

Drone-Based Environment Perception for Manually Operated Industrial Trucks

Tim Noah Laumann¹,
Torben Mente²,
Carsten Schmidt¹,
Benjamin Küster²,
Ludger Overmeyer^{1,2}

¹ Institute of Transport and Automation Technology
Leibniz University Hannover, Hannover

² IPH - Institut für Integrierte Produktion Hannover gGmbH, Hannover

Manually operated industrial trucks remain indispensable in intralogistics but are prone to accidents caused by limited operator visibility, particularly at intersections and when handling large loads. Existing operator assistance systems are vehicle-mounted or otherwise limited to a direct line of sight, and none provide an elevated, forward-looking monitoring perspective. This paper presents the concept and initial results of a drone-based operator assistance system for the anticipatory environmental monitoring of manually operated forklifts, in which a drone equipped with a fish-eye camera flies a head of the forklift to detect lanes, estimate its pose, and identify hazards. The concept is implemented on an experimental setup using a stationary sensor package in place of the drone. The results demonstrate promising wireless network performance with regard to latency requirements, as well as promising lane detection with an mAP@50 value of 0.986 and person detection with a precision value of 95.2% using YOLO-based networks. The latter also reveals challenges in more crowded scenes. Marker-based pose estimation further enables a correctly oriented representation of the surroundings on a schematic bird's-eye-view map, achieving a maximum positional deviation of 5.0 cm and a rotational deviation of 4.3°. Validation with an actual, moving drone as well as the acquisition of larger, more diverse datasets remain subject to future work.

[Keywords: Environmental Perception, Industrial Truck Safety, Operator Assistance System, Drone-Based Monitoring, Intralogistics]

1 INTRODUCTION

Manually operated industrial trucks remain an indispensable part of intralogistics. Even as automated systems become increasingly intelligent, intralogistics continues to rely on human flexibility and creativity [1]. People remain superior to automated systems in terms of their quick comprehension and adaptability [2]. However, manual transport processes are prone to accidents. In 2024 alone, 18,035 people were injured in accidents involving industrial trucks in Germany, with 41% of the injured not operating the industrial truck themselves but rather being struck, crushed, or run over by it [3]. This resulted in 322 new disability pensions and eight fatalities.

Some of these accidents are attributable to limited visibility during the operation of industrial trucks [4]. The driver's visibility is particularly limited when handling large loads and in intersection areas, which is why these areas are of great importance for accident prevention. People can be obscured or suddenly appear in front of or behind the vehicle without the operator noticing. Currently, there is no operator assistance system (OAS) in the intralogistics context that enables platform-independent and autonomous monitoring of these areas. Therefore, there is a need for an adequate solution that is particularly suitable for small and medium-sized enterprises (SMEs). For SMEs, high investments in networked and automated systems, such as automated guided vehicles (AGVs), always pose a challenge. These systems are currently too inflexible due to their dependence on specific use cases, which is why the flexible use of manually operated industrial trucks remains of great importance [5]. A drone-based OAS for environmental monitoring can extend the operator's field of view to ensure safer operation of manually operated industrial trucks and prevent accidents.

Existing approaches to OASs primarily focus on the immediate vicinity of the load or use ultra-wideband (UWB) and similar technologies to locate industrial trucks and people. To monitor the immediate surrounding area, sensor technology such as fork-tip cameras or a rear-view camera is mounted on the vehicle, and visual cues or audible signals are used to assist in guiding the vehicle [6, 7]. These systems are therefore limited to fixed mounting positions on the industrial truck and consequently cover only a portion of the area in front of or behind the truck. UWB technology requires active tags to all relevant industrial trucks and individuals, facilitating radio-based positioning [8, 9]. However, people not equipped with a tag are not detected and continue to pose a safety risk.

For this reason, the present article proposes a drone-based system for monitoring the surroundings of industrial trucks. This system displays people and objects on a map for the truck operator. The drone's mobility enables it to operate in any desired location, providing a comprehensive view that extends beyond the immediate surroundings of the industrial truck. In comparison with UWB technology, it possesses the additional advantage of eliminating the necessity for objects and people to be equipped with tags, thereby enabling independent detection.

2 STATE OF THE ART

A variety of OASs are currently employed to ensure the safe operation of industrial trucks. In addition to systems that actively intervene in vehicle control, there are systems that provide the operator with additional information or warn of objects in the surrounding area. The following is a brief description of some of these systems.

2.1 COMMERCIALLY AVAILABLE SYSTEMS

The commercially available OASs for industrial trucks can be divided into four functional categories. The categories are as follows: camera-based safety zone extension and object detection, Light Detection and Ranging (LiDAR)-based distance measurement and object detection, radio-based positioning using wearable tags, and optical warning projection.

Camera-based systems, including Toyota SEnS+ [10], Hyster's Pedestrian Awareness Camera [7], Jungheinrich's addedVIEW [11], and STILL's Smart Assist AI camera system [12], utilize artificial intelligence (AI) algorithms to classify persons and alert the forklift operator through both auditory and visual means. In certain instances, the systems intervene in the vehicle's controls, such as automatically reducing speed. A distinguishing characteristic of all camera-based systems is that they are permanently mounted on the forklift, thereby restricting their coverage to the direct field of view of the respective camera position.

LiDAR-based systems employ a laser to scan the forklift's surroundings, thereby measuring the distance to objects. The Crown ProximityAssist system is designed to detect objects in the forklift's projected path and respond by slowing the vehicle down [13]. The ELOprotect ACS is a safety system designed to provide autonomous collision protection for narrow-aisle forklifts [14]. It utilizes two laser scanners mounted on the vehicle to ensure operational safety. These systems have been demonstrated to provide robust distance measurement. However, their functionality is constrained to a fixed, vehicle-bound line of sight.

Radio-based location systems use UWB technology to detect persons and objects. The implementation of this technology requires the provision of tags to all participants, facilitating precise location tracking. Examples of such systems include the Linde Safety Guard [9], Jungheinrich's zoneCONTROL assistance system [15], and the STILL Smart Assist Collision Control [12], among others. The tracking system's capacity to function through obstacles is noteworthy, as it operates in environments where line-of-sight is not a viable condition. This technology offers distinct advantages over conventional vehicle-mounted cameras and laser scanners. However, this system necessitates that every person or object at risk be equipped with a tag, as persons lacking tags will not be detected. In areas where unauthorized persons may enter, or if an operator forgets to wear the tag, the operator of the industrial truck must continue to drive by visual observation.

Optical warning projections, such as Linde BlueSpot [16], are utilized to indicate the direction of travel for industrial trucks by projecting a blue light spot onto the ground. This approach entirely eliminates object detection and functions exclusively as a preventive measure by notifying third parties of an approaching vehicle. Consequently, responsibility is transferred from the operator of the vehicle to external individuals, who can react upon seeing the light spot.

A commonality among the aforementioned systems is that sensors must be either mounted on the industrial truck or worn by the individual being protected. It is evident that none of the systems under consideration are equipped with the capability to monitor the surroundings from an elevated, forward-looking perspective. Such an approach would enable the detection of otherwise challenging-to-see areas, including intersections and zones obscured by shelves, at an early stage, regardless of the industrial truck's line of sight.

2.2 SYSTEMS FROM RESEARCH

The scientific community has also developed several approaches to operator assistance for industrial trucks. These approaches primarily rely on sensor fusion and deep learning. As part of the PräVISION research project, Börold et al. developed a manufacturer-independent, retrofittable assistance system that fuses 2D and 3D image data to de-

tect people and collision risks in the critical zone of manually operated forklifts [17]. A comparison of the reliability of results obtained by deep learning approaches with those obtained by traditional image processing methods demonstrated the former to be significantly more reliable. The operator is alerted to people in the surrounding area via tactile feedback, thus avoiding sensory overload caused by additional audiovisual information channels. A notable disadvantage of this approach is the inability to ascertain the precise location of the detected individual.

Kim et al. published a multispectral benchmark dataset consisting of RGB and thermal images for automated forklifts and introduced SSD-2.5D, a detector that identifies people and simultaneously estimates their distance from the vehicle to enable collision avoidance independent of weather and lighting conditions [18]. Furthermore, Sun et al. developed an enhanced SSD architecture specifically for forklift warning systems, designed for high detection speed and accuracy under challenging conditions such as darkness or rain [19].

Sang-ngenchai et al. adopted a complementary strategy by employing YOLOv8-based pose estimation, which focuses not on environmental factors but rather on the behaviors exhibited by vehicle operators themselves. This approach enables the early detection of potentially hazardous driving behaviors [20]. Kao et al. demonstrated reliable person and object detection at distances of up to 20 m by fusing camera and 1D LiDAR data. However, it should be noted that this method was originally developed for road vehicles rather than industrial trucks [21].

Overmeyer et al. developed an innovative driver assistance system using augmented reality [22]. Strategically positioned cameras visually eliminate physical obstructions, such as the lift mast or loads, thereby creating an unobstructed field of vision. The operator is thus enabled to see through these components. However, a notable disadvantage of this approach is the high latency of the overlaid images.

All of the research projects previously mentioned, including the commercial systems discussed in Section 2.1, with the exception of the UWB technology, have a commonality. The sensor technology is vehicle-mounted and, thus, dependent on a direct line of sight from the industrial truck. The elevated, forward-looking surveillance perspective proposed in this paper represents an investigative novelty.

3 SYSTEM ARCHITECTURE

A drone-based OAS for monitoring the environment such as the one presented in this paper must meet a wide range of requirements, some of which are outlined in the following section. Subsequently, the concept for the OAS is presented.

3.1 SYSTEM REQUIREMENTS

To reduce critical driving situations and accidents, this research project investigates predictive environmental monitoring for manually operated industrial trucks, specifically forklifts. The focus is particularly on the early identification of hazards in areas with limited visibility, such as intersections. In addition to reducing the number of accidents, such a system is intended to relieve the operator of some of the monitoring tasks, although it will not assume any safety-critical functions. Responsibility for operating the vehicle and preventing accidents remains with the forklift operator. Such a system faces a wide range of requirements that influence its functions and technical design. Some of these requirements were developed in collaboration with, or verified by, an industry-affiliated project advisory committee. Others stem from previous projects at the Institute of Transport and Automation Technology and Institut für Integrierte Produktion Hannover gGmbH aimed at expanding the field of view of industrial truck operators [23].

The system must be able to detect vulnerable objects with varying priorities. Unprotected pedestrians are subject to the highest risk of injury and therefore have the highest priority in detection. Slightly lower priority is assigned to mobile and protected individuals, such as other manually operated industrial trucks, passenger cars, hand pallet trucks, or bicycles. Stationary obstacles, such as pallets on the travel path, are assigned the lowest priority due to their lower risk of injury and lack of dynamic movement. The OAS must not distract the forklift operator from their work tasks during everyday use. In the event of an imminent collision, the OAS should issue a warning in addition to its normal functionality. Data protection in accordance with the General Data Protection Regulation (GDPR) must be ensured. No personal data may be processed, and no conclusions may be drawn about individual persons or their work practices. In research operations, consent is therefore obtained from participating test subjects. In addition, data processing for purposes such as AI training is performed locally. Publication of the recorded data requires prior anonymization of the data.

The sensor technology required for the system must be capable of enabling comprehensive environmental monitoring and therefore must have a wide field of view (FOV). Furthermore, the sensor technology must be capable of precisely localizing detected hazards in 3D (with a positional deviation of less than 25 cm) and classifying them. In order to achieve the objective of researching a self-sufficient, platform-independent system, a mobile sensor platform is required, for which a battery life of approximately 30 min is targeted. The system should be retrofittable and not exceed a total cost of €3,000 to remain feasible for SMEs. The OAS must be capable of real-time operation. Since such a mobile, radio-linked system is technically complex and presents many sources of latency, increased latency can lead

to safety-critical situations that undermine confidence in the system and thus reduce its acceptance in real-world deployment. In the absence of specific standards, an end-to-end latency of 200 ms, based on the guidelines for camera-based display systems in passenger cars, is targeted as an approximate threshold not to be exceeded [24].

The system is also subject to safety requirements. In the event of a system failure, such as a loss of connection between the sensor platform and the forklift, the forklift operator must be notified immediately. Furthermore, in this case, depending on the selected carrier platform, the system must remain in a position that poses no danger to people, e.g., by returning to the forklift or performing an emergency landing. The mobile sensor platform must be equipped with suitable sensors and algorithms to autonomously prevent collisions with itself, its surroundings, or people, even without a connection to the forklift. As the system initially operates exclusively indoors, it falls outside the scope of EU airspace drone regulations. Occupational safety requirements for operation near personnel apply instead.

In addition, clearly identifiable, level lanes for the forklift are assumed, which are delineated by markings on the sides. There should be no intervention in the forklift's control system or electronics to perform, for example, braking maneuvers or trajectory planning.

3.2 CONCEPT

Based on the requirements for the technical system and the functionality of the OAS, a concept is developed. This is shown in Figure 1.

A drone is used as the carrier platform for the sensor technology used for environmental monitoring. Compared to a ground-based platform, this reduces the risk of collision with other traffic. In addition, the elevated sensor position offers advantages for extensive environmental sensing, e.g., across obstacles. Challenges in this regard are seen in battery life, increased noise emissions, and the risk of injury caused by, e.g., crashes. To minimize the risk of injury, a lightweight protective cage is to be mounted around the drone, designed to have as little impact as possible on weight and, consequently, flight time. Currently, the sensor package is mounted stationary on a tripod to emulate the drone's hovering. A fisheye camera with a 190° FOV is used to extract semantic information for segmentation and classification. To more precisely localize objects and to potentially improve the 6D pose estimation of the forklift, a 96-line LiDAR sensor with a hemispherical FOV is additionally installed. The system is intended to rely on image data alone in order to keep the system's cost and the drone's weight as low as possible. In addition to improving battery life and thereby reducing noise emissions, this also lowers the risk of injury to bystanders. The LiDAR sensor thus serves as an optional sensor modality in case image

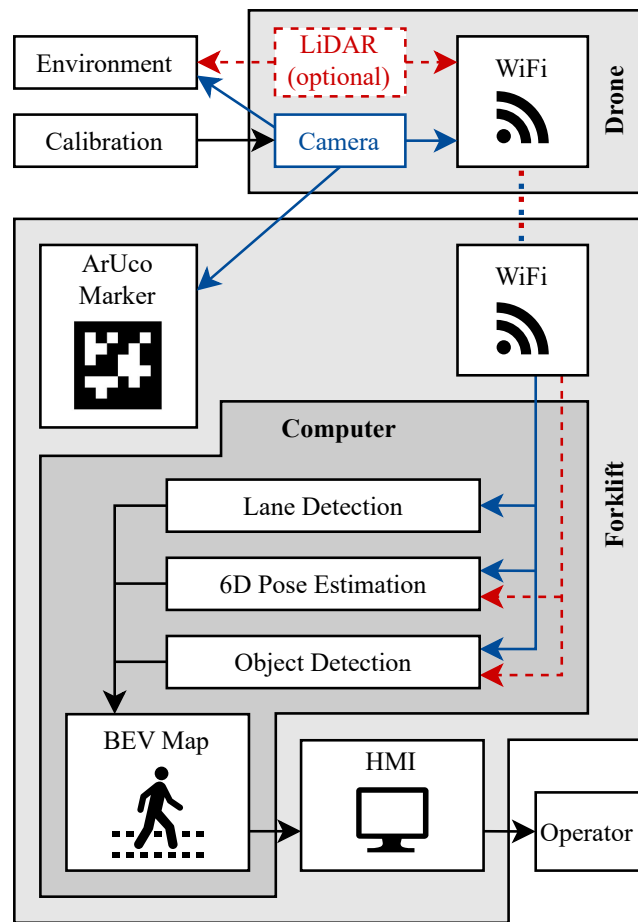


Figure 1: Concept for the system architecture of a drone based OAS for environment perception

data alone prove insufficient for the technical implementation. A Raspberry Pi 5 serves as an interface to transmit the recorded data to the forklift via a WiFi interface. Transmission is performed using User Datagram Protocol packets, since these are provided by the LiDAR sensor by default and represent the technical standard for latency-critical applications. As a further central component, a computing platform with AI-accelerator capabilities, based on a CUDA-enabled Graphics Processing Unit, is set up on the forklift. As a framework the Robot Operating System 2 (ROS2) is used for handling sensor data streams and the processing of them through individual nodes. Several nodes running in parallel process the various data. One program component detects the lanes in the environment based on the camera images. A further program component is intended to enable precise 6D pose estimation using the image data. This is required for the transformation rule from the drone to the forklift, in order to visualize detected hazards to the forklift operator with positional accuracy. By means of an ArUco marker mounted on the forklift's roof, the drone is able to local-

ize it. The transformation rule between the ArUco marker and the forklift coordinate system is known, which allows the transformation rule between the drone and the forklift to be determined. A further subprogram performs object detection and classification, likewise based on camera images, according to the priorities defined in Section 3.1. For pose estimation and object detection, the LiDAR point cloud can be incorporated to increase precision in cases of insufficient accuracy. The data from the three subprograms are required to compute a visualization of the hazards in the environment and to display this to the forklift operator via a human-machine interface (HMI). This representation is provided in the form of a computed schematic map using generic pictograms, shown from a bird's-eye view (BEV).

4 IMPLEMENTATION AND INITIAL RESULTS

The presented system architecture is built as an experimental setup. This setup is presented in the following, followed by first results from the data processing.

4.1 EXPERIMENTAL SETUP

For data acquisition and the execution of initial practical tests, an experimental setup is built based on the presented concept. Figure 2a) shows the experimental setup schematically.

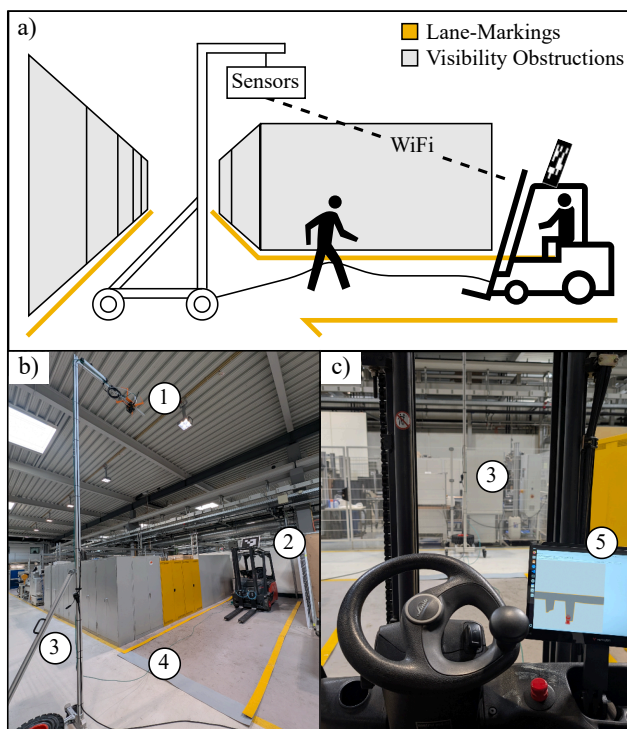


Figure 2: Experimental setup - a) schematic, b) exterior view, c) view from the cab

Here, the sensor package (cf. Fig. 2b - 1) is initially not mounted on a drone but rather stationary on a tripod (cf. Fig. 2b,c - 3). This can be moved to represent dynamic situations. The recorded data can be transmitted to the forklift (cf. Fig. 2b - 2) via two paths. On the one hand, wirelessly via WiFi to test the transmission under real-world conditions, or, on the other hand, via a wired Ethernet connection to rule out network issues during the development of the other subsystems. Additional, flexibly placeable floor markings allow various lane topologies to be replicated (cf. Fig. 2b - 4). Cabinets, shelving units, and movable walls serve as visibility obstructions, which can be adapted to the replicated lane topologies. As the HMI, a 14-inch touch display is mounted in the forklift's cab (cf. Fig. 2c - 5). For the 6D pose estimation of the forklift, a 60 x 60 cm ArUco marker is mounted on the roof of the forklift.

4.2 EXAMINATION OF THE WIRELESS CONNECTION

As part of the experimental setup, the wireless connection is examined in more detail. This represents a critical component to the real-time capability of the system and the usable amount of data. The experimental setup employs WiFi modules (2.4 GHz) with a theoretical maximum possible transmission rate of 400 Mbit/s, establishing a dedicated point-to-point connection between the sensor package and the forklift. To identify the bitrate usable in practice, measurements are carried out with different target bitrates in various positions and movement conditions. Figure 3 shows the results for eight different target bitrates (5, 10, 20, 30, 40, 50, 75, and 100 Mbit/s) under the configurations "static close = sensor package hovering 3 m above the target PC", "static far = sensor package additionally positioned 5 m in front of the target PC" and "dynamic far = sensor package additionally moving laterally and rotationally relative to static far".

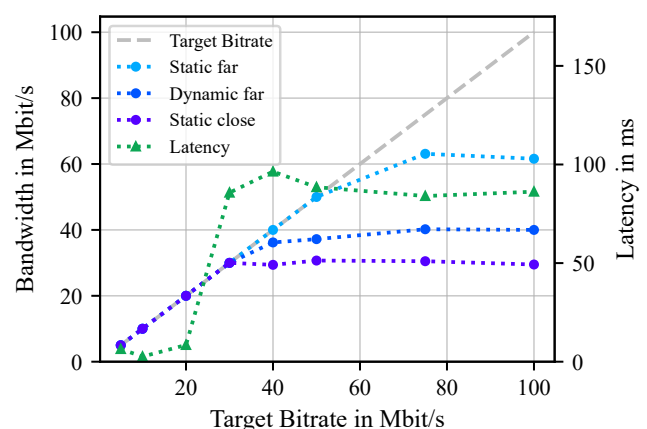


Figure 3: Measurements of network quality between the drone and the forklift

Through these experiments, the throughput limit is identified based on the breakdown of the transmission rate under

“static close.” Here, a stable transmission rate of more than 29.4 Mbit/s can not be achieved. At the same time, upon reaching this limit, the transmission latency increases from an average of approximately 10 ms to an average of 90 ms, which alone would already account for 45% of the targeted 200 ms end-to-end latency budget. Possible causes limiting throughput, to be investigated in future measurements, include suboptimal WiFi antenna alignment, high utilization of the 2.4 GHz frequency bands in the test environment, as well as shielding and reflections caused by metallic objects. Since the LiDAR sensor used requires a bitrate of approximately 22.5 Mbit/s, only approximately 6.9 Mbit/s remains available for the transmission of image data. This underscores the effort to realize environmental monitoring using image data alone. Nevertheless, to increase network quality, the use of higher WiFi standards is being pursued.

4.3 LANE DETECTION

A central component of environmental sensing is the detection of areas drivable by the forklift. Reliable detection of these areas is required not only for the orientation of the forklift operator on the computed schematic BEV map, but is also intended to serve as the basis for navigating the autonomously flying-ahead drone. The drone is intended to fly ahead along the center of lanes at a defined distance and to hold position in detected intersection areas. The representation of lane centerlines and intersection areas is intended to be realized through a graph to facilitate navigation. Once the forklift reaches an intersection and the turning direction becomes apparent, the drone advances to the next graph node. If the forklift deviates from the lane, the drone returns along the lane centerlines and follows it directly above. The computation of a representative graph is not yet implemented at this stage and will be investigated in future work.

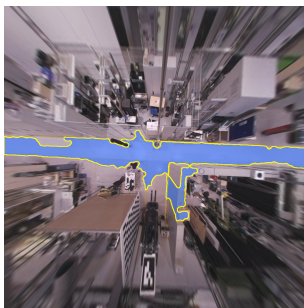


Figure 4: Example of faulty lane detection with classical image processing

For the schematic representation in the BEV map, the entire drivable area must be segmented. Initial attempts using classical image processing proved insufficiently reliable. In particular, fluctuating lighting conditions, partially occluded lanes, contamination, and other color-similar objects led to errors in lane detection as can be seen in Figure 4. A machine-learning-based approach is therefore chosen for lane segmentation. First the image is undistorted

using the known camera calibration matrix and distortion parameters into a rectangular, straight-line image with a FOV of 150°. A YOLO11n-seg network segments the

lanes on this image. For this purpose, a small initial dataset of 129 images with various replicated lane topologies is recorded at the experimental setup to assess feasibility, split into 80% training (103 images) and 20% validation data (29 images). The validation and training data originate from the same test environment. The results of the network trained in this way appear promising and allow for more robust segmentation than with classical image processing. A mean Average Precision (mAP)@50 value of 0.986 and an mAP@50-95 value of 0.885 are achieved on the validation dataset (higher values indicate better performance, mAP@50-95 applies stricter overlap criteria than mAP@50). To rule out overfitting of the network, a larger dataset with higher variance in environment and lane topology must be recorded in the future.

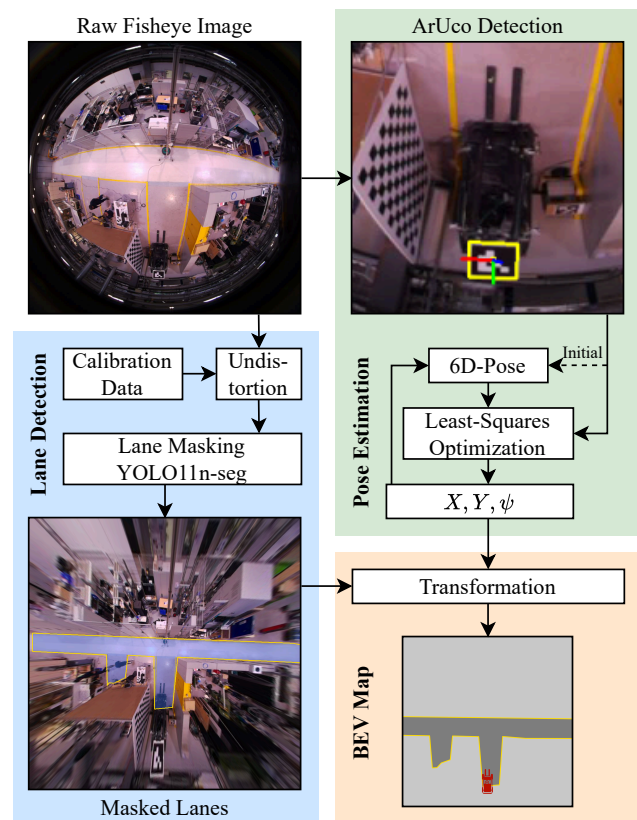


Figure 5: Software architecture for BEV map generation based on ArUco pose estimation and AI-based lane detection

Figure 5 shows the architecture of the various software components for generating a positionally accurate overview map. The segmented lanes must be transformed into the BEV map so that they are displayed to the forklift operator in the appropriate position. A symbol on the overview map indicates the position of the forklift on the lanes. The forward direction of travel corresponds to the vertical axis of the map. For the correct rotation of the segmented lanes,

particularly during cornering or angled travel, the yaw angle ψ of the vehicle is additionally required.

For this reason, a precise estimation of the position (X, Y) and ψ of the forklift relative to the drone is required. Therefore, an ArUco marker is mounted on the roof of the forklift and is detected in the camera image. First, a coarse 6D pose is determined from the undistorted marker corner points using a Perspective-n-Point (PnP) algorithm, then refined through nonlinear least-squares optimization. For a candidate value, the 3D coordinates of the marker corners are computed based on the known mounting geometry and projected directly into the distorted image using the calibrated fisheye camera model, as well as the camera pose, which will in future be provided dynamically by the drone telemetry. The deviation between these predicted and the actually detected image coordinates of the marker corners is iteratively minimized until the pose converges. For subsequent images, the refined pose of the previous image serves as the initial value, meaning that the initial PnP step is generally only required once. Since the entire processing chain operates continuously on the distorted fisheye image, pose estimation remains robust even at wide viewing angles near the edge of the image. In comparison, under prior undistortion, the marker is no longer fully visible even at shorter distances.

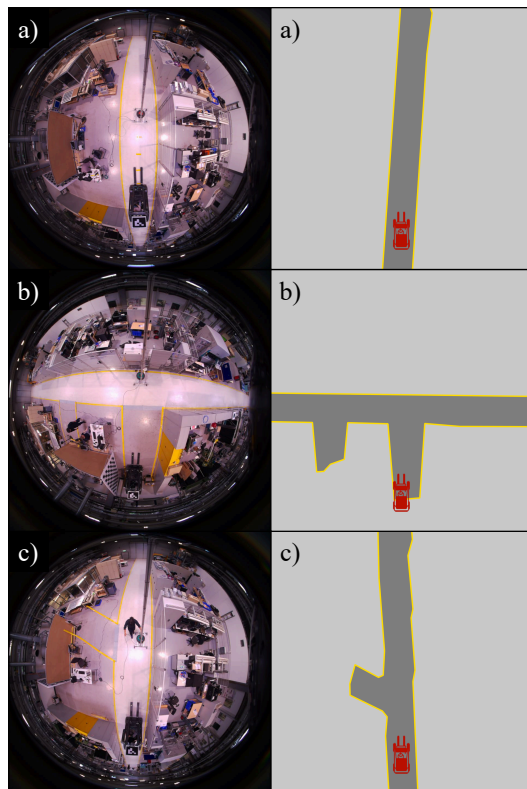


Figure 6: Computed BEV maps for various input images at different rotations and topologies

Based on the estimated position of the forklift (X, Y, ψ) and the lanes segmented in the camera image, these can be transformed into the forklift's coordinate system and displayed with correct rotation. A systematic evaluation against manually measured reference positions across 31 test configurations yields a maximum positional deviation of 5.0 cm and a maximum rotational deviation for ψ of 4.3° , well within the 25 cm accuracy requirement defined in Section 3.1. Figure 6 visualizes various input images and their computed, transformed environment BEV maps for the case recorded statically using the tripod. As can be seen in all examples, the lane area is segmented and displayed in the correct position relative to the forklift. However, a slight rotation of the lane relative to the forklift can be observed in Figures 6a and c. A possible reason for this is the transformation between the ArUco marker and forklift coordinate systems, which is currently only measured manually. In examples 6b and c, it can furthermore be seen that the edges of the lanes are not always straight and/or parallel. Through larger datasets, more extensive training, and modifications to the segmentation network, as well as post-processing of the detected polygons, an improvement in segmentation quality is intended to be achieved in the future.

4.4 PERSON DETECTION

The initial step in displaying persons on a map for the operator of the industrial truck is to identify them. The initial approach utilizes existing datasets for person recognition. Since the camera uses a fisheye lens and its mounting at an approximate height of 4 m, the feasibility of employing existing datasets is under evaluation. The fisheye lens, with a FOV of 190° , provides excellent coverage of the area being observed. However, due to the characteristics of the camera and its position, subjects are occasionally distorted and oriented differently within the image. For instance, they may appear in an inverted orientation, rotated by 90° , or only perceptible from a superior perspective.

In this instance, YOLOv8l from Ultralytics is employed as the person detection model, and its efficacy is evaluated using recorded data. A visual analysis of the dataset reveals the presence of three individuals throughout the entire dataset. The dataset comprises a total of 1,017 frames, of which every third frame is manually annotated, resulting in a total of 339 images being evaluated. These images are used to evaluate the model's detection quality. The metrics evaluated include the number of detected people per image and the true positive (TP), false positive (FP), and false negative (FN) rates. To detect people who are rotated, each frame is evaluated after being rotated by 90° , 180° , and 270° . The confidence threshold is set to 0.25, and the intersection-over-union threshold for evaluation is set to 0.5.

Table 1 presents a comprehensive summary of the aggregated results concerning person detection. Among the total of 888 detections, 845 are classified as TP, while 43 are de-

terminated to be FP. Additionally, 172 of the 1,017 expected detections (339 frames $\times$ 3 ground-truth persons per frame) are missed, resulting in an FN rate of 172. This leads to a precision of 95.2%, accompanied by a recall of 83.1% and an F1 score of 88.7%. The lower recall, as compared to the precision, indicates that undetected individuals occur at a significantly higher frequency than false detections.

Table 1: Key metrics from the analysis of person detection across all analyzed frames

Metric	Value
Analyzed frames	339
Ground truth (people per frame)	3
Total detections	888
True positives (TP)	845
False positives (FP)	43
False negatives (FN)	172
Precision	95.2%
Recall	83.1%
F1-Score	88.7%
Complete detection frames (FN = 0)	204/339 (60.2%)
Error-free frames (FN = 0, FP = 0)	188/339 (55.5%)

Figure 7 provides a general overview of the results at the frame level. Out of the 339 frames, all three individuals are correctly detected in 188 frames (55.5%) at the manually annotated position without any false positives or false negatives. This is exemplified in Figure 8a. In an additional 16 frames (cumulative: 204, 60.2%), all subjects are detected, including false positives. In a total of 107 frames (31.6%), one person is not detected. In 19 frames, two people are not detected (5.6%). In nine frames, none of the three people are detected (2.7%). Figures 8a through 8c offer a more detailed analysis of specific frames.

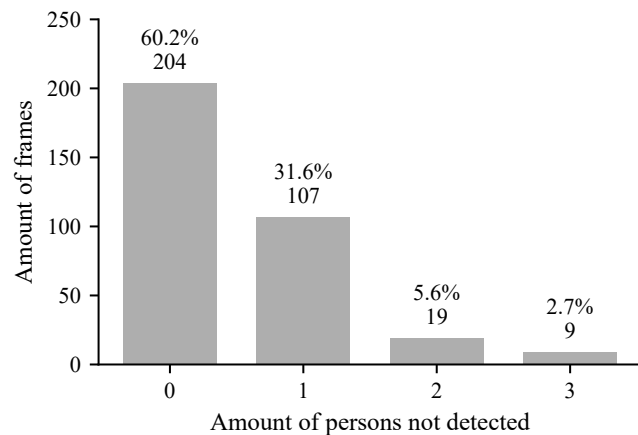


Figure 7: Distribution of frames by number of unidentified individuals per frame

The results demonstrate the general viability of employing the YOLOv8l model for person detection. Detection performance depends on the separation of persons in an image and their distinctiveness from the background. This is particularly evident in Figure 8a, where all the persons stand out clearly from the floor of the environment and are separately visible. Recognition errors occur particularly when people are in a confined space, overlap, and do not stand out clearly from the background. This is demonstrated in Figures 8b and c. A comparison of the two images reveals that the subjects have become more distant from one another, thereby enhancing their recognition.

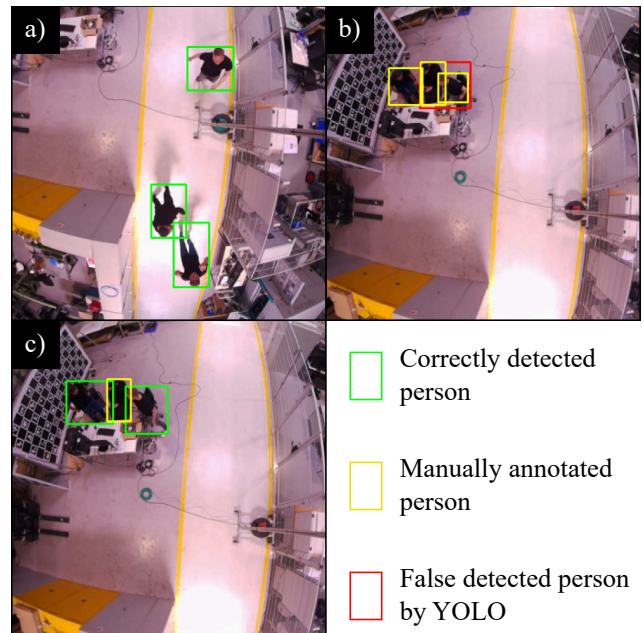


Figure 8: a) Correct frame showing all detected persons, b) No persons are correctly detected, c) Two of the three persons are correctly recognized

5 CONCLUSION AND OUTLOOK

This work addresses the identified gap by employing a novel drone-based OAS for monitoring the surroundings of industrial trucks. An experimental setup is developed in which the sensor technology, currently mounted on a movable tripod to emulate a drone, transmits data to the forklift both wirelessly and via a wired connection. An ArUco marker mounted on the forklift's roof enables pose estimation. Initially implemented algorithms perform AI-based image processing for lane detection and person detection. For both trained YOLO networks, suitability for the system is demonstrated using small, self-recorded and annotated datasets. In addition, an ArUco marker-based pose estimation is implemented to enable the correctly oriented display of a schematic map from a BEV.

Initial measurements of network latency yield a value of 10 ms when remaining below the maximum available bandwidth, which represents a favorable result with respect to the target of an overall latency below 200 ms. The available bandwidth of 29.4 Mbit/s would be largely utilized by LiDAR data alone, at approximately 22.5 Mbit/s, which underscores the goal of relying solely on image data. The AI-based lane detection, which utilizes a YOLO11n-seg network, achieves an mAP@50 value of 0.986 while showing instabilities with the lane edges. For the detection of hazard sources, only the highest priority class (unprotected persons) has been investigated so far. The YOLOv8l network used for this purpose yields a precision of 95.2% and recall of 83.1% while people close together or hardly distinguishable from the surroundings pose a challenge. The pose estimation enables a correctly oriented transformation of the computed BEV map, achieving a maximum positional deviation of 5.0 cm and a maximum rotational deviation of 4.3° across 31 test configurations, well within the 25 cm accuracy requirement.

As the next step, a drone will be procured to validate the developments under real dynamic, wireless conditions. This includes a detailed analysis of drone telemetry and subsystem latencies to assess the system's real-time capability. The acquisition of larger, more diverse datasets is planned to improve detection accuracy and generalize beyond the current, relatively homogeneous test environment. Detection of the two lower priority classes (protected persons and static obstacles) will also be addressed. The two YOLO networks will be merged into one, complemented by post-processing for interpolation and prediction during detection failures.

For drone navigation, an algorithm will be developed that generates a simplified graph from the detected lanes. Furthermore, the control of the drone via ROS2, as well as hazard mitigation for bystanders during flight, will be investigated. A protective cage for the drone will also be implemented. The project continues to pursue the use of image data alone. If this proves insufficient in the future, sensor fusion with LiDAR data remains possible. The anonymization of recorded data prior to publication, as well as local data processing, must be further examined and ensured in the future to guarantee compliance with the GDPR.

6 FUNDING

This article is published as part of the research project "Mobile, environment-monitoring sensor system for industrial trucks to increase safety and work efficiency in intralogistics (MouSe)" with the funding code 01IF23654N. The project is funded by the Federal Ministry for Economy and Energy pursuant to a resolution of the German Bundestag using Industrial Collective Research (IGF) funds.

REFERENCES

- [1] M. Braun, D. Marrenbach, and O. Scholtz, "Präventive Arbeitsgestaltung in der Intralogistik - Menschliche Arbeit als Grundlage von Flexibilität und Adaptivität intralogistischer Systeme," *wt Werkstattstechnik*, vol. 110, no. 6, pp. 456 – 460, Jun. 2020. [Online]. Available: <https://doi.org/10.37544/1436-4980-2020-06>
- [2] D. Spath, R. Ehmann, M. Braun, B. Kille, D. Marrenbach, and O. Scholtz, *Der Mensch in der Intralogistik: Unternehmensstudie von IWL AG und Fraunhofer IAO zur präventiven Arbeitsgestaltung in der Intralogistik im Rahmen des BMBF-Projekts PREVILOG*. Stuttgart, Germany: Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, 2019. [Online]. Available: <https://doi.org/10.24406/publica-fhg-299732>
- [3] Deutsche Gesetzliche Unfallversicherung (DGUV), "Arbeitsunfallgeschehen 2024," Deutsche Gesetzliche Unfallversicherung e.V. (DGUV), Referat Statistik und Datenanalyse, Berlin, Germany, Tech. Rep., Oct. 2025. [Online]. Available: <https://publikationen.dguv.de/widgets/pdf/download/article/5157>
- [4] F. Zelt, S. Butz, K. Schwarzmann, and K. Stöckl, *Handbuch für Gabelstaplerfahrer*. Bonn, Germany: Berufsgenossenschaft Handel und Warenverteilung (BGHW), 2012. [Online]. Available: https://www.komnet.nrw.de/KomNet/files/30978/requ/Handbuch_f%C3%BCr_Gabelstaplerfahrer.pdf
- [5] D. Diaz Torres, "Entwicklung eines Vorgehensmodells zur Flexibilisierung und Verbesserung der Transporte auf dem Werksgelände durch den Einsatz autonomer mobiler Roboter," Ph.D. dissertation, Technische Universität München, 2023. [Online]. Available: <https://mediatum.ub.tum.de/1662575>
- [6] VdSM – Verband der Siegerländer Metallindustriellen, "Die intelligente Gabelzinke: VETTER Industrie GmbH," *Wirtschaft regional – Siegerland-Olpe-Wittgenstein*, no. 3, pp. 4–5, 2019. [Online]. Available: <https://www.vdsm.net/wronline/wp-content/uploads/2019/02/wr0319.pdf>
- [7] Hyster Europe, "Pedestrian awareness camera: Solution brochure," Product brochure, Hyster-Yale UK Limited, 2025, accessed: 2026-08-05. [Online]. Available: https://www.hyster.com/globalassets/coms/hyster/emea/documents/industry-solutions/hyster-reaction/0071het4bc001_e-engb.v1r0_042025_pedestrian-awareness-camera-brochure.pdf
- [8] E. Sun and R. Ma, "The UWB based forklift trucks indoor positioning and safety management system," in *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2017, pp. 86–90. [Online]. Available: <https://doi.org/10.1109/IAEAC.2017.8053982>

- [9] Linde Material Handling GmbH, “Linde Safety Guard: Das Assistenzsystem, das Gefahren erkennt, bevor sie entstehen,” Product brochure, Linde Material Handling GmbH (KION Group), Aschaffenburg, Germany, 2023, accessed: 2026-07-29. [Online]. Available: https://www.linde-mh.de/media/Global-Content/08_Downloads/Safety/Linde_Safety_Guard_Broschuere_DE.pdf
- [10] Toyota Material Handling, “SEnS+ smart environment sensor plus,” Product webpage, Toyota Material Handling, Columbus, IN, USA, accessed: 2026-07-29. [Online]. Available: <https://www.toyotaforklift.com/sens-smart-environment-sensor-plus>
- [11] Jungheinrich AG, “Das passende Gabelstapler-Assistenzsystem für Ihr Lager,” Product brochure, Jungheinrich AG, accessed: 2026-07-29. [Online]. Available: <https://media-live2.prod.scw.jungheinrichcloud.com/resource/blob/2075862/cb4cf61852dc2ce1b8e8a916372ffb1e/checkliste-assistenzsysteme-data.pdf>
- [12] STILL GmbH, “STILL Smart Assist: Intelligente Sicherheit für Ihr Lager,” Product brochure, STILL GmbH, Hamburg, Germany, accessed: 2026-07-29. [Online]. Available: https://data.still.de/assets/products/Vehicles/Safety_new/pdf/Smart-Assist_DE_Brochure.pdf?mod=1772539953&download=1&s=e9ab0c26f0476fb04e4a1e33f371aaaa
- [13] R. LeBlanc, “Crown equipment launches ProximityAssist to help prevent forklift collisions and reinforce training,” Forklift Revolution, May 2025, accessed: 2026-07-29. [Online]. Available: <https://www.forkliftrevolution.net/crown-equipment-launch-proximityassist>
- [14] ELOKON GmbH, “ELOprotect ACS – Kollisionsschutz für den Schmalgang,” Product webpage, ELOKON GmbH, Tornesch, Germany, accessed: 2026-07-29. [Online]. Available: https://www.elokon.com/de/intralogistik-sicherheitssysteme/eloprotect-acskollisionsschutz?utm_source=direct&utm_medium=direct&utm_campaign=direct-website-traffic
- [15] Jungheinrich AG, “Assistenzsystem zoneCONTROL: Die modulare Lösung für höchste Sicherheit im Lager,” Product brochure, Jungheinrich AG, accessed: 2026-07-29. [Online]. Available: <https://media-live2.prod.scw.jungheinrichcloud.com/resource/blob/826740/9be6052ddf23974b5d0ca05fd703abc8/produktvorteile-zonecontrol-data.pdf>
- [16] Linde Material Handling GmbH, “Linde BlueSpot: Blaues Licht für mehr Sicherheit,” Product webpage, Linde Material Handling GmbH (KION Group), Aschaffenburg, Germany, 5 2026, accessed: 2026-07-29. [Online]. Available: <https://www.linde-mh.de/de/Ueberuns/Innovationen-von-Linde/Linde-BlueSpot.html>
- [17] A. Börold, M. Freitag, A. Lang, W. A. Günthner, A. Sherman, J. Tödter, and H.-P. Kany, “Abschlussbericht zum Vorhaben: PräVISION – Methodenentwicklung zur präventiven Steigerung der Arbeitssicherheit an Flurförderzeugen mit Umsetzung eines Assistenzsystems durch Fusion und Analyse von 2D- und 3D-Bilddaten,” Deutsche Gesetzliche Unfallversicherung (DGUV), Tech. Rep. FP-0379, Nov. 2018. [Online]. Available: https://www.dguv.de/projektdatenbank/0379/neu_2018_11_28_praevision_abschlussbericht.pdf
- [18] H. Kim, T. Kim, W. Jo, J. Kim, J. Shin, D. Han, Y. Hwang, and Y. Choi, “Multispectral benchmark dataset and baseline for forklift collision avoidance,” *Sensors*, vol. 22, no. 20, 2022. [Online]. Available: <https://doi.org/10.3390/s22207953>
- [19] E. Sun, X. Ma, and M. Li, “Improved SSD based pedestrian detection algorithm for forklift active warning system,” in *2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. Chongqing, China: IEEE, Dec. 2022, pp. 1523–1528. [Online]. Available: <https://doi.org/10.1109/IMCEC55388.2022.10019947>
- [20] A. Sang-ngenchai and M. Nakazawa, “Smart warehouse safety: Computer vision for forklift driver monitoring in warehouse setting,” in *2025 IEEE International Conference on Consumer Electronics (ICCE)*. Las Vegas, NV, USA: IEEE, Jan. 2025. [Online]. Available: <https://doi.org/10.1109/ICCE63647.2025.10930064>
- [21] I. Kao, Y. Yian, J. Su, Y. Lai, J. Perng, T. Hsieh, Y. Tsai, and M. Hsieh, “Design of sensor fusion driver assistance system for active pedestrian safety,” *CoRR*, vol. abs/2201.09208, 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2201.09208>
- [22] L. Overmeyer, L. Jütte, and A. Poschke, “A real-time augmented reality system to see through forklift components,” *CIRP Annals*, vol. 72, no. 1, pp. 409–412, 2023. [Online]. Available: <https://doi.org/10.1016/j.cirp.2023.03.010>
- [23] A. Poschke, L. Jütte, B. Küster, and L. Overmeyer, “User acceptance of see-through augmented reality in intralogistics,” *Logistics Research*, vol. 18, no. 1-2, pp. 7–22, 2025. [Online]. Available: <https://doi.org/10.1108/LORE-01-2025-0003>
- [24] International Organization for Standardization (ISO), “ISO 16505:2019 – road vehicles – ergonomic and performance aspects of camera monitor systems – requirements and test procedures,” International standard, Geneva, Switzerland, Jul. 2019.

Tim Noah Laumann, M.Sc., Research Assistant at the Institute of Transport and Automation Technology (ITA) at the Leibniz University Hannover.

Tim Noah Laumann was born in 1999 in Hameln, Germany. He pursued a dual B. Eng. in mechatronics at CLAAS SE GmbH through Baden-Württemberg Cooperative State University in Stuttgart. He earned his M. Sc. in mechatronics and robotics at Leibniz University Hannover, specializing in systems engineering. Since completing his master's degree in 2025, he has been working in the Transportation Group at the ITA, where he specializes in sensor data processing for environmental perception and automation.

Address: Institute of Transport and Automation Technology, Leibniz University Hannover, An der Universität 2, 30823, Garbsen, Germany
Phone: +49 1523 762-0549, Fax: +49 511 762-4007
E-Mail: tim.laumann@ita.uni-hannover.de

Torben Mente, M.Sc., Project Engineer, IPH – Institut für Integrierte Produktion Hannover gGmbH

Torben Mente was born 1993 in Celle, Germany. Between 2011 and 2017 he studied industrial engineering at Leibniz University Hannover, specializing in production technology and production management. Since 2021, he has been working at IPH in the Production Automation department.

Address: IPH – Institut für Integrierte Produktion Hannover gGmbH, Hollerithallee 6, 30419, Hannover, Germany
Phone: +49 511 27976-236, Fax: +49 511 27976-888
E-Mail: mente@iph-hannover.de

Carsten Schmidt, M.Sc., Research Assistant at the Institute of Transport and Automation Technology (ITA) at the Leibniz University Hannover.

Carsten Schmidt was born in Willich, Germany, in 1993. In 2018 he passed the dual study program in cooperation with AUMUND Fördertechnik GmbH and Rhein-Waal University of applied sciences in Kleve, Germany. In 2021 he received his M. Sc. degree in mechanical engineering from Leibniz University Hannover. He currently is head of the Transportation Group at the ITA at Leibniz University Hannover. His research interests mainly include bulk material handling, belt conveyors and alternative drive concepts. Mr. Schmidt has been appointed to the Standards Committee "Elastomer-Technik (NET)" of the German Institute for Standardization (DIN) from 2022 to now.

Dr.-Ing. Benjamin Küster, Head of Production Automation Department, IPH – Institut für Integrierte Produktion Hannover gGmbH

Dr.-Ing. Benjamin Küster (*1988) studied industrial engineering at Leibniz University Hannover, specializing in pro-

duction engineering and business informatics. He has been working as a project engineer in the field of production automation since November 2014. He has been Head of Production Automation since September 2017.

Prof. Dr.-Ing. Ludger Overmeyer, Managing Director, IPH – Institut für Integrierte Produktion Hannover gGmbH, Head of Institute of Transport and Automation Technology (ITA), Leibniz University Hannover

Prof. Dr.-Ing. Ludger Overmeyer, born in 1964, studied electrical engineering at the University of Hanover and then worked as a research assistant at the Laser Zentrum Hannover e.V. (LZH). After obtaining his doctorate in engineering, he worked as a manager in the field of research and development in mechanical and plant engineering for the semiconductor industry. Since 2002, he has headed the ITA at Leibniz University Hannover. In 2007, he took on the role of managing partner of IPH.