single image with the best characteristics for the number plate's recognition, every event, when the number plate's recognition was performed returned a vehicle's recognition confidence rate higher than 99%. However, at the start of the experiment not every recognized vehicle had its LiDAR data synchronised correctly - the results not only had error regarding the synchronisation rate that was in the range between 67.52 and 99.8% (this includes cases such as 99.96% recognition rate and 82.44% synchronization rate), the LiDAR itself returned various erroneous values for the vehicle speed and/or dimensions; therefore, it was noted that a particular vehicle (by the number plate) may have different classes assigned by LiDAR data processing algorithm. LiDAR also had problems with classification of vehicles sometimes mistakenly classifying certain vehicles such as a car with trailers.

These results were affected by multiple factors, identified during the validation period, and include precipitation, illumination during the day-night cycle, network throughput and traffic intensity, hardware settings.

Precipitation. The most obvious factor is the precipitation in the form of raindrops and snowflakes. Video cameras use changes in contrast against background, and when any such change is detected, initiates recognition event's creation. Raindrops and snowflakes introduce the change of contrast by themselves, thus false positive events (Fig. 5) are being created; this, in turn, heavily increases the amount of data that is unduly processed. Similar problems were detected with animal activities in the range of the sensor operation, such as multiple birds flying in the area of operation.

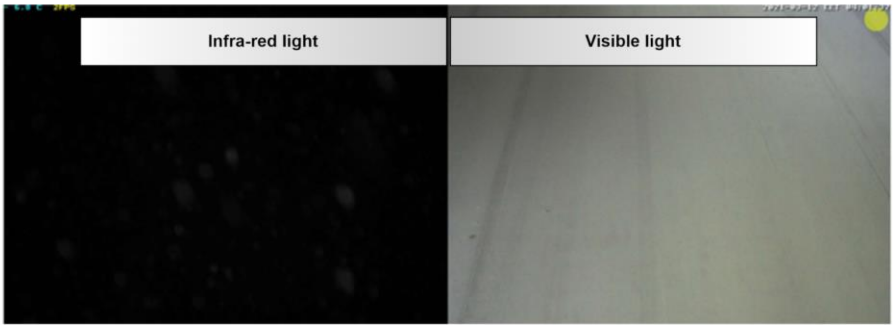

Fig. 5. **Screenshot of false positive event initiation by snowflakes in two light spectres**

The way to alleviate these effects is to reduce the contrast and sensitivity of video cameras, and to decrease the frames per seconds transmitted. It must be noted that the particular combination of these parameters depends on the environment, and must be selected case-wise. In case of Jelgava the calibration led to decrease of false positive event initiations from 42% (244 out of 582 total events) to 16% during heavy rain (74 out of 461 total events) for a period of 40 minutes, and 3% (8 out of 279 total events) during mild rain for a period of 20 minutes. To achieve the contrast changes to capture a new frame were reduced in several increments from 20% to 5%. As an alternative method, instead of decreasing the number of images taken, the rain drops and snowflakes can be removed using precipitation removing algorithms that require data centre with suitably high performance. For instance, [36] used the two stage conditional GAN-based algorithm for de-raining images and achieved significant improvements in image recognition capabilities. Similarly, such approach can be used in conjunction with contrast calibration to achieve higher data synchronisation rate.

Illumination. During the periods of sunrise and sunset the number of unsuitable images varied between 6% and 34% (the number of dropped images versus good ones in 1 hour) depending on the traffic intensity. For instance, at sunset, number plates reflected light and number plate symbols had reduction in contour precision. Illumination during night time also introduced creation of false positive events. The light reflection from wet/icy asphalt causes contrast changes so impressive that the camera algorithm considers any sequential changes as insignificant for the event's initiation; thus, keeping only the first image. This image, however, may be inadequate for the vehicle's number plate recognition (Fig. 6). In addition, it was found that during the full moon phase (if clear sky at night), infrared light produced images with much more accurate outlines of number plates, and thus these images should be used for recognition.

Fig. 6. **Example of an image that cannot be used for number plate recognition**

The effects of illumination at different times of the day can be significantly reduced by taking images in both visible and infrared light at the same time and performing vehicle's number plate recognition using both images while keeping the highest quality image for data synchronisation.

Network throughput and traffic flow intensity. During the validation the amount of data transferred was around 20 GB per day, which correlates with the traffic flow intensity. The higher the traffic flow intensity, the more data has to be transmitted, which causes heavy Internet traffic and/or interruptions in data transmission. This can also cause data loss during breaks, where connection is lost during transfer itself. On workdays, around 120 or more images may be taken for each vehicle during traffic congestion hours, but less than 5 images of the same vehicle would be taken during other hours. In order to reduce the total amount of data in addition to contrast changes that put a higher contrast threshold for image creation, frames per second were also reduced: initially, from 10 to 6 frames, and later from 6 to 4 frames; anything lower resulted in infrequent images or multi-vehicle images that are not usable for data synchronisation. In addition, the coordinate plane for the cameras' operating range was reduced (i.e., the area the object occupies for it to be considered detected (Fig. 7)) from initial 20% to 5%.

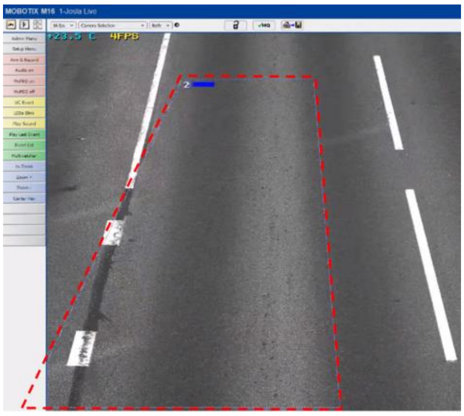

Fig. 7. **Example of camera's operation range configuration**

Hardware settings and synchronisation aspects. The coordinate system used for data synchronisation was configured to monitor road lanes. However, the road itself is a 4-lane two-way road, thus occlusion from opposite directions noticeably increases the amount of observable vehicles. The number plate recognition algorithm builds objects around the detected number plate, regardless if it is the vehicle's front or backend, thus introducing so-called data collision noise, which complicates tracking procedures required for detection of traffic rule violation (Fig. 8). In order to alleviate such scenarios each lane had its own coordinate plane segment assigned. This, in addition to the already reduced camera operation range resulted in reduction of multi-vehicle images from by 24% (from 17% to 4% of total images with "noise" during a 4-day period).

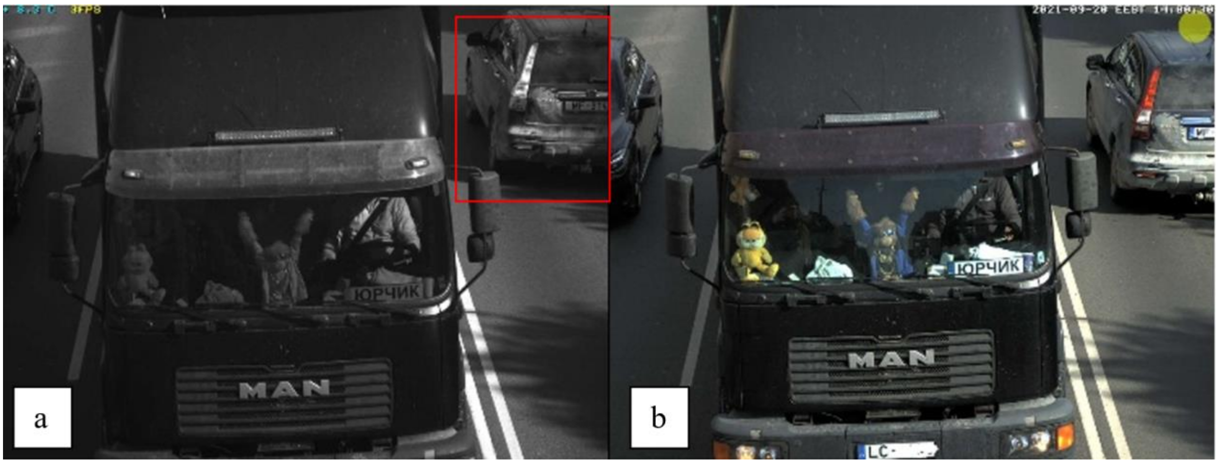

Fig. 8. **Example of data collision "noise":** a – truck number is not visible, while opposing direction car's is (red square); b – both number plates are visible

Even though the 3D LiDAR scanner and video cameras were connected to the same network, the delay between taken images and LiDAR XML file exists. Using a convolution neural network to determine the mean time delay, i.e., 30-60 s for particular date time was insufficient as there were still a lot of cases of mismatching. The implemented data synchronisation method uses timestamps as one of the fusion factors, thus the correct timestamps of both LiDAR XML file and images were required. In order to provide completely identical timestamps the NTP (Network Time Protocol) server was added to the initial hardware architecture. The configuration of the NTP server requires the calibration of hardware. The calibration in Jelgava was performed using filming video cameras' operating area manually with a mobile phone in parallel to the operational system. This data was used as an alternative source of vehicle detection and was manually compared to the data synchronisation results. In result, the data synchronisation algorithm and hardware had been reconfigured.

Summary. As a result of the improvements, 97% mean data synchronisation rate was achieved. The most noticeable difference was the error of speed detection using LiDAR: from initial 8% (first 3 days) to less than 0.5% at later stages. For instance, during a 4-day period, a total number of 60381 vehicles were detected, where only 69 vehicles had their speed incorrectly measured, which returned an error of 0.114%, whereas the traffic violation of speeding was identified for 2850 vehicles, or 4.72%. The results depict the necessity to evaluate each particular aspect of any traffic surveillance system both individually and as a system component. The experiment also proved that it is close to impossible to predict all variable pre-production phases and extended configuration and calibration is required. In order to develop similar traffic surveillance, the following guideline can be followed:

1. *Sensor placement selection*. Placement of the video and LiDAR monitoring devices must be chosen so that they encapsulate as many surveillance tasks as possible, because: a) overlapping Field of Views provide necessary data that can be fused, b) multiple video cameras pointing in the same direction can be used as a redundant system's component, and c) correct position alleviates occlusion issues raised by natural objects like trees. The setup in Jelgava successfully used four video cameras uniformly distributed across 4 lanes, with one camera having its FoV overlap the parking area, providing images to capture violations. Similarly, overlapping FoVs proved to be successful in [37].

2. *Network throughput*. Available network providers must be analysed, and if possible, cable connection between monitoring devices and data centres should be used. The amount of data transmitted must correlate with a particular surveillance task and traffic intensity, that is, a) instead of 60fps video to verify parking violation one image every two-three minute will suffice, and b) if heavy traffic is predicted, lower frame rate may be more appropriate. This will reduce any unwanted delays of data acquisition instances and will greatly help with data management. In case a lot of additional sensors are used, the network architectures and data transmission protocols must also be analysed and selected. For instance, multi-hop networks proved to be successful, when implementing Ad-hoc On-demand Distance Vector [38] and Dynamic Source Routing algorithms [39].
3. *Real-world configuration*. The video and LiDAR monitoring devices must be configured during initial stages of real-world operation. The contrast and sensitivity of video cameras must be either set up static to accommodate a full day cycle, or have it dynamically updated according to environmental illumination and other external factors. Dynamically adjusting sensitivity during precipitation allows to reduce the number of likely unsuitable images [40]. However, every new configuration must be validated and calibrated anew.
4. *Traffic intensity*. Make a statistical analysis of traffic intensity at a particular intersection and dynamically adjust configuration based on time and day: vehicle may change lanes and the traffic flow rate may change several times per minute, resulting in a requirement to dynamically rearrange the image sequences from camera detection; therefore, it is necessary to make a dynamic adjustment to the instantaneous speed of the traffic flow and the delay coefficient. The best approach is to find an effect the traffic flow has on the video and LiDAR data acquisition frequency. This can be performed by using neural networks [41].
5. *System's redundancy*. In order to maximise the accuracy of solution in addition to overlapping Field of Views, additional sensors (infrared sensor, inductor embedded in asphalt), which cause the camera to take additional pictures when the vehicle is in its field of view, may be used. In case night time monitoring is required an infrared sensor must be installed; although these sensors suffer performance loss in foggy conditions [42].
6. *Calibration*. After the installation of LiDAR, perform the manual on-site validation of the actual vehicles' recognition operation and configure the equipment in accordance with accuracy. If possible, automatic calibration must be implemented [43].
7. As the privacy of every vehicle's owner is protected by GDPR in order to implement a similar solution on public roads the consensus from local authorities is required.

From the engineering standpoint, the implementation of IoT has achieved two aspects: improved the security – automatic notification of parking or speeding violation; and automation of registering traffic related events with associated data about each individual vehicle and traffic in general. In regard to extensibility and future work, the solution uses a centralized processing data centre and a single observation cluster. This cluster, containing monitoring hardware can be multiplied across different intersections and/or intersection parts, thus creating a network that is usable for creating a city wide intelligent transportation system.

Conclusions

1. During practical experimentation, the most noticeable performance effecting factors were hardware settings (up to 8% erroneous speed detection using LiDAR in 3 days, up to 17% images with data collision "noise"), illumination (up to 34% of images unsuitable for recognition in 1h) and precipitation (up to 42% in false positive events in 40 minutes). The degree of their effect directly relates to the task on hand – while parking violation can be realised with limited resources and computing power, real-time traffic surveillance with vehicle number plate's recognition requires optimisation of data exchange and security protocols and performance aspects of machine learning algorithms.
2. Each factor can be tackled on its own by various existing methods; however, these methods are not developed in mind of being used in combination, and, therefore, the particular stack of algorithms and methods must be determined on-site and be based on complexity of particular tasks.
3. Based on the literature review and experiment results there is currently no ideal setup or combination that could be used for every task. Instead, the general guidelines should be used as the

first step when developing a similar traffic surveillance system. In addition, implementation of the proposed or similar solutions must take into consideration local policies (for example, data protection or city policies), existing infrastructure (availability of fast network, camera and sensor placement) and funding as LiDAR, cameras and processing information system can add to significant costs for local governments.

Acknowledgements

The research leading to these results has received funding from the project "Competence Centre of Information and Communication Technologies" of the EU Structural funds, contract No. 1.2.1.1/18/A/003 signed between the IT Competence Centre and Central Finance and Contracting Agency, Research No. 1.12 "Multi-object detection and tracking in vehicle traffic surveillance: fusing 3D-LiDAR and camera data".

Author contributions

Data interpretation and analysis, N.B., G.V.; writing and original draft preparation, N.B.; review and editing, G.V.; data acquisition and communication with the development team of WeAreDots Ltd., I.M. All authors have read and agreed to the published version of the manuscript.

References

[1] Nižetić S., Šolić P., López-de-Ipiña González-de-Artaza D., Patrono L. "Internet of Things (IoT): Opportunities, issues and challenges towards a smart and sustainable future", Journal of Cleaner Production, vol. 274, 2020.
[2] Alam F., Mehmood R., Katib I., Albeshri A. "Analysis of Eight Data Mining Algorithms for Smarter Internet of Things (IoT)", in Procedia Computer Science, 2016, vol. 58, pp. 437-442.
[3] Putra A. S., Warnars H. L. H. S. "Intelligent Traffic Monitoring System (ITMS) for Smart City Based on IoT Monitoring", in 1st 2018 Indonesian Association for Pattern Recognition International Conference, INAPR 2018 - Proceedings, 2019, pp. 161-165.
[4] Chen N., Chen Y. "Smart City Surveillance at the Network Edge in the Era of IoT: Opportunities and Challenges", 2018, pp. 153-176.
[5] Luo W., Xing J., Milan A., Zhang X., Liu W., Kim T. K. "Multiple object tracking: A literature review", Artificial Intelligence, vol. 293. 2021.
[6] Xia R., Chen Y., Ren B. "Improved anti-occlusion object tracking algorithm using Unscented Rauch-Tung-Striebel smoother and kernel correlation filter", Journal of King Saud University - Computer and Information Sciences, 2022.
[7] Vu A. K. N. et al. "Few-shot object detection via baby learning", Image and Vision Computing, vol. 120, 2022.
[8] Agustina J. R., Galdon Clavell G. "The impact of CCTV on fundamental rights and crime prevention strategies: The case of the Control commission of video surveillance devices", Computer Law and Security Review, vol. 27, no. 2, pp. 168-174, 2011.
[9] Patil P. W., Dudhane A., Chaudhary S., Murala S. "Multi-frame based adversarial learning approach for video surveillance", Pattern Recognition, vol. 122, 2022.
[10] Kumar V., Subramanian S. C., Rajamani R. "A novel algorithm to track closely spaced road vehicles using a low density flash lidar", Signal Processing, vol. 191, 2022.
[11] Pawar K., Attar V. "Deep learning based detection and localization of road accidents from traffic surveillance videos", ICT Express, 2021.
[12] Adamová V., Boroš M. "Effective Placement of Video Surveillance System Using 3D Scanning Technology for Traffic Safety", in Transportation Research Procedia, 2021, vol. 55, pp. 1665-1672.
[13] Zhao Y., Wu W., He Y., Li Y., Tan X., Chen S. "Practices and a strong baseline for traffic anomaly detection", in IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2021, pp. 3988-3996.
[14] Minnikhanov R., Dagaeva M., Anikin I., Bolshakov T., Makhmutova A., Mingulov K. "Detection of traffic anomalies for a safety system of smart city", in CEUR Workshop Proceedings, 2020, vol. 2667, pp. 337-342.

[15] "Adaptive Recognition: Vidar in action", 2022. [online][14.03.2022] Available at: https://adaptiverecognition.com/vidar/vidar-in-action/#Request_an_offer
[16] "Traffic Control & Surveillance System (TCSS)", 2022. [online][14.03.2022] Available at: https://ev-dynamic.com/traffic-control-surveillance-system-tcss/
[17] "SensoTec: City Surveillance Solution", 2022. [online][14.03.2022] Available at: https://sensortec-eu.com/city-surveillance-solution/
[18] Felici-Castell S., García-Pineda M., Segura-Garcia J., Fayos-Jordan R., Lopez-Ballester J. "Adaptive live video streaming on low-cost wireless multihop networks for road traffic surveillance in smart cities", Future Generation Computer Systems, vol. 115, pp. 741-755, 2021.
[19] Finogeev A., Finogeev A., Fionova L., Lyapin A., Lychagin K. A. "Intelligent monitoring system for smart road environment", Journal of Industrial Information Integration, vol. 15, pp. 15-20, 2019.
[20] Fu L., Yu H., Juefei-Xu F., Li J., Guo Q., Wang S. "Let There be Light: Improved Traffic Surveillance via Detail Preserving Night-to-Day Transfer", IEEE Transactions on Circuits and Systems for Video Technology, 2021.
[21] Bahnsen C. H., Moeslund T. B. "Rain removal in traffic surveillance: Does it matter?", IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 8, pp. 2802-2819, Aug. 2019.
[22] Pramanik A., Sarkar S., Maiti J. "A real-time video surveillance system for traffic pre-events detection", Accident Analysis and Prevention, vol. 154, 2021.
[23] He Y. et al. "Vehicle Traffic Driven Camera Placement for Better Metropolis Security Surveillance", arxiv.org, 2017, [online][14.03.2022] Available at: https://arxiv.org/abs/1705.08508
[24] Han Y., Hu D. "Multispectral fusion approach for traffic target detection in bad weather", Algorithms, vol. 13, no. 11, pp. 1-13, 2020.
[25] Asborno M. I., Burris C. G., Hernandez S. "Truck Body-Type Classification using Single-Beam Lidar Sensors", Transportation Research Record, vol. 2673, no. 1, pp. 26-40, Jan. 2019.
[26] Zhang J., Lin X. "Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing", International Journal of Image and Data Fusion, vol. 8, no. 1. Taylor and Francis Ltd., pp. 1-31, Jan. 02, 2017.
[27] Li J., Zhang X., Li J., Liu Y., Wang J. "Building and optimization of 3D semantic map based on Lidar and camera fusion", Neurocomputing, vol. 409, pp. 394-407, 2020.
[28] Faouzi N. E. El and Klein L. A. "Data Fusion for ITS: Techniques and Research Needs", in Transportation Research Procedia, 2016, vol. 15, pp. 495-512.
[29] Faouzi N. E. El and Klein L. A. "Data fusion in intelligent traffic and transportation engineering: Recent advances and challenges", in Multisensor Data Fusion: From Algorithms and Architectural Design to Applications, CRC Press, 2017, pp. 563-594.
[30] Kenk M. A., Hassaballah M. "DAWN: Vehicle Detection in Adverse Weather Nature Dataset", Aug. 2020, [online][14.03.2022] Available at: http://arxiv.org/abs/2008.05402
[31] Manogaran G. et al. "Multi-Variate Data Fusion Technique for Reducing Sensor Errors in Intelligent Transportation Systems", IEEE Sensors Journal, vol. 21, no. 14, pp. 15564-15573, 2021.
[32] Schneider S., Himmelsbach M., Luettel T., Wuensche H. J. "Fusing vision and LIDAR - Synchronization, correction and occlusion reasoning", in IEEE Intelligent Vehicles Symposium, Proceedings, 2010, pp. 388-393.
[33] Wu T. E., Tsai C. C., Guo J. I. "LiDAR/camera sensor fusion technology for pedestrian detection", in Proceedings - 9th Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2017, 2018, vol. 2018-Febru, pp. 1675-1678.
[34] Zhao G., Xiao X., Yuan J., Ng G. W. "Fusion of 3D-LIDAR and camera data for scene parsing", Journal of Visual Communication and Image Representation, vol. 25, no. 1, pp. 165-183, Jan. 2014.
[35] Vitols G., Bumanis N., Arhipova I., Meirane I. "LiDAR and Camera Data for Smart Urban Traffic Monitoring: Challenges of Automated Data Capturing and Synchronization", in Communications in Computer and Information Science, 2021, vol. 1455 CCIS, pp. 421-432.
[36] Wang J., Gai S., Huang X., Zhang H. "From coarse to fine: A two stage conditional generative adversarial network for single image rain removal", Digital Signal Processing: A Review Journal, vol. 111, Apr. 2021.
[37] Rashid H. T., Ali I. H. "Multi-Camera Collaborative Network Experimental Study Design of Video Surveillance System for Violated Vehicles Identification", in Journal of Physics: Conference Series, 2021, vol. 1879, no. 2.

[38] Yan S., Chung Y. "Improved Ad Hoc On-demand Distance Vector Routing (AODV) Protocol Based on Blockchain Node Detection in Ad Hoc Networks", International Journal of Internet, Broadcasting, vol. 12, no. 3, pp. 46-55, 2020.
[39] Zhang D. gan Liu S., Liu X. huan Zhang T., Cui Y. ya, "Novel dynamic source routing protocol (DSR) based on genetic algorithm-bacterial foraging optimization (GA-BFO)", International Journal of Communication Systems, vol. 31, no. 18, Dec. 2018.
[40] Jiang S., Babovic V., Zheng Y., Xiong J. "Advancing Opportunistic Sensing in Hydrology: A Novel Approach to Measuring Rainfall With Ordinary Surveillance Cameras", Water Resources Research, vol. 55, no. 4, pp. 3004-3027, Apr. 2019.
[41] Khazukov K. et al. "Real-time monitoring of traffic parameters", Journal of Big Data, vol. 7, no. 1, Dec. 2020.
[42] Gulati I., Srinivasan R. "Image processing in intelligent traffic management", International Journal of Recent Technology and Engineering, vol. 8, no. 2 Special Issue 4, pp. 213-218, 2019.
[43] Hu Z., Ma W. "Self-calibration of Traffic Surveillance Camera Systems for Traffic Density Estimation on Urban Roads", limos.engin.umich.edu, vol. 1, no. 412, pp. 1-8, 2020.