Journal of Electronic Research and Application, 2026, Volume 10, Issue 4
http://ojs.bbwpublisher.com/index.php/JERA
Online ISSN: 2208-3510
Print ISSN: 2208-3502

A Multimodal Sensor Fusion Approach for Real-Time Detection of Electric Vehicles in Elevators

Fenglin Zhang*

College of Emergency Management and Safety Engineering, North China University of Science and Technology, China

Corresponding author: Fenglin Zhang, 18230170785@163.com

Abstract: This paper presents a multimodal sensing-based system designed to improve the accuracy and real-time performance of detecting electric vehicles (EVs) attempting to enter elevators. The system architecture integrates a data acquisition layer, an analysis and processing layer, an intelligent decision layer, and an application layer. By employing sensors such as flexible tactile arrays and visual cameras, the system captures multimodal data. Detection results are generated through a global scene context information extraction module and a local spatio-temporal feature extraction module. Experimental outcomes indicate that the proposed system achieves a mean average precision (mAP) of 95.60%, operates at 31.5 frames per second (FPS), and maintains a model size of 26 MB. Under occluded conditions, it attains average mAP@.5 and mAP@.5:.95 scores of 91.2% and 80.9%, respectively, surpassing other existing detection methods. The results demonstrate that this multimodal sensing system can effectively mitigate safety risks associated with EVs entering elevators, offering high practical utility and reliability.

Keywords: Electric vehicle prohibited from entering the elevator; Data acquisition; Intelligent decision; Multimodal sensors; Local scene

Online publication: May 21, 2026

1. Introduction

E-bikes are popular due to their convenience, but battery quality issues and improper maintenance lead to frequent spontaneous combustion accidents [1]. Transporting e-bikes into buildings via elevators for home charging further increases fire hazards [2–4]. Therefore, effective automatic detection technology is needed to monitor e-bikes entering elevators, ensure safe elevator operation, and protect lives and property [5,6].

This paper proposes an intelligent detection system based on multimodal sensing, fusing data from image, sound, temperature, and tactile sensors to overcome single-sensor limitations. The system extracts global scene context and local spatio-temporal features, improving detection accuracy, reducing model overhead, and enhancing real-time performance. It effectively addresses the safety risks of EVs entering elevators.

2. Literature review

Manual elevator monitoring is labor-intensive, costly, and prone to fatigue-related errors. Automated methods have been explored. He proposed YOLOv5s-TBCA for electric motorcycle detection, which is fast but degrades under occlusion [7]. Zhang et al. developed a lightweight YOLOv5s-based method using MobileNetV2 and CBAM, but it has dataset limitations [8]. Su *et al.* presented a YOLOv5s-based lightweight detector for e-bikes, yet it was only validated on a self-constructed dataset [9]. Li *et al.* introduced YOLOv3-G with GhostNet to reduce parameters, suitable for edge devices but not general scenarios [10].

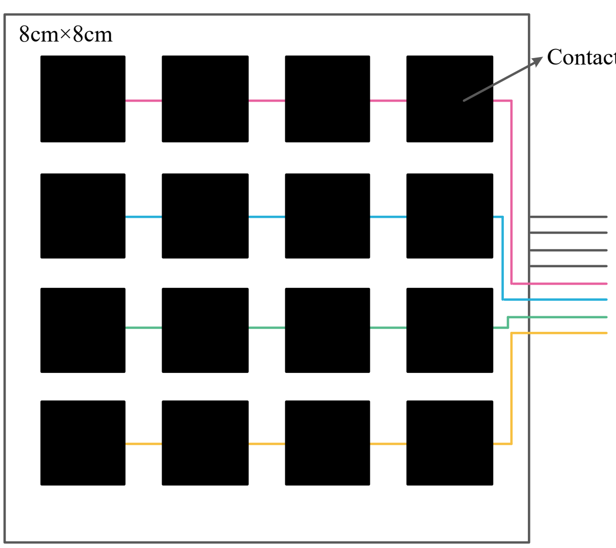
In summary, existing methods still face challenges: insufficient accuracy under occlusion, lack of real-world image complexity, and high computational overhead hindering real-time deployment. This paper aims to address these key issues.

3. Multimodal sensing design

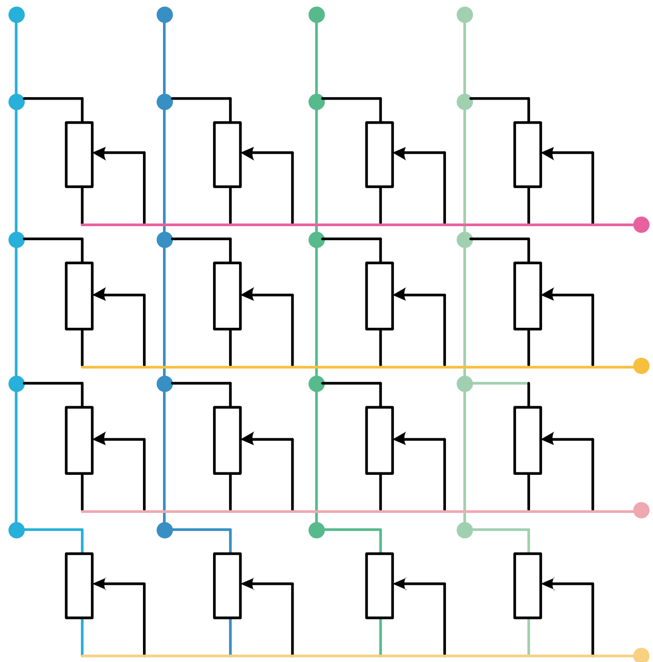
3.1. Sensor principle

Currently, there are a wide variety of sensors, and their outputs can break through the limitations of a single sensor to form a more comprehensive and three-dimensional environmental awareness [11,12]. Therefore, in the sensor perception layer designed in this paper, the flexible tactile sensors used are 4 × 4 flexible piezoresistive array tactile sensors with strong conductivity, with a high degree of softness of the canvas, in which the array distributed flexible pressure sensors are made by precision printing technology, transferring nano force-sensitive materials and silver paste, etc. to the flexible film substrate, and then drying and curing, and the array distributed flexible pressure sensors are produced As shown in **Figure 1**.

Multimodal sensor array structure as shown in **Figure 1(a)**, when there is an object in the range of this multimodal sensor, the resistance at the contact will become smaller, thus generating a voltage change, the A/D chip can read the analog voltage value to get the size of the pressure change. The piezoresistive tactile sensor has a range of 1–50kg for a single contact, a static resistance greater than 1MQ, very little hysteresis and drift, and no electrostatic discharge or electromagnetic interference.



(a) Multimodal sensor array structure.



(b) Equivalent circuit of multimodal sensor.

Figure 1. Array distributed flexible pressure sensor.

3.2. Sensor design

Distributed flexible tactile sensing arrays are utilized to obtain deformation information about the shape of the contact area, this scheme is based on the fact that multimodal sensing arrays can recognize the depth information and shape information of the object pressed in the contact area and can capture the deformation information generated by the object with a camera to get the geometric shape of the object [13]. The designed sensor consists of a camera, a support frame and tricolor LED beads consisting of red, green and blue, as well as a sensing layer with embedded markers and peripheral circuitry [14].

The conceptual design of the sensor is shown in **Figure 2** and consists of camera, support frame, LED beads and perception layer as the hardware part.

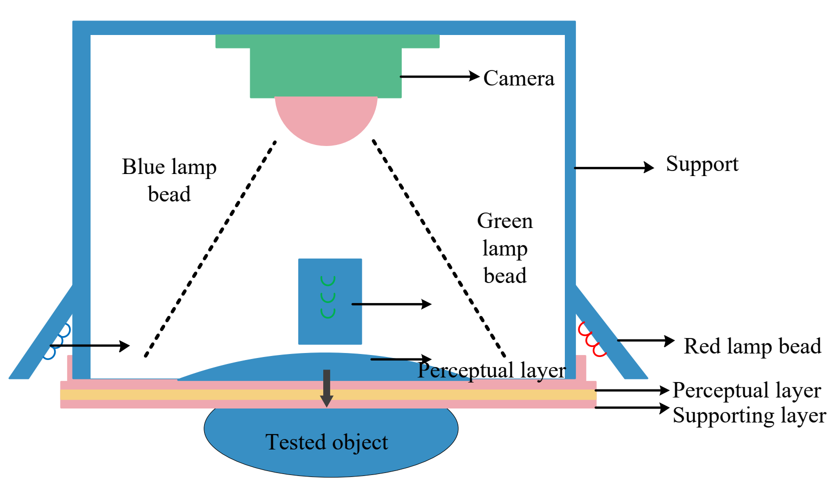


Figure 2. Sensor conceptual design.

The sensing layer consists of 2 layers of canvas laminated multimodal sensing arrays, **Figure 3** shows the structure of the sensing layer, fixed in the designed square frame, and on one side there is a dense array of 2D markers, which can better reflect the shape of the connected object. The fisheye camera is used for taking pictures, which has a wider shooting angle and increases the sensing area, and the LED beads of various colors inside the sensor are distributed on the three-side bracket at 60°, tilted to the surface of the sensing layer, which is used to illuminate the whole sensing surface.

Figure 3. Perception layer structure.

4. Design of intelligent detection system for electric vehicle elevator prohibition

4.1. Elevator entry detection system architecture system

The architecture of the wheat field weed detection system based on multimodal information fusion includes the architecture of the elevator no-entry detection system, as shown in Fig**ure 4**. The elevator linkage control layer can feedback the intelligent determination results to the system, thus realizing the integrated function of EV detection, warning and forbidden entry control [15].

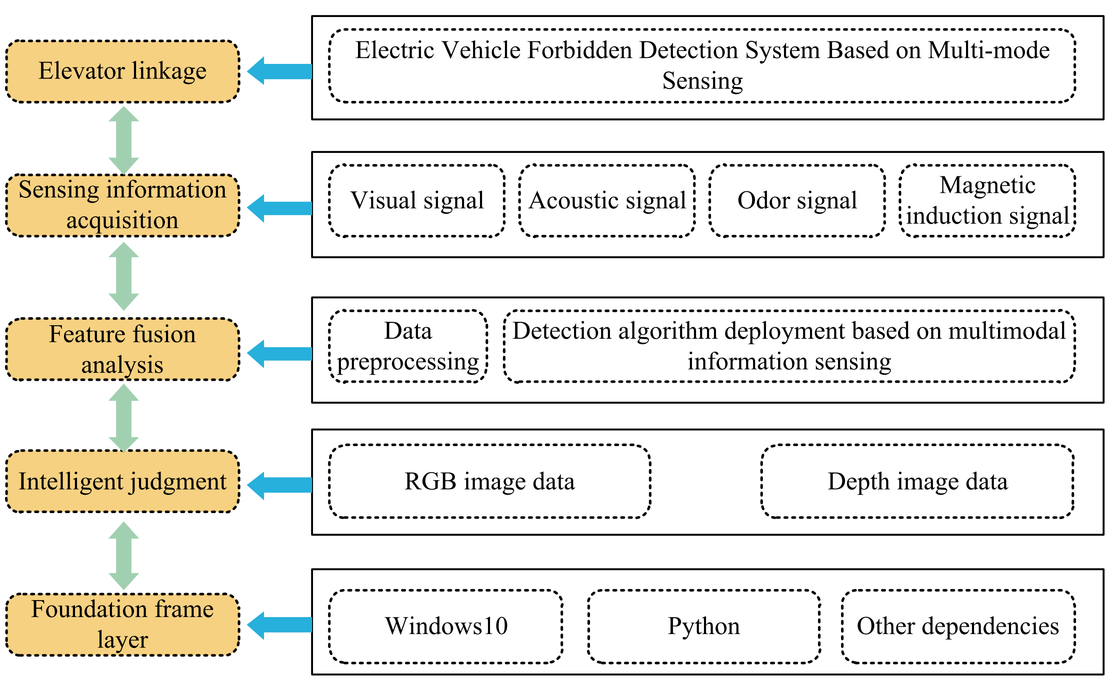


Figure 4. Elevator entry detection system architecture.

4.2. Intelligent detection model construction

4.2.1. Problem description

In this paper, we define the forbidden entry of an EV in an elevator as an optimization problem associated with multiple sources of modal information inputs, i.e., $P\left(A_i^{t+n} \middle| C_{li}, P_i, L_i, C_g, S\right)$, which is measured by audiovisual and haptic multimodal responses and is rescaled to the range of [0,1], so that for each target EV i, given a sequence of video frames observed from elevator monitoring for m time-steps and the relevant information about the EV, based on the multimodal sensors it is possible to identify the out that the EV $t+n$ moment enters the elevator [16–18].

The information sources used to detect the EV entering the elevator include, the positional trajectory information $B=\left\{b_i^{t-m}, b_i^{t-m+1}, \cdots, b_i^t\right\}$ of the target object EV i represented by the 2D bounding box of the elevator, the attitude key point information $P_i=\left\{p_i^{t-m}, p_i^{t-m+1}, \cdots, p_i^t\right\}$ of the target EV i, the local environment context information $C_{li}=\left\{c_{li}^{t-m}, c_{li}^{t-m+1}, \cdots, c_{li}^t\right\}$ around the target EV, the global environment context interaction information $C_g=\left\{c^{t-m}, c^{t-m+1}, \cdots, c^t\right\}$, and the multimodal information $S=\left\{s^{t-m}, s^{t-m+1}, \cdots, s^t\right\}$ of the target EV, where m is the history of the observed time.

4.2.2. Equipment calibration

The data collected by surveillance cameras and multimodal sensing are highly complementary, and fusion of the two data can not only overcome the deficiencies of surveillance cameras in environmental sensing, but also obtain richer information about the observed targets.

4.3. Evaluation indicators

When comparing the detection results of different intelligent detection systems, the commonly used average detection accuracy (AP) is used as an indicator to measure the detection accuracy, which indicates the detection accuracy of each category, and AP can be a better measure of the model's detection ability.

5. Analysis of test results

5.1. Dataset construction and preprocessing

In order to train and test the intelligent detection system, it is necessary to construct a dataset containing images of electric cars in an elevator. In this paper, the dataset is constructed from the following two ways respectively:

(1) Field shooting and network video crawling, a large number of videos of scenes inside elevators in different time periods and under different lighting conditions are selected and image frames are intercepted from them, totaling 1500 frames;

(2) Public dataset PaddlePaddle EV detection dataset totaling 2000 sheets. The public dataset is labeled using Labelme, so the relevant image frames obtained from field shooting and web crawling are also labeled using Labelme, and the labeling information includes target category and bounding box. By the above method, a dataset applicable to the environment inside an elevator was constructed.

In the data preprocessing stage, this paper adopts data enhancement techniques. The above data enhancement techniques can expand the data set and improve the data processing capability of the detection system.

Meanwhile, occlusion is an important factor that affects the accuracy of target detection. In addition to device calibration, combined with border information, image information, key point information, and global, this paper also simulates the occlusion scenario through the above Mixup and Cutout data enhancement techniques, so as to improve the recognition accuracy of the occlusion situation inside the elevator.

Hardware configuration includes a USB camera, Knowles microphone array, MQ-138 odor sensor array, Hall magnetic array, and Jetson Nano processing unit.

5.2. Comparison of testing performance

5.2.1. Performance metrics for different algorithms

In order to further confirm the advantages of this paper's multimodal sensing over the current mainstream algorithms in the detection of EVs prohibited from entering the elevator, the performance of this paper will be compared by comparing the performance of the frames per second (FPS), the system overhead, the mAP, and the AP values of each category, on the visual-only detection (YOLOv8), the vision + acoustic fusion, the vision + acoustic + scent fusion, and the multimodal four-target detection system. The performance metrics of different algorithms are shown in **Table 1** and the results are compared:

Table 1. Performance indicators of different algorithms

Detection method	mAP (%)	AP (%)	FPS	Model overhead (MB)
YOLOv8	86.8	88.5	27.4	604
Vision + Acoustic Fusion	77.6	80.2	26.2	105
Vision + Acoustic + Smell Fusion	89.5	81.9	30.8	302
Multimodal Sensing	95.6	94.5	31.5	26

5.2.2. Detection and identification results

In order to test the feasibility and stability of the multimodal sensing-based electric vehicle elevator access prohibition system, a large number of system experimental tests have been conducted on this access prohibition system, which include field shooting and web video crawling, and public datasets. A total of 3500 no-entry targets were tested, of which 1200 no-entry targets were electric bicycles, 500 no-entry targets were motorcycles, and the rest were normal images, and the detection and recognition results are shown in **Table 2**. The correct rate of its recognition is 94.6%, the false detection rate is 1.37%, and the leakage rate is 4.03%, which fully proves the feasibility and stability of the two-wheeled vehicle elevator prohibited entry system in detecting and recognizing in different environments through the test.

Table 2. Detection and recognition results

Detection target category	Accuracy	False detection rate	Missed detection rate
Electric bicycles	94.17	1.67	4.16
Motorcycle	94.00	1.20	4.80

6. Conclusion

In this study, a multimodal sensing-based EV entry detection system in elevator is developed, visualization and interaction layer, and application layer, and improves the detection accuracy while ensuring compliance with the real-time requirements of the project. The experimental results show that the system achieves 94.5%

AP and 95.60% mAP for EV forbidden elevator detection and identification, which are better than other detection algorithms. In addition, it is verified that the intelligent detection system based on multimodal sensing has a certain stability, and the obtained results can be combined with the elevator control system to reduce the occurrence of EV illegal entry. In future research, combining multi-sensors will be considered and methods such as utilizing impulse neural networks will be explored to further improve the computational efficiency and accuracy of the model.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Liao F, Correia G, 2022, Electric Carsharing and Micromobility: A Literature Review on their Usage Pattern, Demand, and Potential Impacts. International Journal of Sustainable Transportation, 16(3): 269–286.

[2] Liu W, Liu H, Liu W, et al., 2021, Life Cycle Assessment of Power Batteries used in Electric Bicycles in China. Renewable and Sustainable Energy Reviews, 2021(139): 110596.

[3] Li Y, Chen Q, Ma Q, et al., 2022, Injuries and Risk Factors associated with Bicycle and Electric Bike use In China: A Systematic Review and Meta-Analysis. Safety Science, 2022(152): 105769.

[4] Wang L, Wu J, Liu M, et al., 2021, Sociotechnical View of Electric Bike Issues in China: Structured Review and Analysis Of Electric Bike Collisions Using Rasmussen's Risk Management Framework. Human Factors and Ergonomics in Manufacturing & Service Industries, 31(6): 625–636.

[5] Zhao Z, Li S, Wu C, et al., 2023, Research on the Rapid Recognition Method of Electric Bicycles in Elevators based on Machine Vision. Sustainability, 15(18): 13550.

[6] Cao F, Sheng G, Feng Y, 2024, Detection Dataset of Electric Bicycles for Lift Control. Alexandria Engineering Journal, 2024(105): 736–742.

[7] He Y, 2024, Automatic Detection of Electric Motorcycle Based on Improved Yolov5s Network. Journal of Electrical and Computer Engineering, 2024(1): 4889707.

[8] Zhang Z, Yang X, Wu C, 2024, An Improved Lightweight YOLOv5s-Based Method for Detecting Electric Bicycles in Elevators. Electronics, 13(13): 2660.

[9] Su J, Yang M, Tang X, 2024, Integration of ShuffleNet V2 and YOLOv5s Networks for a Lightweight Object Detection Model of Electric Bikes Within Elevators. Electronics, 13(2): 394.

[10] Li Y, Zhang M, Wang J, et al., 2024, Elevator E-Bike Detection based on Improved YOLOv3. International Core Journal of Engineering, 10(5): 47–51.

[11] Duan S, Shi Q, Wu J, 2022, Multimodal Sensors and ML-Based Data Fusion for Advanced Robots. Advanced Intelligent Systems, 4(12): 2200213.

[12] Kullu O, Cinar E, 2022, A Deep-Learning-Based Multi-Modal Sensor Fusion Approach for Detection of Equipment Faults. Machines, 10(11): 1105.

[13] Qiu Y, Wang Z, Zhu P, et al., 2023, A Multisensory-Feedback Tactile Glove with Dense Coverage of Sensing Arrays for Object Recognition. Chemical Engineering Journal, 2023(455): 140890.

[14] Li S, Gao X, Xie Z, 2023, Binocular Underwater Measurement with Multicolor Structured Light. IEEE Journal of Oceanic Engineering, 49(2): 649–666.

[15] Lai S, Yang M, Wang R, et al., 2022, Remote-control System for Elevator with Sensor Technology. Sensors & Materials, 2022(34).

[16] Gong C, Liu J, Han Y, et al., 2021, Safety of Electric Vehicles in Crash Conditions: A Review of Hazards to Occupants, Regulatory Activities, and Technical Support. IEEE Transactions on Transportation Electrification, 8(3): 3870–3883.

[17] Vlachou E, Vlachou V, Efstathiou D, et al., 2024, Overview of IoT Security Challenges and Sensors Specifications in PMSM for Elevator Applications. Machines, 12(12): 839.

[18] Zarikas V, Zholdasbayeva M, Mitra A, 2023, Statistical and AI modeling of UK and France Elevator Accidents and their Violating Safety Rules. Journal of Applied Engineering Science, 21(4): 999–1015.