Security-Oriented Artificial Intelligence-Embedded Image Processing for Lift Car Interiors

Tevfik Keleş[1], Mehmet Selçok[1], Eren Kalay[2], Adem Candaş[2], C. Erdem İmrak[2*]

[1] Emlak Konut Asansör Sistemleri San. ve Tic. A.Ş., Istanbul, Turkey
[2] Istanbul Technical University, Elevator Technologies Laboratory, Istanbul, Turkey

Keywords: artificial intelligence, image processing, passenger security, car capacity.

Abstract. In lift systems, in-car cameras have become widely used for both passenger counting purposes within the cabin and for security reasons. It is known that security systems in lifts generally remain limited to basic security measures. Current applications include emergency buttons, audible warnings, and intercom use; these solutions offer a reactive approach that allows users to intervene after an event causing danger in the lift has occurred. The developed web-based interface enables the monitoring and tracking of passengers in the car, small children, pets, belongings, and potentially hazardous situations. The integration of this system with the building security system will be studied later.

1 INTRODUCTION

Lifts are indispensable components of high-rise structures, facilitating efficient vertical mobility and substantially decreasing transit time between floors. They have become an essential part of daily life, offering great convenience while also presenting certain safety risks. Their widespread adoption has also given rise to new safety concerns. Detecting the presence of humans, animals, and objects within a lift car is crucial for both lift traffic analysis and the safety of passengers and building residents. While existing literature primarily focuses on estimating passenger counts within lift cars, there is limited research regarding the detection of children, unauthorised individuals, pets, or unidentified objects.

Estimating the number of passengers within a lift car is essential for lift traffic analysis and for optimising the efficiency of lift usage within a building. Accurately estimating lift occupancy allows for the analysis of passenger flow patterns across different periods, enabling a more precise determination of the actual load range. Furthermore, monitoring car capacity serves as a preventive measure against overloading.

Conventional lifts monitor the total load capacity rather than the distribution or specific mass of individual loads. Common methods for this task include the use of load cells mounted on suspension ropes or placed under the lift car. Furthermore, excessive occupancy may compromise passenger safety and create additional hazards.

Kumar et al. [1] developed a computer system for maintaining the load balance of a lift by counting the number of people inside the lift car before operating it, in which a Haar-like algorithm was implemented. Hashimoto et al. [2] demonstrated that infrared sensors can be utilised to detect both the direction and the number of passengers entering and exiting the lift. Ma et al. [3] proposed a passenger counting approach in the lift based on channel state information (CSI) measurements using commodity Wi-Fi. Vision-based methodologies [4–8] predominantly rely on image recognition techniques for crowd estimation. Fan et al. [9] proposed a person counting methodology based on morphological outline connection area analysis, which was subsequently validated using real-world lift surveillance footage.

While other methods focus on estimating the number of passengers inside the lift car, Shen and Chu [10] deployed GPU-based edge devices in a four-story building to count individuals entering specific areas. In contrast to other studies in the literature, Al-Sharif and Barney [11] developed a mathematical approach known as the "I-S-P method" (Inverse Stops-Passengers), which infers the number of passengers in a lift car based on the number of stops made in either the upward or downward direction. Unlike the I-S-P method, this study employs an image-based methodology to facilitate accurate passenger estimation.

2 PASSENGER COUNTING TECHNOLOGIES

Passenger counting technologies – whether applied to public transit (metros, trains, buses) or lift cabins – are primarily categorised into two distinct methodologies: vision-based (image-based) and non-vision-based (sensor-based) approaches.

In recent years, non-vision-based methodologies have gained significant traction, being broadly categorised into device-based and device-free approaches [12]. Non-vision-based methodologies incorporate a diverse array of technologies, such as physical sensors, wireless radio frequency (RF), network signals, and acoustic waves to facilitate accurate occupancy estimation within a given environment. Nevertheless, these methodologies often encounter spatial limitations and demonstrate a heightened susceptibility to inaccuracies when monitoring dense pedestrian flows.

Sensor-based methodologies utilise a diverse range of sensing technologies to perform people counting. Traditional sensor-driven crowd estimation techniques, such as infrared sensors, pressure sensors, and mobile-based systems, often mandate specialised infrastructure or require proactive user participation, complicating their implementation. These requirements often prove intrusive and impede the seamless integration of such systems across diverse environments.

In contrast to sensor-based methodologies, image-based approaches provide significant advantages for occupancy detection, as they are not subject to spatial constraints and can capture comprehensive environmental data within a single frame. Video-based surveillance and image processing techniques provide a practical and non-intrusive approach to occupancy detection, achieving high levels of accuracy at a relatively low cost [13]. While image-based methods provide broader spatial coverage compared to their non-image-based counterparts, they also tend to offer superior reliability in counting accuracy. This study concerns a system that integrates artificial intelligence-based image processing and object recognition technologies.

3 DETECTION IN THE ELEVATOR

The proposed methodology is most effective in scenarios where cameras are positioned to provide a plan view of the monitored area. The challenge of occupancy estimation within lift lobbies and lift cars represents a specific case of the broader problem of object counting in complex and dynamic environments. A prerequisite for counting is the initial detection and localisation of objects within the frame. The counting process begins with the identification of individual objects as they enter the monitored area. The counting process does not necessarily depend on object recognition or precise localisation. Instead, the system must simply detect the presence of entities that do not belong to the normal background environment. The system is designed to facilitate efficient route planning, allowing users to reach their destinations via the fastest possible paths while minimising lift wait times and lobby congestion.

3.1 Monitoring and tracking in the car

The developed web interface enables the monitoring and tracking of passengers in the car, small children, pets, belongings, and potentially hazardous situations.

By utilising the developed system, precise detection can be achieved for all occupants inside the lift car, encompassing potential fall detection, smoking inside the cabin, objects that pose a danger in enclosed environments (e.g., firearms, knives, and prohibited materials), and abandoned items (such as boxes, parcels, bags, and suitcases). Additionally, real-time alerts can be transmitted to report various negative occurrences or noteworthy anomalies.

The integration of real-time data collection and visualisation allows operators to maintain constant situational awareness of the lift environment, thereby accelerating the identification of critical incidents and streamlining emergency response procedures.

3.2 System design

The proposed smart lift system is designed primarily for high-rise building applications, including high-traffic commercial and residential complexes, as well as environments where occupant safety and operational comfort are paramount. With the aim of improving safety within lift systems, camera-based image processing and artificial intelligence techniques are utilised to identify passengers, objects, animals, and ongoing events inside the lift car.

Figure 1 depicts representative configurations of the lift car utilised in the system, demonstrating the empty state, the presence of passengers and animals, and the condition containing exclusively objects.

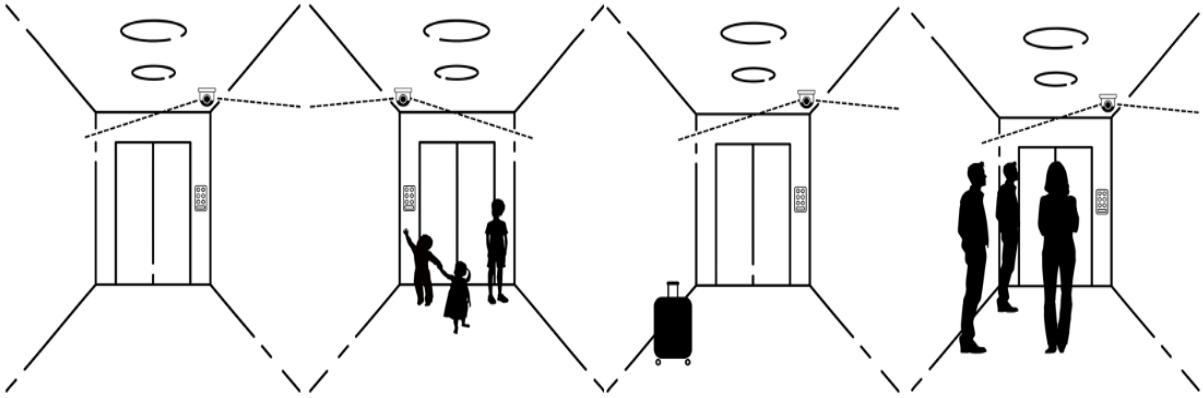


Figure 1 Lift car configuration and possible situations

The smart lift system, integrated with artificial intelligence-based image processing and object recognition technologies, comprises at least one wide-angle camera capable of operating in low-light conditions to capture real-time images within the cabin; at least one embedded computing board that processes the collected images, detects objects and individuals, analyses various scenarios, and executes at least one object recognition algorithm; at least one scalable database where all processed data is securely stored; a lift control panel to which the processed data is transmitted via a CANBUS or MODBUS communication line; and at least one web interface functioning as a management panel, enabling authorized users to monitor real-time events, review historical logs, and intervene in emergencies.

A representative schematic diagram illustrating data processing and the destination routing of emergency alerts within the lift cabin is presented in Figure 2.

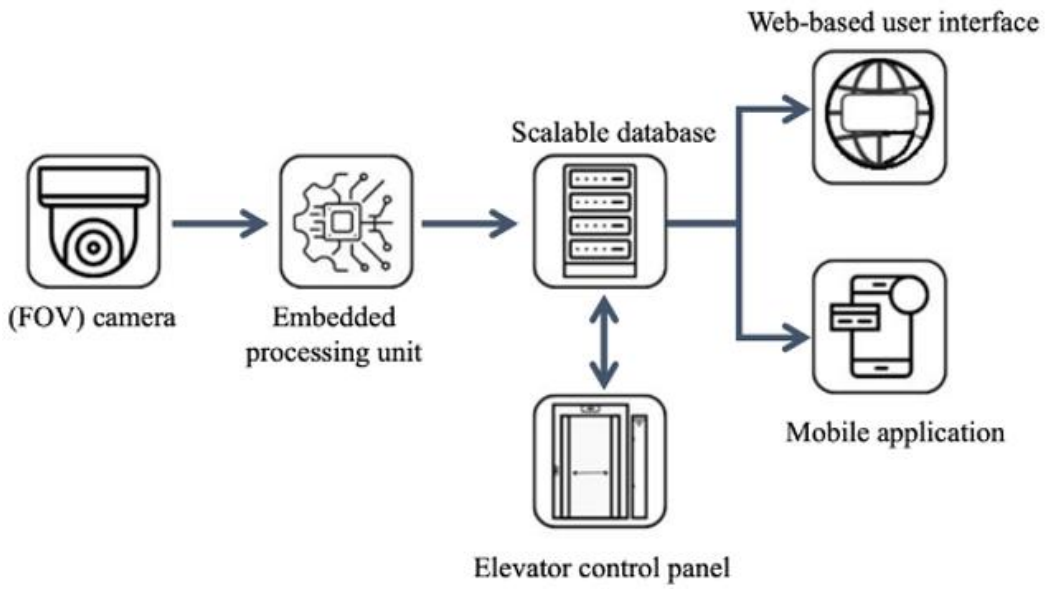


Figure 2 Representative schematic diagram of the system

A wide-field-of-view (FOV) camera with low-light capability is used for real-time image acquisition inside the lift car. The system necessitates an embedded computing board configured to handle image processing, perform human and object detection for scenario analysis, and run the designated object recognition models. A scalable database is integrated into the architecture to ensure the secure storage and efficient management of all processed logs. Within the system configuration, a web-based graphical user interface (GUI) is integrated to allow authorised personnel to track live system events, examine historical records, and execute emergency response measures (Figure 3).

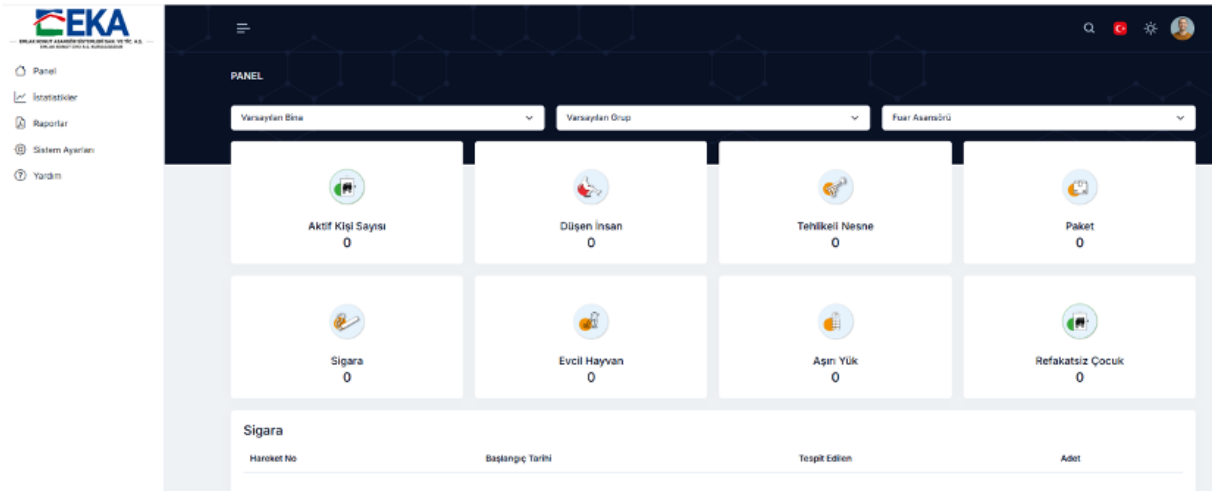


Figure 3 Web-based user interface

Additionally, a mobile application is included as a software component, allowing authorised personnel to monitor incidents and execute remote countermeasures if required. The images obtained from the wide-angle camera are processed using the YOLOv8 [14] object detection algorithm, together with the OpenCV [15] and NumPy libraries [16] for image processing. Face-blurring algorithms are applied for privacy protection, while psycopg2 [17] and PostgreSQL [18] are used for database interaction. InsightFace is employed for authentication and face recognition. The authentication/face recognition software identifies individuals and evaluates whether two separate faces belong to the same person. The integrated AI algorithms can be updated over time to improve

predictive accuracy in new environments. As a result, system robustness and precision are progressively enhanced over time. Figure 4 depicts the procedural flowchart governing the developed computer vision algorithm. The reliability of the proposed algorithm is evaluated through experimental trials.

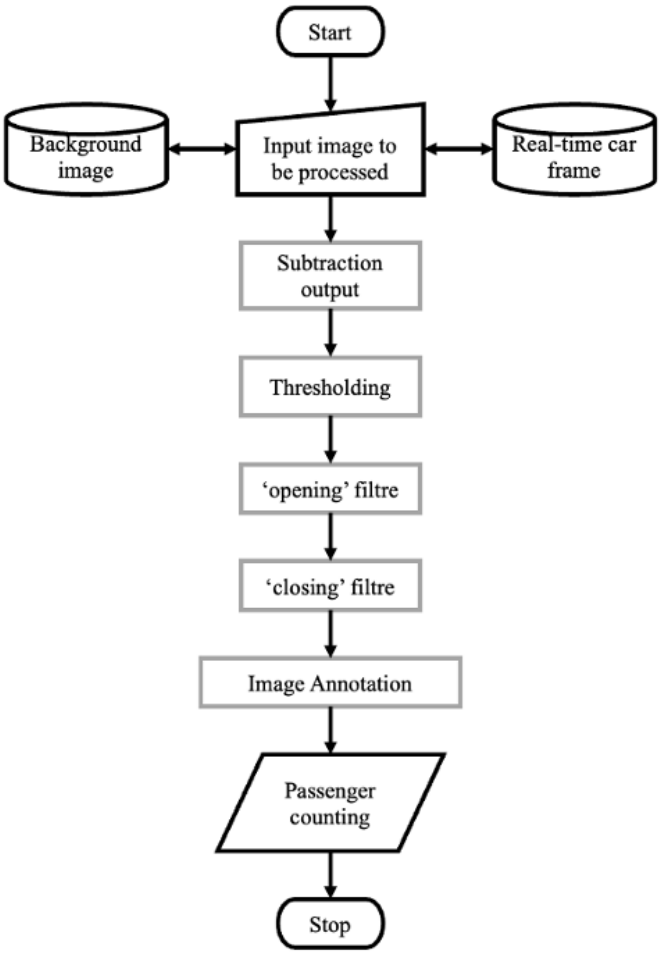


Figure 4 Flowchart of the proposed image processing algorithm

Using both camera-based monitoring and the load weighting system gives the best result because it provides the current number of passengers in the car. In the proposed framework, the visual data representing three occupants within the lift car are captured by the camera and processed in accordance with the steps outlined in the procedural workflow. Additionally, system validation is performed by cross-referencing the computer vision outputs against the weight signals extracted from the load cells under the car floor. Utilising 0.5-second sampling intervals enables the system to infer passenger boarding and alighting activities based on weight fluctuations between these steps. In practice, it is unlikely for multiple individuals to enter the cabin within a 0.5-second window in a single-door lift. The defined threshold values were compared with the difference between steps, and the number of passengers in the elevator car was accurately determined for each step. Figure 5 shows the data obtained from a test run simulating four passengers of different weights getting on and off the example system over a period of two minutes. It shows the weight-step graph processed through a moving average filter where each step represents a 0.5-second weighted average generated using the software developed for the 4-person car configuration selected for the test suite. This cross-validation strategy provides an additional reference for assessing the consistency of the image-based counting results. By comparing visual detections with load-cell variations, possible mismatches caused by occlusion, lighting conditions, or temporary detection errors can be identified.

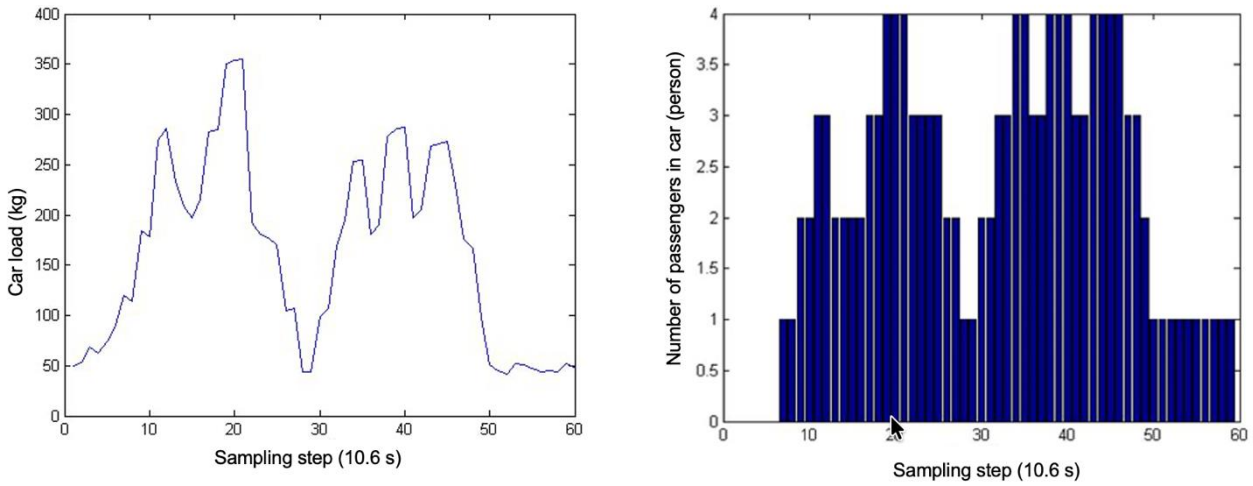


Figure 5 Plot of passenger weight and car capacity count over 0.5 s sampling step

The ESPROS/VOL infrared camera is configured for monochrome imaging to detect the occupancy rate by cross-referencing captured frames against the baseline reference (empty) image. Based on the pixel-intensity variation compared to the empty image, the embedded internal processor transmits the extracted data to the computer over an RS232 serial interface; this transmitted metric specifies the occupancy rate. Furthermore, the proprietary software module generates the occupancy rate as a percentage (%) and estimates the passenger count using this ratio. By means of its proprietary software, the system outputs the occupancy rate as a percentage, which is then utilised to approximate the passenger count. A test suite incorporating a four-person car was employed to evaluate car capacity determination through weight measurement and camera image processing. For this purpose, signals and data from four load cells installed on the car floor were transferred to the computer via the USB-based data acquisition card (MiniLAB 1008) after filtering and amplification processes using an operational amplifier, and numerical values were obtained through the developed software.

Additionally, the image captured by the car-mounted camera was transferred to a computer and calibrated using the CEDES ESPROS software package; the resulting occupancy rate screens provided by the software for one, two, and three passengers are shown in Figure 6.



20 % occupancy rate for 1 person | 31% occupancy rate for 2 persons | 42 % occupancy rate for 3 persons

Figure 6 Occupancy rate for different number of passengers

The weight data, sampled at 200 Hz via a data acquisition card, were passed through an average filter and downsampled to 0.5-second intervals, thereby eliminating undesired artifacts such as noise and oscillations. The camera is fixed at an elevated location to maintain a minimum distance of 2.5 meters from the ground. As an example, the extraction of the passenger metric for three individuals through image processing techniques is presented in Figure 7.



Figure 7 Object detection output for a car with three passengers

The proposed model was trained using a 10,000-image dataset. Initially, 300 custom lift-environment images were uploaded to Roboflow [19], a computer vision platform that provides tools for image annotation, dataset management, preprocessing, augmentation, and dataset export for model training. Using this platform, rotation, noise addition, scaling, and mirroring were applied to expand the dataset to 10,000 samples. The repository was subsequently divided into 75% training, 12.5% testing, and 12.5% validation subsets [20]. The proposed architecture operates on an event-driven basis rather than continuous surveillance, restricting video acquisition solely to critical, restricted, or risk-bearing incidents. Pursuant to data privacy standards, a face-blurring mechanism is applied to support the anonymisation of individuals within the logged footage. Initial model optimisation was executed via an NVIDIA P4000 GPU for 100 cycles; however, the preliminary outcomes proved suboptimal. This performance shortfall necessitated the exploration of alternative repositories hosted on Roboflow, leading to the integration of roughly 10,000 supplementary samples from diverse sources. Subsequently, a second training phase was conducted for the same number of epochs using the augmented dataset. The YOLOv8s architecture demonstrated optimal outcomes upon completing a 100-epoch execution phase. Consequently, four primary evaluation metrics were used to assess the performance of the YOLOv8s architecture.

The compiled repository integrates open-source video sequences with custom-captured photographs, encompassing a diverse spectrum of human postures and locomotor activities such as ambulation, falling, inclination, sitting, and crouching. The procedural workflow for identifying fall incidents is illustrated in Figure 8.

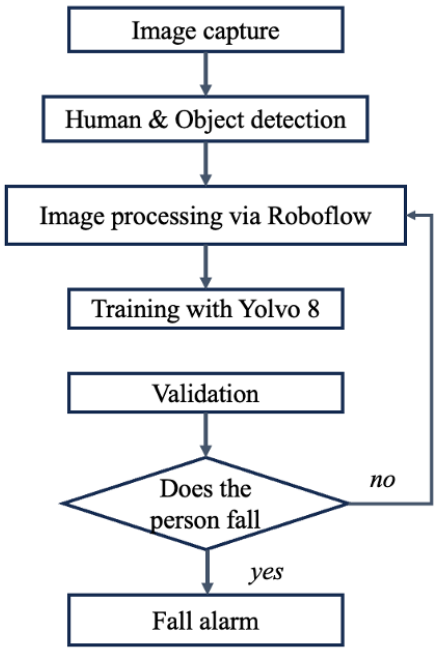


Figure 8 Flowchart of fall detection

These performance indicators confirm that the model is highly effective at recognising human falls. The model demonstrated a continuous decline in both training and validation losses, indicating stable learning. In the event of an anomaly within the developed system's cabin, the lift reroutes to the closest floor and opens its doors at a secure level; meanwhile, the existing emergency communication framework is activated to perform a GSM call.

4 CONCLUSION

Integrated with artificial intelligence-based image processing and object recognition technologies, this smart lift system is developed for high-rise structures – such as commercial and residential buildings operating under heavy traffic – and areas where user safety and comfort are prioritised. By utilising computer vision for lift occupancy estimation, this study addresses the inherent limitations and shortcomings of conventional counting methodologies. Engineered for areas with high-density lift traffic, such as high-rise structures, malls, hotels, hospitals, industrial facilities, and commercial complexes, this smart lift system integrates AI-based computer vision and object recognition technologies, offering a solution with practical applicability.

The practical viability of the proposed system is validated by its fall detection performance (85% precision and 82% recall), indicating that its integration into conventional lift controllers can support faster emergency response. The developed website interface enables the monitoring and tracking of passengers in the car, small children, pets, belongings, and potentially hazardous situations. Future work will focus on integrating the system with building security systems.

REFERENCES

[1] N. V Rajeesh Kumar, G. DhanaSekar, and M. Dennis, “Application of face detection system for passenger counting in lifts using haar features”. *ARPN Journal of Engineering and Applied Sciences*, Vol. 11, No. 13, 8336-8341 (2016).

[2] K. Hashimoto, M. Yoshinomoto, S. Matsueda, Katsuya Morinaka, and N. Yoshiike, “Development of people-counting system with human-information sensor using multi-element pyroelectric infrared array detector”. *Sensors and Actuators A: Physical*, Vol. 58, No. 2, 165-171 (1997). doi.org/10.1016/S0924-4247(96)01400-8.

[3] X. Ma, W. Xi, Z. Chen, H. Hao, and J. Zhao, “ECC: Passenger Counting in the Elevator Using Commodity WiFi”. *Applied Sciences 2022, Vol. 12, Page 7321*, Vol. 12, No. 14, 7321 (2022). doi.org/10.3390/APP12147321.

[4] R. Delussu, L. Putzu, and G. Fumera, “Scene-specific crowd counting using synthetic training images”. *Pattern Recognition*, Vol. 124, 108484 (2022). doi.org/10.1016/j.patcog.2021.108484.

[5] S.F. Lin, J.Y. Chen, H.X. Chao, and others, “Estimation of number of people in crowded scenes using perspective transformation”. *IEEE TRANSACTIONS ON SYSTEMS MAN AND CYBERNETICS PART A-SYSTEMS AND HUMANS*, Vol. 31, No. 6, 645-654 (2001).

[6] J.D. Nichols, L.L. Bailey, A.F. O’Connell Jr., et al., “Multi-scale occupancy estimation and modelling using multiple detection methods”. *Journal of Applied Ecology*, Vol. 45, No. 5, 1321-1329 (2008). doi.org/10.1111/j.1365-2664.2008.01509.x.

[7] Y. Wang, Z. Ma, X. Wei, S. Zheng, Y. Wang, and X. Hong, “ECCNAS: Efficient Crowd Counting Neural Architecture Search”. *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 18, No. 1s, 1-19 (2022). doi.org/10.1145/3465455.

[8] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, Single-image crowd counting via multi-column convolutional neural network. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 589-597 (2016).

[9] H. Fan, H. Zhu, and D. Yuan, "People Counting in Elevator Car Based on Computer Vision". *IOP Conference Series: Earth and Environmental Science*, Vol. 252, 032131 (2019). doi.org/10.1088/1755-1315/252/3/032131.

[10] T.C. Shen, and E.T.H. Chu, "Edge-Computing-Based People-Counting System for Elevators Using MobileNet–Single-Stage Object Detection". *Future Internet*, Vol. 15, No. 10, 337 (2023). doi.org/10.3390/fi15100337.

[11] L.R. Al-Sharif, and G.C. Barney, *The Application of the Inverse S-P Method in Deriving Up Traffic Patterns*. UMIST/Manchester (1991).

[12] M. Seyyedi, A. Candaş, and C.E. İmrak, Artificial Intelligence Embedded Image Process-Based Fatigue Life Determination on Wire Ropes Subjected to Bending Loads. *15th Symposium on Lift and Escalator Technologies, 2024* (2024).

[13] A.J. Schofield, T.J. Stonham, and P.A. Mehta, "Automated people counting to aid lift control". *Automation in Construction*, Vol. 6, No. 5-6, 437-445 (1997). doi.org/10.1016/S0926-5805(97)00022-8.

[14] G. Jocher, A. Chaurasia, and J. Qiu, "YOLO by Ultralytics v8". (2023). doi.org/doi.org/10.5281/zenodo.3908559.

[15] G. Bradski, "The opencv library.". *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, Vol. 25, No. 11, 120-123 (2000).

[16] C.R. Harris, K.J. Millman, S.J. van der Walt, et al., "Array programming with NumPy". *Nature*, Vol. 585, No. 7825, 357-362 (2020). doi.org/10.1038/s41586-020-2649-2.

[17] F. Di Gregorio, D. Varrazzo, and T.P. Team, "psycopg2: Python-PostgreSQL Database Adapter". (2026).

[18] PostgreSQL Global Development Group, "PostgreSQL". (2025).

[19] P. Yozgatli, Y. Acar, M. Tulumen, et al., "Fall Detection in Passenger Elevators using Intelligent Surveillance Camera Systems: An Application with YoloV8 Nano Model". (2025). doi.org/10.32388/YJZ5XW.

BIOGRAPHICAL DETAILS



TEVFİK KELEŞ serves as an R&D Electrical Systems Specialist at Emlak Konut Asansör (EKA). He completed his undergraduate education in the Department of Electronics and Communication Engineering at Doğuş University in 2018. Throughout his career, he has taken part in numerous R&D projects in the fields of automotive, micromobility, and most recently elevator electronics. His area of expertise focuses on transforming elevator systems into more innovative and technologically advanced structures through software, electronics, and artificial intelligence technologies. He also has experience in the development and implementation of R&D projects.



Mehmet SELÇOK serves as the R&D Electrical Systems Manager at Emlak Konut Asansör. He graduated from the Mechatronics Department of Sakarya University and the Business Administration Department of Anadolu University. Throughout his career, he has actively taken part in processes related to elevator control panels, electronic boards, safety systems, testing, and commissioning. He served as Production Manager until 2024, after which he began leading the Electrical Systems R&D team at EKA. He currently continues his work in the fields of elevator electronics, safety applications, system integration, and innovative engineering solutions.



Dr. Eren KALAY is a research assistant at Istanbul Technical University (ITU), specializing in mechanical engineering. He received his B.Sc., M.Sc., and Ph.D. degrees in Mechanical Engineering from ITU in 2014, 2017, and 2023, respectively. His research interests include computer-aided engineering, the Discrete Element Method, materials handling, and artificial neural networks.



Dr. Adem CANDAŞ is an Assistant Professor at Istanbul Technical University (ITU). He received his B.Sc., M.Sc., and Ph.D. degrees in Mechanical Engineering from ITU in 2010, 2013, and 2021, respectively. His research interests include fracture mechanics, computer-aided engineering, numerical analysis, materials handling systems, and lift systems.



C. Erdem İMRAK is a full-time Professor at Istanbul Technical University (ITU). Professor Imrak received B.Sc., M.Sc., and Ph.D. degrees in Mechanical Engineering from ITU in 1990, 1992, and 1996 respectively. He has carried out research into computer-aided engineering, CAD/CAM, numerical analysis, materials handling, and especially lift systems. Currently, his activities include an Honorary Member of ASYAD; a Member of the Safety, Education & Training Committee of ELA.