# Elevator E-Bike Detection based on Improved YOLOv3

Yichen Li, Meina Zhang[a], Jiayu Wang, Yutong Xing, Yuxuan Liu, Xinyuan Wang

School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan 114001, China

[a]zhangmeina@163.com

## Abstract

**Given the high parameter count of existing electric bike detection models in elevators, which complicates deployment on edge devices with limited computational capacity, this paper proposes an improved electric bike detection algorithm for elevators based on YOLOv3. By substituting the original network architecture with the lightweight GhostNet as the backbone for YOLOv3, we reduce both the model's computational load and parameter count, while minimally impacting detection accuracy. The enhanced YOLOv3-G model achieves comparable detection accuracy to the original YOLOv3, with a 47.5% reduction in parameters and achieving approximately 0.047 seconds per frame in detection speed.**

## Keywords

**E-Bike Detection; Yolov3; GhostNet.**

## 1. Introduction

Amid the advancements toward smart cities, smart elevators stand out for their potential to significantly enhance operational efficiency and safety[4]. Central to realizing this potential are the intelligent detection systems powered by edge computing. However, traditional object detection algorithms, including the widely acclaimed YOLOv3, often exceed the computational limits of edge devices used in smart elevators due to their demanding resource requirements[2].

Addressing this challenge, our research focuses on optimizing YOLOv3 for deployment within smart elevators. By replacing its backbone network, Darknet-53, with a more efficient GhostNet architecture, we aim to significantly reduce the model's computational footprint. This modification not only ensures compatibility with edge computing devices but also maintains high detection accuracy essential for elevator safety.

## 2. Introduction to Model

### 2.1 Yolov3

YOLOv3, an acronym for "You Only Look Once version 3", is a pivotal advancement in real-time object detection technology, proposed by Joseph Redmon and Ali Farhadi in 2018[3]. This model is distinguished by its unique approach to object detection, performing predictions directly from full images in one evaluation, which significantly boosts processing speed. At the heart of YOLOv3 lies the Darknet-53 network, serving as its backbone. Darknet-53 is composed of 53 convolutional layers, enhanced with residual connections to facilitate deep network training without succumbing to the vanishing gradient problem. This architecture enables YOLOv3 to accurately detect objects across different scales within an image, marking a significant leap forward in the efficiency and effectiveness of object detection systems.

### 2.2 GhostNet

GhostNet, proposed by Huawei's Noah's Ark Lab[4], is a lightweight neural network that offers a more streamlined architecture compared to Darknet-53 without compromising accuracy. The Ghost module is primarily divided into three parts: the first part consists of 1x1 convolutions; the second part involves depth-wise convolutions; and the third part concatenates the compressed feature maps with redundant features. The Ghost Bottleneck module, a fundamental component of GhostNet, comprises two sections. The first section serves as the main body, while the second section acts as a residual edge responsible for transmitting feature information and enhancing network performance. Figure 1 displays the structure of the Ghost Bottleneck.

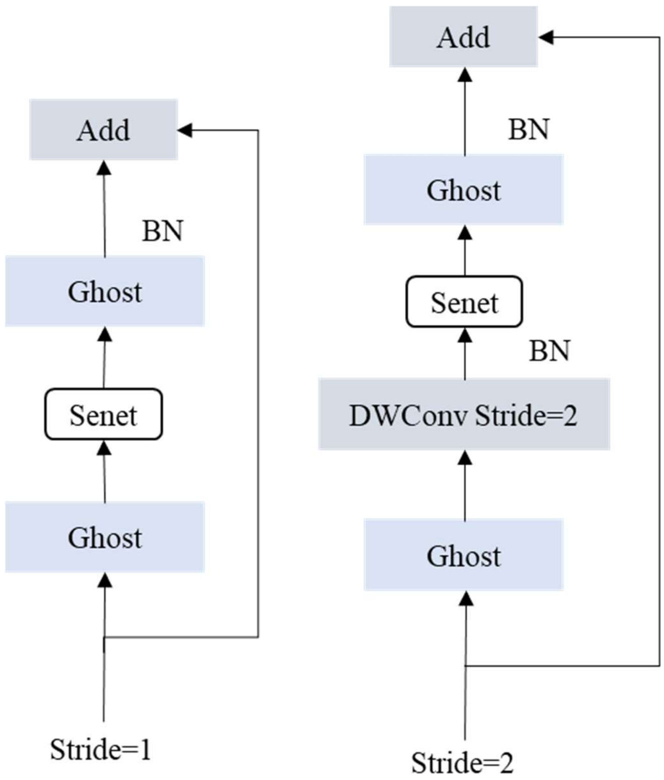


**Figure 1.** The structure of the ghost bottleneck

### 2.3 Improved Network

This paper improves upon the YOLOv3 detection model by employing GhostNet as the backbone network for feature extraction, which reduces the model's parameter count and computational load while minimizing the loss in model accuracy. The enhanced YOLOv3 detection model is named YOLOv3-G.

## 3. Experimental Result and Analysis

### 3.1 Data Collection and Preprocessing

This project employs methods such as online searches, web scraping, and video segmentation to collect images of electric bikes inside elevators, resulting in a curated selection of 1500 images. To expand the dataset size and enrich the image features, three data augmentation techniques were applied: Gaussian blur, random rotation, and horizontal flipping. The augmented dataset consists of 3500 images in total, with 2800 designated for the training set and 700 for the validation set. The effects of data augmentation are illustrated in the figure below.

ISSN: 2414-1895
DOI: 10.6919/ICJE.202405_10(5).0008

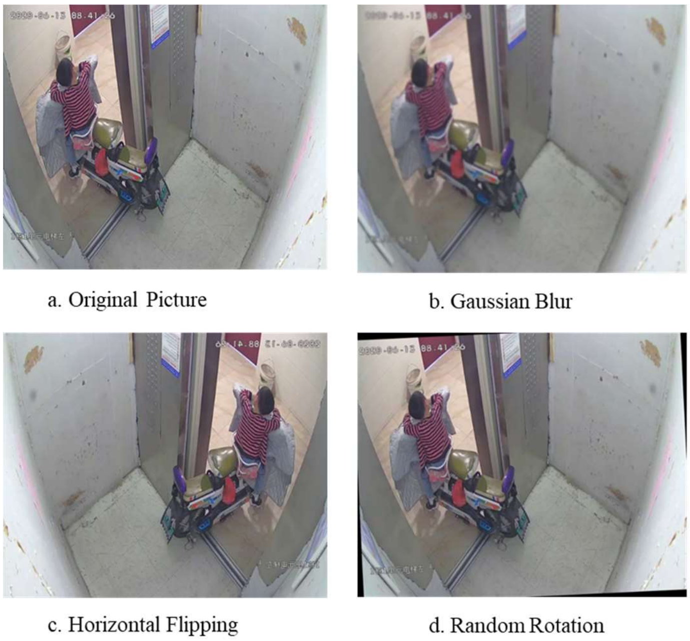


**Figure 2.** Data augmentation

For the collected data, a portion was manually labeled using the labelImg annotation tool, followed by automatic labeling of data using a script. The principle behind automatic labeling involves reading each image and performing object detection with the YOLOv3 model, resulting in a list of outcomes that include coordinates of the bounding boxes and class information. The create_tree function is called to generate the basic structure of an XML file, including details such as filename and dimensions. The pretty_xml function is used to refine the XML structure, making its output more readable. Finally, the entire XML tree structure is written into the corresponding XML file, ensuring each image is matched with its coordinate document. The visualization results of the dataset are shown in Figure 3. The top-left image displays the distribution of class categories, the top-right image shows the distribution of bounding boxes, the bottom-left image illustrates the distribution of bounding box centroids, and the bottom-right image presents the size distribution of the dataset.

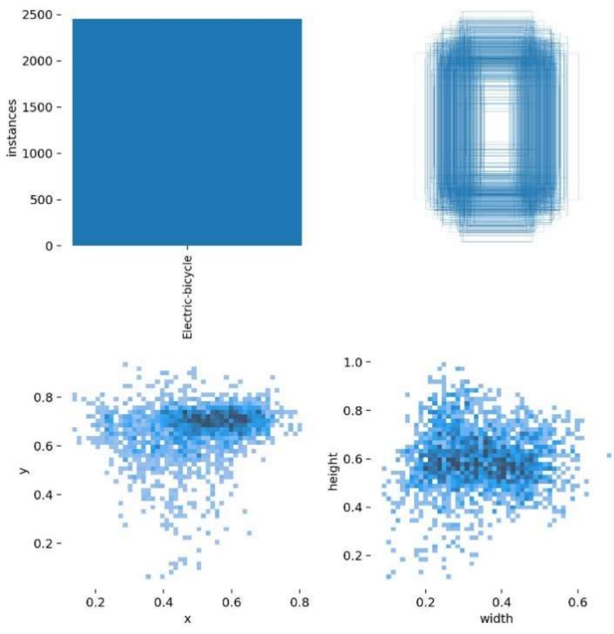


**Figure 3.** Visualization results of the dataset

ISSN: 2414-1895 DOI: 10.6919/ICJE.202405_10(5).0008

### 3.2 Experimental Environment and Parameter Setting

The experimental environment for this study is detailed as follows: The experiments were conducted using PyTorch version 2.0.0 on a system running Python 3.8 under the Ubuntu 20.04 operating system. The hardware setup included an NVIDIA RTX 4090 GPU with 24GB of memory and a single instance of a 12 virtual CPU Intel(R) Xeon(R) Platinum 8352V CPU operating at 2.10GHz for computation.

For this experiment, the network input was set to 640x640, utilizing the ADAM optimizer and incorporating transfer learning to obtain pretrained weights. The initial learning rate of the model was set to 0.001, with a batch size of 16 and a total of 200 iterations.

### 3.3 Experimental Result

The performance of the improved YOLOv3 is depicted in the figure, achieving a mAP@0.5 of 0.95 and a mAP@0.5:0.95 of 0.69. During the inference process, the speed of processing video frames is approximately 0.047 seconds, demonstrating that the algorithm has been significantly improved in terms of detection speed while maintaining detection accuracy, making it suitable for deployment on edge devices. In summary, the YOLOv3-G model balances inference speed and accuracy well, making it suitable for detecting electric bikes inside elevators. Figure 4 shows the inference results of YOLOv3-G; it can be observed that the algorithm effectively identifies electric bikes, validating the effectiveness of the improvements made in this paper.


**Figure 4.** The results of detection

## 4. Conclusion

YOLOv3-G represents a pivotal advancement in deploying efficient real-time object detection within smart elevators, leveraging the lightweight GhostNet architecture to meet the challenges of edge computing[5] limitations. Through targeted data augmentation and rigorous testing on advanced hardware, the model demonstrates enhanced performance and speed, marking it as an effective tool for improving safety and operational efficiency in smart urban infrastructures.

## Acknowledgments

Fund:2024 Innovation and entrepreneurship training program for College Students.

## References

[1] Xie Yinlin. Research on improvement of elevator monitoring system under the background of artificial intelligence [J]. Smart City,2023,9(01):23 -25.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202405_10(5).0008

[2] Huang Peng, Fang Zhiming, Zhu Man, et al. Research on Detection of Electric Bicycle in Elevator based on SSD Network [J]. China Work Safety Science and Technology,2023,19(02):167-172.

[3] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. arXiv preprint arXiv: 1804.02767, 2018.

[4] Han K, Wang Y, Tian Q, et al. GhostNet: More features from cheap operations. //2020 IEEE/C-VF Conference on Computer Vision and Pattern Recogniton(CVPR)[C]. IEEE,2020:1577-1586.

[5] Zhang Yan, Hu Mingxing. Edge under the calculation of the learning challenges and progress [J]. Journal of digital technology and applications, 2023, 9 (10) : 160-162.