sciendo

Transport and Telecommunication, 2023, volume 24, no. 1, 13–21
Transport and Telecommunication Institute, Lomonosova 1, Riga, LV-1019, Latvia
DOI 10.2478/ttj-2023-0002

LIDAR AND AI BASED SURVEILLANCE OF INDUSTRIAL PROCESS ENVIRONMENTS

Maik Groneberg[1], Olaf Poenicke[1], Chirag Mandal[1], Nils Treuheit[1]

[1]Fraunhofer Institute for Factory Operation and Automation IFF
Magdeburg, Germany, Sandtorstr. 22
maik.groneberg@iff.fraunhofer.de
olaf.poenicke@iff.fraunhofer.de
chirag.mandal@iff.fraunhofer.de
nils.treuheit@iff.fraunhofer.de

The paper describes a system approach to use LiDAR sensors for capturing dynamic point cloud data in industrial process environments and to interpret the captured scenes with AI based object detection. The object detection is used to distinguish between humans and other mobile objects in safety relevant workspaces. Several AI methods relevant for such application are analysed. One method is applied with annotated test data and evaluated concerning its accuracy.

Keywords: LiDAR, dynamic 3D point cloud, object detection, AI, industrial application, synthetic data

1. Introduction

The development of new LiDAR sensors is pushed by the general trend of autonomous vehicles, as the principle of LiDAR sensors is seen as a key enabler for such applications. Due to these developments performance of LiDAR sensors is increasing while costs are decreasing. This of course is also relevant for other application areas such as the capturing and analysis of dynamic 3D point clouds in industrial environments – e.g., to monitor process environments for safety purpose (Hünermund *et al.*, 2021). In order to actually interpret such scenes and situations, relevant objects must be reliably detected and separated within the 3D scene.

The paper describes the system approach for a LiDAR and AI based surveillance system and the first evaluation of AI methods for a specific application scenario. Following the related work (section 2) and based on the description of two different industrial application scenarios (section 3), the paper specifically describes how the system concept was worked out (section 4) and what data are used for training and evaluation of AI methods (section 5). In section 6 it describes how AI methods are used to detect humans and to differentiate other objects within the LiDAR data. Furthermore, the first evaluation results are summarised. The paper concludes in section 7 with an outlook on future work.

2. Related Work

The section of related work focusses on three different aspects, as the research described in this paper is integrating different technical aspects. Thus, the following sub-sections focus on related work in the fields of safety related applications, the use of AI methods for object detection and classification and finally specifics of 3D data processing.

2.1. Safety related Applications

The development of safety concepts and technical surveillance of automated cells is not only a recent topic. The requirement to design robot cells more flexible emerged at the early days of human-robot collaboration. Based on that, requirements emerged to also make automated cells fenceless, which have no focus on human-robot collaboration. Thus, novel technical approaches were required to ensure the functional safety of such systems. LiDAR technology was already employed some time ago to detect objects in a robot cell and to accordingly adapt the robot control (Karlsson *et al.*, 2000). Early LiDAR technologies were, due to a limited field of view and a lower sampling rate, only partially suitable to identify all possible risk factors. Only with the technical advancement, these limitations became obsolete.

2D cameras offer some benefits – with increasing progress in the field of artificial intelligence, the use of 2D cameras emerged in recent years for localization of humans and other objects (Tan and Arai, 2011). However, 2D imaging basically enables detection and identification. Accurate spatial perception is only possible to a limited extent or by fusion of different technologies (Frese *et al*., 2014).

There are different vision-based solutions available on the market for industrial application environments. Systems such as Pilz Safety Eye (using 2D cameras) and VeoRobotics (using TOF-based 3D imaging) (Moel, 2022) enable robust and certified detection of mobile objects in the working environment of automated machines. Another system for spatial 3D surveillance is INXPECT which is based on radar technology. However, these systems are not able to differentiate between different object classes and therefore do not enhance the efficiency of automated systems.

2.2. AI Methods for Object Detection

In recent years there have been several important research publications focussing on the use of AI methods for object detection and classification, also with focus on point cloud data as a basis. (Li, 2017) is one of the earliest papers regarding object detection within point cloud data. In this paper the researchers propose the extension of 2D fully convoluted networks for object detection on images to 3D point cloud data. A comprehensive review of the recent progress made in the deep learning methods for point clouds is given by (Guo *et al*., 2020). It covers 3D point cloud classification, 3D point cloud object detection, tracking and 3D point cloud segmentation. It also presents comparable results on several publicly available datasets along with insightful observations and a lot of future research directions. This paper suggested the use of *PointPillars* as a network for object detection. The performance of *PointPillars* is better than other networks with similar architecture. The details about the network architecture are discussed in section 6 of this paper.

The original *PointPillars* paper (Lang *et al.*, 2019) proposes its novel encoder architecture which enables inference at 62 Hz, which is much faster than other networks. Another relevant network, *Complex YOLOv4* (Simon *et al.*, 2018), aims to design an object detector which operates really fast in a production environment and optimizes parallel computation. *SE-SSD* (Self-ensembling Single Stage Detector) (Zheng *et al.*, 2021) contains a student and teacher single shot detector, in which an effective intersection over union (IoU) matching strategy is designed to filter soft targets from the teacher SSD and a consistency loss is formulated to align student predictions with them.

The research work (Petschnigg and Pilz, 2021) describes object detection on the factory floor to increase the safety of the staff on the floor. In the paper, the researchers have worked on the uncertainty estimation of the point cloud segmentation in factory planning. As part of future work, uncertainty estimation will be used for point cloud object detection to increase the reliability and explainability of the application.

2.3. 3D Data Processing

Although new AI methods for object detection in point clouds are becoming more and more relevant, conventional methods that segment dynamic partial point clouds by identifying and subtracting the background still play an important role. This is due to the required computing power and due to the reliability of such systems. By using different filters and hybrid background models, the negative influence of noise (which plays a significant role in point clouds) onto the result of background subtraction could be reduced (Braham *et al.*, 2014).

Some AI methods extract specific information from the point clouds – e.g. a voxel-based representation can favour the use of AI methods and accelerate inference. Furthermore, a voxel grid can serve as a starting point for a segmentation of the point cloud and thus also represent an alternative for the detection of dynamic objects (Hünermund *et al.*, 2021).

3. Application Scenarios

The developments described in the paper aim to differentiate between different mobile objects and humans in industrial workspaces. Accordingly, the aim is not only to detect movements in the workspace, as other safety systems do, but to distinguish between humans and different object classes. By means of such a differentiation, the application scenarios described below can be technically supported in order to further increase the efficiency and productivity of automated and semi-automated systems. For the scenarios described, the authors are working on the development and testing of LiDAR-based applications in currently ongoing projects.

3.1. Scenario 1: Workspace Surveillance for a fenceless Order Picking Cell

In this scenario the system is intended to monitor the external safety zone of an autonomous and fenceless robot cell for picking of automotive components. The internal safety zone is covered by conventional safety laser scanners. The cell is accessible on 3 of the 4 sides for autonomous and conventional vehicles supplying the cell with load carriers, which are stored at dedicated locations. On the fourth side, only an AGV can pass to convey the picked parts. No other vehicles or individuals shall enter the cell. Since the three sides must be accessible from the outside, conventional safety systems are not suitable for safeguarding the robot cell, as the cell would need to stop every time that new load carriers are delivered.

The task for the LiDAR-based system surveilling the external safety zone is to not only detect but also to analyse moving objects in the external safety zone. Information about the class, position, direction of movement and velocity of the objects enables the robot control to react to different objects in a different way. This prevents unnecessary emergency stops, as the robot control can enable supplying vehicles to enter without interrupting the robot operation. If a human or unknown object is detected, the robot operation shall slow down or come to a full stop. Thus, a robust differentiation between humans, known objects (supply vehicles) and other unknown mobile objects is required, to ensure an efficient picking process.

3.2. Scenario 2: Workspace Surveillance in a Foundry Turntable

In contrast to scenario 1, this scenario does not deal with hazards from the robot itself, but with hazards from molten metal in a foundry. In many permanent mould foundries, so-called foundry turntables are used. On these tables, the moulds are arranged in a circle along the process chain, with the actual pouring of the molten metal into the mould taking place at one position.

Due to the accessibility of the turntable, it must be ensured that humans inside are not in the safety-critical area of this casting station, as splashes of the molten metal can cause hazards. Due to the process with the table turning after each process step, the safety-critical area changes continuously in relation to the positions on the turntable.

Since other objects can also be temporarily placed in the inner area of the turntable (e.g. due to parallel work) and access must be enabled throughout, the LiDAR system approach offers the advantage here that humans can be differentiated from other objects in the interior. This can potentially provide a safeguard for the dynamically changing safety area.

3.3. Relevant Aspects of the Application Scenarios

In the scenarios described above, as well as in other potential fields of application, the core task is to robustly detect and locate humans in the workspace of an automated system. The position of a detected human must be evaluated in relation to a concrete hazard. In this respect, a desired safety device must operate with very low latency. With regard to the connection to a robot PLC, for example, typical query cycles to safety systems are 100ms.

Both scenarios are about functional safety, as defined for example according to *EN ISO 13849*. As described at the beginning of section 1 and further substantiated in the following sections, AI methods are used in the developed approach to detect humans and other objects within the dynamic 3D point clouds. To date, there are no established standards and norms on the basis of which the use of AI in connection with functional safety can be certified. However, first individual standards exist (ANSI, 2020) and the use of AI is an important topic that is to be anchored in future standards.

The development work described in the following sections focuses on the application of the described scenario 1. However, the requirements and work on the robust classification of humans and their differentiation from other mobile objects are directly transferable to scenario 2 and other similar application scenarios.

4. System Concept

As is clear from the use cases, the aim of the LiDAR system considered in this paper is not to monitor machines or processes themselves. Rather, it is about monitoring the safety-critical workspaces in which human traffic is permitted under certain conditions. Using the scenario described in section 3.1 as an example, a solution approach is presented below.

4.1. System Architecture

The proposed system consists of one or more LiDAR sensors distributed in the workspace. Each sensor has assigned an edge computing unit for processing the point cloud data. The edge devices are connected to an IPC via Ethernet. This IPC serves as a central interface and aggregates or filters all incoming and outgoing signals of the LiDAR sub-systems and forwards the relevant ones to the PLC. Furthermore, it serves as an interface for the configuration and monitoring of the overall LiDAR systems (see Figure 1).

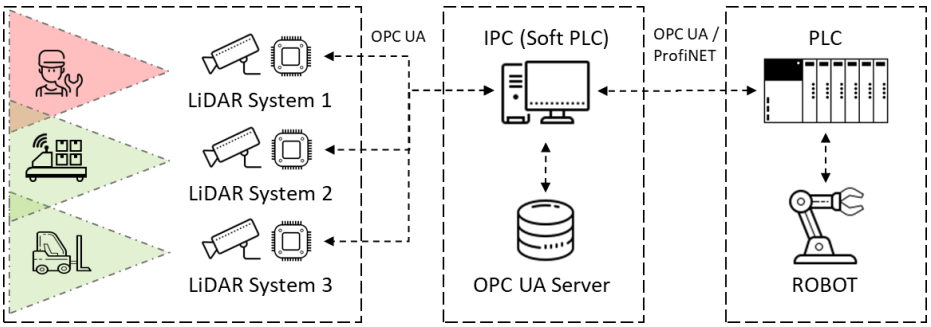


Figure 1. Schematic representation of the components and communication interfaces which constitute the hardware system

The arrangement of the sensors must be defined in such a way that the fields of view (FOV) of the sensors are at least adjacent or overlapping (see Fig. 2. left). However, the processing of the sensor data is sensor-specific, which means that a registration of the sensors is not necessary for the sensor data processing during runtime (see Fig. 2 right). However, registration is important for validation of the sensor orientations and the required coverage of the overall system.

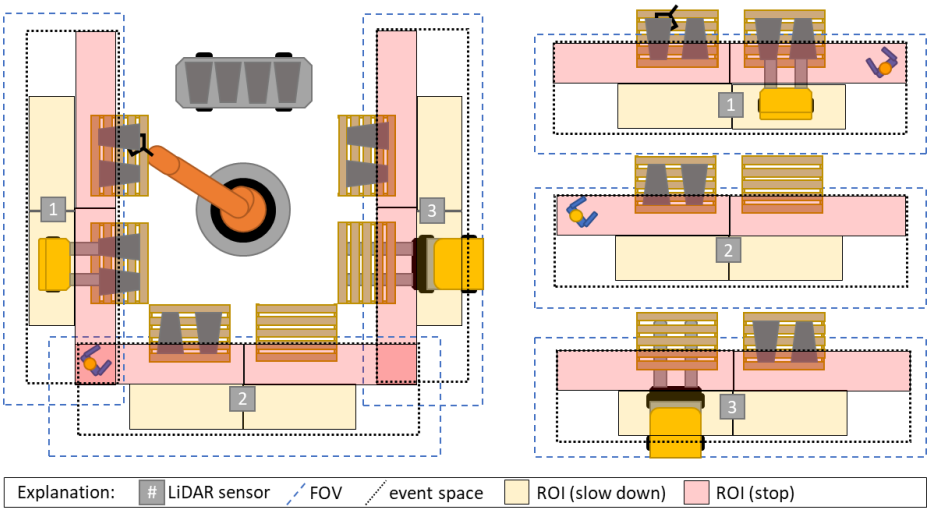


Figure 2. Schematic representation of the robot cell layout (left) and the distinctive sensor-specific views (right)

Each LiDAR sub-system (sensor + edge computing unit) functions autonomously and is not influenced by the functioning of neighbouring systems or dependent on their results. This has various advantages: the chosen AI methods are independent of the number and arrangement of the sensors used. Furthermore, one sub-system achieves computes consistent point clouds with regard to the number and distribution of points. Using sub-systems also enhances the flexibility to integrate the solution in a wide variety of scenarios.

For each sub-system an interface enables configuration of various parameters (e.g. event space, safety areas (ROI)). An event space is defined within the FOV (generally oriented downwards). All 3D point clouds located in the event space are included in the further data processing. Within the event space, ROIs of any number, shape and with different safety levels can be defined (see Figure 2). If people or objects enter an ROI, events are triggered that can be forwarded to the PLC. Furthermore, it can be defined under which conditions an ROI is triggered (e.g. depending on the object class).

Based on the differentiation of humans and other objects, rules can be defined within the PLC to enable movements of defined objects in safety critical areas but to slow down (yellow area in fig. 2) or stop (red area in fig. 2) the robot if humans or unknown objects enter the critical ROIs. In overlapping FOVs and event spaces the sub-systems trigger events independently. In this way, redundant information about one and the same object is generated, which should not be transferred directly to the robot's PLC. For this reason, the central IPC is taking over the clearing of event information.

4.2. Design and Functionality of the Application Modules

In this section, the data processing of the LiDAR systems is briefly described. The schematic process shown (Figure 3) includes all tasks from data acquisition to information transfer via OPC UA within a target cycle time of 100ms. After the point cloud has been recorded, it is transformed from the LiDAR coordinate system into a uniform coordinate system. This way any tilted angles of the sensor orientation towards the floor are compensated. The advantage of this approach is that the floor is always horizontally oriented, which enables the extraction of the floor from further data processing. Another effect is that the coordinates of the points of an object are independent of the sensor position, especially the height.

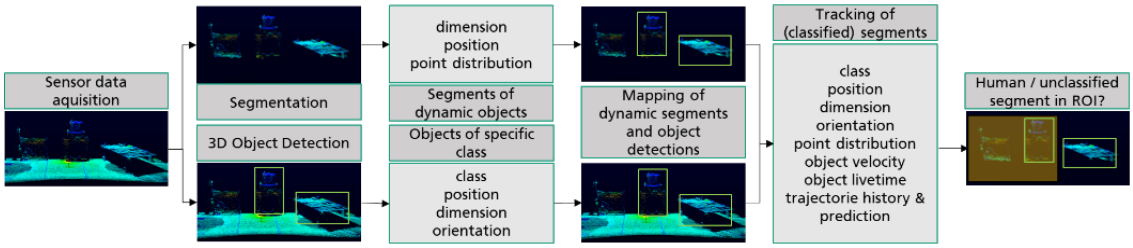


Figure 3. Exemplary processing pipeline for a LiDAR system, consisting of a 3D AI object detection and a parallel, non-AI-based object detection, using background subtraction, to ensure functional safety claims

The transformed point cloud is then processed in two parallel process streams. Firstly, a background subtraction is applied to the raw data, whereby all moving objects can be identified as point cloud segments. The quality can be positively influenced by various filters, such as noise filters. The output of the background subtraction are segments with specific properties, such as the dimension, position and the point distribution. In the second processing stream, the 3D object recognition is performed on the raw data. AI methods which are considered for this are discussed in chapter 6. The AI provides the position, dimension, orientation and a confidence for the different objects of a class that were detected in a frame (3D point cloud integrated over 100ms).

Both processing streams are then connected again to search for matches between the point cloud segments (background subtraction) and the object detections (AI) in terms of position and dimension. The corresponding objects and their properties are summarised and tracked over their lifetime (visibility in the FoV) and enriched with further information. Thus, the speed, the trajectory, as well as the gain-filtered confidence can be decisive in the evaluation of the objects with regard to hazardous situations. In the last step, the objects movements within ROIs are checked.

If a human or a segment to which no class could be assigned is detected within a ROI, a referring safety event is triggered. As mentioned, each LiDAR sub-system acts autonomously and only processes point clouds of its own sensor. However, an exchange of information via OPC UA is not excluded. This allows information about already known/tracked objects to be exchanged, which makes tracking across multiple LiDAR sub-systems possible.

5. Dataset for AI Application on Scenario 1

Before the AI methods for 3D object detection were tested for the use case of scenario 1, recorded data from the application environment had to be processed in an appropriate form. Due to the specific point clouds and objects in the application environment, free available models that were trained with data sets from autonomous driving (e.g. *KiTTI* or *nuScenes*) could not be used. Thus, the real process data of scenario 1 were manually annotated with a specially developed programme. The format of the labels is based on the *KiTTI* format for 3D object detection. 3D bounding boxes were set with a position, orientation, dimension and flags for occlusion and truncation. Fig. 4 shows an example of a point cloud of humans with (right) and without truncation (left).

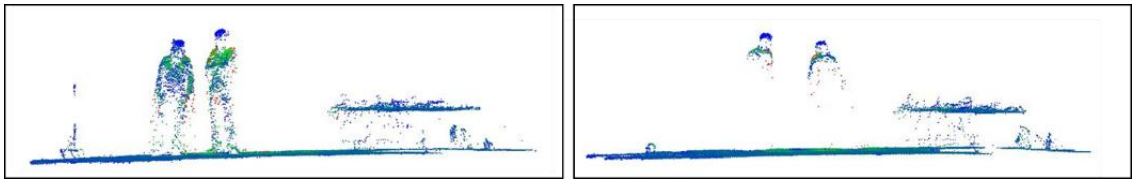


Figure 4. Exemplary point cloud of a human group and a sequence wagon with truncation (right) and without (left)

This enables a targeted evaluation of the training process with regard to the quality of the object point clouds. So far, 10,248 annotated LiDAR frames have been generated, containing 31,323 objects across all object classes, which are distributed across the five classes of load carrier, human, sequence car, cyclist and AGV as shown in Table 1. The available dataset was only divided into two parts, the training dataset (70%) and the validation dataset (30%), due to the small amount of data.

Table 1. Structure of the annotated LiDAR data used for AI evaluation

Description	Frames	Carrier	Humans	Seq. Car	Cyclist	AGV
Training	7174	8830	8917	3736	50	2269
Validation	3074	2934	1793	994	14	1786
Overall	10248	11764	10710	4730	64	4055

Obviously, some classes occur more frequently than others in the real process environment and are therefore more often contained in the data. This needs to be compensated in the future through targeted data recording and annotation. Since manual annotation of the data is very time-consuming and costly, approaches are tested to artificially enrich the data. This is where data augmentation and synthetic data generation can help, as the use of synthetically generated data can significantly reduce the necessary amount of annotated real-world data (Borkman *et al.*, 2021). First tests were carried out with *ISAAC Sim*, which is a simulation environment from *Nvidia* that makes it possible to simulate LiDAR sensors and entire 3D modelled processes, thereby automatically generating annotated training data (see Figure 5 comparing a real-world point cloud and a synthetic point cloud).

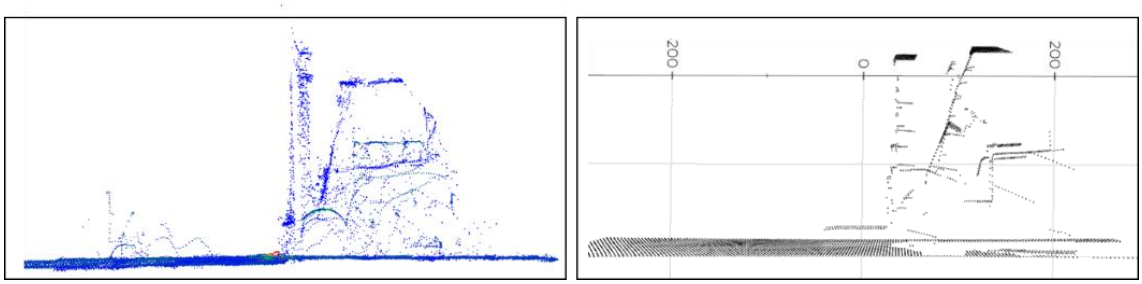


Figure 5. Reals point cloud of a forklift (left) and a synthetic generated with Nvidia´s ISAAC Sim (right)

The evaluation of different AI methods as described in section 6 was carried out based on the annotated real-world data summarised in Table 1.

6. Selection and Evaluation of AI Networks

Initially, the selection of suitable networks was done, using the top performers of the *KITTI 3D benchmark* (Geiger *et al.*, 2017). Special attention was paid to the quality of the results and the performance during inference. Three networks, which were already mentioned in section 2.2 were finally selected:

PointPillars (Lang *et al.*, 2019) which consists of 3 main components. A feature encoding network to convert a 3D point cloud into a sparse pseudo image. Then a 2D convolutional network that acts as a backbone to convert the pseudo image into a high-level representation and finally an object recognition head that recognises and predicts 3D boxes.

The *SE-SSD* (Zheng *et al.*, 2021) which is built on two main components – Teacher Single Stage Detector and Student Single Stage Detector. Both the single stage detectors comprise of a sparse

convolution network (*SPConvNet*), multi-task head (*MT_Head*) and a Bird's Eye View convolution network (*BEVConvNet*).

Finally, the *Complex YOLOv4* (Simon *et al.*, 2018) which comprises of the following components. The backbone of the architecture is a *CSPDarknet53*. The middle portion of the network has two sub parts, a Spatial Pyramid Pooling which can generate a fixed-length representation irrespective of image size/scale and a Path Aggregation Network (*PANet*) which helps in boosting information flow in an instance segmentation framework which is proposal-based. The head of this network comprises of a *YOLOv4* network.

The aforementioned networks were adapted, parameterised and trained with regard to the available data. Initial results on the performance of *Complex YOLOv4* are already available and described subsequently. The training was performed for 300 epochs on the *Complex YOLOv4* model and then the test was performed on the validation set. Since more objects were found as contained in the data set, a method was implemented to merge detections that belong to the same class and have a very high similarity. For this purpose, a box loss was defined, which is calculated with the following equation.

$$boxL = \sqrt{locL^2 + dimL^2 + rotL^2 + iouL^2 + giouL^2 + ciouL^2 + diouL^2}. \quad (1)$$

Since the network is not able to predict the Z-position and -dimension, and instead simply uses predefined values for classes, these values were not included in the calculation of the loss metrics. Intersection over union (*IoU*) is an essential metric for evaluation of bounding box predictions, which are used here in different forms as described by (Wang and Song, 2021) (*iouL*, *giouL*, *ciouL* and *diouL*). Location loss (*locL*), dimension loss (*dimL*) and rotation loss (*rotL*) are normalised values for position and dimension error on two axes, as well as yaw rotation. The box loss was calculated for each possible combination among the predictions of one LiDAR frame. Assuming that only one prediction per frame can be nearly identical, the two boxes of one class with the strongest match were selected. If a match is found and the box loss is smaller than the defined threshold, the two detections are merged into one. This resulted in two subsets of the predictions P(1.4) and P(2). The subsets are identified by the threshold used to compare the similarity of the detections. For *Complex YOLOv4*, the resulting errors of predictions and filtered predictions in relation to the ground truth are presented in Table 2.

Table 2. Evaluation of the predictions of Yolo-Complex with different filters

Class	**GT**	**P**	$\mathbf{e^P}$	**P(1.4)**	$\mathbf{e^{P(1.4)}}$	**P (2)**	$\mathbf{e^{P(2)}}$
Carrier	2934	2934	2	2936	2	2936	2
Cyclist	14	15	1	15	1	15	1
AGV	1786	1859	73	1859	70	1848	62
Human	1793	1801	8	1800	7	1800	7
Seq. Car	994	997	3	995	1	995	1
Overall	7521	7608	87	7605	81	7594	73

The error reflects the deviation per class between the (filtered) prediction (P) and the ground truth (GT) information. Since in each of the attempts, despite filtering, all objects contained in the ground truth were detected, the recall is always 1. However, precision could be increased from P = 98.86% to P(1.4) = 98.89% and P(2) = 99.04%, respectively. The filtering reduced the false detections by a maximum of 16.09%. Some metrics are presented in Table 3 in order to provide an indication of the accuracy of the box predictions. The merge in the different filtering approaches has no notable influence on the loss metrics due to the small number of filtered values. The review of the false predictions showed that many of these false predictions for AGV consist of point clouds that can be assigned to AGV and Carriers. Such behaviour could be due to the fact that AGVs and Carriers often co-occur during transport processes and occupy a similar space in the underlying 3D point cloud data.

Table 3. Exemplary averaged loss values per class

Class	**Location loss**	**Dimension loss**	**Rotation loss**	**IoU loss**
Carrier	0.002665	0.006679	0.011309	0.005874
Cyclist	0.003723	0.053525	0.014873	0.033284
AGV	0.006686	0.009158	0.013783	0.018670
Human	0.002628	0.003964	0.007364	0.010592
Seq. Car	0.001551	0.003702	0.004418	0.005163

The results show a slightly increased dimension loss for cyclists. This is probably due to the small amount of available data. Yaw rotation detection can be helpful in analysing time series. Since this has a very low loss, the results are reliable and can give insight into the direction of movement of the objects, despite the frame-based evaluation.

7. Conclusion and Future Work

The paper described the current status of developments for a LiDAR and AI based system for surveillance of industrial process environments. The modular approach of the system potentially allows its use in various application fields. However, the use of AI for detection of object classes requires application specific training. The first training and evaluation of an AI model was described for one application scenario and shows promising results.

As described in section 6 several AI models seem relevant to the targeted application context. Thus, the next step will be to also train and test further AI models (*PointPillars* and *SE-SSD*) and to evaluate and compare the test results. To also further enhance data sets the relevance of data augmentation and of synthetic data generation was discussed in section 5. This will also be an important task to develop the existing approaches for synthetic data generation further and to evaluate the performance of the AI models with different shares of real and synthetic data.

Another relevant task is to enable the use of different types of LiDAR sensors, as these differ concerning their FOV, their resolution and also their scan patterns. The test data used and described in the paper were generated with *Livox* sensors which employ a so-called pseudo random scan pattern. As other LiDAR sensors usually employ regular linear scan patterns, it needs to be analysed in how far the scan characteristics have an influence on the AI based object detection.

Acknowledgements

Part of the developments and evaluations described in the paper was carried out in the project ‘LiDAR-basierte Arbeitsraumanalyse’ funded by the Federal State of Saxony Anhalt and the European Commission within the EFRE and REACT-EU program. Funding reference: FKZ 2104/00066.

References

1. Hünermund, M., Schütz, A., Groneberg, M. (2021) LiDAR-Sensoren. In: Sackewitz, Michael (ed.) Leitfaden zur optischen 3D- Messtechnik. Fraunhofer Verlag, Stuttgart.
2. Karlsson, B., Karlsson, N. and Wide, P. (2000) A dynamic safety system based on sensor fusion. *Journal of Intelligent Manufacturing*, 11(5), 475-483.
3. Tan, J.T.C., Arai, T. (2011) Triple stereo vision system for safety monitoring of human-robot collaboration in cellular manufacturing. In: *2011 IEEE International Symposium on Assembly and Manufacturing (ISAM)*, IEEE, 1-6.
4. Frese, C., Fetzner, A. and Frey, C. (2014) Multi-sensor obstacle tracking for safe human-robot interaction. In: *41st International Symposium on Robotics, ISR/Robotik 2014, VDE*, 1-8.
5. Moel, A. (2022) How Collaborative Is Your Manufacturing Application? *Quality Magazine*, 61(2), BNP Media.
6. Li, B. (2017) 3D fully convolutional network for vehicle detection in point cloud. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 1513-1518.
7. Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L. and Bennamoun, M. (2020) Deep learning for 3D point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(12), 4338-4364.
8. Lang, A.H., Vora, S., Caesar, H., Zhou, L., Yang, J. and Beijbom, O. (2019) Pointpillars: Fast encoders for object detection from point clouds. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 12697-12705.
9. Simon, M., Milz, S., Amende, K., Gross, H.M. (2018) *Complex-YOLO: Real-time 3D Object Detection on Point Clouds*. arXiv preprint arXiv:1803.06199.
10. Zheng, W., Tang, W., Jiang, L. and Fu, C.W. (2021) SE-SSD: Self-ensembling single-stage object detector from point cloud. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 14494-14503.
11. Petschnigg, C. and Pilz, J. (2021) Uncertainty estimation in deep neural networks for point cloud segmentation in factory planning. *Modelling*, 2(1), 1-17.
12. Braham, M., Lejeune, A., Droogenbroeck, M. V. (2014) A physically motivated pixel-based model for background subtraction in 3D images. In: *International Conference on 3D Imaging (IC3D), Liège, Belgium*, 6-7.
13. Hünermund, M., Groneberg, M., Schütz, A. (2021) Fast Connected Components Object Segmentation on Fused Lidar and Stereo-Camera Point Clouds with Visual-Inertial-Gimbal for Mobile Applications Utilizing GPU Acceleration. In: *Reliability and Statistics in Transportation and Communication*, Bd. 195. Cham: Springer International Publishing (Lecture Notes in Networks and Systems), 73–83.

14. ANSI B11.TR10-2020: Functional Safety of Artificial Intelligence for Machinery Applications (2020).
15. Borkman, S., Crespi, A., Dhakad, S., Ganguly, S., Hogins, J., Jhang, Y. (2021) *Unity Perception: Generate Synthetic Data for Computer Vision*. Online: http://arxiv.org/pdf/2107.04259v2, pp. 6-7.
16. Geiger, A., Lenz, P., Stiller, C., Urtasun, R. (2017) 3D Object Detection Evaluation 2017. Karslruhe Institute of Technology, last checked 04.08.2022, available online at http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=3d.
17. Wang, X., Song, J. (2021) ICIoU: Improved Loss Based on Complete Intersection Over Union for Bounding Box Regression, *IEEE Access*, 9, 105686-105695.